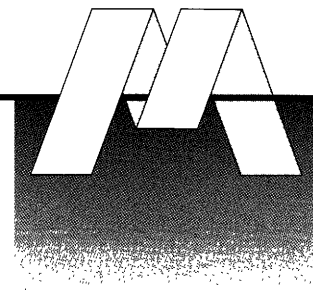


COMMISSION
OF THE EUROPEAN
COMMUNITIES

SCIENCE
RESEARCH AND
DEVELOPMENT



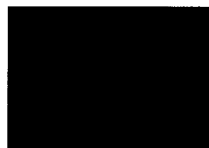
MONITOR

SPEAR

**A Guideline for Survey-Techniques
in Evaluation of Research**

Research evaluation

EUR 14.339 EN



Commission of the European Communities

A Guideline for Survey-Techniques in Evaluation of Research

Authors : S. Lahlou, *et al* (CREDOC)
: R. van der Meijden (STB-TNO)
: M. Messu (CREDOC)
: G. Poquet (CREDOC)
: F. Prakke (STB-TNO)

MONITOR/SPEAR Coordinator
: F. Sand (CEC-DG XII)

March 1992

**Published by the
COMMISSION OF THE EUROPEAN COMMUNITIES
Directorate-General
Telecommunications, Information Industries and Innovation
L-2920 LUXEMBOURG**

LEGAL NOTICE

Neither the Commission of the European Communities nor any person acting on behalf of the Commission is responsible for the use which might be made of the following information

Cataloguing data can be found at the end of this publication

Luxembourg: Office for Official Publications of the European Communities, 1992

ISBN 92-826-4093-0

Catalogue number: CG-NA-14339-EN-C

© ECSC—EEC—EAEC Brussels • Luxembourg, 1992

Printed in Belgium

Contents

INTRODUCTION

- | | |
|--------------------------|------|
| 1. Objectives | p. 5 |
| 2. Method | p. 6 |
| 3. Purpose of the report | p. 6 |

I. PROCEDURES FOR THE SET-UP OF A QUESTIONNAIRE

- | | |
|---|-------|
| 1. Why use a questionnaire ? | p. 8 |
| 2. Criteria for evaluation of EC programmes | p. 8 |
| 3. What kind of survey ? | p. 10 |
| 4. Using sampling techniques | p. 13 |
| 5. Language to be used in questionnaire | p. 15 |
| 6. The use of data already available | p. 15 |
| 7. Confidentiality | p. 16 |
| 8. Response | p. 16 |
| 9. Initial briefing survey-organisations | p. 21 |

II. PROCEDURES FOR ANALYSIS AND PRESENTATION

- | | |
|----------------------------|-------|
| 1. Analysis of data | p. 24 |
| 2. Presentation of results | p. 25 |

III. QUESTIONNAIRE TECHNIQUES AND RULES

- | | |
|-------------------------------------|-------|
| 1. Three basic concerns | p. 31 |
| 2. The nature of the questions | p. 35 |
| 3. The form of the questions | p. 37 |
| 4. Question formulation | p. 44 |
| 5. Examples of question formulation | p. 51 |
| 6. The actors of an evaluation | p. 54 |

CONCLUSION

- | | |
|--|-------|
| The three phases of drafting a questionnaire | p. 60 |
|--|-------|

Annexes

- | | |
|--|-------|
| Annex 1 : Example of procedures for using internal data : AMPERE of DG XII | p. 63 |
| Annex 2 : List of the evaluation reports analyzed | p. 71 |

INTRODUCTION

1. Objectives

The R&D-programmes of the European Communities are periodically evaluated by an evaluation committee. This committee, consisting of independent experts, is appointed by the Commission, and is based on the proposals made by the Evaluation Service. Most evaluations use questionnaires to gather information needed for the evaluation. Most of the time this information gathering is subcontracted to a third party (e.g. a consulting or research institute). These institutes develop the questionnaires independently on an ad-hoc basis. The fact that a new questionnaire is developed for every evaluation, even though a substantial amount of information requested is the same for the different programmes, results in a large variety in the content and lay-out of the questionnaires. This heterogeneity in the questionnaires hinders cross evaluation of the different studies. It also results in higher costs for each individual evaluation due to design costs and new procedures for coding and analysis. The same aspects can be mentioned for the survey reports ¹. There is also no standard procedure for analysis and presentation of the results. In some reports tables are used. Other reports only show graphics. There is also no consistency in the presentation of the data : in round figures or in percentages. In some survey reports the use of a questionnaire is not even mentioned.

The objective of the programme MONITOR/SPEAR is to improve the tools for the evaluation methodology. The aim of this study is to supply the Commission with a guideline for more standardized procedures for evaluation-surveys. Standardization does not mean that there is one uniform methodology for evaluation-surveys. **It does mean that in comparable situations comparable techniques will be used.**

The guideline can also be useful for other directorates of the European Commission, which have to deal with evaluation surveys, and also for other organisations which have to deal with evaluations (eg. national ministries).

The study has been jointly executed by the Centre for Technology and Policy Studies - TNO (STB-TNO) and the Centre de Recherches pour l'Etude et l'Observation des Conditions de Vie (CREDOC).

¹ The survey reports are made by an independent organisation on behalf of the evaluation committee to help them to complete the evaluation.

2. Method

The guideline will be divided into three parts. The first part concerns the procedures for setting up a survey. In this part we shall deal with aspects, such as, whether to use a written, telephone or face-to-face questionnaire, types of questions, language, etc. The second part will deal with the analysis and presentation of the data. The final part concerns the questions asked. We will suggest standard questions which can be used for most programmes and how to deal with programme specific questions.

The information needed to compile the guideline is only partly available in the evaluation reports of earlier studies on behalf of the Commission. Therefore, most of the information in this guideline is based on the expertise of STB-TNO and CREDOC in the field of evaluation-surveys. Additionally there have been some interviews with programme-managers, evaluation-secretaries and an external research institute concerning the usefulness of standardization.

The guideline also includes comments for the use of complementary information, already available within the Commission, that can be used for evaluation. In DG XII for instance, AMPERE data can be used for the questionnaire and for the evaluation itself.

3. Purpose of the report

The guideline will not be an ultimate guide for the setting-up of evaluation surveys. There is too much variety in the content and participants of European R&D-programmes to strive for an ultimate guide. **This guideline will help the evaluation team to make consistent choices** on the best way to carry out the field-work needed and the best practice for analyzing and presentation procedures. Future experience with the guideline will offer opportunities to update the guideline towards the specific needs that the Commission may have.

The guideline is aimed at persons who have to deal with surveys but have no special knowledge of survey-techniques. In this respect, the guideline can be useful for the members of the evaluation panels and for the persons within the Commission coordinating the evaluations. Furthermore, the guideline can be useful for those who have to occasionally deal with surveys, and for persons and organisations not working for the Commission, but who have to deal with evaluations.

To assist the reader in using this guideline, each chapter will start with a short summary on the highlights.

I. PROCEDURES FOR THE SET-UP OF A QUESTIONNAIRE

Summary

- * Standardization of evaluation-procedures will reduce cost and time and facilitate cross-evaluations between different programmes (p. 8).
- * Criteria for evaluation used in a questionnaire must be simple, and easy to answer for the respondent (p.9).
- * The type of interviewing depends on the information to be collected, the time-period for fieldwork and the size of the population.
In-depth interviews are useful when little knowledge is available and the information is complex (p. 10).
Structured face-to-face interviews are useful for more lengthy questionnaires and, when there is enough information available, to structure the questionnaire (p. 11).
Telephone interviews are very useful for screening the respondents or for very simple interviews (p. 12).
Written questionnaires are useful if the respondent does not have the information directly available (p. 12).
- * Sampling techniques are not very useful for most EC-programmes because in most evaluations the entire population, which is often very limited, will be interviewed (p. 13).
- * For most R&D programmes, the questionnaire can make use of simple and common English. For programmes aimed at small enterprises or agriculture the questionnaire must use native languages (p. 14).
- * It is advisable to include a priori in the questionnaire data available directly from programme management. It shortens the time needed to fill in the questionnaire and its personalisation will have a positive impact on the response rate (p. 15).
- * In respect of confidentiality, we can follow one of three procedures (pp. 16-17).
- * There are several ways to increase the response rate (pp. 18-20) :
 - Keep the questionnaire brief and ask only the relevant aspects.
 - Make use of experienced interviewers only when conducting personal or telephone interviews.
 - Prenotification of the interview by telephone screening or by letter.
 - Make a written questionnaire visually appealing with a logical routing, and use only simple and common words.
 - Enclose a new written questionnaire if a reminder is send out.
- * We advise against the use of monetary incentives or presents aiming to increase the response-rate in evaluation surveys (p. 20).
- * The briefing of the organisation that will do the survey must be clear about aspects such as timing, size of the population, type of respondent, etc. (p. 21).

1. Why use a questionnaire ?

The European R&D-programmes show a great variety in content and participants. There are programmes which stimulate complex technological research with just a few participants (e.g. the evaluation of Nuclear Reactor Safety Programme in 1986, report no. 18, concerned only 33 projects), which can only be judged on their merits by experts. Methods like peer review and physical visits are not much of a problem. Other programmes (e.g. SCIENCE or Brite) have many participants. Especially for those programmes with a lot of participants, or projects, who or which cannot all be visited (physically and in terms of time and money), the evaluation team will need to conduct a survey to gather the information needed.

Despite the great variety, it is not necessary to develop each survey and questionnaire on a pure ad hoc basis. This guideline will allow the evaluation team the possibility to **use standard procedures** for the setting-up of the survey and the questionnaire. Standardization of the procedures has several advantages. First, it will **reduce the costs** involved in developing a questionnaire. Secondly, it will **reduce the time** needed for the development. Thirdly, it will **facilitate cross evaluations** of different programmes. This last aspect is of great importance (also expressed by the programmes-managers within the Commission) because each programme will be evaluated in a comparable way and on comparable aspects. This will increase the fairness of the evaluation.

In this part we will discuss the following questions :

- criteria for evaluation;
- what kind of survey;
- interview the total population or use sampling techniques;
- the language to be used;
- confidentiality;
- the use of the input of information already available;
- how to deal with non-response;
- briefing of the survey-organisation.

2. Criteria for evaluation of EC programmes

A programme is generally evaluated on three levels of objectives; 1) the overall objectives of the Framework Programme, 2) the general objectives of the specific programme and 3) some specific objectives within the programme, corresponding to

special projects or sub-programmes (eg. The general objective of SPRINT is the stimulation of innovation and technology transfer. One specific objective is to support the development of science parks). The first level of objective is the most difficult to measure while the third is the easiest. **In a questionnaire one must restrict oneself to the objectives that are measurable (see table below) and also recognizable to the respondents.** A second restriction is that a programme must be evaluated on the aspects that can be influenced by the programme-managers or the Commission. One complaint by the programme managers is the fact that programmes are also evaluated on procedures used by the Commission. It is common knowledge that participants will always complain about the procedures (too much bureaucracy, uncertainty, lengthy procedures, etc.), which can cause a negative evaluation. The fact is that these complaints can be checked on the internal data available to see whether are reasonable or not (see Part I.6). However, the most important goal for evaluations is to judge to what extent the Programme has achieved its objectives.

The table below shows an overview of different types of programmes, the intended objectives and possible criteria that can be used in a evaluation-questionnaire.

<i>Type of programmes*, their objectives and examples of measurable criteria to be used in the questionnaire</i>		
Type of programme	Objectives	Criteria for questionnaire
Stimulate scientific research	Scientific progress	number of presentations on conferences, number of publications, to what extent would the research have taken place without the programme...
Stimulate pre-competitive research	Technological progress	Criteria depends on the testability of the specific objectives within the programme
Demonstration projects	Diffusion	Did diffusion take place, to how many organisations, what type of organisations, which countries...
Stimulate international cooperation	Increased cooperation	Increase in number of international contacts and in which countries...
Stimulate the use of large installations	Facilitate the use for other researchers	Number of users per installation, would research have taken place without this programme, users from which countries...

* Each of these types of programmes is not exclusive from another.

Although the listed projects will not cover all types of European R&D-programmes and some programmes are a mixture of different types, the listing shows examples of criteria that one could possibly use in a questionnaire. The criteria must be simple and easy to answer for the respondent, and equal for all respondents.

3. What kind of survey ?

Basically one can choose between four types of information-gathering : the in-depth interview (a personal interview using only a rough guideline), the face-to-face interview (a personal interview using a structured questionnaire), the telephone interview and the written interviews. The first type is called the informal method for data-gathering. The last three types of interviewing are formal methods which use a uniform and (semi-) structured questionnaire, a standardized method of the interview itself, and have the possibility for sampling. Each type of interview has specific advantages and disadvantages and can be used in specific situations. Cultural differences between countries can also influence the type of data-gathering in different countries (*see 3.4.*) We shall give a brief description of each type of interview, the main (dis)advantages and in which situation one can use a specific type.

3.1. In-depth interview

The in-depth interview (or the informal method for data-gathering) makes use of a roughly defined interview guideline, which only lists the aspects to discuss. **This type of interview is used when the subject is very complex and there are no possibilities for pre-defined measurable criteria.** It must be conducted by an experienced interviewer. It should not be used when the data is not directly available to the respondent.

Advantages :	<ul style="list-style-type: none">- Rich information- Flexibility during the interview
Disadvantages :	<ul style="list-style-type: none">- Costly- No equality between interviews- Limited number of respondents- No routine-procedures for data-processing available
Use :	<ul style="list-style-type: none">- Conceptualizing measurable criteria- Formulation of hypothesis

This type of interview is useful for programmes with a small number of participants and complex projects.

3.2. Structured face-to-face interview

Structured face-to-face interviews are useful when there is sufficient information for the development of a structured questionnaire. **It is used when the duration of the interview is expected to be long, the interview requires material to be shown or the questionnaire is complex and the respondent may require an explanation.** The data asked for must be immediately available from the respondent.

- | | |
|------------------------|---|
| Advantages : | <ul style="list-style-type: none">- Allows complex questionnaires- Allows the presentation of materials- Structured procedures for data-processing possible- Sampling techniques possible- Control over timing of the fieldwork |
| Disadvantages : | <ul style="list-style-type: none">- Costly- Possibility of interviewer-bias- Confidential/private questions will hardly be answered- Limited flexibility in interview |
| Use : | <ul style="list-style-type: none">- Complex questionnaires of a non-confidential nature- Questions concerning attitude; spontaneous reactions are wanted, or the order of the questions is important |

For a few years there has been an increasing use of computer interactive interviewing in the face-to-face interviews. Interviewers use laptops for the interviews. Use of these new techniques has several advantages, such as complete control of the routing in the questionnaire, making complex routings possible, and increased speed of data-collection and -processing ².

² See : Bje de S.E., I.A.L., Stoop and K.L.M. de Vries, "CAI software ; an evaluation of software for computer assisted interviewing", VOI uitgeverij, Amsterdam, 1989.

3.3. Telephone interview

Telephone interviewing has become quite popular in recent years. It is relatively cheap and due to the high penetration of telephones in companies and households there is hardly any drawback on the sampling aspect. However **telephone interviewing has a limited use and is often used wrongly.**

Advantages :	<ul style="list-style-type: none">- Relatively cheap- Allows sampling techniques- Control over timing of fieldwork- Standardized procedures for data-processing
Disadvantages :	<ul style="list-style-type: none">- Limited length of interview (about 10 to 15 minutes)- Only simple straightforward questions to be used- Limited explanation possible- No use of presentation material possible- Requires high telephone penetration
Use :	<ul style="list-style-type: none">- Identification/screening of respondents in complex organisations- Reminder for written questionnaire- Gathering of simple or additional information

3.4. Mail survey

The last type of “interviewing” is the use of the written questionnaire. This form is especially popular in the academic world. The written questionnaire is perceived as the cheapest way to interview. But if the reminder cost and the effects of non-response are considered, the price-quality ratio may turn out to be very low, particularly when the selection and the writing of questions has not been subjected to careful preparation. Another important aspect, especially in international studies, is the reliability and speed of postal services, which can differ greatly between countries. Therefore the use of **written questionnaires must be considered with caution.**

Once chosen, one should be extremely careful with the visual aspects in the questionnaire and the complexity of the routing. Therefore written questionnaires must be thoroughly pre-tested.

Advantages :	<ul style="list-style-type: none"> - (Sometimes) cheap - Sampling techniques are possible but there is hardly any control over the response so the original sample can be completely ruined. - Pre-explanation possible - Standard procedures for data-processing available
Disadvantages :	<ul style="list-style-type: none"> - Low response rate - No control on information given - No control over timing fieldwork - Unclear who will fill in questionnaire (delegation)
Use :	<ul style="list-style-type: none"> - Gathering of administrative data - Gathering of confidential or sensitive data ³

Many studies require a combination of interviewing techniques. If the criteria for evaluation are not known a few in-depth interviews may supply the necessary information. In a few face-to-face interviews these concepts for criteria can be tested and a telephone or written interview can be the final survey. Another frequent combination which is used is a telephone screening for respondents followed by a written questionnaire. A telephone screening can be extremely useful when a survey has a 'split-run' character (this means that different versions of questionnaires are used in the same study, e.g. different questionnaire for participants and non-participants in a programme). The type of combination to be used depends on the knowledge already available and the type of data to be gathered.

4. Using sampling techniques

Especially when there is a large population, sampling techniques can be considered . For large populations it is not always necessary to question all the respondents. Advanced use of statistics can reduce the number of respondents needed to draw conclusions. By using sampling techniques it is possible to question a proportion of the population more intensely than would be possible for the entire population. It is beyond the scope of this report to fully discuss the methods for sampling, but a few guidelines can be given.

³ For the collection of confidential or sensitive data, there may be a difference between cultures (countries). In the northern EC countries, these types of data are preferably collected by a mail survey. In latin-orientated countries this may not be the case when one never wants to write these types of data on paper.

- Sampling is useful when the population is large (at least a few hundred).
- Sampling is more easy when the population is homogeneous.
- When the population is heterogeneous, stratified sampling techniques should be used.
- Sampling should be considered when the importance of each individual project is small.
- Using samples requires more expertise in the data-processing and analyzing parts of a study.

Depending on the information needed and the importance of each individual project (e.g. costly demonstration projects), an evaluation team may want information on every project, regardless of the number of projects. A last remark about the size of the sample when one intends to use statistics. **The size of the sample must be large enough to justify statistics.** How much is enough cannot be justified for all types of surveys. It depends on the heterogeneity of the population. We recommend a size of at least a few hundred before using advanced statistics.

For most EC-Programmes sampling will not be useful due to the small number of participants. In some programmes, the limited number of participants can even cause problems in the field of confidentiality and respondents anonymity.

5. Language to be used in questionnaire

The language to be used is especially important for the written questionnaires otherwise there is no control over the respondent's understanding of the question.

In most Commission evaluations, the language to be used is dependent on the level of respondents' schooling. Most questionnaires are in English. This may be valid for questionnaires aimed at international operating companies but is doubtful if the questionnaire is aimed at individual scientists or small companies, who do not actively use English and therefore are unable to interpret the exact meaning of the questions. Using one language implies that the questionnaire must be pre-tested in different countries to measure the respondents' comprehension. This pre-testing is more complex than testing questionnaires in the national languages and partly eliminates the cost-reductions when one language is used.

We recommend using questionnaires in other than the native language only when the questions are relatively simple, there are no or a few open-ended questions and the respondents have a reasonable level of education. In all other situations we advise the translation of the questionnaire into the national languages. Regardless of the language used in written questionnaires, it is always advisable to use only the most common words.

Most EC-programmes are aimed at international operating firms or scientists ; so for most evaluations a simple questionnaire in English is quite possible. Evaluations of programmes aimed at the support of small and medium enterprises or agriculture should always consider the possible use of the native language.

6. The use of data already available

There are two ways to use the data which are already available within the Commission (e.g. data from the contract, or additional data which is gathered within the programme). The first possibility is to insert personalized data in the questionnaire, which will save the respondent time because he or she must only check this data. Examples of data that can be used in this respect are :

- * Name, address, etc.
- * Type of contract, subject, discipline, etc.
- * Other participants within the contract

As this allows personalized questionnaires, and hence enhances response rates, we advise using it whenever possible.

Another way of using internal data is in the phase of data analysis, by merging internal data with survey data. Annex 1 shows this procedure.

7. Confidentiality

As for any data collection, evaluation surveys raise the question of confidentiality. File builders must ensure that the information stored cannot be detrimental to those concerned.

It is very useful to have a database concerning projects and their evaluations, so that, in future years, it will be possible to make further data processings, and, for instance, compare past and future evaluations of various phases of one programme, etc. Data is valuable (and expensive) material, destroying them is an irremediable waste, because we cannot know how useful they may be to our successors. On the other hand, data might be used for other purposes than the ones the respondents were aware of. Although data from evaluation questionnaires is generally harmless, all precautions must be taken to avoid any possible misuse.

Deontology codes of survey institutes vary from country to country. In some places, the files must be deleted shortly after processing. In others, each individual recorded in the files must have free access to data concerning him, and is entitled to correct it. In others, cession of nominative computerized files is submitted to a special commission.

Destruction of the files is probably the most secure way to avoid future misuses, but some of us think this is too drastic and costly. On the other hand, inserting the names of respondents with the questionnaires in AMPERE is probably not advisable in some cases.

The following procedures must be respected :

No individual results may be published from a file, only aggregated figures.

No nominative files must be given (or sold, or lent) by the organism who collected them to another organism without authorization by the respondents concerned.

Respondents must be warned that their answers will be used in a survey. A special addendum must be inserted in the questionnaire, mentioning that their answers will only be used for aggregated results.

All personnels handling or processing nominative files must be bound by "statistical secrecy" (they are not allowed to communicate any nominative data).

We suggest, for the time being, to choose between the three following procedures for evaluation surveys :

1. The Commission *is* the explicit receiver of the questionnaires. For instance, if the questionnaire is mailed, the return address must be the CCE. Data collecting is explicitly done by the EC for its internal use.

In this case, the Commission can keep the nominative questionnaires for further studies, and even transfer the nominative results to AMPERE, when necessary. There is no property transfer of nominative files outside the collecting organism. The survey can be executed and the statistical treatments realized by an external consultant or institute when needed, but there is no confidentiality problem, since its deontology obliges it to statistical secrecy, and the data is processed *for* the collecting organism and not for anyone else.

The drawback of this procedure is that respondents might be biased in some cases, because they know that the Commission will have their nominal individual responses. We recommend it for questionnaires where the risk for this bias is low.

2. The Commission is *not* the receiver of the questionnaires. In this case, data collecting is explicitly done by an external consultant or institute, under its own responsibility. The following two sub-cases apply :

- 2.1. The external institute finally hands over nominative files to the Commission after exploitation. In this case, the respondent must be explicitly warned that this will be done. This is similar to case 1.

- 2.2. The external institute either transfers all *non-nominative* files to the Commission after exploitation, or none at all. In this last case, the advantage is that the respondent might feel more free to answer, because he knows that the institute will not transfer nominative data to the Commission. However,

the institute must be obliged, by contract with the Commission, to keep the nominative data for 10 years in case it is needed for a further evaluation necessitating the tracing of individual organisms (e.g. network studies).

Drawback : A further study must then be conducted with the same institute. One must make sure that this institute will live long enough.

As long as no general procedure has been adopted by the Commission, **each panel must take a decision in his own case, after considering the specificities of the programme and evaluation.** We suggest that the question of confidentiality be discussed and that the panel's choice be formally expressed in the minutes of the meeting.

8. Response

8.1. Reasons for non response

The response rates are, as stated earlier, heavily dependent on the type of questionnaire used for the survey ⁴.

A second aspect that influences the response rate is the **expected involvement of the respondents** in the subject of the survey (if some types of respondents have, whatever the reason may be, a high motivation for answering the questionnaire, one can expect that the responding population can be different from the original population. This may affect the results if these different types of respondents think differently on certain aspects).

A third aspect is the **perceived length of the questionnaire** (especially for written questionnaires). Therefore questionnaires must be brief and only concern the relevant aspects. Other "interesting" items should be avoided.

A fourth aspect that negatively influences the response is when the respondents have very **busy time schedules** (e.g. executives). This type of respondents will only answer very short questionnaires (unless they are very concerned in the addressed subject).

⁴ Although it is very likely that the response rate for a specific type of questionnaire will vary for different countries (cultures) we do not have reliable data on this matter.

A last aspect that can be mentioned is if the respondents have functions that are interesting to many surveyors but where the population is relatively small (e.g. data-processing executives within large companies). This type of respondents suffer from “survey-fatigue”.

8.2. Remedies

There are several options for trying to enhance the response rate. The methods differ for the type of questionnaire used.

The response for in-depth and face-to-face interviewing can be enhanced by **sending a letter** which clearly states the reason for the survey and the principle of the study. This letter must be sent shortly before the interviewers try to make the appointments. The interviewers must be thoroughly instructed on the objectives of the study.

The research-institute doing the survey should only make use of experienced interviewers. Universities are especially notorious for using cheap but unqualified students as interviewers. This enable them to offer surveys at a low price, but with a questionable price-quality ratio. For these kinds of surveys it is advisable that the evaluation-panel secretary is present during the interviewers’ verbal instruction given by the research institute responsible for the survey.

The response rate for telephone interviewing is more difficult to enhance by a previously sent letter. More important is the **quality of the interviewers**. Telephone interviewing requires adept skills in explaining the reasons for the study in a very short time and convincing the respondent that his response is conducive to the study. Again the research institute should only make use of qualified interviewers who have been well instructed. The evaluation panel secretary should not only be present at the instruction but also visit the interview site for at least one day, in order to obtain a clear view of the quality of the fieldwork.

There are several possibilities for enhancing the response rate of a written questionnaire. The best possibilities are the **use of a good accompanying letter** and the **sending of a reminder**. Most of the time this reminder is sent out without a new questionnaire, which has already been lost by the respondent. So the effect is minimal. One should always enclose a new questionnaire with the reminder. The disadvantage of this method is that it does not give any insight into the reasons for refusal and there is little control over response and time. As an example, one can take the BRITE-

evaluation of 1988 (*report no. 25*). Even after the second mailing there were only 177 useful responses out of 450 participants, which means a response rate lower than 40%.

Another possibility for enhancing the response rate is a **short telephone screening** before sending out the questionnaire (with an accompanying letter). In this screening one can announce and explain the reasons for the survey. Thereby foregoing sending out questionnaires to respondents who will not cooperate, and the reason for refusal is known. It is the best but relatively costly way of enhancing the response rate. It is also extremely useful for split-run surveys or if the respondent is not known by name. This last aspect is quite important. Written questionnaires are only useful if the respondents are known by name. Questionnaires directed to the 'head of', 'Person responsible for ...', 'Management', etc. are doomed to have a very low response rate, making the data useless.

Prenotification by letter also has a positive effect on the response rate ⁵.

The last possibility, which we will mention, is the **telephone reminder**. One can either phone the respondent and urge him/her to send in the questionnaire or try to do the interview by telephone. This can be very effective if non-response was minimal (in respect of costs) and the importance of a high response beneficial to the results. This is especially the case if the population is very small. e.g. the large Installation Programme has a very limited number of organisations which participate. An evaluation should include all organisations.

We do not advise the use of gifts, monetary incentives, lotteries, etc. These actions are quite common in consumer research but are not to be used in industrial and evaluation research. The respondent must participate because he or she is convinced that it is important to fill in the questionnaire, not because of possible benefits. In some organisations, the personnel is not even allowed to accept gifts. **The best incentive in this respect is to send the respondents the report afterwards.**

If a written survey has a high non-response rate it is advisable to conduct a telephone non-response survey in order to find out the reasons for non-response and whether the non-response group differs from the response group, which may have an effect on the results.

⁵ The effect of prenotification for written surveys is described by : Murphy P.R., D.R. Dalenberg, and J.M. Darley, "Improving survey responses with postcards", *Industrial Marketing Management*, Vol. 19, nr 4, november 1990.

9. Initial briefing survey-organisations

When the evaluation panel has decided that a survey is necessary and that it will be sub-contracted to a third party, there must be criteria for the selection. Good selection goes along with a good briefing to this third party. What kind of information does a survey-organisation need to make a good proposal and do a good job ? The briefing must contain at least the following information :

- Size of population per country (number of participants).
- Indication of type of participants and heterogeneity.
- Indication of complexity of the information wanted.
- Type of fonctionnaries to be questionned.
- Timing and deadlines.
- Background information about the programme.
- Requirements for the survey-report (see Part II.2.).

This briefing is clearly not enough for a detailed definite survey-design or the exact wording of the questionnaire. However it is sufficient for developing a rough survey-design and outlining of a research proposal.

***II. PROCEDURES FOR ANALYSIS AND PRESENTATION**

Summary

- * The Commission does not need advanced software for analyzing the data. Most evaluations do not use more than straightforward frequencies or simple tables (p. 24).
- * Computing price-production ratios can be useful for cross-evaluations between programmes and historical evaluations (p. 25).
- * The reports concerning the survey must meet the following minimum requirements (pp. 26-30) :
 - Background information (e.g. the objective of the study).
 - Population, realised sample and possible influence of non-response on the results.
 - Data gathering methods used.
 - Presentation of the factual results.
 - Indication of reliability and significance of the results.
 - Questionnaire.
- * In order to ensure that the reports will meet these requirements we have suggested a standard lay-out for this report (p. 29).

As mentioned in the introduction, standardization is not only useful for the procedures of setting-up of a survey, but also for the analysis and presentation. In this respect we will not give recommendations on punching and coding procedures, because this will be the responsibility of the organisation conducting the survey. If they are using computer-aided survey techniques no punching procedures are necessary. In the case of the processing of written questionnaires, double punching, which can almost eliminate punching errors, should be recommended, but it is difficult for the Commission to control. The same can be said for coding procedures. It is debatable if a questionnaire should be coded by one person who would have an overall view of the answers thereby facilitating his/her comprehension of open-ended questions. Others suggest that one person should code just one question for all questionnaires because then there is a specialization on that particular question. We believe these decisions are the responsibility of the research institute.

In this part we will concentrate on two aspects. The means which can be used if the Commission (the evaluation team within SPEAR) wants to analyze the data itself (also) and some standard procedures and means for the content of the evaluation reports.

1. Analysis of the data

If the evaluation team has sub-contracted a large scale survey to a third party it is to be expected that this party supply the team with the data (although most professional organisations are committed to the ESOMAR/ICC-code⁶ which states that all data will be given to the principal anonymously) (*see I.7.*). The data supplied is normally MS-DOS Ascii-files, which can be analyzed with many types of software. It goes beyond this study to report on all types of software but we will give a few examples. Most evaluation studies do not require specific statistical operations (e.g. advanced statistical tests, multi-dimensional scaling, multivariate analysis or advanced data-manipulation). Most evaluation reports present little more than straightforward frequencies. For easy tabling we can recommend special software for easy data handling (e.g. A-cross or Surveytab) or just database-programmes (e.g. DbaseIV, Foxbase or Oracle). The skills required for this software are easy to learn and the software can be used on any common MS-DOS computer system. However for large data-sets the use of a powerful PC can be very useful.

For more advanced analysis it is almost impossible to make recommendations. It is dependent on the needs and the skills of the persons who need to operate the system. Programmes such as SPSS, SYSTAT and SAS offer a lot of possibilities (e.g. advanced data-manipulation, multi-dimensional scaling, factor-analysis, etc) but are an 'overkill' for easy tabling. They also require more specialized skills, from the researcher as well as the person who operates the software. It must be stated that a good analysis, by making use of cross-tables and filter-techniques, on a simple questionnaire provides better results than a straightforward analysis on a complex questionnaire. In other words when one develops a questionnaire one has to already think of the analysis outline.

Another important aspect for the analysis is the use of available internal data. Internal data can be used for the construction of criteria that can be useful for comparing evaluations for the same programme in the future. Sometimes these criteria can also be useful for comparing the evaluations of different programmes. Useful are price-production ratios within the more scientific orientated programmes (the ratio is computed by the division of the objective criteria by the costs of the programme). Criteria that can be constructed are:

- * number of publications -----> price per publication
- * number of citations -----> price per citation

⁶ The address of ESOMAR (European Society for Opinion and Marketing Research) is : J.J. Violastraat 29, 1017 JP Amsterdam, Netherlands, telephone 020-6642141.
ICC : International Chamber of Commerce

- * number of presentations -----> price per presentation
- * number of conferences organized -----> price per conference
- * number of conference attendants -----> price per attendant

Programmes that stimulate international cooperation could compute the price per international contact that has been formed within the programme.

Although these price ratios are just rough indicators of the cost-effectiveness of the programmes it is quite useful to have a general idea about the goals achieved and the costs involved. The ratios use for cross-evaluation between programmes may also be limited to programmes that are reasonably comparable in objectives. However they are extremely useful for historic evaluations ⁷. It could give some clue whether there has been any improvement in the ratios after a programme has been adjusted due to recommendations in an earlier evaluation. For long term programmes which are evaluated on a regular base, we advise to use as much of these ratios as possible and to use the same ratios in future evaluations. However, it must be stressed that these ratios are of special use for the evaluation by the evaluation committee. They cannot be picked out of an evaluation report and used as an indicator for price-production out of it's original context.

It must be noted that in computing ratios that are based on internal data and data from a questionnaire, there should be a correction for the non-response, otherwise the non-response is of great influence in these price-ratios, which renders the comparisons useless.

2. Presentation of results

We recommend some guidelines for the presentation of the survey reports. Indeed, in the analysis of the previous evaluation reports (*see annex*), it is remarkable how each report differs in content and presentation. There is no homogeneity in the reports whatsoever. In most reports some essential data (e.g. the number of contractors or questionnaire used) are quite difficult to find or even missing. What elements should be mentioned in the final evaluation report ? A checklist for these elements has been developed by ESOMAR and the ICC. In 1976 they jointly developed a code for the behaviour of research organisations. In this code there are guidelines on how to deal

⁷ For most EC-programmes historical evaluations are not suitable because the programmes are redefined drastically on a regular basis or have short life cycles.

with privacy and ownership of the material, but also the minimal requirements for a report based on research. These requirements are also applicable to the reports concerning the surveys on behalf of the evaluation panel. The requirements are broken down in four main aspects:

A. Background

- Name of principal, contractor and possible subcontractors.
- Objective of the study.

B. Population and sample

- Description of the population and realised sample.
- Relevant indicators of the realized sample (e.g in terms of geography, size, sic-code, etc.).
- Used sampling techniques and, if applicable, weighting methods.
- Non-response and the possible influence of non-response on the results.

C. Data-gathering methods

- Description of the methods used for the data-gathering.
- Description of instruction method of the interviewer corps and methods used for quality control.
- Date and period of the fieldwork.
- References of other information sources.

D. Presentation of results

- Relevant factual results.
- When using percentage the base must be indicated and a clear indication of figures that have been adjusted by weighing techniques.
- Tables should contain sufficient explanatory information (e.g. title, identification of variables, etc.) to be readable independent of the text.
- General indication of the reliability and significance of the results.
- Questionnaire and other relevant documentation.

It is worthwhile to see how these elements are covered in previous evaluation reports.

2.1. Background

In most evaluation reports these elements are covered quite reasonably. Most reports mention the assignment by the Commission and on which elements the study concentrates.

2.2. Population and sample

Most evaluation reports do not give a clear description of the population and the realized sample. Most reports give information about the objective of the programme and the budget and sometimes the distribution of the budget (e.g. distribution over countries, type of activity, etc.). For a survey it is important to know :

- how many participants (population) there actually are ;
- if sampling techniques are used and if so which techniques ;
- if stratified samples are used ⁸, how the respondents are distributed over the strata.

A reasonable description can be found in the evaluation of the geothermal energy programme (*report no. 2*), although it is necessary to examine different chapters for the total population, the sample sent out and the response rate. It was also surprising to read in the chapter concerning methodology that the whole population has been mailed, but in the annex concerning mailing it becomes clear that only half of the population has actually received mail, without any indication why. There was no indication either that non-response (only one quarter had responded) was of any influence on the results. Also in the BRITE evaluation (*report no. 25*), where the response is less then 40%, there is no indication if the population of respondents differs from the total population.

The BEP/BAP-evaluation (*report no. 32*) deals rather well with the description of the population and final response and also what kind of conclusions one can draw in relation to the response figures for each particular item.

Clearly it is necessary that in future survey-reports, there must be one standard chapter in which these aspects are described.

⁸ A stratified sample means that one uses unequal sampling percentages in a stratum. A stratum is a subset of the population defined by variables such as size or sic code. Stratified sampling means for instance that one wants to sample all big companies, only half of the middle size companies and only 5% of the small companies. Stratified sampling is very common in industrial research.

2.3. Data-gathering

The methods of data-gathering are described rather well in most reports. In most reports the evaluation panel does a lot of data-gathering themselves and the use of a (written) survey is mentioned. Because most surveys use a written questionnaire there is no need for information on the interviewers. With respect to the written questionnaire we did not find any information concerning quality control of the data gathered.

2.4. Presentation of results

The presentation of the results is generally poor in the evaluation reports. Most reports do not contain the actual questionnaire that has been used. Many reports simply show marginal frequencies, sometimes in percentages and sometimes in round figures. Information on reliability and significance is almost always lacking. The presentation in the Brite evaluation however by showing graphics and the basis for these graphics is rather good. Information on reliability, significance and influence of non-response would have made this study almost complete.

We consider that most evaluation reports lack most elements which must be included when survey techniques are used. Using surveys without this information is almost useless and certainly not reliable. The Commission should demand that in every survey-report, this information is included in a standard format suggested on the following page.

STRUCTURE OF SURVEY REPORT

Executive summary

Chapter I : Methodology

- 1.1 *Background and objectives*
- 1.2 *Methods used*
- 1.3 *Population and Sample*

Population:	Total number of contractants
	Sample population
	Response

If the sample is stratified, e.g. activity or geographically this information must be given in a matrix. Such matrices can look like this:

		POPULATION			
		Country			
		A	B	C	Total
		n = ...	n = ...	n = ...	n = ...
Type of organisation					
1		%			
2		%			
3		%			
Total	100%				

The same matrix must be given for the sample and for the final response. Only then is one able to judge fairly if the response is representative for the total population.

1.4 *Interpretation*

Here there must be information about the reliability of the figures presented and the possible influence of non-response on these figures.

Chapter II : Results

The total results can be presented in tables or graphics, but whatever the method used it must be clear on what basis the percentages are computed (e.g. does the % include non-response, do-not-know answers or omissions, or if percentages are based on subgroups it must be clear how large this subgroup is).

If figures are based on adjusted figures these weighing principles must be mentioned including the reasons for adjustment.

This chapter must also indicate the relevance of the results.

ENCLOSURE : QUESTIONNAIRE

2.5. Quality control and continuous improvement

We recommend that within the SPEAR evaluation service there should be at least one person who can judge evaluation reports on these aspects. This “methods-expert” should inform the evaluation team on the required format of the report and what elements should be mentioned about the survey during the briefing. The report must also be checked if these elements have been reported concisely. Only then is it possible to reach some consistency in the different evaluations. It also improves the quality and the value of an evaluation when it is clear what procedures have been followed.

The methods-expert should also be given the task of collecting feedback from the use of this “Guideline for Survey Techniques in Evaluation Research” in actual practice. This should lead to an updated and improved edition about every two years.

III. QUESTIONNAIRE TECHNIQUES AND RULES

Summary

- * A good questionnaire must take into account upstream, information to be gathered and downstream, the analysis constraints ; this implies a lot of technical recommendations (p. 33) and elementary rules to set up a questionnaire (pp. 34-35).
- * The nature of the questions depends on the object of the survey (pp. 35-37) and their form has a direct influence on the way to make the questionnaire operational (pp. 37-44).
- * The formulation of the questions must be precise, concise and neutral and follow a carefully ordered plan. The question must include an instruction, the most explicit possible, on how to answer (pp. 45-51). A table summarizes the main instructions (pp. 52-53).
- * The respective role of the actors of the evaluation with respect to the questionnaire must be clearly defined : if definition of the objectives of the programme come within the competence of management of the concerned programme, selection of the key issues of the evaluation comes in the scope of evaluation panel, while formulation and exploitation of the questionnaire are to be entrusted to a specialist. (pp. 54-58).
- * Respondents must be identified (p. 56) and the questionnaire previously tested (p. 57).
- * It is absolutely necessary to conserve a memory of the survey, in order to allow secondary exploitations, like as comparisons between evaluations, in the future (p. 58).

1. Three basic concerns

Any questionnaire may pursue different objectives, for instance :

- . description of a population
- . estimation of magnitudes
- . verification of assumptions.

Its role is thus to **translate the objectives of the study or the research in specific questions**, and to do this, it must favor communication by the survey of information which is expected from it.

This presupposes that before drafting the questionnaire, the objective of the study has been clearly defined, or, if it is original research, a choice made of objectives of the research and specification of the data which are expected from the survey. In other words, the questionnaire as such follows logically from the objectives of the study or the research.

For this reason, it can only be drafted after careful consideration of the objectives. This drafting is certainly one of the most tricky tasks of a survey. Another is the analysis of the data collected.

This leads to a general rule :

A good questionnaire must take into account upstream, information to be gathered and downstream, the analysis constraints.

Obtaining the best possible information upstream means precisely defining the information to be collected, facilitating the task of the survey, making the survey credible by appropriate introduction, being attentive to the length and the structure of the questionnaire, formulating proper questions, providing for the processing of non-responses.

It is generally accepted that drawing up a questionnaire must meet three fundamental concerns. A questionnaire must be :

- **valid.** A questionnaire must authorize, via the responses that it permits, collection or representation of the objects that best conforms with reality.
- **reliable.** A questionnaire is faithful when it is used by different surveyors and provides, if it is circulated to the same people, the same responses.
- **operative.** A questionnaire is operational if its use and the interpretation of the answers can take place under proper conditions and satisfy the expectations of the person who has designed it.

These are primary preoccupations, i.e., the points on which the vigilance of the assessor must be exercised in the first place.

Practically, when drawing up a questionnaire, absolute attention must be paid to the **structure of the questionnaire**. The formulation of the questions, their arrangement, the "aesthetic" presentation of the whole, are points which, when they are neglected, can invalidate the results, or, at the very least, make the task of the assessors much more difficult.

The proposed viewpoint in this guide to work out questionnaires for evaluation in the EC is that of the postal questionnaire. This questionnaire is thus self-administered, i.e. respondent is alone when facing the questionnaire : this implies severe requirements for **clarity** and **accuracy** in its working out and in the formulation of questions.

It is important that the designers of a questionnaire should constantly have in mind, firstly, certain technical recommendations concerning the general economy of a questionnaire and a postal questionnaire in particular and, secondly, the rules related to the drafting of questions. These recommendations and these rules are given in the boxes below.

TECHNICAL RECOMMENDATIONS

The appearance of the questionnaire must be carefully designed : it should not resemble an administrative form, and for instance be printed on poor quality paper. The questions raised must be typographically differentiated from the instructions accompanying them.

The **identification elements** will have been, whenever is possible, entered first (see above I.6.), and will be repeated on the questionnaire sent to the respondent. Questionnaires will thus be as far as possible individualized.

The overall economics of the questionnaire will meet the imperatives of **concision, precision and coherence** explained below.

Questions that meet no precise information objective (or any particular presumption defined) must be systematically eliminated. Questions **"just to see"**, particularly those that take the appearance of open questions without precoding simply must be discarded

Priority of place will be given to closed questions with a forced choice. However, the respondent must be given the possibility of expressing his opinion as he wishes on any particular sensitive aspect of the evaluation of the programme. Thus, certain open questions, carefully situated in the body of the questionnaire, will permit the respondent to express himself freely.

In all cases, the order concerning the way of formulating the response will be clearly indicated.

When use is made of **multiple choice questions** and to limit, from the respondent, the echo-effect of the order of presentation of the responses, a different presentation of the order of the possible answers in the questionnaires will be provided to the different respondents ("split ballot" technique).

Account must be taken of the fact that **the respondent knows the totality of the questionnaire** : for instance, we must avoid trying to obtain responses to one question, when the previous (or the next) question, in their formulation, proposes possible answers to this question.

Lastly, **the answers proposed for closed questions or a selection will be precoded.** This greatly facilitates work during the analysis phase of results. In particular, in certain cases it should be clearly specified for non-reply of the respondent with the corresponding codification.

Now, elementary mandatory rules are deduced from the principle that we must obtain the best possible information while at the same time making the respondent's task easier.

ELEMENTARY RULES TO DRAFT AN EVALUATION QUESTIONNAIRE

Rule 1 : Before formulating questions, the list of information to be gathered must be carefully drawn up.

Rule 2 : A presentation text of the questionnaire must make the survey credible by clearly explaining its objectives.

Rule 3 : As a matter of principle, a questionnaire is always too long. To be effective, the questionnaire must constantly arouse and maintain the interest of the respondent.

Rule 4 : The initial question must explicitly take up the objective.

Rule 5 : The order of the questions must comply with a certain logic : if possible, references backwards and subject changes should be avoided.

Rule 6 : Questions relating to the same subject should be grouped together.

Rule 7 : Factual questions should precede opinion questions.

Rule 8 : Tricky questions must be placed at the middle of the theme to which they are related and be led up to by easier questions.

Rule 9 : A question must contain one idea and only one to avoid confusion.

Rule 10 : Terms should be used which are those of the current language of the recipient, even if they must be technical.

Rule 11 : A question must not be ambiguous.

Rule 12 : Sentences should be short, use no negatives and particularly no double negatives (e.g. : Do you not think that it is not necessary, etc.).

Rule 13 : A question must be objective and not influence the survey by its wording.

Rule 14 : It should be avoided that a question could influence the response to the following question, i.e., creating a halo effect.

Rule 15 : For tricky questions, an open question is often preferable.

Rule 16 : Avoid wording that implicitly calls on the hierarchy of social values.

Rule 17 : Alternate questions where the same general opinion results both in positive responses and also in negatives responses.

Rule 18 : The questionnaire should be drafted in the language of the respondent, which does not systematically mean translating it. Certain concepts, expressions or linguistic habits specific to a scientific discipline and in current use will be retained : the evaluation questionnaire is addressed to a predefined population and, in principle, well known.

Rule 19 : A questionnaire can always be improved. Consequently, it is a necessity to test it beforehand on a sample of the population to be surveyed, in order to check that the questions are popularly understood, that it is easy to get a position on the scale of the responses given, that the order of the questions is logical, that certain questions are not embarrassing or do not interfere in any way, that a tiredness effect does not appear quickly.

The drafter of a questionnaire must always choose between different types of questions, between different formulations, etc. **The choices to make are not indifferent. They are conducive to the success of the questionnaire.** At the same time, each of these choices must meet the specific requirements of the evaluation for which the questionnaire has been designed. Let us envisage problems that may be raised for each type of question taken individually.

2. The nature of the questions

Classically, a distinction is made between factual questions, questions on actions carried out, questions on hypothetical behaviour and opinion questions.

2.1. Factual questions

In the factual questions, the person surveyed is asked to give an opinion on real facts, that are not necessarily concrete, but always susceptible to having been the subject of his experience or his knowledge.

Thus, factual questions permit single answer responses, since, as a general rule, the facts to which they refer exist or do not exist. In the same way, there is much less inaccuracy or indecision in the answers. It is better to start with these questions than with opinion or hypothetical behavioural questions.

In an evaluation questionnaire which is trying to determine, for instance, the socio-economic effects of the implementation of a particular programme, the respondent may be requested to list the “address book” that he has been able to build up during this programme.

It should already be noted that certain factual questions on identification can cause negative reactions or non-acceptance, if they are introduced into the questionnaire without precautions. In an evaluation questionnaire, this may be the case of the question suggested above. Such questions must be “introduced” and possibly located elsewhere in the body of the questionnaire.

Lastly, all questions relating to facts are not necessarily good factual questions. In order for them to be so, it is necessary to ensure in advance that the respondent is probably concerned by the fact in question, i.e., that he is in a position to understand its reality.

2.2. Questions on actions carried out

These questions may be taken for a subset of factual questions. They ask the respondent whether or not he has carried out a particular action. This, for an outside observer, can be considered as factual. Here again, the responses obtained are generally sincere and a refusal to reply is rare, except when dealing with taboo areas. In ordinary questionnaires, these fields may range from sexuality to abstention from voting. In R&D evaluation questionnaires, this could possibly be relating to more or less legal practices, getting round or evading administrative or legal rules.

2.3. Questions on hypothetical behaviour

Questions on hypothetical behaviour, also called **intention or attitude questions**, have the purpose of requesting the person interrogated what he would do if placed in a hypothetical situation. It should be noted that between the intention and the attitude, there may be a difference of viewpoint. Attitude requires objective items, where intention is supplied by the subjective component.

Questions on hypothetical behaviour are often used to indirectly study supposed opinions of the people questioned. In all cases, it should never be forgotten that a behavioural intention cannot be mistaken for effective behaviour.

As part of the evaluation of R&D, it may be desired to research and measure the attitude of respondents with respect to any particular problem, or any particular orientation. The use of questions on hypothetical behaviour, and as we will see later, attitude scales, must be carefully distinguished from research on factual information and for information research relating to opinions.

2.4. Questions of opinion

If it is necessary to differentiate opinion questions from attitude questions or hypothetical behaviour, because it is understood that the opinion relates to a lesser level, and less easy to rationalize, than attitude. As a result an opinion appears much less versatile. For this reason, the opinion is fundamentally written in time, at the moment it is expressed, and in its context.

If opinion is to be related to its context, it appears obvious that the expression of this opinion may be sensitive to the formulation of the question and to the actual context of this question. This leads us to say that this type of question must be accompanied by careful thinking as to the context.

Thus, dealing with the overall and terminal opinions of participants in a programme that may be "exploratory", to the objectives that are after all "vague", to the conditions of implementation that are uncertain and with limited means, can only take place in the form of questions that make explicit reference to this context.

3. The form of the questions

As a general rule, the form of the questions depends on the actual object of the survey. i.e., what information is required and the way in which it can be obtained.

Secondly, the form of the questions has an impact on the operational character of the questionnaire. Depending on the object, certain questions are easier to administer than others. Similarly, at the moment of analyzing the results, certain types of questions will turn out to be easier to deal with than others.

Before making instructions as to the form of questions, let us look quickly at the possible forms.

3.1. Closed questions

Closed questions are used to obtain certain factual information ; to know whether a given opinion is approved or disapproved ; to determine the position of the person surveyed in a range of situations, judgments, etc. Thus, the "conventional" questions relating to sex, marital status, etc., occur in the form of closed questions.

Closed questions are those where the formulation appears at its most simple. Very frequently, but wrongly, closed questions are reduced to the formulation of alternatives such as "yes"/"no", or "true"/"false". This is a restrictive interpretation. Nevertheless, it is frequently used and quite acceptable for closed questions. Furthermore, this form of question can sometimes have the advantage of "forcing a choice", since it only proposes a dual choice where assessment could be wider ranging. The advantage, where it exists, takes place at the moment of the interpretation of the results since it can be carried out as a dichotomy.

To summarize, closed questions considered the whole set of expected answers, and may be with a **forced choice** or a **multiple choice**.

Example of a forced choice :

Did your research proposal form part of a larger existing or planned research programme ?

Yes

No

Example of multiple choice :

What kind of staff is involved in your contract ?

** scientists and engineers*

** postgraduate students*

** undergraduate students*

** technicians*

** others*

In all cases, such questions reduce the initiative of the respondent to preformulated answers by the designer of the questionnaire. This manner, which in certain cases can appear to be reductive, reduces the ambiguity of the answers and ease of analysis.

The discriminating power of these questions means that they will often be used as **filter questions**, i.e., as questions used to distribute respondents between several classes to which a specific series of questions will be applied subsequently. In return, **they can hardly be used to obtain detailed information**. Generally, as soon as the questioning relates to delicate or complex things, the closed question is no longer in a position to provide an expression of a range of answers.

For instance, if we close the following question :

What were your impressions on the Commission staff involved in the research programme ?

by adding two items :

** unknown or few contacts*

** efficient, active*

responses are obtained that are certainly less rich in information than would have been provided by an open question. On the other hand, two quite specific classes of respondents will have been obtained.

3.2. Open questions

In open questions, the possible responses are not supplied. The respondent is thus free to supply the answer he wishes. The form of the question is also very simple as it is sufficient to reserve a limited amount of space at the end of question to permit the respondent to write out his answer in its entirety.

Example :

What was your principal motivation for submitting a research proposal.

.....
.....
.....
.....

Practically, for open questions, it is advised to provide insets of sufficient size, i.e., proportional to the assumed length of the possible answer. These insets consist, in most cases, of a simple blank, or light dots to guide writing, clearly enclosed. **The presentation of open questions must facilitate the materialization of the answer.** Open questions, if they have been suitably formulated, authorize deeper information, particularly on more tricky subjects or those on which possible answers cannot be thought up in advance. **But open questions can only be used among persons who are able to answer.**

If this is not the case, and if the interrogation takes place face to face or by telephone, the respondent will be forced to ask for explanations from the surveyor and thus undergo the influence of the latter, if it is a questionnaire filled in by the respondent himself, he may have a tendency to give up making any answer.

This means that the formulation of open questions is a very delicate exercise. They must be easily understood by the respondent, while remaining accurate and free from ambiguity. This leads us to consider not only the wording used, but the semantic variations of the language used : for instance, the translation of English into the usual idiom, or again, "jargon" or subcultures of the scientific specialities concerned. Ideally, as certain work has shown, both in English and in French ⁹, the formulation of the question must also include the type of instruction which is given to the respondent, the location of the question in the questionnaire, the whole of the context of the question as well as the method of administering the questionnaire itself. We shall return to this (*see below*).

Choosing to introduce an open question in a questionnaire can only take place after having very precisely assessed the advantage of this form of question. Particularly since the processing of answers to an open question calls on greater "subjective" intervention.

⁹ On this point, it is useful to consult

- Payne (S. L.), *The art of asking questions*, Princeton, Princeton University Press, 1951.
- Sudman (S.) and Bradburn (N. N.), *Asking questions. A practical guide to questionnaire design*, San Francisco, Jossey-Bass, 1983.
- Grémy (J.-P.), "Les expériences françaises sur la formulation des questions d'enquête", *Revue Française de Sociologie*, XXVIII, 1987.
- Juan (S.), "L'ouvert et le fermé dans la pratique du questionnaire", *Revue Française de Sociologie*, XXVII, 1986.

It necessarily provides greater interpretation of the answers provided, since they must be divided up into pertinent answer classes. This is post coding. An operation which sometimes takes place a long time after the questionnaire itself has been designed, and which has perhaps been forgotten.

Thus, choosing the formulation of an open question can in no case be a default choice. An open question is never a substitute for another form of question where the formulation raises some difficulties.

3.3. Multiple choice questions

The principle is simple : the respondent is offered a choice of possible answers between which he will have to choose freely. If this range is itself closed, i.e., if the list of answer proposals is halted, we return to the case of **the multiple choice closed question**. In multiple choice questions, the list of possible answers is not limited and enables the respondent to add the response of his choice. Multiple choice questions are thus semi-open questions, or rather “*à la carte*” questions. They are sometimes called “*cafeteria questions*”.

Example of a multiple-choice question :

On what grounds do you think your participation in the programme has been useful ?

- *Establish contact in other countries*
- *Exchange of information and ideas*
- *Initiate new joint research projects*
- *Learn new techniques/methods*
- *Patents*
- *Dissemination of results to the scientific community on a European level.*
- *Large market for exploitation*
- *Encourage a European outlook*
- *Initiate new lines of research*
- *Reach a predetermined goal*
- *Training of young scientists*
- *Others :*
-
-

The value of multiple choice questions is certainly to be able to offer the respondent the possibility of adding free answers alongside the responses that the designer of the questionnaire has judged to be “primary” or “principal”. Thus, there is a **psychological value** since it enables the respondent not to feel enclosed in a set of preformed responses that do not suit him entirely. **Informative interest**, since it authorizes, as in the case of the open question, finer tuning of the response.

Another value of multiple choice questions is that **they facilitate analysis**, since a certain number of possible answers are already provided for. This provides, for this “closed” part of the multiple-choice, a prior job in order for it to properly correspond to the “main” answers so as to avoid too high a concentration of answers on the “open” part which is generally introduced by the “*others*” item. Otherwise, we return to the case of the open question, but under operating conditions that are often less good for the responses.

Multiple-choice questions do not only comprise advantages. Proposals for “closed” answers, while they guide the respondent as to the range over which he should answer, **also risk influencing his response** by suggesting, for instance, responses which he would not spontaneously have thought of, or even, lead him to believe that the responses proposed are more “evident” or more “suitable” than those that he could himself formulate.

3.4. Scales

As is suggested by the term “scale”, these are **techniques introducing a graduation between the sets of replies** and thus between what they denote (characteristics, attitudes, etc.). Two types of scales may be distinguished :

- * simple scales
- * attitude scales

The advantage of being able to have a scale is twofold. This introduces a principle of coherence or even homogeneity, that is stronger between the choice of respondents while increasing the degree of comprehension. And this authorizes, or facilitates, the measurement and calculation of responses.

Simple scales

- **Nominal scale.** This sort of scale is an enumeration of items without any order of magnitude or hierarchy. Each response therefore has the same weight. Example : the nominal scale of nationalities of the EC. This type of scale relates to a closed question with a forced choice.
- **Ordinal scale.** The ordinal scale is used to classify options in an order. The ordinal scale is at the origin of the simplest attitude scales. Thus the scale known as “Bogardus social distance” is of this type. Proposed in 1925 by E. S. Bogardus, it attempts to measure the racist and nationalistic attitude of an individual with respect to another individual of a different race or nationality.

It is presented as follows :

I would readily accept members of a particular race or a particular nationality :

- 1) *as close parents by marriage ;*
- 2) *as personal friends in my club ;*
- 3) *as neighbors in my street ;*
- 4) *as employees in my business ;*
- 5) *as citizens in my country ;*
- 6) *only as tourists.*

This scale is a “distance” scale with respect to the respondent. The distance increases as the questions advance.

As part of an R&D evaluation questionnaire, this type of scale can be very suitable to assess for instance the propensity of research teams to collaborate with each other, by ordering the types of possible collaboration.

To do this, and as an example, the following ordinal scale may be imagined :

As a part of Programme X, I would without difficulty accept :

- 1) *reorientating the work objectives of my team depending on the result supplied by another team in the Programme,*
- 2) *exchanging researchers with one of the teams in the Programme,*
- 3) *transmitting, at least once per quarter, a memorandum summarizing the progress in work carried out by my team,*
- 4) *communicating with the other teams in the Programme in the form of a synoptic memorandum containing the results obtained by my team,*
- 5) *transmitting to other teams in the Programme the list of members in my team and the work that they carry out.*

- **Intervals scale.** Here, not only is the order taken into account, but in addition the effective *distance* may be determined that separates two items. This requires that the intervals should be evaluated from a common unit. Chronological order is a good example of interval scales.

- **Proportional scale.** With the proportional scale, it is also possible to express the *relationship* that exists between two items. The age scale is a proportional scale.

These different scales thus introduce gradations in the responses and, particularly, in the evaluation of these responses. In addition, carried out in a regular manner, they can authorize useful comparisons.

Attitude scales

Many scales have been worked out to provide a measurement of the attitude of respondents in a given field. A certain number of attributes characteristic of this attitude will serve to establish the measurement.

We shall only consider here scales that have already been the subject of a conceptual elaboration and critical assessment. Of course, anybody can design an ad-hoc scale for the study he is in the process of undertaking. But, and notwithstanding of the often high costs of the design, there will be great uncertainty as to its validity. Its major fault will remain -at least for some time- in its absence of empirical validation.

Let us briefly consider some of these scales.

The Osgood semantic differentiator. The Osgood semantic differentiator principle amounts to asking the respondent to position himself between two opposing values which have been proposed to describe a particular "item" (situation, conceptual representation, etc.).

As an example, to determine the attitude of the teams taking parts in a scientific programme concerning the evaluative follow-up carried out by the EC, a scale of the Osgood semantic differentiator type may be designed :

- if it is wished to introduce a general assessment dimension, an opposing scale may be proposed :

useful ----------*-----*----- useless*

- if it is wished to introduce an activity dimension, an opposing scale may be proposed:

stimulating ----------*-----*-----paralysing*

- if it is wished to introduce an efficiency dimension, an opposing scale may be proposed :

productive ----------*-----*----- unproductive*

etc.

The advantage of this technique is to identify the different dimensions that can take place in the formation of an attitude¹⁰. And, the positions taken by the respondent on the binary opposition scale will be deduced, with different distinctive traits.

The Likert cumulative scale. The Likert attitude scales are part of the most widely used psycho-social survey techniques. Each item permits 5 grades ranging from total approval to total disapproval, the median position being occupied by "don't know". The response supplied receives a score : 1, 2, 3, etc. All the responses are summed up and each respondent receives a total score. To uniformize the results, the items that give rather unusual scores, i.e., those that do not represent a sufficient correlation with the global score will be eliminated.

In this way, a truly uniform set of proposals is obtained which all relate in a satisfactory manner to the field studied. The final total score obtained by each respondent determines his position in the attitude scale.

The Guttman hierarchical scale. The Guttman scale is used to class respondents into different attitude levels. The central hypothesis of the scale is that agreeing with the proposal to a certain degree involves agreement to all the proposals with a lower degree. An individual who responds affirmatively to the

¹⁰ It must be noted that the respondent is forced to take position on the issue, if the proposed scale is even, and he keeps the possibility to be neutral, if the proposed scale is odd.

question of knowing if he is more than 1.80 m tall, must also respond affirmatively to the one proposing : 1.75 m, 1.70 m, 1.65 m, etc. Guttman has thus made the same assumption with respect to opinions and attitudes.

The items come in increasing order over a given continuum. The answers to be given are of the binary choice type. Using the so-called scalogram technique, Guttman purifies his scale, that is to say, eliminates items that do not relate to the problem studied.

The advantage of the Guttman scale is that it guarantees the unidimensional character of the items selected to define the attitude. The major value of this scale is that it permits a distinction between "scalable" attitudes and those that are not. In other words, it permits the attribution in certain fields and to certain populations of typical attitudes which are to be found on known scales.

4. Question formulation

4.1. What is a good question ?

It should be remembered, even if this appears to be a truism, that the primary objective of a question is to elicit a response from the interlocutor. From this viewpoint, a good question is one which gives room to the best response, i.e., the one where the informative content is the richest in respect to the object of the study.

A good question must relate to an object in relation to which the answer of the respondent is capable of retaining a meaning for the person in charge of its interpretation, namely, for the evaluator. Then the person surveyed must have a minimum of information, have had the occasion to form an opinion, etc., with respect to the actual object to which the question relates.

The question must include an instruction as explicit as possible, as to the way of providing an answer. The respondent must know without difficulty if he is expected to answer with "yes" or with "no", or to choose one or more proposed answers, to respond like he wants, etc. Since indecision on the respondent's part can very often occur from a refusal to reply which, when it takes place, frequently reduces or even invalidates the informative range of responses supplied. A report by B. Bobe and H. Viala ¹¹ has shown that the study of R&D programme evaluation questionnaires demonstrated a "lack of strictness in the instructions (the person surveyed did not know if he was to reply and how)".

¹¹ Bobe (B.), Viala (H.), Une décennie d'évaluation de R&D à la Commission des Communautés européennes (1980-1990).

As to the formulation of the question, and to comply with the elementary principles, it must be precise, concise and neutral.

Since the purpose is always to obtain the best answer with respect to what is required, the question must be formulated so as to be understood. In order to do this, it is important that it should be drafted clearly, i.e., ensuring that the language used is understandable by the person surveyed, that it removes the ambiguity of terms as far as possible as well as the uncertainty or the vague character of certain semantic references. An expression such as “have contacts” remains completely ambiguous if it is related to the type of relationships that have been developed between bodies.

For instance, the answers to the question :

Did you have contacts with DGs ?

teaches nothing about the nature of these relationships, when they have existed. A multiple choice question should have been added here, proposing a series of answers and terminating by the conventional “other”.

Similarly, many adverbs or forms of expressions of quantity such as “many”, “a little”, “often”, “frequently”, etc., should be removed, except if a subjective evaluation is expected from the respondent, since their comprehension can vary from one person to another, and is dependent on the context : “often” has not the same meaning if it relates to the frequency of meetings between lovers or with the surgeon.

For instance, the following question :

Have you developed contacts with industry in the framework of the programme ?

Often

Occasionally

Rarely

will be of little use if it is really wished to measure the frequency of contacts with industrialists. On the other hand, the question will be more pertinent if the assessor specifies what he means by “often”, “occasionally”, “rarely”. The question can then become for instance :

What was the frequency of your contacts with industry in the framework of the programme ?

Often (3 times/year or more)

Occasionally (1-2 times/year)

Less often

No contact

In other terms, the accuracy of the question is obtained when designing the questionnaire, by specifying the objective which is wished to be obtained by asking the question, and by concrete drafting strictness.

4.2. Linguistic frontiers

Attention should be drawn here to the difficulties raised by the use of a language which is not the mother tongue of the respondent. In part II.5, "Language to be used in questionnaire", it is recommended, when questions are complex and the level of comprehension of the respondents is doubtful, that questionnaires should be provided in the habitual languages of those who will be asked to respond. This is because, and beyond the greater or lesser "mastery" of the foreign language, it is the cultural universe which is associated with it which is often lacking. Each language makes a call on mental structures whose nuances make it rich and provide difficulties. Without speaking of peculiar quirks, the semantics "little nuances" introduced by the translation of questions can have an influence on the answers and their interpretation. A number of precautions are thus to be taken both at the moment of designing the evaluation questionnaire itself and in the formulation of the questions.

Even so, the use of a common language is not ruled out. In certain fields, the disciplinary practice has laid down the use of a single language, generally English. More precisely, this practice has laid down a lexicon and a set of linguistic constructions specific to the discipline, and which are required with the discipline itself. In certain cases, we can even obtain a virtual disciplinary idiom which is understandable by the specialists only, whatever their national and cultural origins. This is to be found for instance with biologists. Here, translation, even if linguistically correct, becomes a handicap for the respondents because they may have lost their own specific cultural "references".

To summarize, as long as a technical language exists which is properly set up, a "jargon" that is disciplinary or professional, and provided that the questionnaire designer himself has assimilated it, it is preferable to use it. But in this case, the questions will be specific, technical, and well formulated for the discipline. Outside this situation, we return to the general case initially mentioned.

4.3. Neutrality of questions

The canonic form of the questions also raises a few problems. Work is being carried out on this point, and some lessons may be learned :

* *Positive or negative form.* Here, account must be taken of a current psychological phenomenon which can be called “the propensity to say yes”. Certain people are in fact “yessayers”, and have a tendency to approve rather than to disapprove their interlocutor. By systematically responding in a positive manner to the questions that are asked of them, they introduce an “acquiescence effect”.

* *Conservative deformation.* This is what American sociologists call an acquiescent psychological phenomenon, but this time in favor of status quo. Thus, questions involving an idea of a change, may receive approval less easily.

* *Tone of wording.* Certain terms used in the formulation of the question may be loaded with an effective meaning that can introduce approval or disapproval. These effective connotations, as for semantic nuances, involve a real modification of the meaning of the question.

* *Illustrative developments.* Practicians of questionnaire surveys know that agreement with a general principle is always greater when it is presented generally and abstractly than when it is accompanied by its possible practical consequences, particularly when the latter are negative. Conversely, positive illustrative developments will have the effect of increasing the rate of approval.

These few examples, which all make greater or lesser call on the psychology of the respondent, illustrate what is sometimes called “the socially desirable character of the question (socially desirable bias)”. **People to whom a questionnaire is submitted tend indeed to give answers that they feel to be, if not desired by the interviewer, at least socially desirable with respect to the current standards and values.**

4.4. The order of succession of questions

The order in which questions are asked may have an influence on the answers and thus on their value. However, the economics of a questionnaire are assessed on the basis of its context, i.e., essentially, the type of subject which it is dealing with, the competence of the respondents, and also some general principles resulting from experience.

The “logic” of the questionnaire does not necessary run from the simpliest to the more complex. Questions involving “elementary” answers are not held to precede questions calling for thinking.

The “logic” of the questionnaire is based on the presumed psychology of the respondent. This is why, **it should begin with questions that can call into question the confidence of the person surveyed** by enabling him, at the outset, to have an idea of what is expected of him. Under the circumstances, on what and how he wishes to be surveyed. These “ice-breaking” questions must arouse the interest of the respondent. This is why he should be made to understand immediately why he is being asked to participate. Here, a short introductory text explaining this, is often well accepted, particularly when the questionnaire is self-administered.

The initial question after introduction : a few criteria to be used

The first question in the questionnaire always plays an important part :

- it must immediately give a precise understanding of the object of the questionnaire;
- it must compel the respondent to answer all the questions.

The initial question must absolutely express as clearly as possible and as accurately as possible what is being asked for. The initial question is in some way the model for the questionnaire.

The criteria for a good initial question are the same as those listed above, added to which are a few qualities of expression and formulation of the subject of the survey, here the fields to evaluate the scientific programme.

Example

Example of an initial question concerning the opinions of contractors on the progress made with a contract :

If the initial question is :

Do you believe that the follow-up carried out by the EC has been performed correctly ?

This question is vague, since what each respondent understands by “follow-up effectively carried out” may vary. Particularly since “follow-up” in this context is not specified. Secondly, making an initial call on the “belief” of the respondent may lead him to think that he is going to be asked something belonging solely to the range of personal convictions where the foundations are not explained. This question can thus be misleading in that the responses expected subsequently will be satisfactory for general moral considerations, prejudices or stereotypes. The assessment of this field must call on something other than vague opinions by the respondent. It must call on factual data, judgments and reasonably “founded assessments”. Here, it is not a question of an opinion survey.

The initial question must take up and formulate the objective of the questionnaire and propose a sufficiently extended but forced choice of answers. For instance -school example- :

The DG. x. (or Service y) was responsible during the conduct of the Programme to which you and your (laboratory, company, etc.), have contributed, for getting information on the progress of your work and helping the development of your relations with the other laboratories and companies, do you think that this has been carried out :

- * completely satisfactorily ?*
- * satisfactorily but could be improved ?*
- * insufficiently though usefully ?*
- * incoherently and of little use ?*
- * completely inefficient ?*

Comment : This formulation proposes opinions that vary between entirely favorable extremes of opinion to entirely unfavorable extremes of opinion, ranging through much more graduated opinions that need some thinking time to be adopted. Even if it is graduated, the range of responses proposed is not strictly linear. It does not encourage the provision of “neutral” or “average” responses, but well informed responses, if it may be expressed in this way. It would even be a good idea for the following question to permit quantification of the number or the frequency of relationships that have taken place between DG x or Service y. Since, at the moment of processing, the cross-correlation of responses to these two questions could be quite enlightening as to the knowledge of the profile of those “satisfied” or “unsatisfied”. One of the assumptions of the questionnaire being perhaps that there is a relationship between the number of meetings and the quality of these, and thus their appreciation.

Thus, the importance will be stressed of constructing the questionnaire in such a way that the answers previously provided do not influence the response to the following question. An attempt will thus be made to avoid “contamination” of subsequent responses.

Ordering of questions : a few criteria to be chosen

In general, the order of succession of questions obeys the following few “rules” :

* **Do not begin with “delicate” questions.** Gaining the confidence of the respondent requires that some easy to answer questions are positioned at the beginning of the questionnaire.

* **“Delicate” questions will be brought about by easier questions.** This does not mean that they must be concentrated at the end of the questionnaire. It is preferable to alternate them with questions whose answers require less effort. Here again, a few transition texts may be useful to relax the respondent.

* **These transition texts,** in addition, make the whole of the questionnaire more harmonious and more coherent. They prevent the respondent sometimes being thrown off course by the succession of questions without evident links between them.

* **Questions relating to the same subject will thus be regrouped.**

* **Factual questions will precede opinion questions.**

* **General questions will precede specific questions : funnel type technic.**

4.5. Order of presentation and the balance of answers

The order between the choice of answers to closed questions, and the balance of these answers are not completely indifferent either, without automatically producing biases. Experiments that have been carried out, have shown that :

*** The longer the list of items proposed, the more a presentation effect exists.** Therefore, and when the questionnaire is administered orally, it is the last item proposed that benefits from this effect, as though it were an echo response. When the questionnaire is self-administered, it is the first item proposed that benefits from the presentation effect (the primacy effect).

“The number of answers given and the balance between positive and negative answers, particularly when they call on the system of values of the respondent, may produce transfer effects of the negative response towards the positive response. S. Juan ¹² has shown that multiplying the positive modalities of the response and by offering only one negative possibility, and thus by proposing an unbalanced response, a much higher positive response percentage was obtained than when the same question was presented in the form of a bipolar balance “yes/no”. **By refining the possibilities of positive presentations of the response, there is a “ranking” effect.** This effect is well known to politicians and marketing specialists, which consists of introducing a greater variety in order to obtain the decision of those hesitating.

In total, we should never forget that the answers always depend, to a certain extent, on the formulation of questions.

5. Examples of formulation of the questions

The following table summarizes and illustrates the main instructions mentioned above.

¹² See : note 9

QUESTIONS LACKING IN PRECISION

What were your impressions of the Commission staff involved in the research programme

What was your principal motivation for submitting a research proposal?

On what grounds do you think your participation in the programme has been useful ? :

- Establish contact in other countries
- Exchange of information and ideas
- Initiate new joint research projects

What kind of collaboration have you had with other scientific task forces involved in the research programme ?

QUESTIONS WITH GOOD FORMULATION

What were your impressions on the Commission staff involved in the research programme

- unknown or few contacts
- efficient, active
- inefficient, inactive

What was your principal motivation for submitting a research proposal?

.....
.....
.....

On what grounds do you think your participation in the programme has been useful ? :

- Establish contact in other countries
- Exchange of information and ideas
- Initiate new joint research projects

- Other :
.....
.....

What kind of collaboration have you had with other scientific task forces involved in the research programme ?

- Exchanges of documents
- Exchanges of researchers
- Visits of laboratories
- Régular contacts by telephone

INSTRUCTIONS TO RESPECT

Close the question

Provide a sufficient space or scroll to make possible development of the response

End multiple choice questions by the possibility of response outside provided items.

Precise the range of responses

QUESTIONS LACKING IN PRECISION

Did you have contacts with DGs ?

- Often
- Occasionally
- Rarely

QUESTIONS WITH GOOD FORMULATION

How many contacts have you had with DGs ?

- more than 10 contacts
- from 5 to 10 contacts
- less than 5 contacts
- no contacts

How many times did you visit DGs ?

What do you think about the evaluation follow-up of the research programme by the Commission ?

What do you think about the evaluation follow-up of the research programme by the Commission ?

- | | | | | | | | | | | | | |
|-------------|--------------------------|---|--------------------------|--------------------------|---|--------------------------|--------------------------|---|--------------------------|--------------------------|--------------|------------|
| productive | <input type="checkbox"/> | * | <input type="checkbox"/> | 2 | * | <input type="checkbox"/> | 3 | * | <input type="checkbox"/> | 4 | unproductive | |
| useful | <input type="checkbox"/> | 1 | * | <input type="checkbox"/> | 2 | * | <input type="checkbox"/> | 3 | * | <input type="checkbox"/> | 4 | useless |
| stimulating | <input type="checkbox"/> | 1 | * | <input type="checkbox"/> | 2 | * | <input type="checkbox"/> | 3 | * | <input type="checkbox"/> | 4 | paralysing |

INSTRUCTIONS TO RESPECT

Formulate the questions with maximum of precision and remain neutral

Introduce subquestions permitting a good understanding of information obtained with the principal question.

Propose an opinion scale based on one or several dimensions

6. The actors of the evaluation

6.1. Who does what ? The three phases to set up a questionnaire

Setting up a questionnaire implies three phases, each of them coming within the competence of a specific actor : The **Programme Management** must define the contents of the programme, the **Evaluation Panel** must define the information required, and the **survey-organisation** must be in charge with formulation and exploitation of the questions.

First phase : definition of the contents

A. Any questionnaire must have one or several clearly defined objectives. The aim is that gathered information, and its subsequent processing, are the best response to initial questions which had justified the whole survey. Defining the objectives of a survey is an important and crucial step. If not done properly, there is a risk of disconnection from the initial intents and initial questions.

Now, a questionnaire, or more precisely information collected by the questionnaire, cannot respond to questions which were not asked in the time of its drafting. In the same way, the interpretation of the results could be depreciated by ignorance of the objectives.

Thus it is important to define objectives of the survey from the outset. That means that the objective(s) of evaluation itself must be specified. **That definition comes within the competence of the management of the programme we try to evaluate.**

B. Only then, we can consider the contents of the questionnaire. It is essentially a matter of listing the key points to be tackled in the questionnaire, of defining as far as possible the nature of information required, of identifying concerned variables, etc. In short, the point is to clearly define the scope of the questionnaire and to make it easily understandable. **Here it is the concern of the members of evaluation panel, to whom the management of the concerned programme must be associated.**

Here, as above, the choosen definitions must be dealt with in a drafting to be recorded in what we call a **“Set-Up of Questionnaire Copybook”**. This copybook will circulate during all the drafting procedure and will be at the same time the memory of the operations and the tool to actualize the objectives and decisions to be taken.

C. To complete the definition of the contents of the questionnaire, it is also important to declare what are the most relevant types of questions, given the defined objectives, the wanted questions and the nature of information to collect. It will be essentially a matter of choice between a questionnaire of opinion, a factual questionnaire, and a mixed questionnaire, knowing the constraints and advantages of each one (*in particular see above III.2.*).

However most parts of factual questions concerning evaluated research programmes and contracting organisations, will be recovered, in the example shown in annex 1, from data base managed by DG XII : "AMPERE". In order to reduce the volume of evaluation questionnaires given to organisations, these factual questions will not be asked again (*see above I.6.*). Possibly, and only if it is of interest for the programme evaluation, we can ask confirmation and adjustments of data already available. The questionnaire to be set up must remain clear, precise, reliable and operative.

Second phase : the formulation of the questions

This whole phase will be in the scope of the survey organisations.

The choice of the question form will be their first concern. For them, it will be a matter of deciding how to obtain the best response, the response which will give the wanted information as precisely and completely as possible. Must we use an open question or a closed question ? Unless we prefer a simple scale ? (*see above III.2.*).

We remember that in general we advise using rather closed questions and simple scales because open questions, especially when they are complex or not precise in their formulation or instruction, create great difficulties in validating the results. Often, there are one or two general ideas only, possibly "vague", and difficult to include in an evaluative process. We suggest also, if evaluators decide nevertheless to use open questions, to provide in this type of questionnaire a previous codification of possible responses. It is necessary to give the person in charge of the exploitation of results, who has not always conceived the questionnaire, a guideline to interpret responses.

Then comes the formulation of the question itself (*see above III.4.1., III.4.2. et III.4.3.*). This exercise requires real qualities of empathy, patience and modesty. It is necessary to give a question careful consideration, to force oneself to propose a lot of formulations, hence, to fly in the face of initial facts in order to adopt the formulation which will best suit the requirements of a "good question".

Following the formulation of the questions, we must think (or think again) of the general economics of the questionnaire (*see above III.4.4. et III.4.5.*). We must choose the order of the questioning which will appear most logical to the respondent, in any case the most satisfactory. The question pitfalls, to test the sincerity of a first response, must be avoided in an evaluation questionnaire. They do not always give the expected results.

Third phase : the test

This is a crucial phase in the setting up of the questionnaire. Ignoring this phase is to risk discovering, after the interview, that the questionnaire is inadequate and that a lot of questions could not be exploited in the correct manner. This aspect of the setting up of the evaluation questionnaire is important, we'll write again about it.

This third phase must always be undertaken by the survey organisation.

6.2. Identification of the respondent

The identification of the respondent is given within AMPERE, in any case when they concern those responsible for the laboratory, the team or the firm with which the contract has been signed. The name of this responsible is on the questionnaire sent to contracting organisation.

For certain questions, it is possibly preferable to collect the response from another person. If this person is personally known, we just have to write his/her name. If it is not the case, we must clearly explain in the instruction what kind of respondent is looked for.

Then, we must think about this aspect when we are conceiving the questionnaire. But generally, it is better not to ask for several respondents. We recommend observing the following rule : only one respondent per questionnaire.

6.3. How to test the questionnaire ?

This phase, as mentioned, is not to be neglected. But there is no well recognized method to do it correctly. Specialized handbooks sometimes advise that the questionnaire is tested on a sub-sample of respondents. Depending on the conditions of the formation of the total sample (random sampling, method of quotas), the sub-sample can be formed by just a part of the total sample or by parts of its different classes. This procedure is very expensive -particularly in time- and encounters some difficulties to be undertaken on a regular basis.

Often we proceed in an empirical and quite satisfactory way. Then, we try to **make an ad hoc test**. It is this procedure which is to be adopted in most cases.

On the one hand, because the homogeneity of the respondents is relatively strong : the evaluation questionnaire cannot be ranked as a questionnaire of opinion for the general public. The recipients of the evaluation questionnaire are sharing the same culture indeed. The culture of their subject at first, but also, without making too strong sociological hypothesis, a general culture that leads to share some knowledges, cognitive skills, values,...

On the other hand, **the test we recommend must be less a general understanding test, than a test focusing on difficulties dealing with particularisms of the intending respondents**. It is the case for the used language. It is the case for legal status of contracting organisations : these status give more or less freedom to the respondent to manage or to organize the contracting center. They make him more or less sensitive to some aspects of the evaluation. It is also the case for the internal principles of management of the laboratories, research teams, etc., which can differ from one country to another and remove the competence fields of the respondent.

Thus, it is necessary to pay attention to what can be the more discriminating for the respondents (language, status of the respondent, subject fields, insertion network, etc.) and to conduct tests with some representatives of lonely sub-groups.

However we need to consider each case individually.

In some cases, a few intending respondents can be asked to be tested. That means that we must state clearly that it is a test, so that they do not think, when they receive the final evaluation questionnaire, that they have already answered it.

In other cases, and because the respondents are not inclined to “play the test game”, it is necessary to get round that obstacle. Because trying to impose the test can lead them to refuse the final questionnaire, which will be contrary to the designed aim. Thus, we can sometimes be content with testing persons estimated as equivalent, that means with the same discriminating characteristics. For instance, it can be a representative of another non contracting laboratory in the framework of the evaluated programme, but well known from another source, or contracting in the framework of another programme.

All things considered, the test must not be a “proof”, an additionnal and discouraging difficulty for the intending respondent. The aim of the questionnaire remains the programme evaluation and must not be compromised by this exercice. The test deals with the form of the questionnaire -even if the latter affects the ground- and a lot of non recipients can judge the form and shed a different light on the quality of the questionnaire itself.

6.4. During and after the survey : what is to be kept ?

Here the question is less a question of the records than that of secondary processing of surveys.

Responses given on paper must be recorded in extenso on a magnetic support and saved. After that, the paper can be destroyed.

Data exploitation will be realized using this magnetic support. The results of this exploitation (data tables : straightforward frequencies, cross-frequencies, hierarchical analysis, etc.) will be the background to elaborate the evaluation judgement. But it is not because this presentation is available that we must destroy the initial data.

These data are a kind of treasure to preserve absolutely. For each secondary exploitation -even for just an easy comparison between two programmes-, it is generally initial data that must be exploited again. The evaluation, which often will take the form of an evaluation report, will show these data in an aggregated form, will process them with processings according to objectives of evaluative research, etc. and “will forget” initial data. A secondary exploitation will need the memory of these data. Hence, the importance of their conservation.

Various operations have to be carried out : the clean up of the data, normalization of the files, elaboration of the codebook or the files guideline, etc ¹³. Without going further, let us remember that the “**Set-Up of Questionnaire Copybook**”, which was circulated during the evaluation process, being at the same time a record of the operation and of the objectives of the evaluation survey, can be completed and can receive the directions for the exploitation and the preserving of the data files. Then this copybook can be used totally as **the record of the survey**.

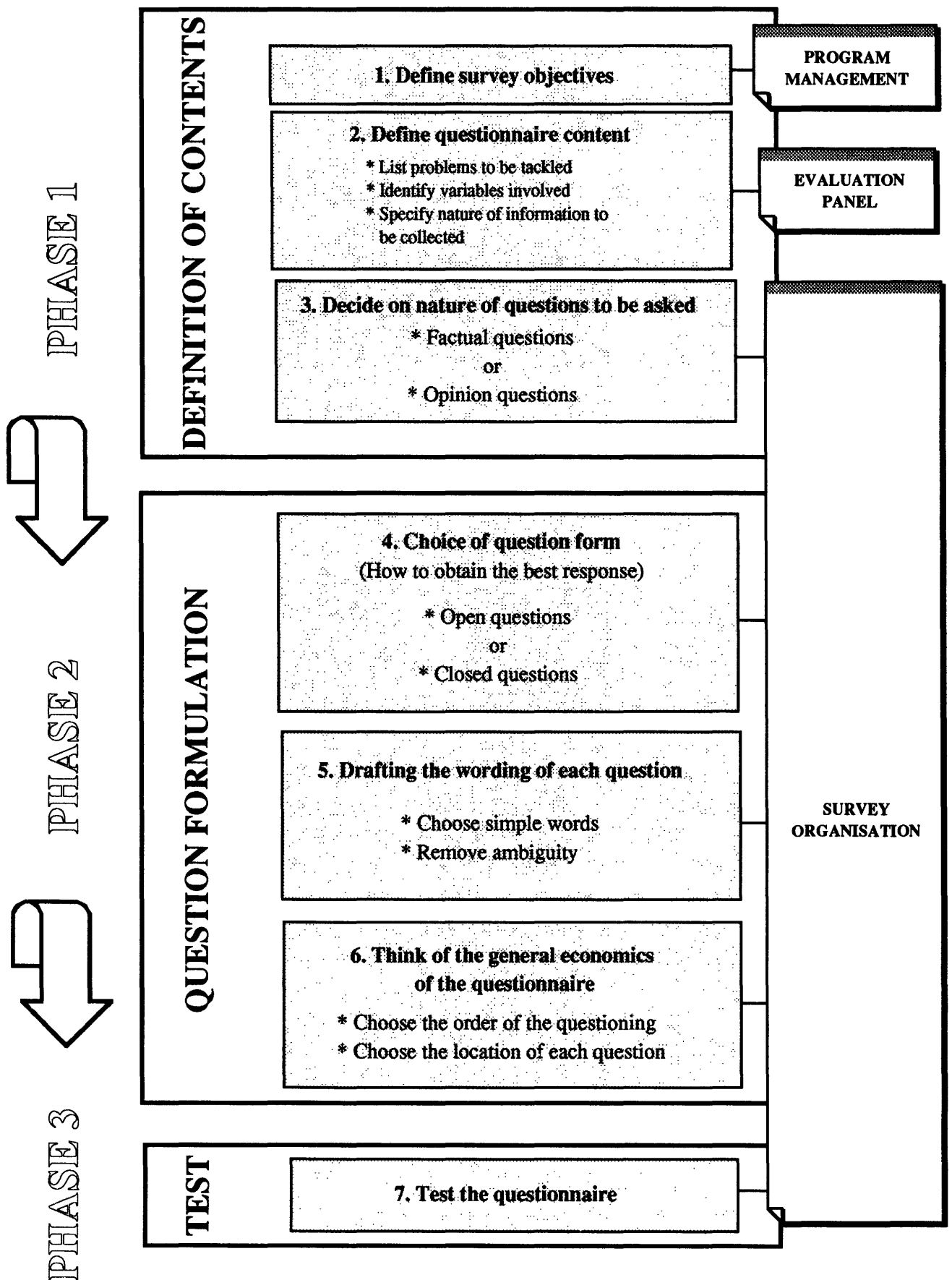
CONCLUSION

It seems to us that the designers of R&D assessment questionnaires must always remember, during each of the phases of drafting the questionnaire, both the principles and the elementary rules that have been recalled above. In addition, as we have seen, drafting a questionnaire must necessarily go through a certain number of phases, the order of which is not indifferent.

The pertinence and efficiency of a questionnaire involves the scrupulous adherence to instructions and the chronology illustrated by the following diagram.

¹³ About this point, see : Frederic Bon, “How to organize statistical software for ecological data processing : the example of French electoral data”, Proceedings of the 1982 IASSIST Conference.

THE 3 PHASES OF DRAFTING A QUESTIONNAIRE



Annexes

Example of procedures for using internal data :

AMPERE Programme of DG XII

1. General principle

The management of research programmes, from the selection of projects up to the final report, is assisted by administrative follow-up by the Commission, during which very many items of information are gathered in a systematic way. These items, available on computerized media, are obviously a source of valuable information for the evaluation. They concern the description of the projects financed (subject, budget, outline of the contractors) and their conduct.

In addition, as it in particular contains various elements such as the detailed description of the contractors and the contract, it shortens the evaluation questionnaires and avoids asking again for the same information which is already available.

In DG XII the available data are in principle in AMPERE, which is an ambitious and complex data base managed by the DG. The user is certain to find :

- standardized information contained in the proposal forms (replies to calls for proposals),
- the standardized information contained in the contract forms,
- the summarized descriptions of proposals and contracts (in full),
- the file administrative items (in AMPERE C).

There are also fields provided for a large quantity of information (for instance the reasons for refusal of proposals, historical information on the preceding participation of the contractors in other programmes, as individuals, organisations or subsidiaries of organisations).

In addition, in each scientific directorate, there is a certain amount of information, in particular on the experts and their assessments, which could be systematically analyzed to give global indicators on the selection procedure (for instance what are the critical principles made in general to the proposals, what has been the average response time,

etc.). It is also possible to extract summaries of programmes that can be analyzed by means of key-words or directly by lexical analysis.

The data available in AMPERE contains at least the information requested in the proposal files filled in by the contractors :

- research title
- name and address of contracting organisations
- topic covered by the research
- number of participants
- summary of the object of the research
- budget (requested/granted)
- beginning and duration of work
- co-financing outside EC
- name and address of scientific person responsible
- breakdown of costs by major headings
- number of months/man of research workers, technicians and others involved in the project
- type of bodies that may be dependent organisations
- etc.

It is therefore not necessary to ask the participants for all this information again : given the contract number or the name of the laboratory, using AMPERE, it is possible to fill in each questionnaire with items that are already recorded.

In addition, these names and adresses make it possible to send personalized questionnaires to the persons or organisations included in the file.

2. The procedure

The survey procedure recommended is the following. This presentation may appear a little detailed and technical to the unfamiliar person, but it is simple for computer personnel and statisticians, and enables the data to be used as efficiently as possible.

A. Analysis of the prior results drawn from AMPERE before even launching a survey. (Optional, at the discretion of the panel).

As an indication, it is possible to get from AMPERE or AMPERE C processing of the type:

- empirical selection criteria of projects used (comparisons, as to amounts, profiles of respondents, topics, etc., and proposals and contracts) ;
- in which way the proposals have been modified when they have become contracts (comparison of proposals selected with final conditions with the corresponding contracts) ;
- to what extent the respondents and the contractors are part of the “customers” of the DG XII (research for the existence of prior contracts for proposals with the DG XII) ;
- what is the average payment time, what are the characteristics of the contracts that have had exceptionally high payment periods ;
- what are the characteristics of the contracts that have had numerous exchanges with the Commission at the accounting level ;
- what are the characteristics of the contracts that have respected costs and time delays provided for in the contracts.

B. List of information needed by the panel which will be divided into 5 major categories:

- a- description of proposers/contractors ;
- b- facts on the contract and its follow-up ;
- c- opinion of respondents on the contract and its follow-up ;
- d- scientific results of the contract ;
- e- socioeconomic effects of the contract

Certain parts of this information may be obtained from the forms or questionnaires. The principle will be to get by questionnaire what is already in the forms held on AMPERE or in the scientific directorates, and to merge these results with those already available.

C. Construction of questionnaires. Information requests of the type c, d, e and possibly of the type a and b will be formulated in the form of questions, assembled in a questionnaire, based on guidelines.

D. Extraction of participants in the programme from a list. The list of proposals of a programme considered will be taken from AMPERE and added to this will be an individual identification key. Also drawn out will be a subfile consisting solely of contractors.

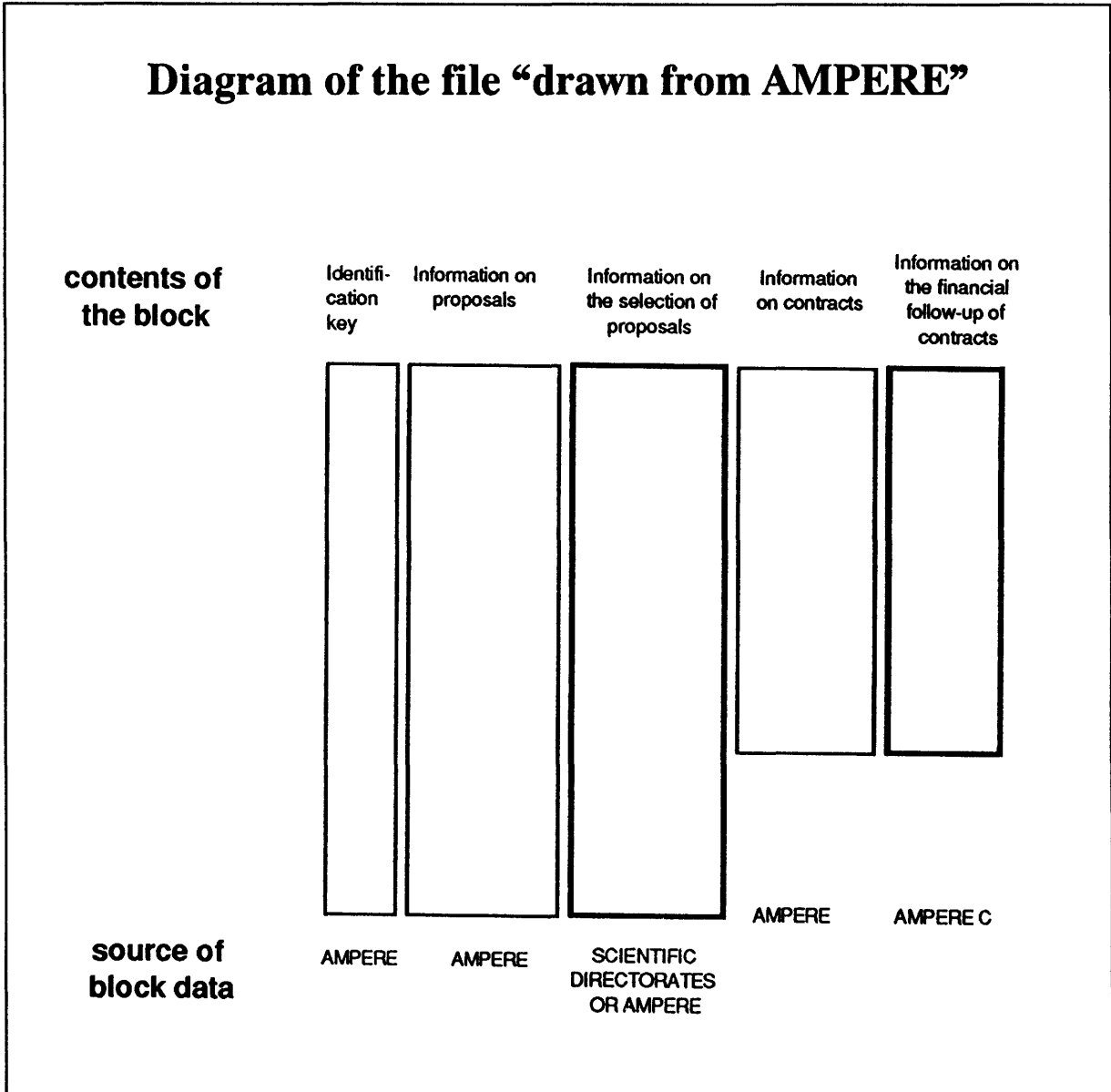
E. Field of the survey by questionnaires. Each of the proposers will be sent (or only contractors) the personalized questionnaire comprising the name of the recipient and an identification key with, possibly, the personalization of certain questions based on the characteristics of the contract.

The questionnaire will thus be drawn up from AMPERE. If possible, a minister paper and window envelope will be used to facilitate mailing (avoiding the use of labels).

F. Recovery of information which the Commission already has. While those questioned send their answers, the secretary of the panel organizes the output of information drawn from AMPERE which will be merged with the responses to the questionnaires.

This file must have the following form : one line per entity surveyed (this may be individuals, laboratories, contracts, participants, depending on the evaluative approach chosen). Each column will contain a value of a variable, drawn from AMPERE, AMPERE C, or possibly from other sources at the disposal of the scientific directorates (for instance, data on the assessments at the moment of selection of the projects or the validation of the final report). The first column will contain the identification key defined in D.

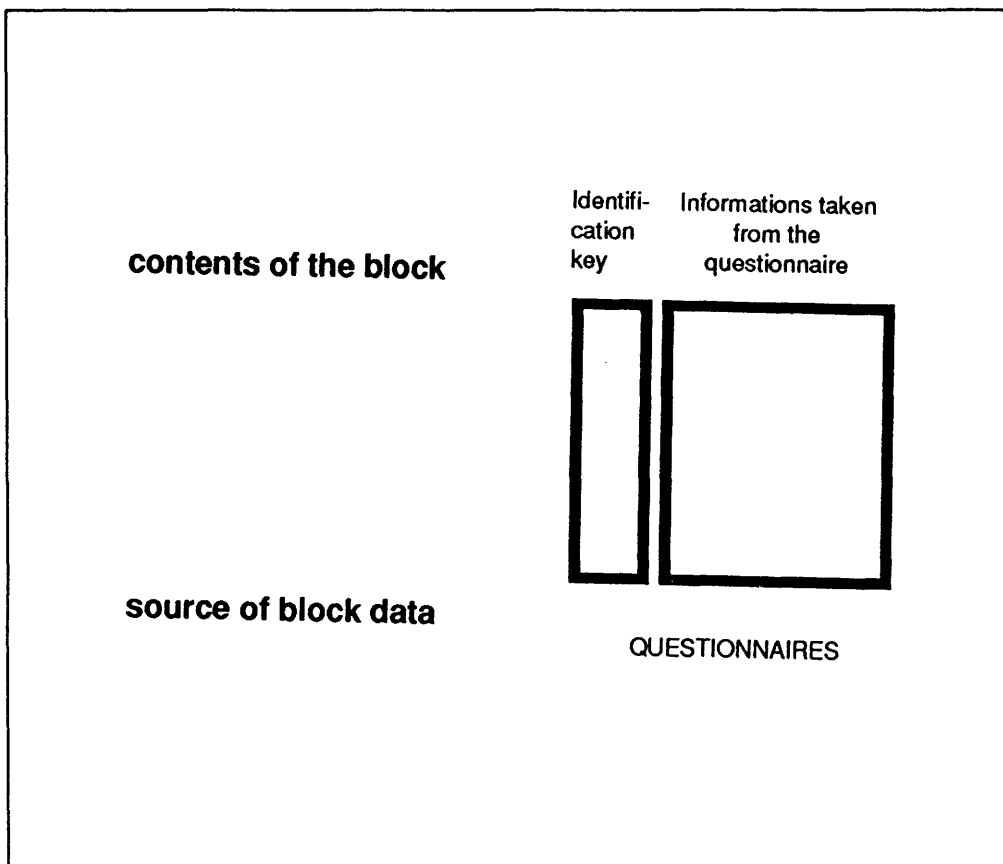
This rectangular file will be in the form of non formatted ASCII, so that it can be used easily by people who will deal with the survey, whatever their software and their processing platform (since the panels in general subcontract the survey to consultants).



NB : The widths of the block on the diagram are not proportional to the quantity of variables that they contain. The length of the columns are unequal since, if each contract comes at least from one proposal, some proposals are refused. It is, however, worthwhile keeping the latter in the file, since they enable comparisons to be made, and the questionnaire may possibly be sent to all proposals for a programme and not only to contractors.

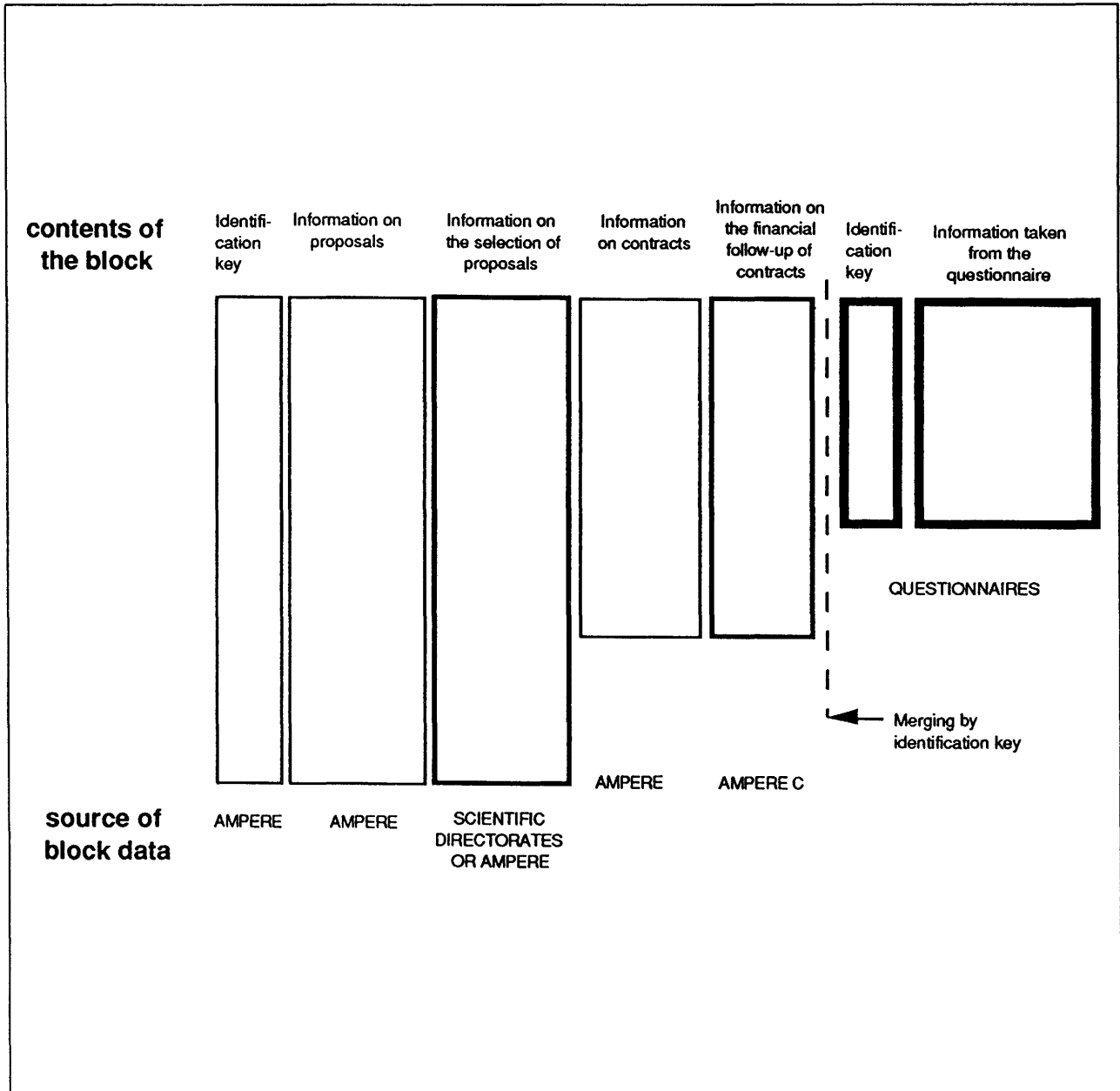
This file must be documented, i.e., for each variable (column), the translation of the codes representing the mode must be supplied. It appears desirable that the mode should be exactly the wording lengths corresponding to AMPERE, with possibly a word of explanation if there are technical conditions. The file should be supplied, as is usual, with the listing of the first lines, and a few simple statistical sorts permitting checking that the user has correctly read.

G. Data entry of survey questionnaires. The questionnaires received will be entered and put in the same form (rectangular ASCII file where the lines are entities surveyed in each column contains an answer mode, the first column containing the identification key that figures on the corresponding questionnaire).



H. File exchange. The personnel responsible for processing the survey then carry out a file exchange. They supply the file coming from the entry of the questionnaires to the panel, and receive the file comprising the information taken from AMPERE defined in F. The file supplied must be, like that taken from AMPERE, suitably documented.

I. File merging. They then merge the files and have for processing, a file permitting comparison of the facts and opinions on the conduct of the contract and enabling them to be related with an extensive description. Processing is made at the request of the panel.



NB : The last block in the file is shortened because of non-response to the questionnaire. The fact of having built up the file on a complete base permits, in the final processing, the evaluation of the nature of the biases introduced by non-responses.

J. Data processing. This aspect is dealt with in Part II.

K. Submission of reports and archiving of data. At the end of the study, the personnel responsible for processing must remit a complete file :

- the complete files in rectangular ASCII form, suitable documents ;
- a survey report describing precisely the procedure used, including the dates, the dates of reply, refusals, instructions for recoding where appropriate.

In this way, it is possible to conserve and make possible the use of the various evaluation surveys.

It is highly recommended to follow this procedure, both to adhere to the general organisation of the processing and archiving of data in DG XII and ensure comparability with subsequent evaluations, and to benefit from a past traced out by professionals. Experience shows in fact that surveys given to subcontractors by panels which, unaware of the practices in the field, return to the procedures proposed by the subcontractors, who then themselves discover the surveys on the EC programmes, most often producing bad rather than good surprises.

LIST OF EVALUATION REPORTS ON RESEARCH PROGRAMMES ANALYZED

1. **Evaluation des sous-programmes communautaires de R et D dans le domaine de l'économie d'énergie et de l'énergie solaire. Evaluation de la Recherche : Rapport n° 1.**
2. **Evaluation des sous-programmes communautaires de R et D dans le domaine de l'énergie géothermique, de la production et utilisation de l'hydrogène et de l'analyse de système : développement de modèles. Evaluation de la Recherche : Rapport n° 2.**
3. **Evaluation du Bureau Communautaire de Référence. Programme d'action indirecte (1975-1978). Evaluation de la Recherche : Rapport n° 3.**
4. **Evaluation du programme communautaire d'action indirecte concernant la gestion et le stockage des déchets radioactifs. Rapport d'évaluation n° 4.**
5. **Evaluation du programme de recherche communautaire dans le domaine de la radioprotection (1976-1980). Evaluation de la Recherche : Rapport n° 8.**
6. **Evaluation du programme communautaire "Formation scientifique et technique". Evaluation de la Recherche : Rapport n° 10.**
7. **Evaluation of the Community's Environmental Research programmes (1976-1983). Research Evaluation - Report n° 14.**
8. **Evaluation of the Community's primary mineral raw materials programme. Research Evaluation - Report n° 16.**
9. **Evaluation of the Community's nuclear reactor safety research programme. Research Evaluation - Report n° 18.**
10. **Evaluation of the COST projects on "Food Technology". COST 90, COST 91, COST 90 bis, COST 91 bis (1978-1987). Research Evaluation - Report n° 21.**
11. **Evaluation of the Community cost-shared research programme on solar, wind and biomass energy and of the Joint Research Centre's programme on non-nuclear energies (1979-85). Research Evaluation - Report n° 22.**
12. **Evaluation of the Community Bureau of Reference 1983-87. Cost-shared research. Research Evaluation - Report n° 23.**
13. **Evaluation of the R & D programme in the field of Non-Nuclear Energy (1985-1988). Research Evaluation - Report n° 24.**
14. **Evaluation of the first BRITE programme (1985-1988). Research Evaluation - Report n° 25.**
15. **Evaluation of Community Research in the field of Management and Storage of Radioactive Waste (Cost-shared research 1980-1986). Research Evaluation - Report n° 26d.**
16. **Evaluation of the Community's Research Programme on Decommissioning of Nuclear Installations (Cost-shared research 1979-1986). Research Evaluation - Report n° 26e.**

17. **Evaluation of the Biomolecular Engineering Programme-BEP (1982-1986) and the Biotechnology Action Programme-BAP (1985-1989).** Research Evaluation - Report n° 32.
18. **Evaluation of the European Advanced Materials Research Programme - EURAM - (1986-1989).** Research Evaluation - Report n° 33.
19. **Evaluation of the R & D Community Programme in primary and secondary Raw Materials (1982-1985). Progress of the Programme (1986-1989).** Research Evaluation - Report n° 38.
20. **Evaluation of the Agricultural Research Programmes (1976-1978, 1979-1983 and 1984-1988).** Research Evaluation - Report n° 39.
21. **Evaluation of the Third Community Programme "RADIOACTIVE WASTE MANAGEMENT AND STORAGE" (1985-1989).** Research Evaluation - Report n° 40.
22. **Evaluation of the SCIENCE/STIMULATION Plans (1983/1985 - 1985/1988 - 1988/1992).** Research Evaluation - Report n° 41.

Commission of the European Communities

EUR 14339 – A Guideline for Survey-Techniques in Evaluation of Research

S. Lahlou, R. van der Meijden, M. Messu, G. Poquet, F. Prakke

Luxembourg: Office for Official Publications of the European Communities

1992 — 72 pp. — 21.0 x 29.7 cm

Science and Technology policy series

EN

ISBN 92-826-4093-0

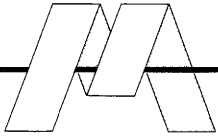
Catalogue number: CG-NA-14339-EN-C

Price (excluding VAT) in Luxembourg: ECU 9.00

This report has been written in collaboration with two consultants, specialist in the field of enquiries. It is based, of course, on their professional skills but relies also, as much as possible, on the experience capitalized within the Commission of the European Communities. It aims at the establishment of optimised procedures to permit a more rational use of this tool represented by enquiries. It should also provide an opportunity to compare efficiently parallel evaluations of different programmes as well as successive evaluations of the same programme.

Initially intended for the Commission Evaluation Services, especially the one at DG-XII, this document results to be interesting for a much broader public than the mere Commission servants. It addresses indeed the basic survey-techniques in a simple manner, accurate and understandable, applying them to a specific case: the Evaluation of the Commission R & D-Programmes.

After a short introduction where the scope and the context of the study are addressed, the report considers first the procedures to be followed for the set-up of questionnaires as also for the analysis and the presentation of the results. It provides then concrete examples to solve the practical problems occurring in the formulation of questions and their adequation to the needs. All this is summarized in a few "technical recommendations" and a set of 19 elementary rules. In an annex, an example for the use of external information in the improvement of the results of the enquiry is proposed.



NOTICE TO THE READER

All scientific and technical reports published by the Commission of the European Communities are announced in the monthly periodical '**euro abstracts**'. For subscription (1 year: ECU 110) please write to the address below.

Price (excluding VAT) in Luxembourg: ECU 9.00

ISBN 92-826-4093-0



OFFICE FOR OFFICIAL PUBLICATIONS
OF THE EUROPEAN COMMUNITIES

L-2985 Luxembourg



9 789282 640937 >