The Eurobarometer: a tool for comparative survey research

Kaase, Max; Saris, Willem E.

Veröffentlichungsversion / Published Version
Sammelwerksbeitrag / collection article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

GESIS - Leibniz-Institut für Sozialwissenschaften

THE EUROBAROMETER

A TOOL FOR COMPARATIVE SURVEY RESEARCH

MAX KAASE AND WILLEM E. SARIS

1.1 Development of comparative survey research in political science

The fourties in the United States witnessed the development and implementation of techniques which were capable of reliably assessing social as well as political attitudes, beliefs and behaviours of large-scale populations through small samples of respondents. The application of probability theory in the development of sampling procedures in connection with the emergence of the standardised personal interview laid the ground for what has since long become a normal tool of social research: surveys with a limited number of respondents, say 2000, producing information that can be reliably generalised to the population from which the sample was drawn, almost independent of the size of that population.

Soon after the end of World War II this methodology spread from the United States to Western Europe where it quickly became the most frequently used instrument in market, social and political research. However, the expertise required to apply this tool intelligently, reasons of cost and not the least a certain conservatism in the academic world for quite a while made surveys a rather scarce phenomenon in sociology and political science. Only as the field of comparative government slowly transformed into comparative politics with its more pronounced interest in political processes and, as one central element in it, citizen orientations, survey data began to assume the kind of important place they now possess in this field (for electoral research as one example see the overview in Thomassen 1994). More than anything else it was the 'Civic Culture' comparative study of political orientations in the United States, the United Kingdom, (West) Germany, Italy and Mexico by Almond and Verba (1963) which revealed the analytical potential of this new methodology for political science.

There are many reasons why this variant of empirical research did not spread much more quickly after the initial success of the 'Civic Culture'. Some of the problems are spelled out succinctly in a book by Szalai and Petrella (1977) discussing five major East-West comparative survey studies which had been conducted in the 60s, partly under the auspices of the Vienna-based European Co-ordination Centre for Research and Documentation in the Social Sciences (Vienna Centre). The experiences from those studies, as summarised in the Szalai and Petrella book, give a good flavour of the theoretical and methodological challenges as well as the practical difficulties this kind of comparative survey research has had to confront (see also Verba, 1971). It is little wonder then that the Political Action study (Barnes et al., 1979)

conducted in seven European nations and in the United States in the 70s (with a replication in three of those countries around 1980; see Jennings et al., 1990) for a long time stood alone in university-based comparative survey research in political science. Since then, the 1981 and 1990 European Values Studies (Harding et al., 1986; Ester et al., 1993) as well as the Beliefs in Government project (Kaase and Newton, 1995) are further hallmarks of comparative survey research and its uses in political science.

What is still missing in this list is the seminal work of Ronald Inglehart on value change in the industrialised democracies of the West (Inglehart, 1977; 1990; 1997). These longitudinal analyses became only possible because Inglehart had been able to influence the content and the availability for secondary analyses of one source that the Commission of the European Communities in Brussels had established in the early 70s as a regular data collection under the supervision of Jacques-René Rabier: the Eurobarometer (Reif and Inglehart, 1991).

The Eurobarometers have been conducted twice a year since the autumn of 1973 in all member states of the European Union (formerly the European Community). As the EU grew in membership, so grew the number of countries in which the Eurobarometer was carried out (at the time of this writing in the spring of 1996 the survey is done in 15 EU member states). The most obvious advantage of these 24 years of regular surveys is its potential to look at change over time; without this data base the Inglehart study of value change would have been impossible, as is true for the analyses in the context of the Beliefs in Government project.

Obviously, the primary interest of the European Commission (as it was renamed after the 1993 Maastricht treaty) in funding these regular surveys was to gain information on the topic most dear to its heart: the process of European unification, as reflected in the attitudes and beliefs of the people in the EU member states (for the most recent, thorough analysis of this field see Niedermayer and Sinnott, 1995). Thus, other information regarded more pertinent for the analytical concerns of the academic social science community could be introduced into the Eurobarometers only in small doses and never covered more than a limited fraction of what one would have liked to find there. Still, there can be no question that by a far margin the Eurobarometers presently are the most valuable data source for comparative survey analysis on Western Europe in political science, a fact witnessed by the enormous number of scholarly publications using these data.

The data of the Eurobarometers can be used in three different ways. The first way is the use of the results in order to make comparisons across time. The second way is the use of the results in order to compare between the different countries. The third is a combination of the two. All three types of use will be illustrated for topics which have been in the Eurobarometer now for more than 20 years.

In figure 1.1 the responses of the people on the questions whether "the membership of the EU was a good thing for your country ?" and whether "your country has benefited of the membership of the EU?". This figure gives the development of the support for the EU by the people in the EU countries from 1981 till 1994. It illustrates clearly that there are considerable changes through time with respect to support for the EU. For legitimacy of the EU policies such data are very relevant. The recent downwards movement gives the politicians of course a lot of concern.

Figure 1.1   Support and perceived benefits of the EU membership from 1981-1994
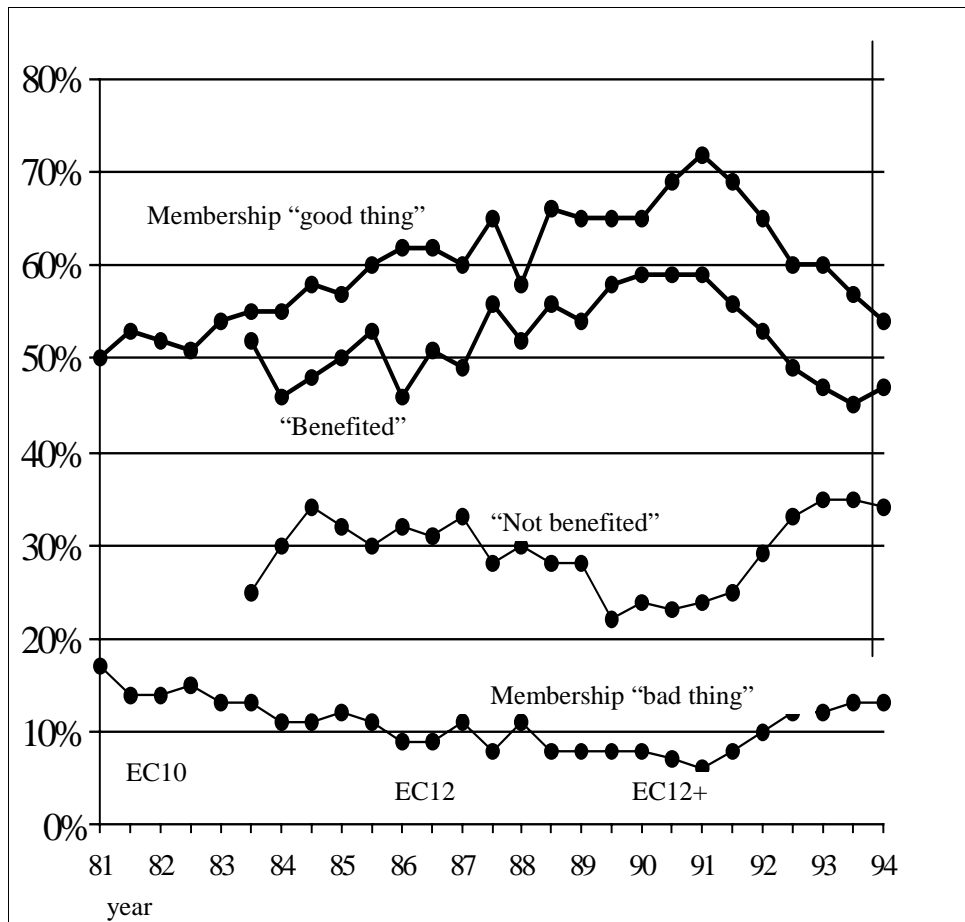
For the illustration of the second type of use we have chosen a very different example namely the answer to the question: "How satisfied are you with your life in general ?" In figure 1.2 the results for this question in July 1994 are given. This figure illustrates nicely the large differences which exist between the different countries with respect to life satisfaction. Also this issue is politically interesting because it is hard to imagine that such differences (if they are real) can exist for a long time in the EU without leading to serious consequences like migration etc.

Finally the third way to use the Eurobarometer data is illustrated by the results on the question " Do you approve the introduction of a single currency in the EU? In figure 1.3 a comparison is made with respect to the answers of the populations of three countries and the EU as a whole for the period 1990 till 1994.

Figure 1.2 Satisfaction with life in general in different countries of the EU

Country	Very satisfied	Fairly satisfied	Not very satisfied	Not at all satisfied	Percentages
DK	97%				3%
L	95%				5%
NL	94%				7%
B	88%				11%
UK	88%				11%
IRL	85%				14%
D	83%				16%
EU12	79%				20%
I	76%				25%
F	73%				26%
E	68%				31%
P	68%				31%
GR	55%				45%

Legend: ■ Very satisfied   ▨ Fairly satisfied   ▦ Not very satisfied   □ Not at all satisfied
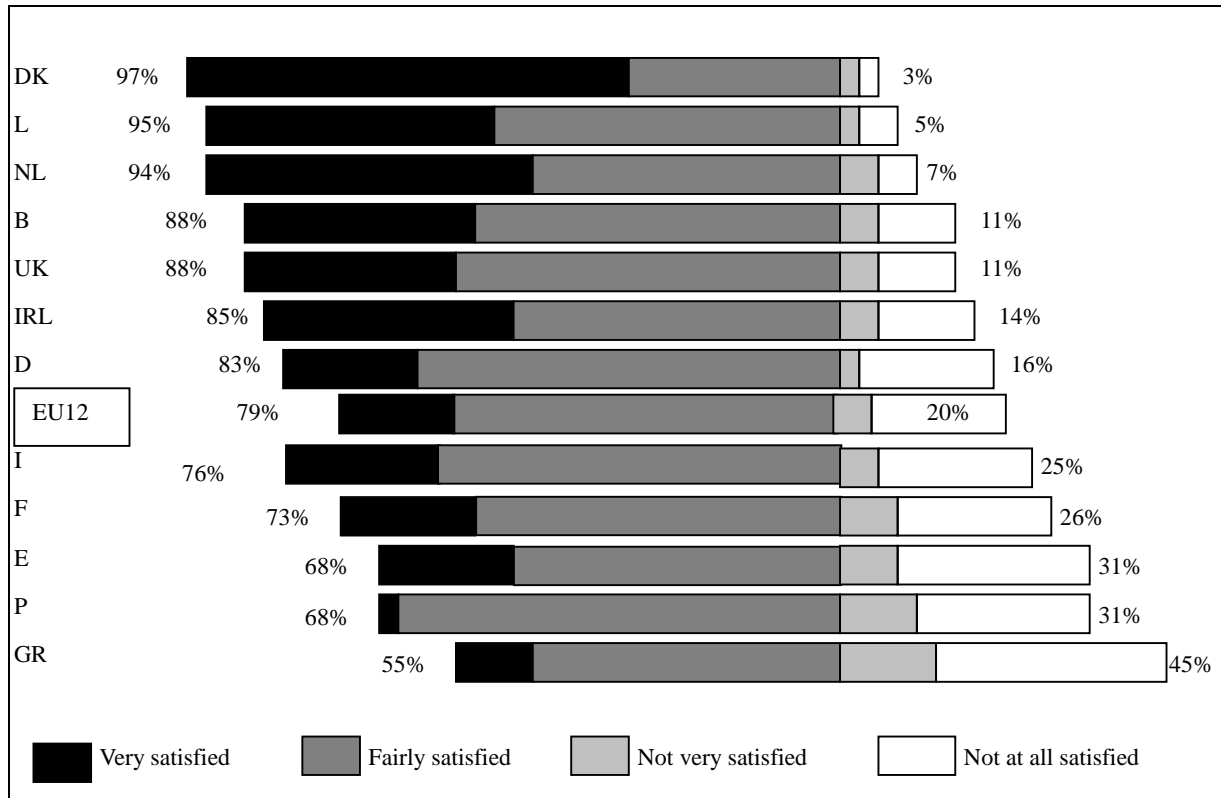
This figure illustrates that all three countries have since 1992 a negative net approval while the EU as a whole is still at the positive side. In figures like these one can see the differences in development of the different countries through time.

As these examples demonstrate this kind of information is very valuable for the European Commission in order to see what the support is for different policies and the EU as a whole in the different countries through time. This information is even so important that the Eurobarometer is undergoing important changes. The information needs of the European Commission are changing in the direction of more short-time data. After having gone through the traumatic - because unanticipated - experience of the narrow margins in supportive post-Maastricht plebiscites in a variety of EU countries (in Denmark, even a second plebiscite became necessary after the rejection of the first), the European commission developed a thorough interest in obtaining information on the swing in public moods in the EU much more quickly than through the inflexible and time-consuming method of face to face interviewing which was used till that time for the Eurobarometer studies.

Figure 1.3   The "net approval " of the single European currency in Denmark, West Germany, UK and the EU as a whole
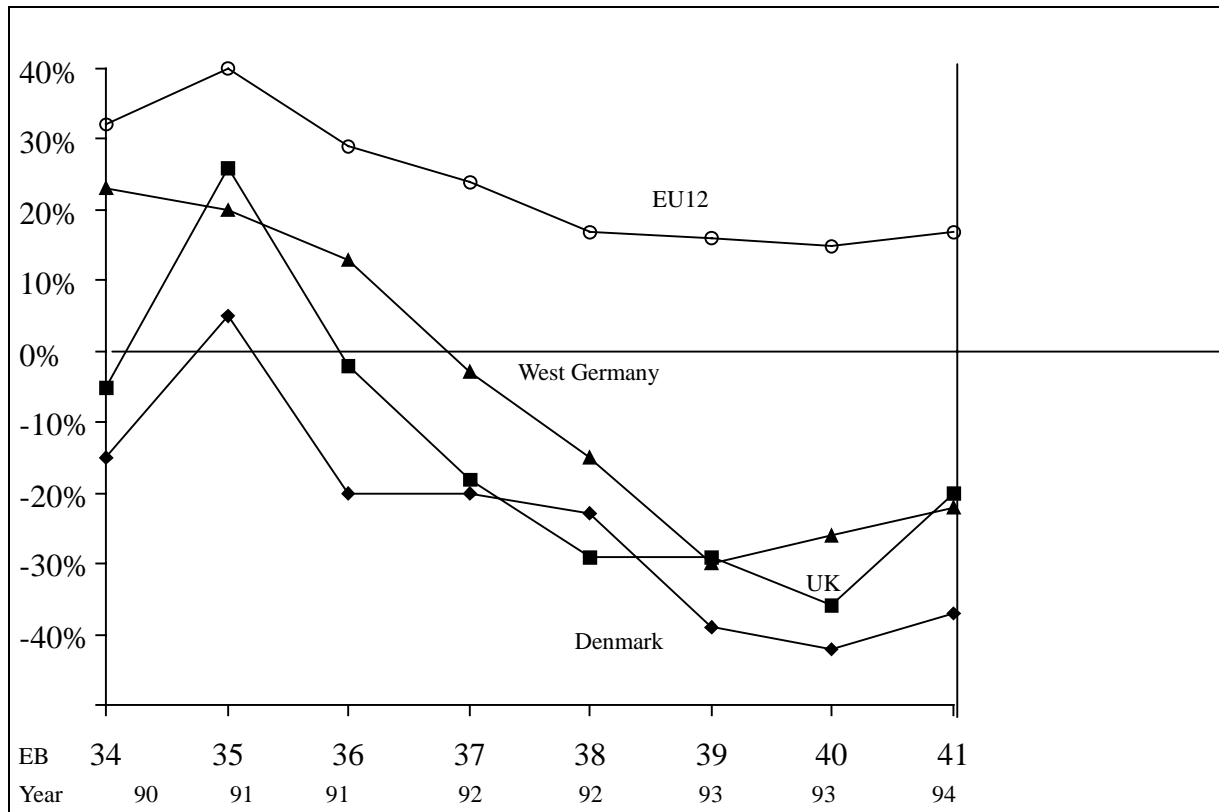
Therefore, the standard Eurobarometer has been extended by a device which reveals changes and new emphases in public moods vis a vis the EU much more quickly than the classical Eurobarometer was able to do. The resulting tracking instrument consists (in principle) of smaller monthly surveys of the people in the member states of the EU. Obviously, this purpose could not be achieved by the traditional face to face interview. Therefore, the method of choice to cope with both problems was to switch to computer assisted telephone interviews.

This choice became possible because - according to Eurobarometer 41 of spring 1994 - only East Germany with about 50 percent, Portugal with about 68 percent and Ireland with about 70 percent are below the 80 percent telephone density mark. In addition, computer-assisted telephone interviewing (CATI) has been perfected to a point where its advantages in being able to do a large number of interviews quickly, to have the collected data immediately ready for analysis and to exert full control over the interviewing process through supervision, are so obvious that this procedure is used in 29% of all quantitative studies in Europe (see ESOMAR, 1996:15).

While the payoffs of using telephone interviewing are apparent, they nevertheless need to be balanced against the requirement to obtain and their ability to produce valid and reliable data

on the populations in question. It is for example questionable if the data remain comparable through time (figure 1.1 and 1.3) moving from one data collection mode to the other. It is also questionable if the results remain comparable across countries (figure 1.2 and 1.3) given that the samples will differ with respect to ownership of telephone.

These are methodological problems which can only be satisfactorily answered by some kind of an experimental research. Therefore, the offer by the Berlin-based FORSA institute under the direction of Manfred Guellner to do a representative telephone survey in the 12 EU member countries in 1994 at the same time the usual face to face Eurobarometer (41) was in the field to test the potential for future telephone interviewing for purposes of the European Commission, was gratefully accepted. On this basis, a complex research design was developed by the scholars contributing to this book to test a set of important methodological issues related to the questions under which circumstances one can expect equivalent and valid information both from face to face and from telephone interviews and how to improve the overall quality of the Eurobarometers.

1.2 Design of the methodological study

Given the relevance of the Eurobarometer data for all people interested in developments in Europe and for social science research, the changes which were expected to occur in the data collection of the Eurobarometer data were sufficient reasons for a number of survey researchers[1] to suggest that methodological research should be done to evaluate the consequences of the change of mode of data collection. The argument for such research was based on the existing knowledge on mode effects. For example Groves (1989) gave 9 possible reasons why one can expect differences between the face to face data collection and the telephone procedures. The reasons given can be condensed as follows:

1. The coverage of the population will be different for face to face interviews from telephone interviews, since those people who do not have a telephone will not be part of the general population from which the sample should be drawn. In Europe, this difference can be substantial because in some areas and countries household telephone density is close to 100% while in other areas and countries the coverage is closer to 50%. In general, it has been found that this bias leads to considerable differences in responses on several dimensions (Groves and Kahn, 1979; Cannel et al., 1987).

2. The field work of the organisations doing the surveys can and usually will be quite different with respect to the interviewers used, their training and supervision, the number of times that a respondent is contacted, and the rules by which a refusal is accepted. Differences in these management aspects will lead to differences in responses and nonresponse and consequently to differences in findings.

3. The mode of data collection itself can also lead to different results. It is possible that people react differently to the same question in a telephone interview than in a face to face

---

1   This idea was brought forward by Roger Jowell, Max Kaase, Hans-Dieter Klingenman, Willem E. Saris and Hermann Schmidt.

interview. For instance, it has been found that open-ended questions result in more elaborated answers in face to face interviews than in telephone interviews. Also, more acquiescence and an extremeness bias might be expected (Groves, 1989). However, the general picture is that these mode effects, after correcting for all other factors, are rather small (De Leeuw and van der Zouwen, 1988).

Also, mode-connected effects are possible, that is effects which might occur due to the fact that changes in the approach are necessary depending on the mode of interview, and that these changes will matter. For example, the use of show cards is not possible in telephone interviews, and as a consequence the procedure for complex questions has to be adjusted. Commonly used in telephone surveys to cope with this problem is a two-step approach where first a small number of crude categories is presented which are later split up into more differentiated ones. The idea here is to obtain the same kind of precision by telephone as in personal interviews where show cards are used to present ten or so categories at the same time. These mode-related changes in the questionnaire can lead to substantial discrepancies in the results, as has been shown by Groves and Kahn (1979), Miller (1984), and Monsees and Massey (1979).

This brief overview indicates that there are three major reasons why a change from face to face to telephone interviews will most likely lead to different results. As indicated above, one can expect that the total difference (T) between face to face and telephone interview responses in percentages or in mean score will be equal to the difference due to coverage (C) plus the difference due to difference in nonresponse (N) plus the difference due to the mode of data collection (M):

$$T = C + N + M \tag{1}$$

This equality should hold exactly for each question separately if populations would be studied. Because samples are used we can not expect a perfect equality. Even if the differences due to coverage, nonresponse and mode would be zero the samples can provide different response percentages because of the randomness of the two samples involved. This difference due to sampling fluctuation can not be studied separately and will therefore be integrated in the effect of the field organisation (M) but later we will test if the differences deviate significantly from zero in order to take this point into account.

Given the possible confusion due to discrepancies in results between the different studies, it has been suggested by those co-operating in writing this book that a methodological study of this issue should be done in order to be able to assess the errors in both procedures more precisely and to find ways to adjust the findings in such a way that the results become comparable. In this chapter, the design will be described which has been chosen by the research group to study the biases in the different approaches and to develop procedures to correct for these biases.

Beforehand, one point should be mentioned: Due to the fact that the Eurobarometer by now has been in use for more than 20 years, one might assume that its findings are valid and reliable, and that deviations from these results in telephone interviews are to be blamed on that method. This is, however, a naive view. Personal interviews have their own errors which can also cause deviations from the population parameters. Of course, this point will be given

some attention in this study, too, but the main emphasis is on the differences in the findings between the two methods.

1.3 The research design

There are two principally different ways to study mode effects. The first approach is experimental and keeps all characteristics of the data collection under control in order to merely study the mode effect (see Ekman, 1965; Mekrabian, 1968; Champness, 1972; Williams, 1974; 1977). Such experimental studies must face the challenge that they are a far distance away from the reality of daily survey research. As Groves (1989) suggests, this methodological approach is mainly aimed at assessing communication effects which is not the primary aim of the present study.

The completely opposite approach, called maximum telephone/maximum personal design, was used by Woltman et al. (1980) and Hochstim (1967); it suggests that one should compare two data collection procedures for the same population and herewith maximise the possibilities for measuring the impact of the two different modes. This can be done with independent random samples or with panel data.

In the present study, the last design could be used because the Berlin-based FORSA research institute offered to collect data for a limited set of questions in all countries through telephone interviewing, while at about the same time the INRA organisation conducted the standard Eurobarometer 41 face to face. This design is a good simulation of the future situation in Europe when two studies will be done about the same time by two different organisations each using a different data collection mode. This design gives the opportunity to estimate the total difference in responses for two specific organisations (T) for all questions present in both studies.

This design, however, has as a major weak point that too many differences will exist between the various approaches, and that one cannot determine which factor possibly causes these differences. Besides, there is the difficulty that all the field work-connected errors are specific for the survey organisations which are involved in the given study. Therefore, should these organisations be exchanged later on, other differences in the results should be expected. Given this lack of strength in the design, also a third approach, a panel element, has been included in this study.

Here, the respondents were first confronted with the normal face to face Eurobarometer questionnaire. In addition, they were asked whether they had a telephone. Those who had a phone were called back after about a week and were asked to respond to a small number of questions already put to them before in the Eurobarometer. This panel design offers better insights into the effects of the two different sources of error, as follows.

First, when the telephone owners and non-owners are compared, an estimate can be obtained of the effect of telephone ownership on the responses to the relevant variables. In this comparison no other variables intervene because the same people are studied and all questions are presented in a face to face interview. So the only possible explanation for

differences is telephone ownership, and thus a good estimate of the coverage error which will occur, is provided (C).

A second effect one can study with this design relates to the mode of data collection since one can compare the answers of the respondents to the same questions in the personal interview and in the telephone interview. This evidence is not so strong as in the case of the comparison of telephone owners and non-owners because there are other factors besides the mode effects which can come into play. One concern has to do with the repeated observations. It is possible that the people want to be consistent by trying to reproduce their answer from a week ago. However, Van Meurs and Saris (1989) have demonstrated that the respondents cannot remember their responses anymore after about 20 minutes of being exposed to questions. In the present study there was at least a gap of one week between the two interviews so that one can expect that memory effects will only play a minimal role, if any.

Furthermore, people may have changed their opinion or behaviour in the time between the first and the second interview. However, for the topics which are asked in the Eurobarometers, such changes are unlikely in a period of one to two weeks unless dramatic events happened but that was not the case in the research period.

A third problem is that the sample of those respondents co-operating in the telephone interview is not the same as the one participating in a normal telephone interview. Selective loss of respondents thus might cause a different group to refuse co-operation after a face to face interview than in a normal telephone interview.

Although these factors all need to be kept in mind, it is nevertheless quite unlikely they will have a strong effect on the existing mode effects. If this assumption is valid, which we think is very likely, then the study of the responses of the same persons in the personal interview and in the telephone interview will provide good estimates of possible mode effects (M).

This design does not allow for an independent estimate of the effects of the fieldwork organisation on the nonresponses (N). However, one could estimate this effect when a combination of a maximum telephone maximum personal design and a panel design is used. The direct comparison of personal interviews with telephone interviews gives an estimate of T. Using the panel design, C and M can be estimated. Using the combination of the two designs, the effect of the difference in nonresponse due to different organisational procedures will be:

$$N = T - C - M \tag{2}$$

It should be mentioned here that the coverage error (C) is an estimate which is for the largest part independent of the organisation that did the research because the effect is determined by the difference between owners and non-owners in the population. This difference will only minimally be influenced by the specific procedure used for data collection, as long as this procedure is not completely flawed.

The same point can be made for the estimate of the mode effect M as was argued above. On the other hand, the estimates of N and T are clearly determined by the organisations which perform the studies. The total difference varies directly with the difference due to

nonresponse which is produced by the two organisations in question. So general statements are difficult to make about these two components although they can be properly assessed for a specific case.

Furthermore, the estimates of the coverage error and of the mode effect can also vary with the questions being used. Telephone non-owners can differ in their opinions on certain questions, and this will lead to differential effects although for other questions the differences can again be very small. In the literature, some questions have been mentioned to be more effected than others, like open-ended questions, questions requiring a heavy cognitive burden such as long questions or questions with a large number of categories, and also questions which are normally asked with a show card, a procedure presently not available in telephone interviewing. Given the effects of the type of question asked, it will be necessary to discuss this problem in the present mode effect study, too.

1.4 The questions used

The questions which have been asked in the different data collections of this study can be subdivided into five groups. Below, an overview is given of the formulation of the questions of the Eurobarometer which have also been chosen for the telephone interviews. Adjustments and small differences in the questions for the telephone interviews are not documented here.

Group 1: The standard Eurobarometer questions

The first group consists of core standard Eurobarometer questions. Obviously, the results from these questions should be equivalent across the different studies. The following questions were chosen from the pool of all possible questions as the most important ones.

Life satisfaction
On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with the life you lead? Would you say you are (READ OUT)

very satisfied
fairly satisfied
not very satisfied
not at all satisfied
——————
DK/No answer

Satisfaction with democracy
On the whole, are you very satisfied, fairly satisfied, not very satisfied or not at all satisfied with the way democracy works in (your country)? Would you say you are (READ OUT)

very satisfied
fairly satisfied
not very satisfied
not at all satisfied
——————
DK/No answer

Frequency of political discussions
When you get together with friends, would you say you discuss political matters frequently, occasionally, or never?

frequently
occasionally
never
——————
DK/No answer

Motivation to persuade others
When you hold a strong opinion, do you ever find yourself persuading your friends, relatives or fellow workers to share your views? Does this happen (READ OUT)

often
from time to time
rarely
never
——————
DK/No answer

Media exposure

About how often do you (SHOW CARD)
a)   watch the news on television
b)   read the news in daily papers
c)   listen to the news on the radio

	every day	several times a week	once or twice a week	less often	never	DK/ No answer
News						
on TV	1	2	3	4	5	6
in daily papers	1	2	3	4	5	6
on the radio	1	2	3	4	5	6

Interest in EU politics
To what extent would you say you are interested in European politics, that is to say matters related to the (EC/EU): a great deal, to some extent , not much, or not at all?

a great deal
to some extent
not much
not at all
——————
DK/No answer

Subjective sense of political information
All things considered, how well informed do you feel you are about the (EC/EU), its politics, its institutions? (READ OUT)

very well
quite well

not very well
not at all well
─────
DK/No answer

Evaluation of EU membership for respondent's country
Generally speaking, do you think that (our country's) membership of the (EC/EU) is (READ OUT)

a good thing
a bad thing
neither good nor bad
─────
DK/No answer

Benefit of EU membership for respondent's country
Taking everything into consideration, would you say that (our country) has on balance benefited or not from being a member of the (EC/EU)?

benefited
not benefited
─────
DK/No answer

These questions have been asked in many Eurobarometers and will also be asked in the tracking studies. It is therefore necessary to know whether differences in responses will occur and whether they can be corrected.

Group 2: Different types of simple questions

There is also a number of questions which have not been asked literally in the same way in each Eurobarometer and will not be asked in each tracking study, but will continue to be asked occasionally in the future. These questions concern opinions on specific problems and knowledge about certain political issues. The content of these questions can be changed, but their format will usually not be changed. Therefore it makes sense to study the effect of these formats in the different modes. A distinction has been made between simple questions which are presented in this section, and complex questions which are covered in the next section.

Opinions
What is your opinion on each of the following proposals? Please tell me for each proposal, whether you are for it or against it. (READ OUT IN ROTATING ORDER)

a) There should be a European Monetary Union with one single currency replacing by 1999 the (national currency) and all other national currencies of the member states of the (EC/EU).
b) The (EC/EU) member states should work towards a common defence policy.
c) Any citizen of another (EC/EU) country who resides in (your country) should have the right to vote in local elections.

d) Any citizen of another (EC/EU) country who resides in (your country)should have the right to vote in European elections.

e) Any citizen of another (EC/EU) country who resides in (your country) should have the right to be a candidate in local elections.

f) Any citizen of another (EC/EU) country who resides in (your country) should have the right to be a candidate in European elections.

g) The (EC/EU) should be responsible only for matters that cannot be effectively handled by national, regional and local governments.

Knowledge of White Paper

Have you ever heard about the "White Paper" by the European Commission in Brussels about growth, competitiveness and employment in Europe?

Yes
No
———————
DK/No answer

Information on European Parliament

Have you recently seen or heard, in the papers, on the radio or on TV, anything about the European Parliament?

Yes
No
———————
DK/No answer

Knowledge about date of the next European elections

Do you know the date on which the next European election will take place in (your country), or not? (IF YES) On which date?

Yes, and the correct date
Yes, but the date mentioned is not correct
No, does not know the date

Only the last of these questions was included in the FORSA study. This means that for most of the questions the coverage errors and the mode effects can just be studied by using the answers of the respondents in Eurobarometer 41 and its panel component.

Group 3: Complex questions

Three more complex questions have been chosen for further study because such questions are likely candidates for mode effects (Groves 1989). These questions concern vote probabilities, reasons why people may not vote, and left-right self-placement.

The question on the probability to vote has a very difficult formulation which can lead to problems in telephone interviews.

The questions on the reasons for non voting can lead to problems not because of the complex formulation of the question but because of the complex response categories which are normally presented on a show card and have to be read out in a telephone interview.

Vote probability
In June 1994, the citizens of countries belonging to the (EC/EU), including the (nationality of the respondent), will be asked to vote to elect members of the European Parliament. If there were such a "European election" tomorrow (for respondents under 18 years old add: "and you had a vote") would you certainly go and vote, probably go and vote, probably not go and vote, or certainly not go and vote? (IF VOTING IS COMPULSORY IN THE COUNTRY; ADD) "if the vote was not compulsory in our country"

Will certainly go and vote
Will probably go and vote
Will probably not vote
Will certainly not vote
Other answer (SPONTANEOUS)
―――――
DK/No answer

Reason for non-participation (only asked of those respondents who are not certain to vote)
What is the main reason why you might not go and vote at the next European elections in June 1994? (SHOW CARD - ONE ANSWER ONLY)

I am not interested in politics or elections
I am not interested in European elections
I lost interest in European matters
I have always been against Europe
I am against even more Europe
Not well enough informed to vote in European elections
I think the result is a foregone conclusion
Other reasons (SPONTANEOUS)
―――――
DK/No answer

Left - right position
This question in face to face interviews is often asked on a ten point scale ranging from extreme left to extreme right, the respondents being provided with a show card with the ten possibilities. For telephone interviews it has been suggested that the use of a two step question would be functionally equivalent (Miller, 1984; Monsees and Massey, 1979). This formulation, however, has also its risks. Sykes and Hoinville (1985) did not find differences, but Groves and Kahn (1979), when looking at other questions, did find that differences can occur.

We have introduced the two-step procedure in a split ballot experiment in the face to face study. If differences were detected already in the identical mode, one would not have to try the same question in a telephone survey because then the question formulation were the cause of deviations and not the mode of data collection. To test this effect, in the face to face

interview one of the following questions has been given to two randomly selected sub-samples.

Split Ballot A
When people talk about politics, the terms " left" and "right" are always used. We would very much like to ask you, whether you put yourself as rather "left" or rather "right"?

Rather "left"
Middle/neither nor (SPONTANEOUS)
Rather "right"
Refusal
———
DK/No answer

When rather 'left'
Please imagine for a moment a scale, from 1 to 5, where 5 means 'very left' and 1 'not very left'. Where would you put yourself?

	not very left				very left
	1	2	3	4	5

When rather 'right'
Please imagine for a moment a scale, from 1 to 5, where 5 means 'very right' and 1 'not very right'. Where would you put yourself?

	not very right				very right
	1	2	3	4	5

Split Ballot B
In political matters people talk of "the left" and "the right".
How would you place your views on a scale from one to ten, where 1 means very much to the left, and 10 very much to the right? You may, of course, use the numbers in between 1 and 10 in order to shade your opinion. (SHOW CARD. DO NOT PROMPT. IF CONTACT HESITATES, ASK TO TRY AGAIN)

left									right
1	2	3	4	5	6	7	8	9	10

No statement
DK/No answer

Group 4: Open-ended questions

Open-ended questions were introduced in the tracking study because they provide information about developments in public opinion without cues by the researchers through the use of response categories.

For the methodological study this type of question is interesting because it has been found that people provide less information in telephone interviews than in face to face interviews on open-ended questions (Groves and Kahn, 1979; Kormendi, 1988). The explanation given is

that telephone conversations are expected to be shorter. The consequence, however, can be that the responses differ in the different studies. Since these questions are of great concern for the tracking study, they have been asked in the two waves of the Eurobarometer 41 panel as well as in the telephone study of FORSA, thereby enabling comparisons between the two data collection modes.

Most important problem in our own country
Generally speaking, what is the most important problem facing (our country) today?

Most important problem in EU
Generally speaking, what is the most important problem facing (EC/EU) today?

Group 5: Background variables

Obviously, also a number of background variables have been asked in all three data collections. Here we mention only those questions which are comparable.

Age
How old are you (YEARS OF AGE)
(FORSA classification: 18-29; 30-44; 45-59; 60 and older)

Household goods
Do you or anyone else in your household own? (READ OUT)

Colour tv
Video recorder
Video camera*
Radio clock*
PC/ home computer
Still camera*
Electric grill*
Electrical deep fat fryer*
2 or more cars
Second home or a holiday home/flat
————

* not asked in both telephone studies

Social class
If you were asked to choose one of five names for your social class, which would you say you belong to? (SHOW CARD)

Middle class
Lower middle class
Working class
Upper class
Upper middle class
Refuse to be classified
Other
————

DK/No answer

After the questions have been specified, the existing options for analysis can be summarised in Table 1.1. This table shows that for the most questions all three kinds biases defined above can be studied. Only for the group 2 questions (simple questions) the analysis is limited to the coverage error and the mode effect because these questions have not been asked in the FORSA study. This table indicates that the chosen design allows for a large number of analyses.

1.5 Plan of the book

The book has started with an introductory chapter discussing the relevance of the topic and the design of the study. In the next four chapters the research design of the face to face and telephone study are evaluated, and the quality of these designs, the differences in sampling and their effects are analysed.

The second chapter will give a description of the research design for both studies. The basis of this chapter are the descriptions given by INRA and FORSA, the organisations which organised the data collections, about their procedures. This chapter provides basic information for anybody who uses the standard Eurobarometer or the tracking studies with respect to the way the data is collected. This description is more detailed than can be found anywhere else in the literature

The third chapter shows that both sets of samples differ considerably from the sets of populations they want to describe, on the demographic variables for which population information is available. Furthermore it is indicated how the different samples can be reweighed using optimal weights. Also, the efficiency of the various procedures and the quality of this reweighing for substantive variables is demonstrated. This chapter provides users of the different EB's with information how the unavoidable deviations in the background variables can be corrected and what the consequences of reweighing on other variables are.

Many people think that face to face studies can not be compared with studies using telephone interviewing. The fourth chapter shows how different the telephone owners and non owners really are with respect to a number of background variables and some substantive variables. The chapter indicates that there are indeed considerable differences. However, these differences will be placed in perspective in the next chapter.

In the fifth chapter of the first part of the book the estimates of the total difference between the two methods are decomposed in the following components: the effects of the coverage error, the nonresponse differences and the mode differences. These components have been estimated for three different countries for which the data are available: Belgium, France and Spain. Other than general opinion has it, the coverage error (due to the differences between owners and non-owners of telephones) is rather small while the differences due to the use of a different mode of data collection and of the differences in fieldwork of the two companies conducting the surveys is rather large. This result is of more general importance than only for this specific study.

Table 1.1 The possibilities for analysis in the study

Questions	EB 41.0	EB 41. Panel	FORSA	Possibilities for analysis
Standard questions				
Life satisfaction	+	+	+	T,C,M,N
Satisfaction with democracy	+	+	+	T,C,M,N
Political discussion	+	+	+	T,C,M,N
Persuade others	+	+	+	T,C,M,N
Political news (3x)	+	+	+	T,C,M,N
Interest in EU politics	+	+	+	T,C,M,N
Informed about politics	+	+	+	T,C,M,N
EU membership	+	+	+	T,C,M,N
EU benefit	+	+	+	T,C,M,N
Simple questions				
Opinions (7x)	+	+	-	C,M
Knowledge of white paper	+	+	-	C,M
Knowledge of parliament	+	+	-	C,M
Knowledge of elections	+	+	-	C,M
Complex questions				
Vote probability	+	+	+	T,C,M,N
Reasons for non-participation	+	+	-	C,M
Left-right 10 point scale	+	+	+	T,C,N
Left right 2 step scale	+	+	-	C
Open ended questions				
National problem	+	+	+	T,C,M,N
EU problem	+	+	+	T,C,M,N
Background questions				
Age	+	+	+	T,C,M,N
Household goods (4x)	+	+	+	T,C,M,N
Social class	+	+	+	T,C,M,N

In the second part of this book the attention goes to the comparability of the results across modes of data collection and across countries.

In the sixth chapter the latent class response model will be introduced, and a test will be done in order to evaluate how the mode of data collection effects the standard barometer questions. This is a very general issue in the literature. In some studies large differences have been found, in others not. Thus, it seems that much here depends on the questions. Therefore, the user has to know whether such effects occur when comparisons are made across studies using different modes of data collection.

The seventh chapter discusses the mode effect for a specific question: the open ended question on the most important problem facing one's own country and the EU. For this analysis the development of a general codebook for different countries has to be discussed first. Then a description of the results will be given comparing the results of the three studies in order to detect if the specification of the different problems is different for the different modes.

In the eighth chapter the possible differences between a 10 point scale for left-right orientation in a face to face and in a telephone interview will be discussed. Next, the consequences of the two step procedure for such questions will be considered on the basis of the available experimental data. This approach has been suggested being most adequate for telephone interviewing and is used in recent studies. In this chapter we show that this was a right or a wrong decision.

In the ninth chapter analyses have been done with respect to the comparability of the results across modes of data collection and across countries. It was found that for some questions comparisons across modes and countries can be made without any problem. For other questions it is found that the results can be compared across countries if the same mode of data collection is used. For other questions, like the frequently used satisfaction questions, it was found that and is explained why comparison across modes and across countries is not possible. Obviously, this result is highly relevant for researchers using these data.

In the last chapters ways are proposed which help to make the results obtained with the different designs as comparable as possible. This has to be done by two separate corrections: one for the sample differences, and one for the response mode effects.

In the tenth chapter procedures are presented to correct for sample differences and mode effects which have been discussed before. This is also illustrated by applying them to the results of the INRA and FORSA surveys.

Finally, in chapter 11, a summary of the results and practical recommendations will be given for users of the questions and data of the standard Eurobarometers and the tracking study with respect to the way in which they should use the data. Furthermore, recommendations will be formulated for correction of some of the questions in order to make them more comparable across modes of data collection and/or across countries.