Combining sources of information to increase survey response rates

Combinando fuentes de información para aumentar las tasas de respuesta a la encuesta

结合信息源，提高调查回复率

Miguel Guinalíu
Department of Marketing Management and Marketing Research,
Faculty of Economics and Business, University of Zaragoza, Zaragoza, Spain, and

Vidal Díaz de Rada
Faculty of Human, Social and Educational Sciences, Public University of Navarre,
Pamplona, Spain

Abstract

Purpose – The purpose of this paper is to show that mixed methods applied sequentially provide sufficient knowledge of topics under study.

Design/methodology/approach – This paper conducted an analysis of a real case using descriptive statistical and regression techniques.

Findings – The present study proposes the use of the so-called "sequential mixed-modes" in survey-based market research. This technique is based on the successive application of different information collection techniques (face-to-face, telephone and internet-based surveys); it offers four fundamental advantages: increased coverage rate; higher response rate; lower costs; and greater cooperation.

Research limitations/implications – In addition to the normal limitations associated with conclusions derived from case studies, the data collection was carried out in only one autonomous community (Spain) and focussed only on one theme.

Practical implications – Firstly, it was found that data collection through mixed sequential modes substantially increases response rates in online surveys. This is particularly important as, in recent years, data collection through the internet has become almost standard market research practice. Secondly, the sample that resulted from the joint use of the three data collection modes more accurately reflected the distinctive features of the universe under study. A third recommendation is that the use of internet-based self-administered surveys is especially promising with individuals with a higher level of education and among students.

Originality/value – The decrease in the response rate is one of the greatest challenges of survey-based market research, given its impact on sample representativeness. This paper addresses this problem and exposes the advantages of the sequential use of mixed modes in the collection of information.

Keywords Online surveys, Face-to-face surveys, Self-administered surveys,
Sequential mixed-modes, Telephone surveys

Paper type Research paper

Resumen

Objetivo – El objetivo de la presente investigación es mostrar que los modos mixtos aplicados de manera secuencial ofrecen un conocimiento adecuado de los temas objeto de estudio.

Metodología – Se realiza el análisis de un caso real mediante técnicas descriptivas y regresión.

Resultados – Este trabajo presenta los denominados "modos mixtos secuenciales" en la investigación a través de encuestas. Esta técnica se basa en la aplicación sucesiva de diferentes herramientas de recogida de información (encuestas presenciales, telefónicas y por Internet), y ofrece cuatro ventajas fundamentales; (1) aumento de la tasa de cobertura; (2) mayor tasa de respuesta; (3) menores costes; y (4) mayor cooperación.

Limitaciones – Además de las limitaciones normalmente asociadas a los estudios de caso, la recogida de información se redujo a una única comunidad autónoma (España) y un único tema.

Implicaciones practicas – En primer lugar, se observa que la recogida de información mediante modos mixtos secuenciales aumenta de manera sustancial la tasa de respuesta en las encuestas online. Esto es particularmente relevante en la situación actual, pues la utilización de Internet casi se ha convertido en un estándar en la investigación de mercados. En segundo lugar, la muestra final obtenida como resultado del uso de tres modos de recogida refleja de manera más ajustada las características distintivas de la población objeto de estudio. Una tercera implicación es que el uso de encuestas online de carácter autoadministrado es especialmente prometedor en individuos de mayor nivel educativo y entre estudiantes.

Palabras clave Modos mixtos secuenciales, Encuestas presenciales, Encuestas online, Encuestas telefónicas

Tipo de artículo Trabajo de investigacion

摘要

目的 – 本文的目的是要证明，按顺序应用的混合方法为研究的主题提供了充分的知识。

文章设计/研究方法 – 本文运用描述性统计和回归技术对一个实际案例进行了分析。

研究结果 – 本文提议使用"顺序混合模式"来调查市场研究。这项技术是基于连续运用不同的信息收集技术(面对面、电话和基于互联网的调查)，它提供了四个基础性优势: 增加覆盖率; 更高的回复率; 降低成本; 和更大的协作。

研究局限性/意义 – 除了常见的案例研究所存在的局限性之外，本文的数据收集仅在一个自治社区(西班牙)进行，并且只集中于一个主题。

实际意义 – 首先，通过混合序列模式收集数据，大大提高了在线调查的回复率。这一点尤其重要，因为近年来，通过互联网收集数据几乎已成为标准的市场研究实践。其次，联合使用三种数据采集模式所得到的样本更准确地反映了所研究社区的鲜明特征。第三个建议是，使用基于互联网的自我管理调查对调查教育水平较高的个人和学生群体具有很大的调查前景。

本文独创性/价值 – 回复率的下降是基于调查的市场研究面临的最大挑战之一，因为它大大影响了样本代表性。本文解决了这一问题，并揭示了在信息收集中顺序使用混合模式的优点。

关键词 – 网络调查，面对面调查，自我管理调查，顺序混合模式，电话调查

1. Introduction

From the end of the past century, to the present day, survey-based marketing research has had to contend with an ever-increasing rate of non-response from individuals approached to take part in scientific studies. This phenomenon has various causes. Firstly, in the case of face-to-face surveys, some authors have pointed to the difficulty of contacting the universe under study in their homes owing to the existence of a large number of empty properties, and the growing number of residential areas with restricted access and security measures (Massey and Tourangeau, 2013). Secondly, telephone surveys present problems based on:

- the high number of unlisted telephone numbers (not included in telephone directories);
- telephone numbers that are never answered;
- the use of devices to screen calls (answering machines and incoming caller ID); and
- the growing substitution of fixed telephones for mobiles, which in most countries are not listed in directories (Callegaro *et al.*, 2010; Lee *et al.*, 2010).

As a consequence of these limiting factors, self-administered surveys have experienced notable growth (see, among others, Díaz de Rada *et al.*, 2019), despite serious doubts about their representativeness with non-probabilistic samples (Dutwin and Buskirk, 2017; Mercer *et al.*, 2017).

Many studies have shown decreases in response rates in different geographical areas and periods of time (Díaz de Rada *et al.*, 2019). For example, the response rate of the European Social Survey revealed that, despite an increase in responses in the first three rounds (Stoop *et al.*, 2010), an analysis of the years between 2002 and 2014 showed a decreasing trend in response rates (Beullens *et al.*, 2018). A similar trend was observed in North America where, between 1996 and 2007, the non-response rate decreased annually by 0.65% in face-to-face surveys and by 2% in telephone surveys (Brick and Williams, 2013). Brick and Williams (2013) showed that the decline was even greater from 2005, with four surveys showing annual decreases of more than 1%, one of them reaching 1.72% (La Encuesta Nacional de Salud – *National Health Interview Survey* – NHIS).

The present study tries to solve the problem of the lack of representativeness caused by reduced response rates through an innovative proposal based on the use of research designs combining various survey modes. This strategy can reduce the uncovered population and increase response rates at low cost (De Leeuw, 2008; Dillman *et al.*, 2014; De Leeuw, 2018). Reduced costs help researchers use specific strategies to increase the cooperation of participants selected (National Research Council, 2013). Thus, this study proposes the use of several survey modes, internationally known as "mixed-modes" (De Leeuw, 2005). This methodology is based on combining the positive aspects of various data collection methods. More specifically, this research discusses the advantages of the so-called "sequential mixed-modes".

In short, the objective of this work is to show that sequentially administered mixed-modes provide sufficient knowledge of topics under study because, among other reasons, the modes are adapted to the characteristics of the different groups of individuals surveyed[1].

The remainder of the study is structured as follows. Firstly, the different mixed-modes are introduced; secondly, the possible combinations of information collection modes are presented. Next, the benefits offered by sequential mixed-modes are discussed. Thereafter, a practical application is presented that empirically demonstrates mixed-modes' advantages over other methodologies. The work ends by synthesising the main conclusions and various implications for market research.

2. Introduction to mixed-modes in survey research

Although the first research uses of various modes took place in the late 1960s (Hoschstim, 1967), it was not until the present century that one could really speak of the use of mixed-modes (De Leeuw, 2005, 2018; Dillman *et al.*, 2014; Sala and Lynn, 2009). The mixed-mode

strategy seeks to take advantage of the positive aspects of the different survey modes, by combining them, to improve the results obtained. Normally, mixed-modes are applied on time-based criteria and on the basis of their level of importance (De Leeuw, 2008, 2018):

- Simultaneous sampling (concurrent design): The different sample segments respond to the survey using one of various modes (face-to-face, telephone, post, Web). Either the interviewees choose the method that suits them best, or the interviewer decides the most appropriate way to collect the information from each segment (Díaz de Rada *et al.*, 2019). For example, different modes might be used for different questions, which eliminates the effect of the interviewer on the interviewee when (s) he is faced with more sensitive topics. This is how the National Institute of Statistics (hereinafter the INE) conducted its Health and Sexual Habits Survey; the "interviewer" figure explained the background to the survey and guaranteed confidentiality, and the interviewee answered the questionnaire on a self-administered basis (Instituto Nacional de Estadística, 2003: p. 3).
- Sequential modality application (sequential): the modes are presented in a sequence (Díaz de Rada and Palacios Gómez, 2013). For example, an online survey might first be used. Thereafter, face-to-face or postal surveys can be used to reach the offline population (Díaz de Rada *et al.*, 2019). This methodology enables response rates to be increased while containing costs.
- Use of a main modality with other, complementary modes. For example, the Andalusian Public Opinion Barometer changed its data collection method from face-to-face to telephone surveys in areas where it detected significant decreases in interviewee cooperation (Instituto de Estudios Avanzados-IESA, 2010; Díaz de Rada, 2019).
- Using all modalities, affording them the same level of importance. For example, Díaz de Rada and Palacios Gómez (2013) used a telephone survey to conduct interviews in an urban area made up mainly of single-family homes, where it was difficult to conduct face-to-face interviews. In a first visit, the face-to-face survey achieved a response rate of 21.2%, a figure that tripled for the telephone survey (61.3%).

Sequential mixed-modes begin collecting information using one mode, then gradually apply other modes to those who did not respond to the first mode (on the basis that this will be a "reduced" sample). To decide the order in which to use the modes, the response rate and/or cost are considered. When the response rate is taken as a reference, it is a question of applying first the mode that produces the most responses, which will involve interviewing less participants through the other modes. When cost criteria are followed, the least expensive mode is first applied to the total sample, followed by other, cheaper modes.

Using mixed-modes has the advantages of reducing coverage and non-response errors and significantly decreasing research costs (Díaz de Rada *et al.*, 2019), particularly when the combinations begin with the most economical modes (Dillman *et al.*, 2014, p. 309) although it is true that sometimes this saving can be low owing to the additional costs generated by the design and implementation of the various modes. Mixed-modes have been used in the present study to improve sample representativeness, based on the larger population covered, and to increase the response rate, all at low cost.

3. Sequential design case study: survey of languages used by the Catalonia-based population
To empirically demonstrate the advantages of collecting information sequentially, this work analyses the results of a study carried out jointly by the *Institut d'Estadística de Catalunya* – Instituto de Estadística de Cataluña (hereinafter Idescat) and the *General Directorate of*

Language Policy; the objective of the study was to identify which languages Catalonian residents speak in their personal spaces (Institut d'Estadística de Catalunya [Idescat], 2015).

3.1 Sample design and data collection process

The questionnaire focussed on identifying which languages are used in Catalonia. The population under study were people 15 years and over living in Catalonia in their main family residence (Institut d'Estadística de Catalunya [Idescat], 2015). A stratified sample was proposed based on territorial areas and large municipalities, where the first-stage units were the census sections, and the residents in each were the second-stage units; a dozen people were chosen from each section. The theoretical sample was set at 7,500 interviewees, which represents a sampling error of 1.36% for the entire sample, at a confidence level of 95.5% (Institut d'Estadística de Catalunya [Idescat], 2015, p. 15).

The data collection process began with a self-administered internet-based (hereinafter online) questionnaire which the sample accessed by clicking on a link in the cover letter provided with the survey. Individuals who did not respond to the questionnaire were then interviewed via a telephone survey. Finally, those who could not be reached by telephone were interviewed face-to-face.

Figure 1 presents the sequential mixed-modes used in this case, including the percentage of the sample reached through each method and the accumulated percentage.

3.2 Responses by modality

The research team achieved the collaboration of 7,255[2] people, of whom almost three out of four (71.4%) were those originally chosen (originals), while the remainder (substitutes) were selected later to address sample reduction caused by non-response. As can be seen in Table 1, in the originals' sample 56.1% of the surveys were conducted online (2.908/5.181), while in the substitutes' sample half of the questionnaires were answered over the phone. One out of every three substitutes responded to the face-to-face survey.

The high number of substitutes who responded through the face-to-face survey suggests the need to carry out a detailed analysis to understand when they joined the sample. However, before carrying out this task, and following Alvira (1991), the first issue to be considered is to what extent the sample is an adequate representation of the universe.
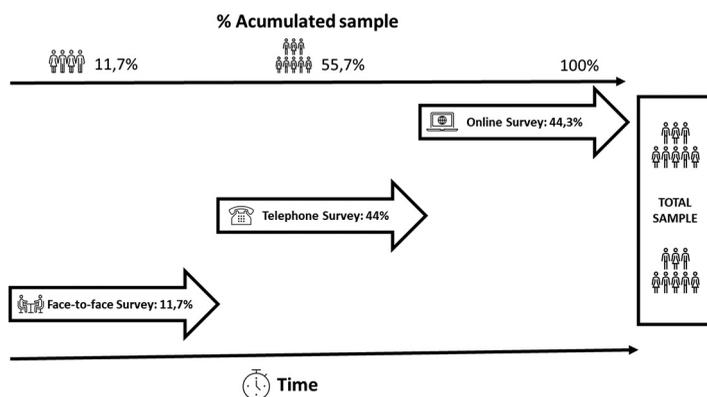

Figure 1.
Evolution of the accumulated sample with sequential mixed-modes (internet, telephone and face-to-face surveys)

Instituto Nacional de Estadística (2013b) estimates that the Catalonian population aged 15 and over is 6,366,783. The survey sample is slightly underrepresented by inhabitants under 34 years of age, which is common in research using general population-based surveys (Stoop, 2005, 2007; Haan *et al.*, 2014; Díaz de Rada and Núñez, 2008). The differences, around 2%, are within sampling error. In the centre of Figure 2 the great similarity between both distributions can be observed, although the number of surveys returned was slightly higher than the distribution of the universe, with differences less than half a percentage point, reaching 0.85% in the group 45–54 years old. Another common feature of general population-based surveys is the overrepresentation of older people, with a magnitude slightly higher than that of young people; this can also be owing to sampling error.

The origin or nationality of the interviewees could be one of the factors influencing this situation. The 963,712 immigrants living in Catalonia in 2013 represented 15.13% of the resident population (Instituto Nacional de Estadística, 2013b), while the analysis of the sample by nationality revealed that 11% of those interviewed were non-Spanish nationals, which suggests that this group is underrepresented; but the difference falls within the margins of sampling error. Although this is a "normal" situation, the underrepresentation is

| | Originals | | Substitutes | | Total | % OF |
	N°	%	N°	%		ORIGINALS
Theoretical sample: 7.492						
Effective sample						
Online survey	2.908	56.1%	304	14.6%	3.212	90.5
Telephone survey	2.148	41.5%	1.047	50.5%	3.195	67.2
Face-to-face survey	125	2.4%	723	34.9%	848	14.7
Total sample	5.181	100.0%	2.074	100.0%	7.255	71.4

Table 1.
Number of surveys achieved by each modality
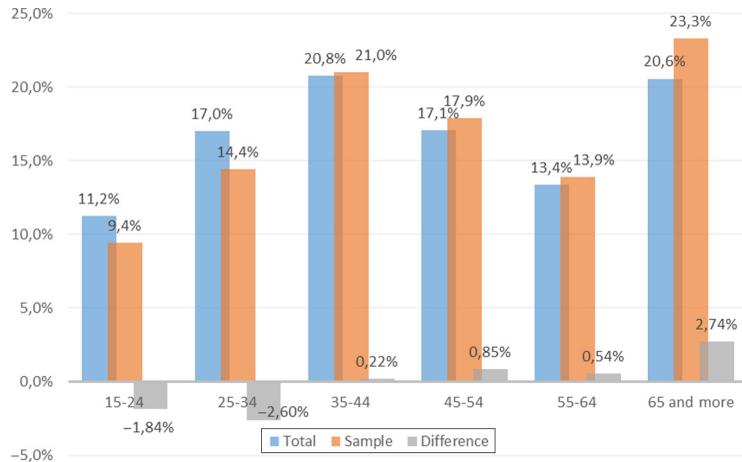
Source: Based on Idescat (2013)

Figure 2.
Representativeness analysis: total sample *vs* age groups

Source: Based on Idescat and General Directorate of Language Policy: Enquesta d'usos lingüístics de la població e INE (2013a)

much lower than has been found in other studies, for example, that of the Sociological Research Centre (hereinafter CIS), which, in the first half of 2018, after contacting 395,886 people, found that 2.2% of households were occupied by non-Spanish nationals (Centro de Investigaciones Sociológicas, 2018).

Finally, it would be interesting to carry out the comparisons with other variables, such as levels of educational studies and economic activity, which is difficult owing to the absence of updated records for the universe. The high sample size of the Active Population Survey (hereinafter APS) and its special design characteristics allow us to regard this source as very similar to the totality of the Spanish population (Díaz de Rada and Núñez, 2008). It should be borne in mind that this is the main household-focussed survey, in terms of sample size, cost and staff employed (Losilla, 2005).

Starting with educational level, in the present study sample there is an overrepresentation of people with primary school studies, and the opposite is the case for secondary and university studies, underrepresented by 3% and 7%, respectively (Figure 3). The economic activity measure makes it possible to differentiate activity level (active/inactive) and employment situation (employed/unemployed). The sample has a higher activity rate (64.1%) than the universe (60.1%); this difference was slightly higher in employment terms. However, these differences disappear when only Spanish nationals are considered. In any case, the differences between the sample and the universe are small, much smaller than found in other studies (Díaz de Rada and Martínez Martín, 2014, 2020) when probabilistic samples have been compared with data about the universe.

3.3 Modes used by the sample

As previously mentioned, nationality had great importance in terms of the mode that participants used to respond. Thus, 45.8% of Spanish nationals responded to the online survey, 45% to the telephone survey, and the remaining 9.2% had to be interviewed face-to-
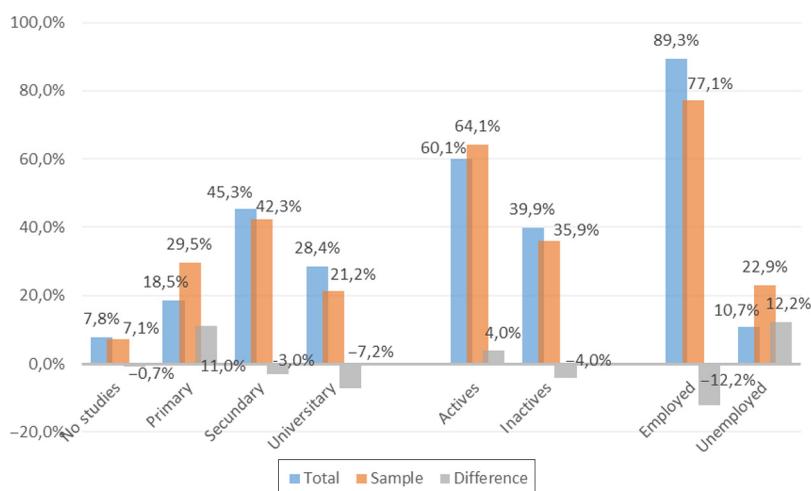

Source: Based on Idescat and General Directorate of Language Policy: Enquesta d'usos lingüístics de la població e INE (2013a)

Figure 3.
Representativeness analysis: total sample *vs* educational level *vs* economic activity

face. Among the non-Spanish nationals, 30% responded to the online survey, 30% to the face-to-face survey and 39% to the telephone survey.

A regression analysis, taking the survey mode chosen as the dependent variable, and five sociodemographic aspects (age, education, economic activity, cohabitation situation, nationality) and normal language spoken as the dependent variables, revealed the great importance of nationality in the mode chosen to respond to the survey. Table 2[3] shows that nationality had a strong, negative effect on choice of the online survey and somewhat less for the telephone survey. It should be take into account that nationality has the highest magnitude and, with normal language spoken, is the only variable significant in both the self-administered and interview-based modes. It should be noted also that the influences that cohabitation situation, educational level and economic activity have in the online survey are not apparent in the telephone survey, while age does have an impact. This is a highly significant model, although the goodness-of-fit measures are not very high; these results are in line with similar research previously carried out in Spain (Riba *et al.*, 2010; Torcal *et al.*, 2006; Buelens and van den Brake, 2015; Couper, 2017; Dillman, 2017; Vannieuwenhuyze, 2014).

Another way to measure the great impact of nationality in terms of cooperation in the surveys is to examine the number of substitutions. The Spanish nationals presented an originals' rate of almost three out of every four individuals (73.8%), a percentage that dropped 20 points in the case of non-nationals (Table 3).

The combination of both aspects reveals the high participation of originals in the online survey, an increase in substitutions in the telephone survey and low participation of

| | | Dependent variable: mode used to participate (Reference category: face-to-face survey) | | |
| | | Online survey | | Telephone survey |
	Coefficients	Wald Test	Coefficients	Wald Test
Cohabitation situation	−0.193	3.581+	0.046	0.216
Age	−0.015	0.111	0.105	5.632[*]
Education	0.303	24.829[**]	−0.027	0.204
Normal language		43.365[**]		13.587[**]
Catalan	0.670	0.370[**]	0.372	0.432
Both equally	−0.111		0.116	0.177
(Ref: Spanish)				
Nationality	−1.301	81.907[**]	−0.779	0.136[**]
Economic activity		[*]		
Employed	0.465	5.049[*]	0.042	0.136
Unemployed	0.542	5.757[*]	−0.101	0.193
Retired	0.202	0.858	0.110	0.213
Student	1.002	11.208[**]	0.370	0.201
(Ref.: domestic work)				0.292
Constant	1.609	15.670[**]	1.760	20.050
Chi square	2544.480			
Pseudo R^2 (Negelkerke)	0.102			
Scale parameter	1.083			
Cases	6.230			

Notes: [**]< 0.01; [*]< 0.05 [+]< 0.10
Source: based on Idescat and General Directorate of Language Policy: Enquesta d'usos lingüístics de la població

Table 2.
Multinomial logistic regression

	Originals' rate(%)	Originals telephone(%)	Substitutes face-to-face	Total
Spanish and dual nationality	73.8	26.2		6.335
Other nationality	53.0	47.0		851
Total sample	71.3	28.7		7.186

Originals' rate and modes(%)		Substitutes(%)	Originals	Total
Spanish and dual nationality				6.335
Online survey		8.7	91.3	2.900
Telephone survey		32.1	67.9	2.848
Face-to-face survey		84.0	16.0	587
Other nationality				851
Online survey		17.4	82.6	258
Telephone survey		38.0	62.0	334
Face-to-face survey		88.0	12.0	259

Source: based on Idescat and General Directorate of Language Policy: Enquesta d'usos lingüístics de la població

Table 3.
Cooperation in the surveys according to mode and nationality

originals in the face-to-face survey, mainly among non-Spanish nationals. Taking into account that 11% of the Catalonia-based non-nationals use Catalan as their normal language (9.2% Catalan only and 1.2% Catalan and Spanish), the administered survey modes were used more among the population that speaks only Spanish.

3.4 Analysis of the originals (without considering substitutions) and typology of those who responded to each mode
Previously, it was noted that the face-to-face survey presented the lowest rate of originals. This statement will now be clarified in this section. The data file does not show the number of reserves used by each sample unit, but a search of the total sample (Institut d'Estadística de Catalunya [Idescat], 2015) revealed that in the online survey 77 sample units were replaced twice, and on 35 occasions, the third reserve was used. These figures were higher for the telephone survey (226 and 55), and higher again for the face-to-face survey: on 273 occasions the second reserve was contacted and, on 210 occasions, the third reserve. That is, 483 reserves were called on to achieve 848 interviews, which explains the low percentage of originals. However, more than the low originals' participation rate, it is surprising that on 210 occasions it was necessary to use three reserves.

Using this information to examine the interviewees' propensity to respond to any one mode could contaminate the findings, because making more substitutions increases the probability of collaboration (Brechón, 2015): the unit that does not collaborate is replaced and, if the substitute does not collaborate, (s)he is also replaced, and so on, with the substitutions stopping when collaboration is achieved. Considering all the surveys collected faced difficulties in terms of the effectiveness of each mode, mainly because of lack of knowledge as to when the reserves were introduced, henceforth the present study will focus on the sample of originals and ignore the substitutes. Proceeding in this way reveals precisely the prospective interviewees' propensity to respond to each mode by establishing exact information about the number of contacts made: once if they responded to the online survey, and twice if they responded to the administered surveys.

For this reason, this section, and the following, will focus on the originals, that is, the 5,181 (56.1%) who responded to the online survey (Table 4).

To analyse the response rate the number of sample units contacted to carry out the 5,181 interviews achieved without substitution must first be identified. Of the 7,492 units forming the theoretical sample, those who responded to the online survey represented a cooperation rate (COOP1[4], American Association for Public Opinion Research [AAPOR], 2016) of 38.9%.

To calculate the response rate of the administered survey, it is necessary to know how many people did not respond to the online survey, that is, 4,584 (7,492–2,908). The 2,148 questionnaires answered by telephone and the 125 completed face-to-face represent a cooperation rate of 49.6%, almost ten points higher than the online survey. In total, 5,181 originals were interviewed, which represents a total cooperation rate of 69.15%.

Limiting the analysis to the originals allows an assessment to be made of the typology of those who responded to each mode. Table 5 shows the sociodemographic characteristics of respondents by mode; highlights are the youth of those who responded to the online survey (almost half – 49.5% – under 44 years), their high level of education (27 % university studies), living in households made up of three people, with an average of 1.2 children, and 69.9% being employed, and greater cooperation from workers and students (almost one in ten being students). Also noteworthy is the greater number of salaried employees and the higher number of Spanish nationals. The online survey was also answered by those groups with the highest incomes: 55.8% had incomes higher than 1,801 Euros per month, while two out of three interviewees had incomes below that figure, 64% of those interviewed by telephone, and 62.4% of those interviewed face-to-face.

The telephone survey, on the other hand, was answered by older people (47 % over 54 years), those with lower educational levels (47 % without studies/with only primary school studies), a greater number of retirees (29 %), and those do unpaid domestic work (twice as many as those who responded online). They also stay in smaller households than those who responded online (2.7 and 3 members, respectively). They also have more children and lower incomes. Some 32.2% earn less than 1,100 Euros per month and a similar amount between 1,101 and 1,800 euros per month.

Those who responded to the face-to-face survey did not show notable differences when the corrected standardised residuals were considered, undoubtedly owing to the small sample size. It was also observed that they had similar characteristics to those who responded to the telephone survey, although they were somewhat younger (36% 55 years and over), there were more casual workers, and almost one in three were non-Spanish nationals. Because of their sociodemographic similarities, that they responded to the survey through an interviewer (as for the telephone interviewees), the low sample size, they will henceforth be grouped in the present study with the interviewees contacted by phone. Thus, two categories of interviewees were created; those who responded to the self-administered survey (online) and those who responded to administered surveys (telephone and face-to-face). The first group had 2,908 interviewees, and the second 2,262; 56.12% and 43.88%, respectively (Table 5).

	Cases	Cooperation rate(%)
Originals:		
Self-administered	2.908	38.8
Telephone and face-to-face	2.273	49.6
Total	5.181	69.15

Source: based on Idescat and General Directorate of Language Policy: Enquesta d'usos lingüístics de la població

		Online	Telephone	Face-to-face	Marginal
Age	16 – 24	*+11.8%	6.6%	8.0%	9.5%
(V = 0.1330**)	25 – 34	+15.7%	11.4%	17.6%	13.9%
	35 – 44	+22.0%	17.7%	25.6%	20.3%
	45 – 54	+19.9%	16.9%	12.8%	18.5%
	55 – 64	13.8%	+16.9%	12.0%	15.1%
	65 and more	16.9%	+30.4%	24.0%	22.7%
Education	No studies	4.6%	+10.0%	8.1%	6.9%
(V = 0.180**)	Primary	23.0%	37.0%	+40.7%	29.3%
	Secondary	39.8%	38.7%	35.8%	39.2%
	University	+27.3%	13.7%	14.6%	21.3%
	Other	5,3%	0,7%	0,8%	3,2%
Household size, average (significant difference)		3,0	2,67	2,9	2,9
Cohabitation situation: partner	Yes	66,9%	65,0%	58,7%	65,9%
(V = 0.031, non-significant)	No	33.1%	35.0%	41.3%	34.1%
Number of children, average (significant difference)		1.2	1.6	1.3	1.3
Economic activity	Employed	+54.1%	45.2%	45.9%	50.1%
(V = 0.136**)	Unemployed	+15.8%	13.3%	20.5%	14.8%
	Retired	16.4%	+29.0%	21.3%	21.8%
	Student	+9.8%	4.9%	5.7%	7.6%
	Unpaid domestic work	3.9%	+7.7%	6.6%	5.6%
Professional situation (V = 0.050**)	Self-employed worker	20.5%	22.5%	22.0%	21.5%
	Salaried employees	+63.6%	59.4%	47.3%	61.4%
	Casual workers	15.9%	18.1%	+30.8%	17.2%
Nationality	Spanish	+91.8%	89.5%	69.4%	65.9%
(V = 0.097**)	Non-spanish	8.2%	10.5%	+30.6%	34.1%
Income level, month (V = 0.156**)	< 1.100 Euros	18.4%	+32.1%	37.6%	24.5%
	from 1.101 to 1.800	25.8%	+32.1%	24.8%	28.4%
	from 1.801 to 2.700	+18.4%	15.2%	8.8%	16.8%
	More 2.700 E	+13.7%	7.3%	6.4%	10.8%
	Does not know or does not respond	+23.7%	13.4%	22.4%	19.4%
Cases		2.908	2.148	125	5.181

Note: (*) Positive corrected standardised residual with a significance level lower than 0.01%
Source: based on Idescat and General Directorate of Language Policy: Enquesta d'usos lingüístics de la població

Table 5.
Sociodemographic characteristics of respondents by mode

The Cramér's V coefficients revealed that the variables which exhibited the greatest difference between the survey modes are educational level (0.180), income level (0.156), economic activity and age (0.136 and 0.133, respectively).

3.5 Explanatory factors in the choice of response mode
Table 5 may provide information that is not entirely accurate if some of the variables shown are influenced by others. Thus, for example, family size may be related to the number of children in a household, and the number of children to the age of the interviewee, as younger people do not usually have children. Age may influence both employment status, given the high rates of youth unemployment, and professional situation, as younger people have more short-term contracts. The use of regression techniques is necessary when assessing the factors that explain participation. This is because regression techniques can analyse the effects of the variables introduced into the model on the dependent variable.

Logistic regression was used to identify the sociodemographic variables that most influenced choice of mode; the mode chosen was used as the dependent variable and the variables listed in Table 5 as the independent variables. That is, age, studies, size, cohabitation situation (partner or not), number of children, economic activity, professional situation and income. Excluded were the professional situation variable (addressed only by the employed interviewees, slightly more than half), number of children (33.4% of the sample had no children) and income level, which 20% said they did not know, or did not answer directly.

The results on the left side of Table 6, which depicts only the significant variables, reveal that only three variables are significantly related to the online response mode: level of studies, cohabitation situation and economic activity. Those without studies or with only primary school studies were less likely to respond via the online self-administered mode; these figures decrease as study levels increase. Those with higher study levels were more likely to use the online self-administered mode.

Couples were also least likely to use the online mode. In terms of level of economic activity, students are 3.03 times more likely to respond using the online mode than those who perform unpaid domestic work, a figure that drops to 1.86 in the case of the unemployed and 1.68 for the employed.

When normal language used is introduced into the model, based on the hypothesis that the Catalan speakers may be more sensitised to linguistic issues and, therefore, present a greater self-selection, the explanatory capacity of the model increases, as does the number of significant variables. Level of studies, cohabitation situation and economic activity showed the same tendencies as in the previous model although with slightly lower coefficients in the groups without studies, the employed and students, so henceforth comments will only be made about the contribution of the three new variables introduced. The variable with the greatest explanatory power is normal language used: Catalan speakers were 1.53 times more likely to respond to the online questionnaire than habitual Spanish speakers, while those who speak both languages to the same extent showed less motivation to respond online.

In terms of age, those under 24 years responded more to the online survey, 2.1 times more than the reference group; this figure remained around 1.6 in those between 25 and 34 years, and dropped to 1.4 among those between 35 and 54 years old. There is a linear relationship between age and tendency to respond to the self-administered survey; the younger the individual, the more (s)he collaborated.

4. Conclusions and implications for market research

Market research every day faces the challenge of profiting from the heavy investments in technical, human and economic resources that it requires. In particular, researchers must deal with individuals who are increasingly saturated and reluctant to participate in surveys. Faced with this situation, it is necessary to adapt the design of tools and use strategies to improve the representativeness of samples by increasing response rate [5].

The present study proposes the use of the so-called "sequential mixed-modes" in survey-based market research. This technique is based on the successive application of different information collection techniques (face-to-face, telephone and internet-based surveys); it offers four fundamental advantages:

(1) increased coverage rate;

(2) higher response rate;

(3) lower costs; and

(4) greater cooperation.

	Coefficients	Standard Error	Wald Test	Odds ratio	Coefficients	Standard Error	Wald Test	Odds ratio
Normal language							**	
Catalan	–,–	–,–	–,–	–,–	0.425	0.066	46.377**	1.53
Both equally	–,–	–,–	–,–	–,–	−0.076	0.131	40.958	0.927
Ref.: Spanish								
Sociodemographic characteristics								
Age							16.135*	
16 – 24	–,–	–,–	–,–	–,–	0.741	0.221	11.179**	2.097
25 – 34	–,–	–,–	–,–	–,–	0.47	0.165	8.059*	1.599
35 – 44	–,–	–,–	–,–	–,–	0.372	0.156	5.715*	1.451
45 – 54	–,–	–,–	–,–	–,–	0.389	0.154	6.405*	1.475
55 – 64	–,–	–,–	–,–	–,–	0.144	0.141	1.036	1.154
Ref.: 65 and more								
Educational level			129.11**				96.597***	
No studies	−1.136	0.144	62.131**	0.321	−0.943	0.155	37.224***	0.389
Primary	−0.998	0.092	118.58**	0.369	−0.937	0.097	92.934***	0.392
Secondary	−0.607	0.084	52.271**	0,545	−0,617	0,088	48.998**	0.54
Ref. University								
Cohabitation situation	−0,201	0.068	8.714**	0,818	−0,283	0,074	14.67**	0.753
Ref.: no partner								
Economic activity			62.357***				15.838***	
Employed	0.518	0.136	14.489**	1.679	0.499	0.152	10.792**	1.648
Unemployed	0.622	0.15	17.3**	1.863	0.629	0.166	14.396*	1.876
Retired	0.128	0.142	0.817	1.137	0.307	0.165	3.461**	1.359
Student	1.109	0.183	36.731**	3.033	0.746	0.236	9.959**	2.109
Ref.: unpaid domestic work								
Constant	0.604	0.17	12.679	1.829	0.206	0.209	0.978	1.229
Chi square	6.269.73				5.849.27			
Pseudo R² (Negelkerke)	0.095				0.104			
Cases	5.181				5.181			

Source: based on Idescat and General Directorate of Language Policy: Enquesta d'usos lingüístics de la població

Table 6.
Variables that influence answering the online survey (logistic regression)

This study presented the main characteristics of sequential mixed-modes from a theoretical viewpoint, and then empirically analysed a real case that allowed interesting conclusions to be drawn. In the case study the information collection began through a self-administered internet-based survey. Those individuals who did not respond to the [6]internet-based survey invitation were contacted by telephone. These two methodologies collected most information (44.3% by the self-administered survey and 44% by the telephone survey). Finally, the remainder of the information was collected through a face-to-face survey.

The analyses carried out demonstrated the preference of certain groups for specific ways of collecting information. For example, it is shown that level of studies, cohabitation situation and economic activity have a decisive influence on the decision to respond online. Thus, through knowing the preferences of specific profiles for particular modes, and by combining modes, it is possible to increase survey response rate, access to certain profiles and, consequently, achieve greater sample representativeness (among others, De Leeuw and Elevent, 2020; Diaz de Rada *et al.*, 2019).

It should be noted that, in this research, the interviewees were not allowed to choose the response mode, which decreases response rates (among others, Medway and Fulton, 2012) but were instead offered the possibility of responding to the questionnaire sequentially through a self-administered online survey, a telephone survey or a face-to-face survey. This sequential offer of options allowed the interviewee to choose the method that best suited his/her preference, which increased the participation of groups that traditionally exhibit lower participation rates. Proof of this is that a "feedback" effect was detected between modes, that is, "that telephone or face-to-face contact with a specific person causes them to choose to answer the online survey" (Biblioteca Técnica de Política Lingüística, 2018), which involves – ultimately – an increase in response rate. Unfortunately, it is not possible to know the profile of those who acted in this way as this situation was not saved in the data file.

4.1 Managerial implications and limitations
The discussions and analyses carried out in this work raise several implications for the management of market research activities.

Firstly, it was found that data collection through mixed sequential modes substantially increases response rates in online surveys. This is particularly important as, in recent years, data collection through the Internet has become almost standard market research practice.

Secondly, it was found that using sequential mixed-modes not only increased response rate, but that it also increased the quality of the sample finally collected. In other words, the sample that resulted from the joint use of the three data collection modes more accurately reflected the distinctive features of the universe under study.

Finally, the effects of certain variables in interviewees' choices of response mode were very interesting. Thus, a third recommendation is that the use of internet-based self-administered surveys is especially promising with individuals with a higher level of education and among students.

The main limitation of the conclusions obtained from this work, apart from those normally associated with conclusions derived from case studies, is that the data collection was carried out only in one autonomous community and on a single topic. Both limitations could introduce specific biases. While studies carried out on the same topic in other geographic contexts have returned results consistent with the present study (Kappelhof, 2015), future studies might verify the stability of the results by examining other geographic areas and by addressing different topics.

Notes

1. At the time of writing another advantage of these modes, hopefully temporary, should be mentioned: that is, the "fear of speaking in person" with strangers owing to COVID-19 (De Leeuw and Elevent, 2020); this fear significantly limits the use of the face-to-face survey.

2. Please note that in the first paragraph of the previous section it was stated that the theoretical sample was set at 7,500 interviewees. The difference between the two is caused by interviews cancelled during the data collection process.

3. This is the information on the number of participants in relation to the number of people contacted (interviewed, rejected, those who could not participate owing to physical problems), which makes it the least demanding measure, the one with the highest response values (AAPOR, 2016).

4. It should be noted that sample representativeness refers to the fact that all the variability of the universe is included in the sample. In this sense, it should be remembered that, although it is desirable to achieve an adequate response rate, some research has indicated that low response rates do not always cause poor sample representativeness. Thus, in one of the first works that analysed the phenomenon, Visser *et al.* (1996) showed that the electoral prediction of a mail poll, with a response rate of 20%, was notably better (average error of 1.6%) than that obtained through a telephone poll with a response rate of 60%, with an average error of 5.2%.

5. As noted by Díaz de Rada *et al.* (2004), the aspect that differentiates self-administered online surveys from telephone and face-to-face surveys is the absence of an interviewer; the presence of interviewers tends to introduce biases (e.g. owing to the "social desirability" effect).

References

Alvira, F. (1991), *La Encuesta: una Perspectiva General Metodológica*, CIS, Cuadernos Metodológicos, Madrid.

American Association for Public Opinion Research (AAPOR) (2016), "Standard definitions. Final dispositions of case codes and outcome rates for surveys (9th edition)", available at: www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf

Beullens, K., Loosveldt, G., Vandenplas, C. and Stoop, I. (2018), *Response Rates in the European Social Survey: Increasing, Decreasing, or a Matter of Fieldwork Efforts, Survey Methods, Insights from the Field*, available at: https://surveyinsights.org/?p=9673

Biblioteca Tècnica de Política Lingüística (2018), "Análisi de l'Enquesta d'usos lingüístics de la població: volum 1, coneixements, usos, transmissió i actituds lingüístics", *Generalitat de Catalunya, Departament de Cultura*, Barcelona.

Bréchon, P. (2015), "Random sample, quota sample: the teachings of the EVS 2008 survey in France", *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*, Vol. 126 No. 1, pp. 67-83.

Brick, J.M. and Williams, D. (2013), "Explaining rising nonresponse rates in Cross-Sectional surveys", *The Annals of the American Academy of Political and Social Science*, Vol. 645 No. 1, pp. 36-59.

Buelens, B. and van den Brake, J. (2015), "Measurement error calibration in mixed-mode sample surveys", *Sociological Methods and Research*, Vol. 44 No. 3, pp. 391-426.

Callegaro, M., McCutcheon, A.L. and Ludwig, J. (2010), "Who's calling? The impact of caller ID on telephone survey response", *Field Methods*, Vol. 22 No. 2, pp. 175-191.

Centro de Investigaciones Sociológicas (2018), "Barómetros de enero a junio, estudios n° 3.203, 3.205, 3.207, 3.210, 3.212, 3.217", available at: www.cis.es/cis/opencm/ES/11_barometros/depositados.jsp

Couper, M.P. (2017), "New developments in survey data collection", *Annual Review of Sociology*, Vol. 43 No. 1, pp. 21-45.

De Leeuw, E. (2005), "To mix or not to mix data collection modes in surveys", *Journal of Official Statistics*, Vol. 21, pp. 233-255.

De Leeuw, E. (2008), "Choosing the method of data collection", in de Leeuw, E., Hox, J.J. and Dillman D. A. (Ed.), *International Handbook of Survey Methodology*, Lawrence Erlbaum Associates y Asociación Europea de Metodología, New York, NY, pp. 113-135.

De Leeuw, E. (2018), "Mixed-mode: past, present, and future", *Survey Research Methods*, Vol. 12 No. 2, pp. 75-89.

De Leeuw, E. and Elevent, A. (2020), "Mixed mode and mixed device surveys: Why, when, and how", available at: https://wapor.org/wp-content/uploads/DeLeeuwMMWAPORSlides.pdf.

Díaz de Rada, V. (2019), "Calidad de respuesta y modo de administración del cuestionario", *Revista Internacional de Sociología*, Vol. 77 No. 1, pp. 1-17.

Díaz de Rada, V. and Martínez Martín, V. (2014), "Random route and quota sampling: Do they offer any advantage over probably sampling methods?", *Open Journal of Statistics*, Vol. 4 No. 5, pp. 391-401.

Díaz de Rada, V. and Martínez Martín, V. (2020), "Diseños muestrales en hogares: Diferencias y similitudes entre muestras probabilísticas y muestras con rutas y cuotas", *Revista Española de Investigaciones Sociológicas*, Vol. 171, pp. 23-42.

Díaz de Rada, V. and Núñez, A. (2008), *Estudio de Las Incidencias en la Investigación Con Encuestas*, CIS, Madrid.

Díaz de Rada, V. and Palacios Gómez, J.L. (2013), "Comparación de las tasas de respuesta en el uso combinado de modalidades de encuesta", *Revista Española de Investigaciones Sociológicas*, Vol. 141, pp. 159-170.

Díaz de Rada, V., Domínguez, J.A. and Pasadas, S. (2019), *Internet Como Modo de Administración de Encuestas*, CIS, Madrid.

Díaz de Rada, V., Flavián, C. and Guinalíu, M. (2004), "Encuestas en internet: Algo más que una simple versión mejorada de la tradicional encuesta autoadministrada", *Investigación y Marketing*, Vol. 82, pp. 45-56.

Dillman, D.A. (2017), "The promise and challenge of pushing respondents to the web in mixed-mode surveys", *Survey Methodology*, Vol. 43, pp. 3-30.

Dillman, D.A., Smyth, J. and Christian, L.M. (2014), *Internet, Mail and Mixed-Mode Surveys*, 4rd ed., Wiley, Nueva York.

Dutwin, D. and Buskirk, T.D. (2017), "Apples to oranges or gala versus golden delicious? Cmparing data quality of nonprobability internet samples to low response rate probability samples", *Public Opinion Quarterly*, Vol. 81 No. S1, pp. 213-239.

Haan, M., Ongena, Y.P. and Aarts, K. (2014), "Reaching hard-to-survey populations: mode choice and mode preference", *Journal of Official Statistics*, Vol. 30 No. 2, pp. 355-379.

Hoschstim, J.R. (1967), "A critical comparison of three strategies of collecting data from households", *Evaluation Journal of the American Statistical Association*, Vol. 62, pp. 976-989.

Institut d'Estadística de Catalunya (Idescat) (2015), "Enquesta d'usos lingüístics de la població 2013, barcelona; generalitat de catalunya, departament de cultura, direcció general de política lingüística", available at: www.idescat.cat/cat/idescat/publicacions/cataleg/pdfdocs/eulp2013.pdf

Instituto de Estudios Avanzados-IESA (2010), *Barómetro de Opinión Pública de Andalucía*, IESA-CSIC, Córdoba.

Instituto Nacional de Estadística (2003), *Encuesta de Salud y Hábitos Sexuales: metodología*, INE, Madrid, available at: www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176785&menu=metodologia&idp=1254735573175

Instituto Nacional de Estadística (2013b), "Estadística del padrón continuo a 1 de enero de 2013: Datos por municipios".

Kappelhof, J.S. (2015), "Face-to-Face or sequential Mixed-Mode surveys among Non-Western minorities in The Netherlands: the effect of different survey designs on the possibility of nonresponse bias", *Journal of Official Statistics*, Vol. 31 No. 1, pp. 1-30.

Lee, S., Brick, M.J., Brown, E.R. and Grant, D. (2010), "Growing Cell-Phone population and noncoverage bias in traditional random digit dial telephone health surveys", *Health Services Research*, Vol. 45 No. 4, pp. 1121-11329.

Losilla, J. (2005), "Recogida de datos de una encuesta dirigida a los hogares: la EPA", course los trabajos de campo en las encuestas del INE. June 21–23",

Massey, D.S. and Tourangeau, R. (2013), "Introduction: new challenges to social measurement", *The ANNALS of the American Academy of Political and Social Science*, Vol. 645 No. 1, pp. 6-22.

Medway, R.L. and Fulton, J. (2012), "When more gets you less: a meta-analysis of the effect of concurrent web options on mail survey response rates", *Public Opinion Quarterly*, Vol. 76 No. 4, pp. 733-746.

Mercer, A.W., Kreuter, F., Keeter, S. and Stuart, E.A. (2017), "Theory and practice in nonprobability surveys: parallels between causal inference and survey inference", *Public Opinion Quarterly*, Vol. 81 No. S1, pp. 250-271.

National Research Council (2013), "Nonresponse in social science surveys: a research agenda", Tourangeau, R. and Plewer, T. (Eds), *Panel on a Research Agenda for the Future of Social Science Data Collection, Committee on National Statistics. Division of Behavioral and Social Sciences and Education*, The National Academic Press, Washington, DC DC.

Riba, C., Torcal, M. and Morales, L. (2010), "Estrategias Para aumentar la tasa de respuesta y los resultados de la encuesta social europea en españa", *Revista Internacional de Sociología*, Vol. 68 No. 3, pp. 603-635.

Sala, E. and Lynn, P. (2009), "The potential of a multi-mode data collection design to reduce non response bias. The case of a survey of employees", *Quality and Quantity*, Vol. 43 No. 1, pp. 123-136.

Stoop, I. (2005), *The Hunt for the Last Respondent. Non-Response in Sample Surveys*, Social and Cultural Planning Agency, The Hague.

Stoop, I. (2007), "No time, too busy: Time strain and survey cooperation", Loosveldt, G., Swyngedouw, M. and Cambré, B. (Ed.), *Measuring Meaningful Data in Social Research*, Acco, Leuven, pp. 301-314.

Stoop, I., Billiet, J., Koch, A. and Fitzgerald, R. (2010), *Improving Survey Response: lessons Learned from the European Social Survey*, Wiley, Chichester, West Sussex.

Torcal, M., Morales, L. and Riba, C. (2006), "Supervisión y control de calidad del trabajo de campo de la encuesta social europea en españa: Evaluación y resultados", *Metodología de Encuestas*, Vol. 7 No. 2, pp. 75-97.

Vannieuwenhuyze, J. (2014), "On the relative advantage of mixed-mode versus single-mode surveys", *Survey Research Methods*, Vol. 8 No. 1, pp. 31-42.

Visser, P., Krosnick, J.A., Marquette, J. and Curtin, M. (1996), "An evaluation of the Columbus dispatch poll", *Public Opinion Quarterly*, Vol. 60 No. 2, pp. 181-227.

Corresponding author
Miguel Guinaliu can be contacted at: guinaliu@unizar.es