

AMERICA AND SPAIN: TWO CULTURES AND TWO DIFFERENT ATTITUDES TOWARDS PUBLIC OPINION SURVEYS¹

JULIETA OJEDA ALBA
Universidad de La Rioja

(Resumen)

En las últimas décadas la utilización de encuestas ha ido en aumento en el mundo occidental. La información que estos sondeos aportan a una determinada sociedad sobre sus propias características y comportamientos es utilizada con fines a menudo económicos, políticos o académicos. Dentro de la investigación por encuestas dos de los aspectos más problemáticos son la elección de la muestra y algunas de las técnicas de recogida de datos tales como el cuestionario.

Para que un sondeo sea útil los resultados obtenidos deben ser fiables, pero sólo lo serán si el proceso se realiza con rigor. La muestra debe ser representativa de la población que se desea investigar. Adicionalmente es necesario determinar si los encuestados pretenden responder con veracidad. Los factores que puede afectar la fiabilidad de la encuesta varían según el país donde ésta se lleve a cabo. El presente artículo discute las distintas actitudes ante las encuestas de dos sociedades: la española y la norteamericana. Se tratan así mismo algunos aspectos de los sondeos que pueden afectar la fiabilidad de los resultados en ambos países.

.....

Public opinion polls have been part of Western life for decades and are now being employed at an increasing rate. Many countries have adopted polls because they are extremely useful for commercial and political purposes.² With them businesses can direct product and service modifications to serve the greatest number of people and themselves well. Also they can offer information upon which individuals can base decisions such as in making career choices, a subject which is of primary importance for university communities. If planned and conducted well, polls offer the knowledge necessary to improve efficiency in society. They unveil for us facts about a civilisation that might otherwise pass unnoticed: what most of us do, the books we read, the music we listen to, etc. Why do we need to gather these data? If for no other reason, because seeing ourselves as others see us, according to Robert Burns, "would from many a blunder free us".³

1. I would like to acknowledge the encouragement and advice of Dean L. Dodge Fernald of Harvard University in the production of this paper.

2. See Robert M. Worcester ed. *Political Opinion Polling* pp. 178-229 about the origin of polling in both America and Spain.

3. Robert Burns verse is:

O wad some Pow'r the giftie gie us
To see oursels as others see us!
It wad frae mony a blunder free us
And foolish notion

Unfortunately, along with these obvious benefits have come the inevitable problems. Many experts suggest that the public has been overexposed to polling and that now a growing resistance or distrust is evolving. This public distrust stems largely from too many incorrect predictions and self-serving analyses. Indeed, in addition to the feeling of distrust the frequency of polling has in many cases annoyed the public, affecting in turn their willingness to contribute to the success of the poll. Within the poll itself, sampling methods and some of the techniques used to gather information, notably the questionnaire, have been the most consistently problematic. I will centre my discussion on those aspects of information gathering which are most affected by the different cultures to which they are applied. There have been several studies on sampling, the wording of the questionnaire, and the reliability of the answers. It is these particular issues that will also be discussed in relation to the differing attitudes of two peoples, the Americans whose culture has Anglo-Saxon roots, and the Spanish whose society is typically Mediterranean.

Sampling is a method used to infer the characteristics of a large number of people by interviewing only a small percentage of them. Its most obvious problem is attaining an accurate representation of the whole group. The investigator must draw accurate conclusions for the whole based upon responses of a very small part of it, for, as Harold L. Nieburg comments: "Humans are limited in their scope. They see only a part of the whole, and they understand only part of what they see."⁴ From this he concludes that we have no choice but to generalise from limited information. This statement does not necessarily imply that low reliability should be expected when using small samples but, on the contrary, he also observes and demonstrates that the use of a scientifically chosen sampling gives us the ability to transcend these limits. The two most important factors in order to obtain valid and meaningful results are that the chosen sample group must be both truly representative and randomly selected.

A serious problem in sampling is that conclusions are often drawn from an even smaller number of respondents than the reported total sample, since there is always a number of individuals who do not respond. Nevertheless the reported results often do not tell the public this whole story. Complicating that, pollsters also attempt to include those nonrespondents by applying a similar distribution of answers to them based upon the demographic data of the total sample. The catch is that according to which part of a sample is missing (or nonrespondent) the results could vary. For example, if all the doctors in a certain sample were away for a convention on the day of a polling, results might change since their responses on certain issues could well differ significantly from other groups and consequently affect the final result.⁵

What also bears watching is that the nonresponse rate has been increasing recently as more people are unavailable (mobile, working two jobs, etc.) or unwilling to participate. One direct reason why unwillingness to participate is dramatically increasing is, among others, telemarketing intrusion. This procedure, whose advantages and disadvantages have been interestingly discussed by Czaja and Blair in their 1996 *Designing Surveys: A Guide to Decisions and Procedure*, has been earning more and more resentment from people who feel intruded upon at home. Also, as Tourangeau and Smith point out, the increasingly

4. Harold L. Nieburg, *Public Opinion: Tracking and Targeting* p. 13.

5. Let's imagine a hypothetical situation in which the issue being tested was, for instance, the control of doctors' fees. Doctors' opinions on this item would likely differ from those of the general public.

sensitive nature of polling questions has resulted in increasing nonresponses; and what is more, as they put it: "the very persons with the most sensitive information to report may be the least likely to report it" (1996: 37). So, as could be expected, the rising of the nonresponse rate alone is leading pollsters to more closely examine their current dubious practice of applying results to nonrespondents in reports.

The most important tool for polling is the widely used questionnaire, a list of questions designed to gather information on a particular subject for later analysis. Questions cannot be written casually since their wording in the questionnaire is a most important and pervasive problem. A fairly worded objective question seems at first to be an easy concept, but the actual construction of such is often elusive. It frequently occurs that in university professor evaluation questionnaires, where all students in the same class are independently filling out the evaluation form, the responses to the same objective question differ greatly. One of the given explanations for this is that it is ambiguity of wording that elicits different answers to a factual question. The questions therefore should be specific enough to avoid any misinterpretations.

Another wording problem is that unscrupulous pollsters often load questions in order to obtain certain desired responses. For instance, in Spain a question about building houses for the homeless might receive more support in certain areas than a loaded question about building houses for gypsies. In like fashion a question in America seeking support for a "defence budget" would receive more support than a loaded question seeking aid for "military spending". So, as we can see, the vocabulary used is determinant here. Also, the words must aim at having the same meaning for everyone and must not have significant connotations. Identifying words with intrusive connotations is by no means an easy task, but it is one that must be attempted. Regional, cultural, and even slang differences in meaning bring many possible problems to wording. Furthermore, the practice of consciously loading a question is a totally irresponsible procedure that has been used and must be eliminated.

Since polls are of little value if not reliable, the most fundamental aspect and main concern of a survey should be its reliability. In addition to the effect of sampling, wording, and vocabulary considerations, misunderstood responses can also undermine reliability. Dichotomous questions, specifically, can generate responses easily misinterpreted. To be against one of the responses does not necessarily imply that you favour the other, nor that a similar position on related topics is valid as may be assumed by some. To cite a case in point, in the November, 1990 ABC/Washington Post poll Americans were asked if they approved or disapproved of President Bush's strategy in the Middle East crisis and 36 % of respondents disapproved; the implications seemed to be that all those people opposed the Government position, but a subsequent question showed that 44% of those respondents were not really against American intervention but, on the contrary, they wanted it done more quickly. Obviously, to increase reliability more attention to possible misinterpretation of results is needed. The role of the public should be considered in the interpretation of poll results. Many aspects of polling are beyond the public's ability to assess, but the wording and placement of questions are relatively easy for people to use in order to judge the reliability of responses. For this reason when the poll results are reported the questions should be made readily available to the public and this is rarely done.

The location and order of the questions in a survey have also been shown to have a dramatic effect on the responses and therefore on their reliability. By framing a question in a definite context the respondent's perception may be altered. For example, it is obvious

that if a preceding series of questions about a political candidate's record reminds respondents of a generally dubious record, a negative response can more easily be generated to a subsequent question requiring an assessment of the candidate's performance. For that reason the same question in different polls may elicit different responses. It is important to remember that such differences in response may not reflect an actual change in opinion, but may be the direct effect of the placement of the question in a differing context. Context is an important consideration for improving reliability.

In regard to the reliability issues aforementioned I would like to first address the different sampling considerations in each of two countries, America and Spain. For accuracy, sample groups must always be scientifically selected and representative. On a national level random sample selection from both countries probably is of comparable value in leading to veracity of results because every individual in each nation has exactly the same chance of being chosen from Government census data. But at regional levels there are other considerations that apply to representativeness. Both countries are heterogeneous but each in a different way. America, being a larger country and, as is well known, a product of multiple immigration, undoubtedly has more subgroups. However, they are more uniformly (although not completely) mixed throughout the entire country: like Nieburg's "chocolate chip cookie" model.⁶

Spain, despite the common external perception, is quite heterogeneous also but with some differences of note to pollsters. Certain subgroups in Spain are much more apt to be geographically concentrated, and found to make up whole provinces or regions. A particular region in Spain may have the characteristic of homogeneity which matches that of a stratified sample. A stratified sample is a subgroup used by pollsters with members alike in terms of one major characteristic. Other characteristics must be totally at random. A random type sample drawn from a population stratum can be a tremendous aid to sampling efficiency if the stratifying characteristic is related to the study and everything else is not. In many regions of Spain local polls in fact deal with a stratified sample. For example, if one wanted to poll the Basques on an E.T.A. issue, the Basque language would be the stratification characteristic and the stratified population would turn out to live in just four provinces. This does simplify some polling tremendously.

On the contrary, the ethnic division by neighbourhoods that is so common in America is almost absent in Spain. For example, a Spaniard from Galicia of Celtic origin and speaking Galician would neither seek nor find a neighbourhood of other Galicians when moving to a far off province; consequently he would blend into the general population. Some neighbourhood differences based on income do exist in Spain but they are not nearly as common or blatant as they are in America. This type of demographic grouping (geodemography) has been commonly used by polling agencies in matters related to consumer patterns.

Nevertheless, Nieburg believes that a different type of grouping or "public taste" grouping is more relevant for polling purposes. For him it is more useful to know what media the respondent is networked to than to know his neighbourhood. Therefore effective polling techniques are determined by each country's social characteristics. It is clear that for a poll to be of any value at all the pollsters must always carefully choose a truly representative sample, and this may require different approaches in different countries as in the case of polling both America and Spain.

6. H. L. Nieburg's *Public Opinion*, pp. 45-46.

Sample choice methods may differ from culture to culture but the selected sample must still strictly govern the choice of vocabulary used in the questions. Words that are loaded for a certain community may be neutral for others, and it is the pollster's job to avoid these dangers. In America, for example, a word like "negro" is full of meanings and would drastically affect the responses of the American people in one direction or another, but it has scarcely any loading in Spain. Other racist words can behave in the same way as would "spic" and "W.A.S.P." On the other hand in Spain different words with racist connotations, like "gypsy", do exist which would provoke little negative or positive reaction in America where the idea of a gypsy is almost a curiosity. In Spain words like this are definitely loaded. In the same way words that refer in a disrespectful way to a person from the South are used in some northern Spanish regions. Examples include "maketo" in the Basque Country or "chamego" in Catalonia. The English counterpart, "southerner", in America has no particular negative connotation nowadays. As we can see, what a pollster writes may not be what the reader reads.

In the political sphere the same thing is true. Let's discuss a term such as "left wing" which used to give American citizens pause because it suggested Communism, which in itself had threatening connotations. In Spain, nowadays, as in the rest of Europe the equivalent term "de izquierdas" is just another political designation and, if anything, it may in certain situations have the positive connotation of intellectualism. Also in the political area a question containing the word "power" referring to the American Government would alert the Spanish respondents in a much more negative way than it would the American.

In the business world there are plentiful examples. The first ones that come to mind are related to the fashion industry. In America clothing labelled "Imported from Paris" can still elicit a disproportionately favourable response, while in Spain these days it has come to be just factual information if not unfavourable (as recent border conflicts have exacerbated) for those against importation. Similar conditions hold with French wine, German cars, and Italian opera in America which do not have equivalent connotations in Spain. So many common words, if unduly used in questionnaires, could influence the distribution of responses in a different way in each country.

I earlier stated that the order of questions in a questionnaire can create context influences and consequently affect its results. These influences caused changed responses to the questions and severely affected the compared results of one poll I conducted in the two countries. I tested this point with a survey of my own which I gave to four different groups of twenty people each in the streets of both an American and a Spanish city.⁷ The questionnaire was about American foreign policy and I centred my study on the question: "Should America give financial support to Central America?". This question had different responses when the placement was changed. Placing it at the beginning of the questionnaire, devoid of context, as I did with my first group of respondents in America, resulted in a low positive response of 15 %. The same question placed after, "Should America sometimes intervene in Central America?" resulted in a much higher positive response rate of 31 %. Obviously the respondents felt pressed to grant financial aid to countries after they had approved of possible intervention. These different results seem to be a direct consequence of question order. However a variation of my own experience is

7. These questionnaires were used in Cambridge, Massachusetts, USA and in Almeria, Spain.

Michaela Wänke's opinion that a change of word order within questions has no *systematic* impact on polling results.⁸

The same operation with the same questions was repeated in Spain. On the first questionnaire the item without content influence yielded a 47 % positive response rate, but, when moved to the same place of context influence as in the American version it produced only a 12 % positive response rate. In this last case the realisation that financial aid may be the first step towards direct intervention seems to have discouraged them from voting for financial aid. In identical questionnaires the response rate to the same question, relocated from a noninfluenced to an influenced position, changed in both countries, but in opposite ways simply because of the context. Not only can context influence response, but the influence can be in different directions, compounding the effect.

How can we be sure that the answers we get from a certain questionnaire are reliable? There are several important problems involved in achieving this most valued goal in polling. The first thing a pollster should make sure of is that the sample he is using is representative and appropriate for the questionnaire. People should not be asked about matters concerning which they have little or no information, since they are unlikely to have a valid opinion. It is obvious that if we give a questionnaire on farming to a group of shop assistants the result would be of little value. Relevancy, or matching a poll to a knowledgeable and interested population, is necessary for reliability but that by itself does not guarantee it.

There are many other factors which contribute to unreliability. A person responding to a questionnaire feels that he is being tested and therefore he tries to appear at his best. If we add anonymity his responses may be freer in terms of self-image, but he would also be free from responsibility. An additional problem is that the responses available for choice may not reflect what he wants to say. Some respondents may be pressed for time or, as indicated above, may be unwilling to participate in the interview. All these circumstances and others are sources of nonattitudes, hence unreliability of responses, and ultimately of the poll.

The presence of nonattitudes is one of the major causes of unreliability. This has been studied in questionnaire responses for two decades. Bishop, Oldendick, and Tuckfarber published in 1980 an inventive study in which they explain how they had introduced a fictitious question in surveys conducted in Cincinnati. The question was: "Some people say that the 1975 Public Affairs Act should be repealed. Do you agree or disagree with this idea?" Although the "Public Affairs Act" was a non-existent organism, over 30 % of the respondents gave an opinion on it. By so doing this percentage declared themselves to be willing to offer answers to unanswerable questions. Probably some of the other answers were of this type or nonattitude answers as well, and we can naturally infer that this kind of respondent is unreliable.

To further test the method of identifying unreliable respondents I designed a questionnaire which included five fictitious questions out of twenty. Then I administered it both in America and Spain to three different samples in each country: 70 high school students, 70 university students, and 30 adults chosen at random from the telephone

8. See Michaela Wänke, "Comparative Judgements as a Function of the Direction of Comparison Versus Word Order", *Public Opinion Quarterly*.

directory.⁹ My goal was to find and compare false question response rates of like groups in the different countries to see if any conclusions about reliability could be drawn.

In the questionnaire I gave to the high school students I purposely gave different directions to one third of the total group in each country. The directions given to the one third segment were designed to make them feel that it would be acceptable if they could not respond to all the questions. The larger segment of about 45 had similar results in both America and Spain and, to be exact, 18 % and 15% respectively had all reliable answers (answered no false questions). The smaller segment which had been given different directions experienced a significant increase of reliable responses in Spain, up to 30 %, but American results stayed almost the same at 19 %. It seems that the American high school students were less influenced by the directions, but the influenced Spanish group attained a higher reliability. However, it appears that overall in both countries, when respondents felt expected to answer to a false question, more of them did just that.

At the university level the percentages of students who did not indulge in answering any of the false questions (hence reliable) in both America and Spain were also similar (20 % and 22 % respectively), but with a very interesting characteristic. In America 85 % of the reliable respondents were women, while in Spain reliability was distributed more equally between the two sexes. This may mean that women in America are less threatened by competition than men and that they are more willing to admit ignorance in a subject. In comparison with the American results the Spanish results are somewhat surprising since American women have traditionally been thought to be assertive and Spanish women have not. In the same vein the respondents in the American survey who were the most unreliable by answering all five false questions (21 % of the total) were mostly men. In Spain, as with the reliable respondents, the percentage of the unreliable was equally distributed among the sexes. The remaining respondents all answered between one and four false questions and the only finding obtained for them is the obvious: that they determined themselves in this way to be unreliable. In further comparison of the total results in America and Spain I noticed a curious tendency of the American respondents to admit ignorance and not answer the first false question, but to indulge in answering other false questions that followed. Perhaps this is a result of American competitiveness. This pattern was not duplicated in the Spanish survey.

The part of my survey based on the telephone directory sampling had what were in my opinion more surprising results in comparing America and Spain. Since America is a far more competitive society than Spain, and the pressure to be informed is much greater, I had expected that throughout the entire survey American respondents would have a much higher overall percentage of false question responses (non reliable) than Spanish respondents who, I thought, would admit their lack of information more freely (reliable). This expectation of a generally higher unreliable American response rate proved not to be valid in these two previous groups. Americans and Spaniards actually had similar percentages except for the one third high school group, and the distribution of the sexes in one of the reliable groups.

9. See questionnaires in the appendix. The participants I used were in America: students of Cape Elizabeth High School, Maine, inhabitants of Portland, Maine, and students of Northeastern University, Boston; in Spain students of the Instituto de Bachillerato Abdera, Almería, inhabitants of this same city, and students of the University of La Rioja, Logroño.

This third sample was most surprising. In fact, my expectations were realised in reverse a second time, as in the higher reliability of American women, and to a much greater degree. The American group had a very high percentage (47 %) of respondents who did not answer any of the false questions (reliable) while the comparable segment in Spain again had a low percentage (16 %). I believe that the high percentage of reliable respondents in the American group resulted because they did not want to waste their time (they did not value the survey) and they thought they would dispose of the interview more quickly if they claimed ignorance. Also, they did not feel that their image was being threatened in answering to an unknown foreigner. The opposite effect might have been at work in the previous questionnaires since they were all either my students or their teacher was present while the interview was being conducted. That would explain a lower American reliability in my previous surveys than in this telephoned group.

Whatever the reasons may be, it is quite clear that the attitudes of both Americans and Spaniards differ in many cases because of their different societies and ways of life. But this difference alone cannot completely explain the sometimes unexpected differences in polling results in the two countries. Reliability seems to be an extremely slippery and unstable quality and becomes even more elusive when a pollster brings his poll to a new society. The rules change! Personal experiences, cultural patterns, and a host of other forces are constantly working on humanity and the resulting opinions and attitudes are not only of individual and collective nature but they travel different paths in different cultures. These patterns of thought and opinion are mercurial to the pollster who tries to understand them in order to suggest courses of action to be implemented in society. When polls must be conducted in different cultures, it is evident that many important and difficult issues must be confronted and problems are exacerbated. Also some basic areas such as sample collection or question wording seem to present little difficulty at first but they are replete with practices that have been proven unreliable.

It becomes evident that a tremendous amount of information germane to the subject and goals of the survey, the populations involved, and the behaviour of people both individually and collectively must be ready and usable before one can hope for fairly reliable results. The results of this particular experiment with five false questions seem to indicate that an explicit directive not to answer items which respondents are not sure of and the practice of conducting interviews in settings that are nonthreatening both encourage as much responsible participation as possible. At least respondent reliability should increase significantly when this is done and the poll results should be more valid.

BIBLIOGRAPHY

- Bishop, Oldendick Tuckfarber & Bennet. 1980. "Pseudo Opinions on Public Affairs." *Public Opinion Quarterly*, Vol. 44. (1980).
- Carlson, Robert. 1975. *Communications and Public Opinion: Public Opinion Quarterly Reader*. New York: Praeger.
- Czaja, Ronald & Johnny Blair. 1997. *Designing Surveys: A Guide to Decisions and Procedures*. Thousand Oaks, California: Pine Forge Press.
- Nieburg, Harold L. 1984. *Public Opinion: Tracking and Targeting*. New York: Praeger.
- Price Vincent. 1992. *Public Opinion*. New York: Sage Publications.

- Tourangeau, Roger & Tom W. Smith. 1996. "Asking Sensitive Questions: the Impact of Data Collection Mode, Question Format, and Question Context". 1996. *Public Opinion Quarterly*. Vol 60, pp. 275-304.
- Wänke, Michaela. 1997. "Comparative Judgements as a Function of the Direction of Comparison Versus Word order". *Public Opinion Quarterly*. Vol 60, pp. 400-409.
- Worcester, Robert M. ed. 1983. *Political Opinion Polling*. New York: St. Martin's Press.