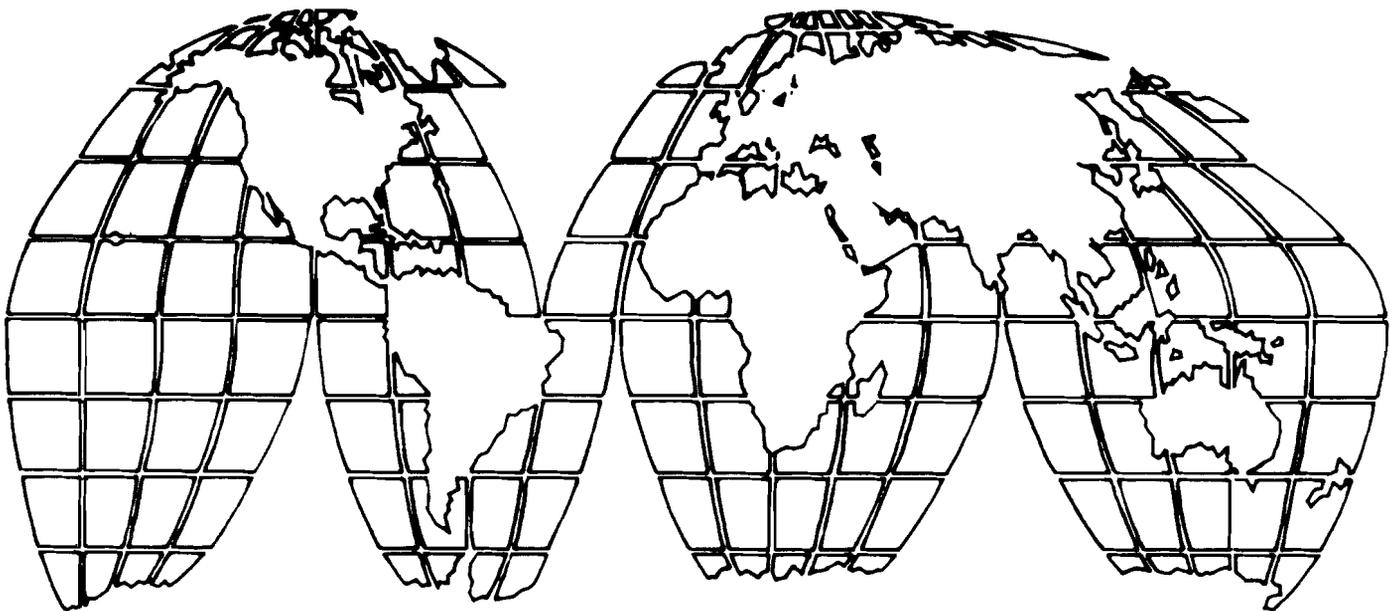


Conducting Mini Surveys in Developing Countries



December 1990

Agency for International Development (A.I.D.)

Washington, D.C. 20523

PN-AAX-249

This report and others in the evaluation publication series of the Center for Development Information and Evaluation (CDIE) may be ordered from

**A.I.D. Development Information
Services Clearinghouse
1500 Wilson Boulevard, Suite 1010
Arlington, va 22209-2404
telephone: (703) 351-4006**

A list of all CDIE evaluation publications is available from

**PPC/CDIE
Room 105, SA-18
Agency for International Development
Washington, D.C. 20523
USA
telephone: (703) 875-4818**

CONDUCTING MINI SURVEYS IN DEVELOPING COUNTRIES

A.I.D. PROGRAM DESIGN AND EVALUATION METHODOLOGY REPORT NO. 15

by

Krishna Kumar, Senior Analyst
(Center for Development Information and Evaluation
Bureau for Program and Policy Coordination, A.I.D.)

U.S. Agency for International Development

December 1990

The views and interpretations expressed in this report are those of the author and should not be attributed to the Agency for International Development.

TABLE OF CONTENTS

	<u>Page</u>
Foreword	v
Acknowledgments	vi
1. Planning Mini Surveys	1
1.1 Description of Mini Surveys	2
1.2 Advantages and Limitations of Mini Surveys	3
1.3 When Are Mini Surveys Most Appropriate?	4
1.4 Conducting Mini Surveys	5
2. Preparing Questions	9
2.1 Wording Questions	9
2.2 Length of Questions	10
2.3 Open-Ended and Closed Questions	10
2.4 Constructing Response Categories	13
2.5 Making Questions Specific	14
2.6 Avoiding Double-Barreled Questions	15
2.7 Framing Questions To Aid Recall	15
2.8 Avoiding Loaded Questions	17
2.9 Wording Sensitive Questions	18
3. Designing the Questionnaire	21
3.1 Order of Questions	21
3.1.1 The First Question	21
3.1.2 Demographic Questions	21
3.1.3 Funnel and Inverted Funnel Sequences	22
3.1.4 Chronological Order	22
3.1.5 Changing Topics	24
3.2 Length	24
3.3 Format	24
3.4 Translation	25
3.5 Pretesting	25
4. Selecting Respondents	27
4.1 Probability Versus Informal Sampling	28
4.2 Probability Sampling Methods	29
4.2.1 Simple Random Sampling	29
4.2.2 Systematic Sampling	30
4.2.3 Stratified Sampling	31
4.2.4 Cluster and Two-Stage Sampling	33
4.3 Informal Sampling	34
4.3.1 Convenience Sampling	34
4.3.2 Judgment Sampling	37

TABLE OF CONTENTS (cont.)

	<u>Page</u>
4.3.3 Snowball Sampling	38
4.3.4 Quota Sampling	38
4.2 Sample Size for Mini Surveys	41
5. Asking Questions	42
5.1 Initial Contact	42
5.2 Using the Questionnaire	43
5.3 Probing Techniques	45
5.4 Recording the Interview	46
5.5 Editing the Interview	47
6. Analyzing and Presenting the Survey Data	49
6.1 Coding Data	49
6.2 Statistical Analysis	52
6.2.1 Frequencies and Percentages	52
6.2.2 The Mode, Median, and Arithmetic Mean	55
6.2.3 Range, Variance, and Standard Deviation	56
6.2.4 Pearson's Coefficient of Correlation	58
6.2.5 Other Statistical Measures	60
6.3 Presenting Data	60
6.3.1 Constructing Tables	60
6.3.2 Presenting Graphics	61
6.3.3 Preparing the Report	62

References

FOREWORD

This report on Conducting Mini Surveys in Developing Countries by Dr. Krishna Kumar is the most recent addition to the Center for Development Information and Evaluation series in A.I.D. Program Design and Evaluation Methodology. It is intended to serve as a guide for those planning and undertaking evaluations as well as other analytical efforts related to development activities. As noted in the text, the mini survey can be valuable in circumstances where large-scale surveys or experimental research design approaches are not feasible, but where systematic empirical information would contribute to the rigor and credibility of rapid appraisal evaluation approaches. This report by Dr. Kumar, while comprehensive, has been written for practitioners, with minimal use of technical terminology and frequent use of examples to illustrate concepts and principles.

John Eriksson
Associate Assistant Administrator
Center for Development Information and
Evaluation
Bureau for Program and Policy Coordination
Agency for International Development
December 1990

ACKNOWLEDGMENTS

I received invaluable advice, comments, and suggestions from Doctors Kurt Finsterburch (University of Maryland), Michael Hendricks (Hendricks & Associates), Charles R. Perry (Bureau of the Census), and Annette Binnendijk and John Eriksson (Center for Development Information and Evaluation) on the earlier versions of this paper. I wish to record my deep gratitude to all of them.

1. PLANNING MINI SURVEYS

...the extensive questionnaire survey with the 30 pages of questionnaire (multidisciplinary, each discipline with its own questions), which if asked are never coded, or if coded never punched, or if punched, never processed, and if processed and printed out, never examined, or if examined, never analyzed or written up, or if analyzed and written up, never read, or if read, never understood or remembered, or if understood or remembered, never used to change action. Rural surveys must be one of the most inefficient industries in the world.

Robert Chambers (1981)

A survey of 20 respondents is better than no survey. For example, in the absence of a survey one only has hunches about who is for or against various policy alternatives facing a decision maker. The sample of 20 at least gives a rough idea of what people are thinking and is better than one's hunches.

Kurt Finsterbusch (1976b)

Surveys are undoubtedly the most widely used method for data collection. In popular belief, the word survey is synonymous with social research--and not without justification. Analysts can hardly think of basic and applied research in social, economic, political, and cultural affairs without relying on surveys in one form or another.

Surveys involve direct collection of information from individuals. The basic element of a survey is a structured questionnaire administered in person, by telephone, or through the mail to respondents who are carefully selected, generally, although not always, on the basis of probability sampling. The responses gathered from the questionnaire are coded and statistically analyzed to draw findings and conclusions.

The popular perception of surveys is that of large investigations involving hundreds and even thousands of respondents generating data on a multitude of variables. Such surveys are undoubtedly costly and time consuming and require an efficient organizational apparatus. However, surveys can also be done on a smaller scale by concentrating on a few variables and using a small sample. These surveys are referred to in this paper as "mini surveys."

1.1 Description of Mini Surveys

Mini surveys, as conceived here, have the following features:

First, they focus on a narrowly defined issue, question, or problem. For example, they address questions such as What proportion of targeted farmers is using the recommended technical package? How do project participants evaluate the services provided by a microenterprise development project? and Are majority of farmers willing to pay user fees to avail themselves of necessary health facilities?

Second, which follows from the above, the number of questions is deliberately kept small, ranging from 15 to 30 in most instances. In this respect, mini surveys differ significantly from traditional household or agricultural survey in which questionnaires may run into several pages. Mini survey questionnaires are designed such that interviews can be completed at most within half an hour.

Third, sample size is kept small and usually ranges between 25 and 70 cases to save time and resources. This is indeed the most important characteristic of mini surveys that distinguishes them from large socioeconomic surveys. The small sample size has several implications about the generalizability of findings, which are discussed later in this section.

Fourth, mini surveys largely, though not exclusively, use closed questions. Such questions list major response categories, and respondents simply identify one or more categories that they consider appropriate. The essential idea is to quantify responses so that statistical analysis can be done rapidly. In this respect, mini surveys differ from key informant interviews or informal surveys in which open-ended questions are used.

Finally, although the use of probability sampling is preferred, informal sampling procedures are also acceptable when the former is not feasible because of time and resource constraints. In probability sampling each unit in the population has an equal chance of being selected so that the resulting sample is representative. As discussed in Section 4, probability sampling is efficient and ensures unbiased findings. Informal sampling, on the other hand, is based on convenience and individual judgment and can therefore be biased.

1.2 Advantages and Limitations of Mini Surveys

The advantages of mini surveys are obvious and require little elaboration. First, unlike other rapid, low-cost data collection methods, mini surveys generate quantitative data. Thus, on the basis of their findings, analysts can say that 50 percent of the women farmers surveyed indicated that the technical assistance provided by the project was valuable or that 40 percent reported that their incomes increased because of their participation in the project. One limitation of qualitative data-collection methods such as key informant interviews, rapid rural appraisal, or focus group discussions is that quantitative data cannot be generated by them.

Second, mini surveys can be completed within 3-7 weeks, which makes them practically the only alternative when quantitative data are needed but not enough time is available to mount a comprehensive survey. For example, when an evaluation team has only about 4 weeks for a field visit to assess the impact of a microenterprise project, it obviously will not be able to launch a comprehensive survey of the local entrepreneurs assisted by the project; however, the team will easily be able to design and implement a mini survey that can produce reasonably credible data for the evaluation.

Third, nonsampling errors tend to be low in mini surveys. One reason is that only a few interviewers are involved, thus they are better trained and supervised. Furthermore, the small sample size and fewer questions reduce interview and coding errors. And the investigator has a better grasp of the data because of the small volume of data involved. The cumulative result is that the overall quality of the data tends to be quite satisfactory in mini surveys.

Finally, mini surveys can be generally managed with relatively low cost. The small size of the sample and of the questionnaire minimizes manpower requirements. In fact, an investigator does not require much outside help and can manage with two or three full- or part-time assistants.

However, despite the advantages described above, mini surveys have several limitations that should be carefully weighed before they are used.

First, the findings of mini surveys are less generalizable than those obtained from large surveys. This is especially true when probability sampling is not used. Analysts cannot be sure that the sample is representative of the population, nor can they compute the sampling error. Even experienced researchers can

make mistakes when they rely on informal sampling. For example, in a survey of rural households in Lesotho, interviewers were asked to visit villages and interview available heads of the households. A significant number of the men in these villages worked in the South African mines for wages much higher than those prevalent in Lesotho; thus such households enjoyed higher incomes and assets. However, because only available heads of households were interviewed, the households whose members worked in South Africa were naturally underrepresented in the sample. Thus findings of this otherwise carefully planned survey were undoubtedly inaccurate.

Second, in many instances the small sample size does not permit an elaborate statistical analysis. For example, if out of 50 farmers in the sample only 8 are women, the analyst cannot make a comparative analysis of behavior of male and female farmers.

Finally, credibility is always a problem with mini surveys. Many policy- and decision-makers may consider (and not without justification) findings from mini surveys unreliable because of their small sample size. The remark is often heard, "So, you are generalizing about the whole project area on the basis of 35 respondents!"

1.3 When Are Mini Surveys Most Appropriate?

At the outset it should be recognized that mini surveys should not be construed as substitutes for carefully designed and efficiently implemented large surveys to study complex social and economic subjects. When rigorous and reliable data from heterogeneous populations are needed for major policy or program initiatives, large sample surveys may be indispensable. However, there are ample situations in project and program settings when the data generated by mini surveys will serve the purpose. Some situations in which such surveys are most appropriate are as follows:

First, when limited time and or resources do not permit or justify the launching of a large sample survey. This is undoubtedly the most important justification for mini surveys. For example, such surveys may be extremely useful for conducting feasibility studies, preparing project papers, assessing beneficiary responses, and preparing terminal and impact evaluations. In such situations the analyst is more interested in learning about broad patterns, trends, and tendencies than in knowing precise measurements. For instance, in evaluating an agricultural project for small farmers, it is often immaterial if the

approval rating among the beneficiaries is 60 or 63 percent. The difference of 3 percent will hardly make any difference in the conclusions and recommendations of the evaluation team.

Second, when the purpose of the survey is to develop questions, hypotheses, and propositions for further testing. In such cases, mini surveys can be a prelude to more comprehensive, large-scale surveys. They can provide reasonable information to sharpen study questions, design relevant questionnaires, and develop sampling strategies.

Third, when some quantitative data are needed to supplement qualitative information. For example, when certain conclusions about the supply of agricultural inputs by private traders have emerged from key informant interviews with concerned project staff, government officials, experts, and a few farmers, but the A.I.D. manager wants further confirmation from the potential users of the inputs. In this particular case, a mini survey can help to measure the perceptions of beneficiary farmers.

Some examples of the situations in which mini surveys were considered appropriate are given in Box 1.

1.4 Conducting Mini Surveys

The following steps are involved in conducting mini surveys.

1. Study objectives: The first step in planning a mini survey is to formulate precisely the objectives of the inquiry by listing study questions. In most instances, the study questions are stated in the Scope of Work prepared by the concerned A.I.D. office. The investigator should discuss with the relevant staff any questions that are unclear or not specific to sharpen the focus of the survey and avoid any possible misunderstanding.

A few precisely formulated study questions help to determine what is and what is not to be covered by a survey. During the planning stage of the survey there is always a temptation to seek more information than that which can be realistically used by managers and policymakers. The focus on study questions curbs this temptation.

2. Review of the literature: The next step is to conduct a review of existing information. Such a review should encompass project or program records and documents, published and unpublished studies, and the statistical data available through public and private agencies.

Box 1. Examples of Mini Surveys

- A. To test the effectiveness of agricultural extension in a South Asian country, a mini survey was conducted. A sample of 60 farmers--30 contact farmers (who were receiving extension advice) and 30 noncontact farmers--was randomly selected in three villages. The findings did not show any significant differences in the adoption behavior of the two categories of farmers. More or less the same percentage of farmers in each group adopted the recommended variety of wheat. The obvious explanation was that the government had been promoting the new variety for the past 6 years and most farmers had become fully aware of its advantages and limitations and, therefore, did not require much extension advice.
- B. A private voluntary organization undertook a mini survey in an African country to discover whether on the basis of a sample of 60 respondents representing 6 villages, hospital users would be willing to pay the cost of using the facility. Respondents were asked to identify factors they would consider in choosing between a paying and a nonpaying hospital. The survey revealed that fees were not a significant determinant of the choice of facility. Rather it was income, quality of the facility, and distance from the facility that were the influencing factors.
- C. In a Southeast Asian country, an evaluation team was examining the impact of a large education project on the quality of primary schooling. The team had access to a variety of information that had been accumulated by the project over a period of 7 years; however, the team wanted to obtain the views on the subject of experts and decision-makers. Therefore, the team designed and conducted a survey of 35 educationists, experts, and administrators who were familiar with the local educational system. Each respondent was given a structured questionnaire, which took nearly 15 minutes to complete. The survey was completed within 10 days. The findings confirmed the documentary evidence that the project contributed to improving the quality of primary schooling.

An effort should also be made to review earlier surveys on similar and related topics. Such a review will provide valuable information on substantive issues and will generate a list of questions that can be used in planning the mini survey. Moreover, the review can alert the investigator to many conceptual, methodological, and logistical problems that he or she might face.

The investigator may at times even find data or information that obviate the need for a new survey. Many agencies and organizations gather data with scant attention to each other's efforts. In fact it is not uncommon to find several surveys being conducted by different agencies on the same topic. As a result there is often a surfeit rather than a shortage of information in many developing countries.

3. Preparing questions: Once the literature review is over, investigators should prepare survey questions, keeping in mind the study objectives. Framing questions is not as simple a task as is commonly assumed. The wording, length, and open/closed nature of a question can significantly affect the nature of the response and therefore require careful reflection. Attention also needs to be given to the recall period; investigators cannot assume that respondents will easily remember and report relevant details, however important those details may seem. Finally, investigators must be careful about how they word questions that may be sensitive in a given sociocultural milieu. Preparation of questions is discussed in Section 2.

4. Designing the questionnaire: The next logical step in this process is to arrange the questions in a carefully crafted short questionnaire that can be easily administered to respondents. All questions should follow a logical sequence to facilitate the interview. The investigators should pretest and revise the questionnaire in light of the pretest findings. Section 3 presents guidelines for designing questionnaires.

5. Sampling: An investigator has to make two choices concerning the sample for a survey. First he or she must choose between probability and informal sampling. As indicated earlier, the former should be preferred over the latter. There are, however, situations in which probability sampling is not feasible because of time and resource constraints. In such cases, the limitations of informal sampling should be recognized and precisely stated in the report. Second the investigator has to decide on the specific sampling technique offered by probability and informal sampling. Sampling is discussed in Section 4.

6. Mode of contact: The investigator has also to determine the way in which the questionnaire will be administered.

Although mail is the simplest and undoubtedly the least expensive method for contacting respondents, it is not a practical course in developing countries for two reasons. First, literacy rates are very low in those countries, especially among populations that are the subject of development interventions. Second, the response rates tend to be extremely low for mailed questionnaires even among the literate people, who have little or no incentive to complete and mail them. Telephone interviews are also out of the question primarily because an overwhelming majority of people do not have access to them. Under these conditions, the only viable method available to surveyors is the individual interview.

Practical guidelines for conducting interviews have to be developed at the outset of a mini survey. The guidelines should cover such topics as initial contact, probing methods, and recording and editing interviews. Section 5 provides general guidance on interviewing.

7. Analysis of data: The final stage in a mini survey is coding and analyzing the questionnaire data. The analysis in mini surveys is invariably limited to simple frequencies, percentages, rates, or at most simple correlations. Some of the issues relating to the coding and analysis of data are briefly discussed in Section 6.

2. PREPARING QUESTIONS

Once two priests were debating whether it was right to smoke during prayers. Both marshalled all kinds of arguments in support of their position without coming to an agreement. So they decided to consult their superiors and meet the next day. When they met, the prosmoking priest said: "My superior told me that it was alright to smoke." "How could it be?" replied the antismoking priest, "My superior was emphatic that it was wrong." "What did you ask him?" "Was it alright to smoke while praying?" came the reply. "That explains it. I had asked whether it was alright to pray while smoking."

A survey tale

As the tale quoted above suggests, preparing good questions is a difficult task requiring more than good grammar. It requires a robust common sense and the ability to empathize with the subject and the social and economic milieu in which the survey will be conducted. Above all it requires a familiarity with the literature on designing survey questions. This section presents general guidelines to help the investigator draft appropriate questions for mini surveys.

2.1 Wording Questions

Obviously, words that are simple, are widely understood, and have clear, precise meanings should be used in phrasing questions. Slang and colloquialisms should be scrupulously avoided because many persons may not understand them, causing both embarrassment and errors. For the same reason, technical terms should not be used unless the sample comprises primarily technical experts.

Often a word that best describes a relevant behavior or concept may not be understood by respondents. In such cases, the ideal course is to give an explanation of the word first, and then mention the word itself. For example, the question, "Should the technical assistance provided by the project be sustained over time?" may confuse many persons not familiar with the word "sustained" in the development context. This question can be better rephrased as "Should technical assistance provided by the project be continued after termination of external funding--that is, should technical assistance be sustained?"

Often words can have multiple meanings in a given context. For example, any, anyone, anybody, or anything may mean "every" "some," or only "one"; fair may mean "average," "pretty good," "not so good," "not bad," "favorable," "just," "open," "according to the rules," or "plain"; and "you" can be singular or plural. Investigators must be extremely careful with and limit the use of such words.

Should investigators use the standard language? Definitely, but because questions are read by interviewers during the interview, norms of spoken language rather than written language are appropriate. In fact, sometimes investigators can achieve better results by even violating the rules of written language--for example, commas, colons, and other punctuation marks should be avoided if they cause a break in the flow of ideas.

2.2 Length of Questions

Questions should be kept short and succinct. Lengthy questions can confuse respondents and cause them to miss the essential point of the question. In fact, the reliability of responses declines with the increase in the length of a question. This is particularly the case when questions focus on opinions, judgments, or attitudes of people.

However, when respondents are to recall old events, longer questions may be more helpful to them for two reasons. First, longer questions provide memory cues and aid recall. Second, because such questions take more time for the interviewer to read, respondents have more time to reflect, which improves the accuracy of responses.

2.3 Open-Ended and Closed Questions

Investigators have to pay careful attention to choosing open-ended or closed questions. Open-ended questions enable respondents to tell their story in their own words. The interviewer reads a question and tries to record the answer verbatim. However, closed questions list major response categories; respondents simply identify one or more categories that they consider to be the most appropriate. Examples of both types of questions are given in Box 2.

Advantages of open-ended questions are obvious. They encourage spontaneity and freedom; respondents are able to use their own language, concepts, and analytical categories.

Box 2. Examples of Open-Ended and Closed Questions

OPEN-ENDED QUESTION

A. During the last year, did you borrow money for your business? _____

If the answer is affirmative, then

B. Please tell me the source(s) from whom you borrowed money.

CLOSED QUESTION

A. During the last year, did you borrow money for your business? _____

If the answer is affirmative, then

B. Who among the following loaned money to you?

1. Relatives and friends _____
2. Money lender/trader _____
3. Cooperatives _____
4. Commercial banks _____
5. Others _____

Open-ended questions also enable respondents to provide additional details they consider relevant or to qualify their statements. Significant unanticipated findings may emerge. For example when responding to questions about credit a respondent may like to say that when interest rates are high he borrows from his rich uncle, otherwise he prefers commercial banks. Such a statement will be recorded in an open-ended question but not in a closed one. Moreover, by giving respondents an opportunity for free expression, open-ended questions make an interview interesting. Respondents are not easily bored. Finally, such questions generate insightful quotes and rich vignettes for the survey reports.

Open-ended questions have several limitations, however. The data generated by them are difficult to code; thus coding becomes more costly and time consuming. Moreover, it is more prone to errors because coders have to interpret answers and then classify them in appropriate categories. It is not uncommon to find two coders coding answers to one question differently, which raises the additional problem of bias. Open-ended questions also require more time because respondents take more time to think and verbalize their responses.

One persistent problem with open-ended questions is that interviewers are inclined to edit answers. They omit portions of answers that do not make sense, condense long worded answers to fit them in the space provided in the questionnaire, or even elaborate points that are not clear. Such editing results in inaccuracies and distortions. Interviewers need considerable discipline and skill to take verbatim notes and to systematically probe respondents when answers are not adequate.

In sharp contrast, closed questions are easy to ask and still easier to record; they therefore do not require highly skilled interviewers. Many respondents also find them less taxing because respondents do not have to recall as much detail or organize their thoughts as much to answer them. Moreover, response categories aid recall. For example, listing consumption items in expenditure surveys facilitates respondents' recall of items the household purchased within the specified time. Furthermore, with closed questions coding is simple and less time consuming. In fact, if a questionnaire is set properly, investigators can key punch the data directly from the questionnaire instead of transcribing the data on coding sheets. Therefore coding errors are fewer in closed than in open-ended questions because the task is simple and straightforward. Finally, because the same response categories are used, the data are comparable.

On the negative side, closed questions lack all the advantages of open-ended questions. They inhibit spontaneity and force respondents to choose from response categories even when they do not agree with those categories. Because closed questions do not provide ample time for reflection and recall, the responses may often be superficial. Thus standardization may be achieved at the cost of oversimplification.

While both types of questions are necessary, the majority of questions should conform to a closed-question format. If it appears that most of the questions designed for a mini survey cannot be answered using a closed format, it will be advisable to use other rapid data-collection methods such as key informant interviews, group meetings, or focus group discussions.

The use of open-ended questions should be restricted to three situations: First, a situation in which answers are sought to questions of why and how. For example, if, in a study on credit, investigators want to know the reasons for respondents' preferences for various sources of credit, a closed-question format may be unduly restrictive. Second, open-ended questions are useful in a situation in which response categories are large or unknown. For instance, if one of the purposes of a mini survey is to seek recommendations for improving the operation of a health project, an open-ended format will be more informative because it is difficult to determine in advance the specific recommendations that will be made by respondents. In fact, a closed-question format will result in some respondents saying "yes" even to recommendations listed in the questionnaire that they had not previously thought about. Finally, an open-ended question format is more appropriate in situations in which the respondents will likely have no information or opinion on an item. In such situations the closed-question format may produce answers where none existed, thus producing inaccurate data.

In sum, while the use of open-ended questions should be kept limited, both types of questions should be used in mini surveys.

2.4 Constructing Response Categories

Response categories for closed questions should be exhaustive and include the full range of possible answers; otherwise the resulting data may not be accurate. If, for example, a question is asked about the sources of prenatal care, all important sources--private prenatal clinic, public prenatal clinic, doctor, traditional midwife, and so on--that are available to the local populace should be listed. This is necessary because most of the respondents tend to choose from the given categories without bothering to suggest new ones, unless they know the subject and feel strongly about it.

In cases where there are more than five possible responses, it is usually preferable to list responses on cards to aid respondents' comprehension. However, this method is often not practical in developing settings because of widespread illiteracy. In such cases, the best solution is for the interviewer to slowly read each category and seek the appropriate response (e.g., yes, no, not applicable, or don't know). For example, in a survey of family planning practices, the question was asked: "Which of the following practices do you use for family planning purposes?" The questionnaire listed 11 response categories (pill, condom, IUD, injection, douche, rhythm, withdrawal, abstinence, female sterilization, male sterilization, other), and interviewers were

instructed to slowly read each category and ask for a yes or no response.

Of course the procedure described above cannot be applied to questions that involve the selection of only one out of several responses--for example questions that require rating or scaling of an event, behavior, or opinion. In an evaluation of a health program, respondents may be asked to rate, on a scale of one to seven, the quality of care provided. Since there is no solution to this problem, investigators are best advised to avoid constructing more than four or five response categories.

In listing response categories, a space should also be provided for responses not mentioned in the question, that is, an "anything else?" category.

2.5 Making Questions Specific

Questions should be as specific as possible for two reasons. First, respondents understand and respond better when questions are specific. Interviewees tend to interpret general questions differently, depending on their background and experience. Take for instance the question "What have been the impacts of structural adjustment programs?" Economists are likely to discuss impacts with reference to inflation, balance of payments, economic growth, and gross national income. Social scientists are more inclined to focus on the effects of the program on living conditions, economic inequalities, and the availability of social services. And political scientists may interpret the effects in terms of institution building, social unrest, and political instability. Thus, the respondents may be talking about different things while responding to the same question. However, when the question is made specific by restricting it to a specific sector (e.g., "What has been the impact of structural adjustment on the balance of payments situation, inflation, and economic growth?") respondents will focus on the same issues and the resulting data will be comparable.

Second, specificity aids recall. For example, in the example given above, if separate questions are asked about the social, economic, and political effects of structural adjustment, respondents are likely to mention items that they might have otherwise forgotten. Thus, listing specific items will refresh their memory.

2.6 Avoiding Double-Barreled Questions

Sometimes investigators combine two or more issues in one question. Consider the question: "Do you think that the government should provide credit to farmers at affordable rates and assist them in getting the improved variety of maize seed at subsidized prices?" Obviously, several questions are being explicitly stated or implied, which tends to confuse the respondent, who may agree with one part of the question, but not with the other. For example, the respondent who does not favor the government providing credit but wants seeds at subsidized rates will not know how to answer the question. Often respondents may answer the first part of the question while ignoring the remaining parts. To avoid such confusion, investigators should phrase separate questions for each issue.

2.7 Framing Questions To Aid Recall

Investigators must pay particular attention to questions dealing with recall. Lapses of memory are more common than many investigators realize. People tend to forget behaviors or events that seem trivial to them or that happened long ago. Recall of multiple events poses additional problems because respondents often mix the events.

Despite these problems, many surveys contain questions requiring vivid details of long-forgotten events and behaviors. The draft "Household Survey Questionnaire" prepared by the World Bank to study the social dimensions of structural adjustment in Africa provides an interesting example. One section of this questionnaire lists dozens of consumption items, ranging from cigarettes to razors, soaps to taxi fares, and sorghum to coconuts, and asks a common set of questions about each of the items. These are given in Box 3.

Unless all respondents kept elaborate records of their expenditures or were gifted with exceptional memory, a majority of them would not likely be able to answer accurately the questions listed in Box 3. How many of us can remember the amount we spent on cigarettes or razors during the past 12 months?

What should be done for questions requiring recall? Three general strategies can be followed. First is to narrow the reference period. Questions should focus as much on the recent past as possible. For example, an investigator is more likely to obtain an accurate answer to the question "How many times did you see an extension agent last month?" than to the question "How

many times did you see an extension agent last year?" However, when the question is about a topic of high importance to the respondent, such as the purchase of a major piece of agricultural equipment or construction of a new house, longer reference periods can be used. On the whole, the guiding principle is that days and weeks should be preferred over months and years.

Box 3. Questions About Each Consumption Item

The following questions were asked for each item listed in the draft questionnaire:

1. Was anything spent on during the last 12 months?
2. Do you buy once a week or more?
3. How many times was bought in the last 12 months?
4. How much was spent on in the last 12 months altogether?
5. How much was spent on in the last 3 months?
6. Were there months when you did not purchase this item?
7. In how many months did you not buy this item?
8. How much was spent on since my last visit?

The second strategy is to ask for average instead of specific time spans. In other words, the question should ask for the prevalent norm rather than the actual incidence of a behavior. For example, in expenditure surveys, investigators ask, "How much meat do you buy every week?" rather than "How much meat did you buy last year?" Some empirical evidence suggests that respondents are able to give relatively accurate information about the average or norm. In many instances, when questions address a short reference period, they can ask about both the norm and actual behavior.

Finally, in some cases, landmarks can be used to refresh the respondent's memory. In many impact evaluations, questions are asked about events that happened several years in the past. Such

questions tend to confuse the respondent who either includes events that happened earlier or focuses on more recent events, omitting events that should have been included. In such situations, reference to an important event (landmark) can be helpful. For example, instead of asking, "Did you see an extension worker during the past 6 months?" the investigator can ask, "Did you see the extension worker since Christmas?," making the reference period more concrete thereby aiding recall.

2.8 Avoiding Loaded Questions

Loaded questions refer to questions that are likely to push the respondent to answer in a certain direction. Such questions undermine the validity of survey data. Loading can take place in many ways. Often expressions such as "Don't you agree" or "Wouldn't you say" push respondents to give affirmative answers. Because of the outright suggestion, people may feel obliged to agree even if they have reservations about the statement. In societies where it is considered impolite to disagree, especially with outsiders, loaded questions can be particularly damaging. For example, the question "Wouldn't you say that the agricultural extension program has benefited farmers?" is likely to evoke a more positive response than the question "What was the impact of the agricultural extension program on farmers?"

Using emotionally charged words, cliches, and appeals to self-respect can also lead the respondent. Consider the wording of two loaded questions for a survey of smallholder farmers in a developing country: "The communist insurgents who are fighting the government believe that farmers should own the land they cultivate. Do you agree with them?" and "The government of this country believes that farmers should own the land they cultivate. Do you agree with its position?" Obviously, the second question will evoke more positive response than the first. The mere mention of the words "communists," "insurgents," or "overthrow" will hamper an objective frame of reference.

A more subtle form of loading occurs when the name of an authority is evoked. The question attributes a statement to an expert, a respected leader, or an established institution and asks the respondent to agree or disagree.

Box 4 presents some examples of loaded questions showing a microenterprise project in favorable light. These questions are meant for the entrepreneurs being served by the project.

Box 4. Examples of Loaded Questions for
an Impact Evaluation

1. Was it not an excellent idea to locate the microenterprise project in this district?
2. Experience shows that such projects contribute to the growth of small business enterprises. Won't you say that this project had similar effects in this district?
3. Do you disagree with the prevalent view that by providing subsidized credit, the project has helped many small entrepreneurs?
4. Several experts believe that in the absence of such projects, small businesses in the area would have faced many technical and financial problems. Wouldn't you agree with this assessment?
5. Are you in favor of continuing the project so that both credit and technical assistance are provided to the local businesses thereby improving the economy?

Loaded questions should be avoided with the possible exception of situations (discussed at some length below) in which data are being gathered for sensitive topics. Moreover, intelligent respondents easily see through such questions, which will further undermine the credibility of the survey.

2.9 Wording Sensitive Questions

In project and program settings, as in other situations, investigators have to ask sensitive questions about people, organizations, or the respondents themselves. However, some respondents do not like to answer such questions directly and either evade the questions or give inaccurate answers. For example, some farmers may not like to admit that they are not using the new variety of maize seed that has been vigorously promoted by the government, or educated mothers may not like to admit that they often visit traditional healers to treat their sick children. Also, people do not like to divulge information that they

believe will sully the image of an organization or an individual. Investigators have to be extremely careful in framing sensitive questions to obtain accurate answers. Some of the strategies employed for framing sensitive questions are discussed below.

First, questions can be worded to convey the impression that the concerned behavior or incident under question is not as unusual as it appears. For example, the question about visits to traditional healers can begin with a statement such as, "Experience has shown that even scientists, doctors, and highly educated people consult traditional healers in the wake of family illness. Was there any time during the past 6 months when you went to see one?" Thus by indicating that educated and respected people also visit traditional healers, the investigator minimizes the implicit threat of the question.

Second, the question assumes that a particular behavior or event happened or happens. For example, in the case just described, the investigator can ask, "How many times did you visit the traditional healer during the past 6 months?" In surveys on smoking habits, pollsters have used the question "How many cigarettes do you smoke each day?" Although this approach is often used in surveys, it poses two problems. First, it assumes that the respondent, or for that matter anyone else, engages in the activity, an assumption that may be resented by respondents. For example, nonsmokers may be irritated by the implicit assumption that they smoke. To deal with this problem, a provision is made for "none" category. The second, and more important problem, is that the question may lead to overreporting of the behavior or event; for example, even mothers who had not visited traditional healers may say they have.

Third, the name of an authority whom respondents are likely to trust can be used to obtain accurate answers to sensitive questions. For example, in a survey on the use of technical assistance provided to microenterprises, the following question can be asked: "Now the Ministry of Industry concedes that there have been serious problems with the current technical assistance program, particularly with its extension activity. Did you face any problem in obtaining technical assistance from the project?" By invoking the name of the Ministry of Industry, the investigator seeks to reassure the respondent that the problems in the technical assistance program are recognized at the highest levels and thus it is right to talk about them.

Fourth, investigators can word a question in a way that minimizes the sensitivity of what is being asked. That is why the phrase "Did you happen to . . ." is often used in public opinion polls. In a humorous vein Barton (1958, 67) suggests that investigators ask "Did you happen to kill your spouse?" rather

than "Did you kill your spouse?" Although it is doubtful that the reformulation of the question will evoke a more truthful answer in the case of killing a spouse, the essential point is that questions can be phrased to reduce their saliency in the eyes of the respondent. For this purpose not only the wording but also the tone in which a question is put to the respondent is important.

Finally, in some cases the question itself can advance reasons for respondents not doing the desired things. A good example: "Many things that are beyond one's control can come in the way of making regular payments to the lending institution. Was there any time during the past 12 months when you were not able to make your repayments to the agricultural credit bank?"

3. DESIGNING THE QUESTIONNAIRE

The same care and thought that are given to the wording of individual questions must be given to the construction of the total questionnaire. The tasks of both interviewers and respondents should be made as easy and enjoyable as possible. The respondent, after all, is doing you a favor; and a well-designed questionnaire makes the interviewer's job easier and improves the quality of data obtained.

Seymour Sudman and Norman M. Bradburn (1988)

After a set of appropriate questions is crafted, the next logical task is to present the questions in a short questionnaire that interviewers can present to respondents. This task is not as simple as it first appears. It involves arranging questions in a logical sequence, developing a suitable physical format for the questionnaire, and pretesting the questionnaire to identify and resolve problems. Each of these tasks requires careful reflection and thought.

3.1 Order of Questions

3.1.1 The First Question

The first question should be simple, but important and non-threatening, stimulating the respondent's interest in the survey. Boring or complex questions at the outset of interviews adversely affect the respondent's willingness to cooperate with the interviewer.

An investigator might consider starting with an open-ended question on an important issue. Open-ended questions allow people to express themselves freely and thus tend to be more interesting. However, such questions are helpful only if respondents are well-informed and articulate; otherwise the respondent may feel threatened by the question.

3.1.2 Demographic Questions

Many investigators routinely include demographic questions pertaining to age, employment, marital status, and even religion

in the beginning of the questionnaire. Such questions are generally superfluous in mini surveys because demographic variables are rarely used in analyzing the data. They not only take time from the interview, they also intimidate many respondents who do not like to divulge this kind of information. As a general rule, therefore, demographic questions should be avoided except when required for analytical purposes. And in such cases, the demographic questions should be put at the end of the questionnaire.

3.1.3 Funnel and Inverted Funnel Sequences

In the survey literature, "funnel sequence" refers to moving from generality to specificity; general questions are successively followed by increasingly specific questions in a questionnaire. Funnel sequence is especially useful when investigators want to ascertain from the opening questions the respondent's perspective or frame of reference.

Box 5 gives an example of funnel sequence. The first question presented is very general and permits respondents considerable freedom in discussing privatization in the agricultural sector. The second question is slightly restrictive in that it asks about the privatization of one type of parastatal, that is, those supplying agricultural inputs to farmers. The third question, which asks about progress toward privatization made by a specific parastatal, has still a narrower focus. The last question designed to obtain respondent's satisfaction with the pace of privatization for a parastatal is undoubtedly the most specific of all the questions.

If an investigator believes that most of the respondents have not thought about the subject and therefore may not be able to give thoughtful answers to general questions, the investigator can use the "inverted sequence" to ask questions; that is, he or she can reverse the funnel sequence, asking specific questions first followed by increasingly general questions. The advantage of the inverted funnel sequence is that it enables respondents to think through a topic before verbalizing their responses. Both funnel and inverted funnel sequences can be used in mini surveys.

3.1.4 Chronological Order

In obtaining historical information, investigators should pose questions that address events in a chronological or reverse chronological order. For example, questions about respondents' experience with technical assistance from a microenterprise

Box 5. Funnel Sequence on Privatization of
Agricultural Input Supply

The following questions are from a hypothetical questionnaire on privatization of parastatals dealing with agricultural input supply. The proposed respondents for the survey are government officials, project and program staff, and outside experts.

1. How is the privatization program going in the agricultural sector?

2. Do you think that the Government has made progress in transferring ownership of agricultural input supply parastatals to the private sector?

Yes _____ No _____ Don't Know _____

3. What progress has been made in privatizing the fertilizer corporation that markets fertilizers to farmers?

4. Are you satisfied with the progress of privatization in the fertilizer corporation?

project may begin with the most recent experience and work backward to earlier periods, or vice versa. Chronological order is helpful in aiding recall because it forces respondents to describe the sequence in the time period under consideration.

3.1.5 Changing Topics

Often a mini survey covers more than one topic. For example, a survey designed to examine the impact of an international training program is likely to include questions on such topics as the selection of trainees, overseas training experience, reentry, placement of trainees, and the contribution of graduate trainees to institution building. The simple rule is that all relevant questions on a topic should be grouped together. For example, the questionnaire will group all questions on selection process or the reentry of trainees at one place in logical order.

A short transitional sentence in the questionnaire can help investigators switch topics--for example, the transitional statement, "So far we talked about the working of the child survival program; we will now ask a questions about its impacts" can lead the respondent to the next topic in the questionnaire.

3.2 Length

The mini-survey questionnaire must be short and succinct containing 15 to 30 questions, taking no more than 30 minutes to complete.

3.3 Format

The general guidelines for the physical format of the questionnaire are relatively simple: A booklet format is preferable to loose sheets, which can be easily lost or misplaced; the questionnaire cover page should provide space for the name of the interviewer, the name and address of the interviewee, and the time, date, and place of the interview. Each question should be numbered on the left margin for easy identification and each page should be numbered; the questionnaire should also include instructions for the introductions and for probing the respondent and recording the respondent's answers and, when appropriate, the nonverbal behavior of respondents.

A questionnaire must provide plenty of space to record answers. Often questionnaires leave only two or three lines for recording responses to open-ended questions, which forces the interviewer to condense responses thereby undermining the validity of the data. Certainly, economizing on paper in such cases is not very productive.

Questionnaires can be formatted on a personal computer or an electric typewriter. Because of the relatively small sampling size in mini surveys, it is more economical to photocopy questionnaires than to print them, thus saving time and resources. If photocopying facilities are not available, the questionnaire can be mimeographed.

3.4 Translation

Often questionnaires prepared in English have to be translated into a local language. Errors in translation can distort the meaning of questions, resulting in inaccurate data. Therefore, translations should be done by persons who are fluent in both languages and who have a strong background in survey research.

The investigator must brief the translator in considerable detail about the overall objectives of the proposed survey and explain each question, the rationale for the question, and the type of information the question is supposed to generate. The time spent on such explanations is more than rewarded because the effort produces a better translation leading to fewer errors. Time permitting, the draft translation can be given to another local expert to translate back into English. The comparison of the two versions will help identify possible errors, which can then be corrected by the translator.

3.5 Pretesting

The investigator should carefully pretest the draft questionnaire by conducting between 5 and 10 interviews. The number of interviews will depend on the complexity of the questionnaire and the composition of the target population. The people interviewed for pretesting must have backgrounds and experience similar to the intended respondents. For example, if the questionnaire is designed for entrepreneurs receiving technical assistance from a project, only such entrepreneurs are to be included in pretesting. When a survey is likely to cover many categories of respondents, at least one respondent from each category should be included. Pretesting should focus on both individual questions and the questionnaire as a whole.

With regard to individual questions, the investigator should pay particular attention to the following.

First: Is the meaning of the question clear to respondents? Simply because a question has been carefully prepared does not mean that it is correctly interpreted by the respondent. Despite an investigator's best efforts, misinterpretations occur because of conceptual and linguistic barriers between the investigator and the respondent. Intelligibility can be easily determined by looking at the answers and by asking the respondents how they interpreted the given question.

Second: Do respondents have difficulty in answering the question? That the meaning of a question is clear does not ensure it can be easily answered. In their enthusiasm for obtaining maximum information, many investigators overlook the problem involved in accurately answering a question. For example, an expenditure survey asked heads of households to state how many yards of cloth their families had purchased during the past 12 months. Obviously, only a few could answer this question, and their replies were suspect. Investigators can sense trouble when respondents do not answer a question or take considerable time to answer it.

Third: Are the response categories appropriate? Often, the investigator may find that some response categories are superfluous or that additional ones are needed.

Fourth: Is there an acceptable level of variation in responses? Obviously, investigators should suspect the usefulness of a question when all respondents give the same answer. For example, if all respondents indicate that they have "benefited" from the intervention, the investigator should reconsider the question, and possibly, revise it. One option will be to use four response categories (very much, fairly, little, or no) to obtain more precise answers.

At the level of the questionnaire, an investigator should pay particular attention to the following.

First: Does the questionnaire read smoothly? The flow of the questionnaire is important because it will be read by the interviewer and not by the respondent in most cases. Second: How much time does it take to administer the questionnaire? Pretesting helps the investigator find out how much time a questionnaire requires. If it takes more than 30 minutes, the questionnaire must be shortened. Third: Does the questionnaire sustain the respondents' attention? If respondents looked bored or indifferent, the questionnaire should be suitably revised by adding or deleting questions, improving the language, and further training interviewers.

4. SELECTING RESPONDENTS

It is true that large-scale population surveys generally make use of probability sampling schemes, but this does not mean that large samples and full population coverage are necessary requirements for such sampling.

Christopher Scott (1987)

As in other types of surveys, a critical issue in conducting mini surveys is the selection of respondents. This section describes both probability and informal sampling procedures that can be profitably used by investigators in the field. Because mini surveys primarily use simple statistical analysis, the examples given here are based on rates and percentages. Moreover, the treatment of different sampling techniques is both brief and elementary. Investigators who do not have a background in sampling theory are strongly advised to consult experts before finalizing their sampling strategies.

The essential concept underlying sampling is that a large aggregate of people, organizations, households, or other units can be accurately examined by carefully scrutinizing a subset of the aggregate. The subset selected to study the aggregate is called a sample, and the aggregate a population or universe. Thus, for example, all small entrepreneurs who have received technical assistance from a microenterprise project constitute the population, whereas those actually selected for interviews for the survey are called the sample.

A few other concepts relevant to sampling must be briefly explained here. First, a sampling frame is the complete list of units from which a sample is drawn. In the case given above, for example, the list of entrepreneurs who have received technical assistance is the sampling frame. Second, estimator refers to the formula used to draw inferences from the sample for the whole population. Finally, sampling bias indicates the difference between the estimated value from a sample and the value computed from the entire population. For instance, if the average income of the entrepreneurs calculated from a sample is \$400 and the true average income computed, including all entrepreneurs in the study, is \$425, the difference of \$25 represents the sampling bias.

4.1 Probability Versus Informal Sampling

The two basic methods of sampling are probability and informal sampling. In probability sampling, each unit in the population has an equal chance of being selected for the sample. There are several variations in probability sampling, but all share a common trait; the selection of units for the sample is carried out by chance procedures and with known probabilities for selection. Informal sampling, on the other hand, uses convenience or common sense rather than mathematical reasoning. For example, an investigator selects 30 farmers who are available for interviews or visits 20 health centers that are regarded as "typical" by her or by other experts.

There are two widely held misconceptions about probability sampling that often lead some investigators to advocate using informal sampling.

First that large samples are required for probability sampling. This belief is totally unjustified and is rooted in a misunderstanding of the sampling theory. In fact, statistically valid generalizations can be made with a relatively small sample. A simple example will illustrate. Suppose key informant interviews indicate that 60 percent of young women have been practicing one of the contraceptive methods promoted by the government. The concerned A.I.D. official wants to find out if this is true. The investigator conducting this survey knows that the concerned official will be satisfied if the survey demonstrates with 90 percent confidence that the percentage of women using contraceptive methods is not below 50 percent. In this case, a sample of only 39 women users will provide the needed evidence.

Second, that the size of the sample depends on the size of the population; therefore larger samples are required for larger populations. This assumption is at best only partially correct because sampling error is determined by several factors, including sample fractions (the proportion of the sample to the population). In fact, beyond a certain point, an increase in the sample size only marginally contributes to a reduction in the sampling error. Therefore the sample sizes needed to study large or small populations are almost the same. For example, the sample size needed to estimate the birth rate in the small South African country of Lesotho is the same as that needed to estimate the birth rate in China.

As a general rule, investigators should use probability sampling for several obvious reasons. First, it minimizes, though not absolutely prevents, the risk of biased selection. Experience has shown that certain categories of population units

are less likely to be selected for the sample if they are not fully listed. For example, in rural household surveys, households that are inaccessible or remote, those whose members are seasonal migrants, those who belong to ethnic minorities or who comprise a single member, and even those who enjoy high social or political status (because enumerators are intimidated by them) tend to be underrepresented if informal sampling is used.

Second, with probability sampling the investigator can easily estimate sampling error, which indicates the probability of error in estimates for a given sample. For example, if an investigator uses probability sampling to estimate the percentage of women using contraceptives in a project, he or she can confidently say that there is only a 5 or 10 percent probability that the sampling error will exceed ± 10 percent of the estimate. Such an estimation is not possible in informal sampling.

These two advantages contribute to a third advantage, which should not be ignored. The data generated by probability sampling are more credible than those derived from informal sampling. Obviously, the findings of a survey can be trusted more if the respondents are selected randomly other than on the basis of convenience or judgment of the investigator.

However, in development fields, investigators often encounter situations in which probability sampling is not possible because of time, logistics, and resource constraints. In such cases, extreme care should be taken to make the sample as representative as possible, and the limitations of sampling method should be clearly stated in the report.

The widely used probability and informal sampling techniques that can be used for mini surveys are described in the next two sections. Finally, the topic of appropriate sample size is discussed in the last section.

4.2 Probability Sampling Methods

4.2.1 Simple Random Sampling

In a simple random sample, each unit of the population has an equal chance of being selected. This type of sample is easy to design and quite adequate when the population is relatively small.

A simple random sample can be drawn by lottery. Tags bearing names or identification numbers of all units of the

population are put in an urn and thoroughly mixed; then the predetermined number of tags are randomly drawn. Although seemingly simple, the lottery method is cumbersome and time consuming. Moreover, its precision rests on the assumption that the tags have been thoroughly mixed.

A better technique would be to number all the units, using random numbers to select the sample. Most statistical calculators have random-number generators. If the list runs to a three-digit number, then every unit is given a three-digit number (e.g., the number 5 is listed as 005), and three random digits are run off. The population unit with the number selected is included in the sample. If the random number repeats or exceeds the highest number assigned it is ignored. The process continues until the desired sample size is reached. The table of random numbers, generated by Kendall and Smith (1939), can be used if a calculator or personal computer with necessary software is not available.

One common problem for investigators in development interventions is that accurate lists of the populations to be studied are not always available; often the lists are outdated and contain many errors. Thus before constructing a sample, the investigator should carefully examine the available list and make every effort to check and improve its accuracy. The time and resources spent will be more than amply rewarded by the increased reliability of the finding.

The investigator drawing a random sample should not discard units that "do not look right." For example, an entrepreneur who is randomly selected should not be excluded because the investigator considers him or her to be atypical. If individual discretion is exercised, the simple random sample becomes a judgment sample, thereby defeating its purpose. However, if an investigator is absolutely convinced that the entire sample is unusual or peculiar--for example, the sample draws only from one geographical area or social class--the best course is to discard it and start afresh.

4.2.2 Systematic Sampling

Systematic sampling involves selecting units from a list on the basis of a fixed interval K so that after a random start every K th unit in the list is included in the sample. Suppose a sample of 50 is required from a population of 455 health workers, which means a sample fraction of $50/455$, or one in nine units. In systematic sampling, the investigator will take a random number between one and nine to select the first health worker and

include every ninth worker thereafter. Thus if the initial random number is six, the selected health workers will be the 6th, the 15th, the 24th, and so on.

The list from which systematic selection is made may be a written list, for example, a list of farmers receiving extension advice or young women getting contraceptives from a project, or a proxy list, for example, rows of houses on a street or individual medical records in a file.

Systematic sampling is undoubtedly more convenient than simple random sampling. Consider how much easier it is to take, for example, the 12th name from a document than to number them individually and then draw a sample. Furthermore, if the units of population are listed in an order showing a steady trend, a reduction in sampling error can be achieved by systematic sampling. For example, if farmers are listed according to the size of their holdings or entrepreneurs are listed in order of the magnitude of technical assistance they received from a project, an investigator could use a somewhat smaller sample, say 45 instead of 50 and achieve the same degree of reliability.¹

The most important danger in choosing a systematic sample is the possibility of hitting a cycle. For example, in some cities or towns, the corner houses are more expensive. In such cases, the sample interval in a systematic sample of houses selected from a map (e.g., every 10th house) may coincide with or oversample corner houses, which are likely to be inhabited by more affluent people than other houses on the block. As a result, an assessment of the community's nutritional status based on data from such households may overestimate the quantity and quality of food they consume. Therefore, lists should be carefully examined before the sampling method is chosen.

4.2.3 Stratified Sampling

In stratified sampling, the population is divided into groups, called strata, and then independent random samples are drawn from each stratum. Stratification is especially appropriate when the sample is designed to make estimates or comparisons for subgroups, as well as for the entire population. Depending on the distribution of groups within the population, a simple random sampling of the whole population may not include a sufficient number of cases from the relevant strata that need to be included.

¹For further explanation see Casley and Kumar (1988, 87).

Strata are created in such a way that there is less variance within a stratum than between strata. For example, project farmers can be classified on the basis of the size of their holdings in three strata--large, medium, and small holders--for a survey of adoption rates. Such a classification will obviously be based on the premise that the size of the landholding is related to adoption rates. Farmers with larger holdings are in a better position to take advantage of new technical packages because of their larger assets, greater interaction with extension workers, more intersystematic contacts, and probably higher educational attainment than farmers possessing lesser holdings. If the size of the landholding does not appear to be a valid criterion, an investigator can use some other criteria, such as education, gender, or proximity to the demonstration center, to classify the population. The essential point is that strata should be characterized by homogeneity.

Two types of stratification sampling are possible: proportional or disproportional, depending on the sampling fractions within strata. In proportional stratified sampling, the strata sample sizes are made proportional to the strata population sizes. For instance, if the proportion of large farmers is only 10 percent in the study population, the size of their strata will also be 10 percent of the sample. The problem with the proportional stratified sampling is that the numbers selected for a relatively small group would not be sufficient to permit a satisfactory statistical analysis. For example, from a sample of 80 farmers, only 8 or 9 are likely to be selected. Therefore, if an investigator is interested in making comparisons, more units will have to be sampled from the stratum that comprises a smaller proportion of the population. This is referred to as disproportional sampling because different sample fractions are employed in each stratum.

Since in disproportionate stratified sampling, there are variations in response rates among strata, an investigator should decide how separate results for individual stratum will be aggregated to arrive at an overall estimate. The simplest procedure is to compute the response rate of each stratum, multiply it by the number of units in the stratum (i.e., sampling frame), sum the total for all strata, and divide the sum by the population total.

Constructing stratified sampling requires a complete population list and additional information on the variables that are the basis for stratification.

4.2.4 Cluster and Two-Stage Sampling

In an overwhelming number of cases, populations are unlisted and widely dispersed. For example, investigators rarely find a list of farmers who received technical assistance from an agricultural extension project or women who purchased contraceptives at village depots. Even if such lists were available, the problem exists that the population units likely to be dispersed over the region under review would make simple random sampling both time consuming and costly. Cluster sampling often provides a practical solution in such cases.

Cluster sampling is based on the fact that most population units are clustered in one way or another. For example, farmers served by extension services live in villages, public health professionals work in organizations, and teachers teach in schools. And while it may be difficult to prepare a list of all farmers or health workers or teachers, lists can certainly be prepared of the concerned villages, public health organizations, or schools for sampling purposes. There are two types of cluster sampling: single stage and two or multiple stages.

In the single-stage cluster sampling, clusters are randomly selected; every population unit in the selected clusters is included in the sample. Take, for example, the case of a survey of agricultural extension workers. A project has 50 clusters (extension units) and each cluster has five extension workers. If, to save the cost of transportation and time, an investigator selects ten clusters through simple random sampling and interviews all the extension workers in them, he or she would be using single-stage cluster sampling.

In the second type of cluster sampling, survey sampling is done in two or more stages. For example, in the survey of agricultural extension workers, the investigator may first select say 10 clusters, then, through simple random sampling, he or she will select three out of five extension workers in each cluster to interview.

Cluster sampling has several advantages that make it highly suitable for mini surveys, especially when the sample is drawn from a relatively large geographic area. First, cluster sampling drastically reduces the costs, especially when a survey covers a whole province or country. For example, consider what it will cost to draw a simple random sample of 60 out of 600 medical professionals working in 100 medical centers spread across the country compared with the cost of a cluster sample of 70 to 80 respondents drawn from 10 medical centers. Second, cluster

sampling simplifies the interviewing process, and finally, it saves time.

The major drawback of cluster sampling is the likelihood of increased sampling error because units are selected in a group rather than independently. In many cases, the sample units selected in clusters will not show the same variation as equivalent numbers selected independently, contributing to greater sampling error. Therefore if cluster sampling is used, a slightly larger sample size is needed than in simple probability sample. Usually, a 15-20 percent increase is sufficient for mini surveys.

Any of the four methods of probability sampling described above can be used by an investigator. The choice should be dictated by the nature of the inquiry, availability of the list of the population units, constraints of time and resources, and, above all, the expertise of the investigator. Box 6 provides some examples of the probability sampling procedures.

4.3 Informal Sampling

4.3.1 Convenience Sampling

In convenience sampling, accessibility to sampling units is the prime consideration in the construction of the sample. Only units that can be easily reached by interviewers are included in the sample, which is why investigators from a wide spectrum of disciplines use this method. For example, medical researchers often depend on volunteer subjects, marketing firms tend to rely on people visiting malls or shops, and educators use their pupils for their surveys. In rural surveys, enumerators are often instructed to interview only those respondents who are available at the time of their visits to save time and transportation costs.

The problem with convenience sampling is that it is highly prone to sampling bias. Often certain strata, socioeconomic subgroups, or categories of population units are inadvertently excluded, underrepresented, or overrepresented. For example, if enumerators visit villages during daytime, they will miss farmers who are working in the fields, or when they interview farmers buying inputs at cooperative stores, they will omit those who get inputs from somewhere else--say, from the moneylender. Under such conditions, the findings can be wrong, and investigators might not even be aware of the shortcoming of their samples.

Box 6. Examples of Probability Sampling Procedures

1. An evaluation team was assessing the overall performance of a microenterprise development project, which provided both credit and technical assistance to women entrepreneurs. The team was interested in the assessments of women participants about the effectiveness of the project and its effects on their economic conditions. On the basis of available records, the team prepared a complete list of women entrepreneurs who had received assistance. Out of this list of 327 entrepreneurs, a sample of 30 was constructed by using a table of random numbers.
2. A monitoring unit in an area development project wanted to find out the differences in the credit requirements of men and women farmers in the project area. It had a list of 407 male and 187 female contact farmers (who were receiving extension advice from the project). The list was current and a preliminary check did not identify any inaccuracies. The monitoring unit decided to construct a sample of 60 farmers, half of whom were women. For this purpose, they prepared separate lists for female and male farmers and selected every 13th male and every 6th female farmer. The responses from the two categories of farmers were then compared.
3. The provincial office of a national family planning agency was interested in getting feedback from its staff about its activities. It ran 60 family planning clinics scattered throughout the province and employing 400 full- and part-time workers. The office did not have up-to-date lists of the local employees. Nor did it have adequate resources to visit a large number of clinics. Under these conditions, the investigators conducting the mini survey randomly selected six clinics and then randomly selected 10 workers from each to interview.

As a general rule, convenience sampling should be avoided in mini surveys. However, when the exigencies of a situation dictate its use, extreme care should be exercised. The following three steps can reduce sampling biases in many cases. The first

step is to identify categories of population units that are likely to be eliminated or overrepresented in the sample. Such determination can be based on in-depth interviews with local experts, careful review of past studies, and a continual review of the cases included in the sample. The second step, which logically follows the first, is to design the convenience sampling in a way that the probability of omission, underrepresentation or overrepresentation is reduced. Thus villages could be visited at a time when most of the farmers are inclined to be there and not on their farms. Finally, when possible, convenience sampling can be complemented with other forms of informal sampling to generate a more representative sample. Box 7 illustrates the steps taken by one investigator to reduce sampling bias.

Box 7. How One Investigator Attempted To
Reduce Sampling Bias

A survey was undertaken in remote villages in a South Asian country to study the prevalent attitudes toward family planning among women of childbearing age. Because of limited time and money, the investigator decided to use convenience sampling. Enumerators were expected to visit villages during the day to interview women. In the course of the survey, the investigator realized that the sample was likely to be highly biased because a substantial proportion of women were working outside their villages and were therefore excluded from the survey.

The investigator tried to persuade enumerators to conduct interviews in late evenings, but women enumerators refused as it required overnight stays in unfamiliar areas. She then suggested that at least half of the interviews be held in work places--farms, shops, and small factories. However, it proved to be an extremely difficult task and the interviewers had limited success. Finally, the investigator decided to select additional respondents to make the sample more representative.

4.3.2 Judgment Sampling

In judgment sampling, investigators construct samples on the basis of their own judgment or the advice of experts. This method is more widely used than is generally recognized. One vivid illustration of judgment sampling is surveying "swing communities" (i.e., communities regarded as "representative" of all voters because they have historically voted for the winning presidential candidates during U.S. presidential elections. On the basis of interviews with voters in these communities, political analysts try to predict the outcome of elections. Similarly, educators studying school systems will often select a few schools they consider to be representative of the whole school population. And, evaluators of extension programs generally select a few villages thought to be typical for their interviews with farmers.

The obvious shortcoming of judgment sampling is that the judgment of investigator or expert may itself be biased. What he or she views as typical may not be representative of the typical village, organization, or farmer in the local. This is essentially a serious problem when an investigator is not fully familiar with the study population and is obliged to depend on an outside expert who may consciously or unconsciously mislead him or her.

The relative accuracy of judgment sampling depends on three conditions. First, that the study population is small so the investigator can make informed judgment about the selection of sampling units. For instance, if the number of health centers runs into the hundreds, an expert cannot realistically be able to establish that the centers are representative of the entire population. Second, that the sample size is small. In fact, if the sample size is quite small, judgment sampling may yield better results than probability sampling. For example, if an investigator is interviewing officials in three districts in a province composed of 20 heterogeneous districts, he or she would obtain better results if the sample is chosen on the basis of an expert's choice instead of the vagaries of random chance, which might yield an odd sample. Third, that more than one expert is involved in constructing the sample. For example, an evaluation team conducting a survey of health centers can request several persons to suggest suitable sites and include in the sample only those centers for which there is a general agreement. Such a course will minimize errors arising from individual biases.

4.3.3 Snowball Sampling

Snowball sampling stems from the analogy of a snowball, which starts as a small ball but grows bigger and bigger as it rolls downhill. In snowball sampling, an investigator starts with a few population units, but ultimately ends up with the required sample size.

Snowball sampling is performed in several stages. During the first stage, a few persons who meet the necessary requirements for inclusion in the sample are identified and interviewed. These respondents are asked to suggest the names of additional persons who meet the sampling requirements and should be interviewed. The second stage involves interviewing the persons (or some of them) identified by the first respondents. The process is repeated until a suitable sample size is constructed.

Snowball sampling is most appropriate in situations where it is necessary to reach small, specialized populations that are not easily visible and can only be located with great difficulty. Box 8 gives an example of the use of snowball sampling in the evaluation of participant training programs.

One obvious limitation of the snowball sampling should be recognized: Respondents are likely to suggest the names of persons who share similar backgrounds, lifestyles, and social and professional orientations. Thus, for example, if the initial farmers contacted by an interviewer are large landholders, it is quite likely that they will suggest the names of other large holders. A good illustration of this problem is provided in Box 9.

4.3.4 Quota Sampling

In quota sampling, the population is divided into various strata and a predetermined number of people or quota is selected for each stratum. The difference between stratified probability sampling and quota sampling is that with the latter the investigator selects respondents within each stratum on the basis of convenience or the judgment of interviewers rather than on probability sampling. Once quotas are established, interviewers are free to include anyone who meets the requirement.

As is the case with stratified sampling, quotas can be established on the basis of age, sex, income, education, location, combinations of these, or any other criterion perceived to be relevant to the purpose of the inquiry. Thus, for example, an

Box 8. Selection of Participants Through Snowball Sampling

A five-nation study was undertaken in Asia to assess the impact of the U.S.-funded participant training programs on the professional advancement of Asian social scientists. The study required a comprehensive survey of trainees who had received their training in the United States during a period of two or three decades.

As the study progressed, the investigator found that records for past trainees were missing in most of the countries. Even when some records were available, the listed addresses were out of date. Presumably, the trained social scientists had moved from the institutions they had initially joined after completing their training. The investigator therefore used snowball sampling. It was impossible to construct a reliable sample frame under these conditions.

He started with a few participants whose addresses were available. During the interviews, each respondent was asked to give the names and addresses of other U.S.-funded trainees. The suggested participants were then interviewed and asked to suggest the names of their fellow trainees. By using this method, the investigator was able to locate a majority of the participant trainees in each country. The investigator succeeded because fellow participants had kept in touch with one another and were therefore able to give him names and current addresses of colleagues.

investigator conducting an informal survey of farming practices in a province may assign quotas for different ecological zones if he or she thinks that farming practices vary significantly among the zones. Similarly, if the purpose is to assess the impact of policy reform interventions on the standard of living of a populace, it may make more sense to establish quotas for various economic strata, because the effects of these interventions are likely to differ for different stratum. For example, suppose an evaluation is planned for a microenterprise project providing technical assistance to local entrepreneurs. The project covers two districts and focuses on handicraft, garment, and food-related enterprises. The evaluation team may consider at least three different criteria--gender, geographical area, and the nature of the business--for which quotas may be assigned. Thus a simple matrix may be developed, as shown in Table 1.

Box 9. An Example of Problematic Snowball Sampling

An investigator used snowball sampling to study the growth of a scientific community in a Southeast Asian country. To his utter dismay, he discovered at the end of his fieldwork that his sample was biased and unrepresentative of the population of scientists he intended to study. The investigator had started with a few scientists who were members of a clique and who referred him to fellow members. As a result, the investigator did not interview members of two other prominent factions of the community. Clearly, the mistake was also caused by the investigator's unfamiliarity with the local situation, and because he did not make sufficient efforts to cast a wider net at the beginning of his fieldwork.

Table 1. Sample for the Microenterprise Survey

Enterprise	District A		District B	
Garment	(M...)	(F...)	(M...)	(F...)
Handicrafts	(M...)	(F...)	(M...)	(F...)
Food	(M...)	(F...)	(M...)	(F...)
Total				

Evaluators will try to establish quotas to make a representative sample. For example, if half of the entrepreneurs under study are involved in garments, efforts will be made to select half of the respondents from this category. On the other hand, if only 25 percent of the businesses are located in district B, only a quarter of respondents will be selected from district B. Although the match may not be perfect, every effort should be made to select quotas in appropriate proportion to their size in the population.

How many variables should be used for establishing quotas? Obviously not many. The problem with using more than three criteria is that specifications become more complex, and interviewers find it more difficult to locate respondents with the requisite characteristics. It is not always easy to locate women entrepreneurs participating in a microenterprise project who are engaged in food-related businesses and reside in a district. But it becomes still more difficult when the interviewers are also asked to select three out of six women who are under age 35.

4.4 Sample Size for Mini Surveys

The sample size in mini surveys is primarily determined by time and cost considerations. If time is limited and respondents are scattered over a wide geographic area not easily accessible, an investigator will have no alternative but to use a relatively small size, say, 20 or 30. In addition to time and costs, investigators should consider the following three factors.

First is the homogeneity of the population. If the population of interest is highly heterogeneous, a larger sample is needed than if it is homogeneous. For example, one of the reasons that in many medical studies, a sample of less than 15 is considered adequate is because human beings are biologically similar. A mini survey designed to estimate the adoption rate of a particular input by farmers with significant variations in the size of their land holdings should have a larger sample than a survey of farmers cultivating more or less similar size holdings.

Second is the number of variables that are to be examined simultaneously. For example, if in the case given above, an investigator also wants to find out the differences in the adoption rates between male and female farmers or literate and illiterate farmers, a larger sample will be necessary.

Third is the degree of precision required. While the relationship between the degree of precision required and the sample size is highly complex and cannot be explained here, it can be safely suggested that time and resources permitting, the investigator should in most cases strive for the upper limit (70 respondents) rather than the lower limit (25 respondents).

5. ASKING QUESTIONS

The farmer's wife was totally aghast by the sudden arrival of an impeccably dressed, elegant woman to her remote hut when she was desperately trying to start fire on a wooden stove. Her hut was filthy, smelling and full of smoke caused by wet wood. She could read sheer horror on the face of the uninvited visitor who was undoubtedly taken aback at what she saw. When the visitor, in an unfamiliar accent, said, "I have to ask a few questions of you," she was speechless for a moment and then asked her son to call his father. Obviously, she had thought that the visitor, an interviewer for a family planning survey, had come to investigate her husband's drinking habits.

Notes of a field supervisor

As the excerpt from the field supervisor's notes suggests, preparing thoughtful questions and compiling them in a questionnaire is not sufficient to generate reliable, accurate data. Interviewers must also present themselves appropriately, establish a rapport with respondents, ask questions in a manner that evokes accurate responses, and, above all, accurately report answers. This section provides general guidance on interviewing techniques.

5.1 Initial Contact

In any interview, the first 30 seconds of contact between the interviewer and the respondent are critical. During this brief period the interviewer and respondent form their first impressions of each other, which condition the ensuing interview. Therefore interviewers should be careful about their overall appearance; they should always dress simply and inconspicuously, respecting the local norms of dress and behavior, even when these customs are inconvenient.

Interviews should be conducted at a time that is most convenient for respondents. For example, if farmers are in the fields during the day, interviews should be held during the evenings when farmers are most likely to be at home. Men and women employed in industrial and service sectors can be better contacted during weekends when they are likely to be relatively free. Interviewers should make appointments with government officials and professionals to avoid scheduling conflicts.

At the start of the interview, interviewers should briefly explain their background, the objectives of the survey, and possible uses of the information provided by respondents. They should also assure the respondents of the confidentiality of information and indicate the time required for completing the interview, which is fortunately no more than 30 minutes for mini surveys. Such remarks should be brief and to the point, and interviewers should honestly answer any question that respondents have at this stage. Box 10 mentions a few questions frequently asked during the initial contact.

5.2 Using the Questionnaire

An interview should be made as pleasant as possible. Interviewers should never give the impression that they are administering a quiz or conducting a cross-examination. Nor should they say anything that would imply approval or disapproval of the respondent's answers. Interviewers should invariably show respect to respondents and a genuine appreciation for their views and opinions.

The interviewer should read each question slowly. The ideal reading pace is two words per second. Studies have shown that a slow, deliberate pace enables respondents to understand the question and formulate a careful reply. When interviewers hurry through questions, respondents tend to follow their example contributing to superficial, even inaccurate responses.

Often some respondents are not sure of their answers and need to be reassured. This is especially the case with people of low socioeconomic status in developing societies, who have not been exposed to survey research. When an interviewer feels that a respondent is diffident and doubtful about the reply, he or she should make a neutral conversational remark, such as, "We are just trying to get people's ideas on this" or "There are neither right nor wrong answers to this question." Such remarks often put people at ease.

Questions should be asked in the order in which they are presented in the questionnaire because, as indicated in Section 2, questions are presented in logical sequence. When interviewers change this sequence, they may inadvertently bias the results.

If a question does not apply to the respondent, it should be crossed out and the reasons for its nonapplicability noted in the questionnaire. This enables the coder or investigator to know

Box 10. Typical Questions and Remarks
Exchanged During the Initial Stage of an Interview

Interviewers frequently encounter a set of common questions at the beginning of an interview. Some typical examples are given below.

1. Question: "Why did you pick me?"

Answer: "The purpose of this survey is to find out the views of people on Since it is not possible for us to talk with every one, we have selected a few persons like you to help us out." (Interviewer can then add a sentence or two about the sampling process.)

2. Remark: "I am too busy. Why don't you leave me out?"

Response: "It will not take more than 15 to 20 minutes of your time. However, if you are very busy now, please tell me the time when I can visit you again. It is extremely important that I have the benefit of your ideas and experience on this important subject."

3. Remark: "I really don't know anything about this."

Response: "We are interested in your opinions and experiences and not in what information you may or may not have. I am sure that you will find questions interesting and will be able to answer them easily. We have asked the same questions from many people who did not have any problem in answering them. I must mention that in a study of this type, there are no right or wrong answers to questions."

4. Question: "Who's behind this?"

"This study is sponsored by the ----- project/program. Its purpose is to get ideas and information that will be helpful in improving its activities."

that the question is not relevant and has not inadvertently been overlooked by the interviewer.

5.3 Probing Techniques

One of the most challenging tasks for interviewers is to obtain complete responses that meet the needs of the survey. Often interviewers encounter situations in which respondents give incomplete or irrelevant answers for a variety of reasons. Some respondents have difficulty understanding the questions; others have problems expressing themselves; and still others may be reluctant to reveal their thoughts or what they know. Interviewers should deal with these situations tactfully and in a way that will not bias responses. Survey experts usually recommend five probing techniques to stimulate fuller, clearer responses.

First, repeat the question. Repeating a question can help when the respondent does not understand the question, misinterprets it, misses the emphasis, seems unable to make up his or her mind, or strays from the subject. Moreover, repetition gives the respondent more time to reflect on the question.

Second, pause for the respondent's answer and convey, through a nod or an expectant look, that you are expecting a fuller response. A pause gives the respondent time to collect his or her thoughts. In many cases, however, the respondent may have nothing more to add and a pause may cause unnecessary embarrassment; the interviewer therefore should be sensitive to these nuances and not probe the respondent any further.

Third, repeat the respondent's reply. Hearing the ideas repeated often stimulates the respondent to expand on the response.

Fourth, use neutral comments or questions, such as "Anything else?" "Any other reason?" "Could you tell me more about your thinking on this subject?" "Why do you feel that way?" "How did this occur?" "When did it happen?" Such questions also indicate that the interviewer is carefully listening to what the respondent is saying.

Finally, gently probe the respondent to clarify inconsistent, contradictory, or ambiguous answers. However, it is important that the interviewer takes the blame for any ambiguity and does not convey the impression that the respondent is unclear or inarticulate. Comments like the following help to clarify a point: "I'm sorry, but I am not sure if I got the point. Would you please repeat it?" or, "I'd like to be sure I understood you correctly. You said that you did not borrow money for your business? Is this correct?"

The success of probing depends on the interviewer's ability to immediately recognize how a specific answer has failed to meet the objective of the question and to frame an appropriate probe.

5.4 Recording the Interview

Interviewers should record responses during the interview and not wait until the interview has ended; interviewers can lose relevant information if they try to remember what the respondent said at the end of an interview.

Responses should be noted down in the language of respondents, keeping the same phrases, grammatical usages, and peculiarities of speech. When interviewers try to summarize or paraphrase respondents' answers, they often unknowingly create a communication gap between the respondent and the investigator. Box 11 provides two examples of how summarization or paraphrasing colors the true intent of the respondent.

Interviewers should not get so involved in note taking that they forget the respondent. One simple technique for holding the respondent's interest while taking notes is to repeat the response. Repeating the response shows the respondent that the interviewer is listening carefully and, as stated earlier, provides the respondent more time to reflect on the answer he or she has given. Often many respondents modify their reply or provide more specifics when they hear interviewers repeat their answers. By repeating the response, the interviewer is also able to confirm the response.

To speed note taking the interviewer can use abbreviations and codes. Interviewers can construct a set of common abbreviations for commonly used terms. For example, they can use "R" for respondent, "DK" for don't know, "P" for project, "E" for evaluation, and the like.

To save time, articles and prepositions can be left out and only key words noted during the interview. Later, when the interviewer is editing the interview, he or she can insert these and the punctuation. For example, a standard probe, "What is on your mind" may be written as "What mind" at the note-taking stage.

In addition to recording the verbal responses of key informants, interviewers should note their nonverbal behaviors, when appropriate. Often facial expressions reveal more than what an informant says. For example, if a woman respondent seems skeptical or uncomfortable responding to questions about the

Box 11. Effects of Summarizing and Paraphrasing Answers

Consider the difference between the verbatim reply and the summary versions reported:

Verbatim: "Yes, indeed! I am positive about it."

Summarized version: "Yes."

In this particular case, the summarized version lacks the intensity of the respondent's reply. A more serious example of distortion is the following:

Verbatim: "I am very upset by the way the technical assistance program has worked in this district. The program has primarily benefited the well-to-do entrepreneurs rather than the poor, struggling self-employed men and women who were supposed to be targeted by the planners of this program. Thus it has totally failed to accomplish its stated objectives."

Summarized version: "Program did not accomplish its objectives."

The summarized version does not capture the respondent's feelings on the subject. Moreover, it even distorts the reply. The essential point that the respondent forcefully made was that the project failed to benefit the "poor, struggling self-employed men and women."

effectiveness of credit delivery to women farmers, the interviewer should make a note of her reaction. Such notes enable the investigator to examine and review the respondent's replies carefully during coding and analysis of the data.

5.5 Editing the Interview

The best time to edit an interview is immediately after the interview or, if that is not possible, the same day. The purpose of editing is to ensure that all answers are correctly recorded.

In particular, interviewers should ensure that entries are legible, "inappropriate" questions marked, probes listed, replies to open-ended questions put in parentheses, and articles and prepositions added.

6. ANALYZING AND PRESENTING THE SURVEY DATA

The manager of a large area development project in a southern African country came to meet with the permanent secretary of agriculture. He brought what he thought was the major achievement of his statistical division, a few neatly typed regression tables. These tables presented regression coefficients for variables predicting the adoption rates for high-yielding varieties of maize in the region. The permanent secretary asked a few general questions and dismissed the manager, promising to look at the tables carefully. As soon as his visitor left, he deposited the tables in his waste basket, breathing a sigh of relief.

An incident reported to the author

The moral of the story described above is obvious. The analysis of survey data in project and program settings should be kept simple. The purpose of the mini survey report should be to communicate and not to impress the reader with the investigator's efforts and statistical skills. Therefore, simple statistical tools that are more likely to be understood by people without statistical expertise are invariably preferred over those that are complex and sophisticated. However, if investigators have to use complex statistical tools, they should make every effort to present the findings in nontechnical language.

This section provides general guidance about the statistical analysis and presentation of survey data. It explains the nature and requirements of coding, describes a few statistical tools that can be used by the investigator (only the most common and simple statistical methods that can be easily computed with a calculator are discussed), and provides a few tips for the presentation of the survey data.

6.1 Coding Data

Coding is indispensable for quantitative analysis. It involves transforming gathered data into categories and translating these categories into numbers. The purpose of coding is to simplify individual responses. For example, suppose respondents' occupations are as follows: farmer, barber, farm worker, blacksmith, general merchant, moneylender, baker, butcher, government servant, primary teacher, mason, and midwife. Because the sample size is small, using so many occupational categories serves little purpose in a mini survey. Therefore, an investigator should

develop a few categories, depending on the purpose of the inquiry. If the investigator is primarily interested in the farming population, he or she may use only three categories: farmers, farm workers, and others. On the other hand, if the purpose of the survey is to understand the problems of microenterprises, a different scheme may have to be followed.

The two methods for coding are known as inductive and deductive. In the inductive method, the data are recorded in as much detail as possible--for example, all the occupations in the community are listed. However, the deductive method involves the use of a predetermined classification scheme that is strictly followed by the coder. For example, if the investigator has decided to use three categories--farmers, farm workers, and others--the coder will classify all the responses on this basis.

Both these methods have their advantages and limitations. One shortcoming of the deductive method is that it does not allow for new ideas and insights. Once the data are coded, the investigator has no freedom to manipulate the data beyond the specified categories without recoding questionnaires. On the other hand, the deductive method saves time and effort. Consider the difference in coding and analysis time if only 4 instead of 15 occupational categories are used. Moreover, the deductive method imposes some order on data by eliminating superfluous or irrelevant details. In contrast, the main strength of the inductive method is that it permits flexibility. The investigator can easily develop new categories even after the data have been coded. For example, if he or she finds that a new category of civil servants is useful, it can be added without any additional effort. The shortcoming of the inductive method is that the investigator might be bogged down with unnecessary details. Moreover, more time is needed for coding the data.

Ideally, both these methods should be used in a mini survey. When investigators know what they are looking for and have a reasonable idea of response categories, the deductive approach is undoubtedly preferable. However, if appropriate categories are not apparent, the inductive method should be used. The inductive method is particularly appropriate for coding responses to open-ended questions.

There are four simple rules for developing a good coding scheme. First, the categories must be mutually exclusive so that each case is classified only once. Second, they should be inclusive in the sense that all the responses are categorized. One example often cited in the survey literature is of the marital classification into three categories: "married," "single," and "divorced." A problem arises if there is a widower in the sample. There should always be a category "other" or

"miscellaneous" to make the system inclusive. Third, investigators should also consider that necessary details (necessary with reference to the objectives of the survey) are not lost. Fewer categories make data neat and manageable, but they also limit the details available to us. Finally, the coding scheme should be related to the purpose and scope of the mini survey.

In some cases, the original information can be presented in actual number. The investigator may code the actual age, size of the household, or the area cultivated. However, in most cases, he or she gives an arbitrary number to a category. Box 12 provides an example of a coding scheme for a few questions.

As indicated in Section 2, a questionnaire is likely to have a few open-ended questions. Coding such questions requires considerable time and effort. The best course is both to numerically code them and record them verbatim for analysis. For example, in a mini survey designed to evaluate a macroenterprise project, the investigator asks for recommendations for improving the quality of technical assistance. He or she can use two steps to code

Box 12. An Illustration of the Coding Scheme

- a. Age
(Code actual age in years)
- b. Gender
Male = 1
Female = 2
- c. Number of Children
(give actual number)
- d. Use of Contraceptives
Yes = 1
No = 2
I don't Know = 8
No answer = 9
- e. Agreement with the statement about the need for family planning
Agree Strongly = 1
Agree = 2
Not Sure = 3
Disagree = 4
Strongly Disagree = 5

the responses. First, all the responses to the question can be recorded in a separate sheet to provide a comprehensive picture. Second, after carefully reviewing the responses, a set of categories for recommendations can be developed and individual responses coded accordingly. Often the perusal of verbatim records at the report preparation stage provides new insights and explanations to the investigator.

6.2 Statistical Analysis

6.2.1 Frequencies and Percentages

The first step in any analysis of the survey data is to construct a frequency distribution by listing all the response categories and counting the number of observations in each of them. The accepted procedure is to list the categories in the left-hand column and the number of observations in the right-hand column.

Nominal categories (categories that cannot be ordered on the basis of their own attributes, like gender, rural-urban residence, religion, and so on), are usually arranged according to the total observations in each category--from higher to lower size. Such categories as "not applicable," "applicable," or "no response" are kept at the end, whatever their size. The intervals into which the data are grouped should be of equal size (e.g., \$001-100, \$101-200, \$201-300, and so on). But in some cases, such as distribution of income or of rent paid, unequal intervals at the high or low margins may be preferred because there might be sudden jumps in the observations. Cutting points for the categories should be determined with reference to the number of observations and the objective of the survey. As a general rule of thumb, the number of intervals should normally not exceed six or seven in a mini survey because of the small sample size.

It is always useful to list percentages in addition to frequencies. The percentage for each category is obtained by dividing the number of observations (f) in that category by the total (N) and multiplying the resulting quantity by 100. Cumulative distribution is obtained by summing the percentages in each category and all the preceding categories.

Table 2 provides an example of a simple frequency distribution giving percentages. A simple glance will show that 60 percent of the respondents cultivate 4 or less hectares of land or that only 5 percent cultivate more than 8 hectares.

Table 2. Size of Farms Cultivated by Respondents

Farm Size (hectares)	No. of Households (f)	Percentage	Cumulative Percentage
0-2	35	35	35
2.1 - 4	25	25	60
4.1 - 6	20	20	80
6.1 - 8	15	15	95
8.1 - 10	5	5	100
N	100	100	

Box 13 illustrates how the percentages can be used to draw relevant conclusions.

Often a change of scale is extremely helpful. Table 3 gives data about the number of credits approved by 10 branches of a credit institution. Column 2 shows the data with 100 removed from each value, and column 3 shows each value as a difference from the mean. Finally, column 4 gives the percentage with reference to the target of 120 credits.

Table 3. Variations of Transformed Data

Branch	No. of Credits Approved	Credit Issued in Excess of 100	Difference From Mean (126-5)	Percentage of the Target (120)
	(1)	(2)	(3)	(4)
1	103	3	- 23.5	86
2	125	25	- 1.5	104
3	117	17	- 9.5	98
4	133	33	+ 6.5	111
5	129	29	+ 2.5	108
6	148	48	+ 21.5	123
7	118	18	- 8.5	98
8	132	32	+ 5.5	110
9	146	46	+ 19.5	122
10	114	14	- 12.5	95

Box 13. Use of Percentages

"Percentages ... can either be very helpful in presenting data clearly or very misleading, depending on the competence (and honesty) of the presenter. Consider the data in this table giving the number of farmers purchasing fertilizer in each of three districts in each of the 3 years.

District	Total farmers	Year 1	Year 2	Year 3
A	10,000	3,612	4,170	4,670
B	14,000	765	1,241	2,073
C	38,000	21,036	20,217	19,416

We can present these data in terms of the percentage of farms in a district purchasing fertilizer in each year (noting that each percentage is independent of the others--they do not sum to 100 in any direction):

Percentage of Farmers Purchasing Fertilizer

District	Year 1	Year 2	Year 3
A	36	42	47
B	5	9	15
C	55	53	51

Or we can present the percentage distribution by district of the farmers buying fertilizer in a year:

Percentage of Fertilizer Purchasers

District	Year 1	Year 2	Year 3
A	14	16	18
B	3	5	8
C	83	79	74
Total	100	100	100

Or we can show the index of the numbers purchasing fertilizer in each district, taking the number 1 as the base (100):

Index of Fertilizer Purchasers

District	Year 1	Year 2	Year 3
A	100	115	129
B	100	162	271
C	100	96	92

The first table shows that fertilizer is most popular (or more available) in districts A and C; it also shows the time trend, but it disguises the dominance of district C in terms of numbers of purchasers. The second table shows the dominance of district C but disguises the trends in the number of purchasers over time. The third table highlights the time shifts more clearly, bringing out the rapid growth in district B and decline in district C while losing the relative importance in absolute numbers or relative incidence of fertilizer purchasers in the various districts. Each such presentation has a role to play, but each presents a partial picture, which in isolation can even mislead."

Often the investigator need not go beyond counting frequencies and percentages. Most of the study questions for a mini survey can be easily answered with these simple techniques.

6.2.2 The Mode, Median, and Arithmetic Mean

In addition to simple frequencies and percentages, measures of the central tendency may be desirable. Such measures include the mode, the median, and arithmetic mean.

The mode is the category or observation that appears most frequently in the distribution. For identifying the mode, the investigator singles out the category containing the largest number of responses. Most distributions have only one modal category in which the observations are concentrated; but, there are distributions in which two categories are nearly or equally prominent. Such distributions are called bimodal. Caution is necessary in using the mode for ordinal data because it may be affected by the way in which values are grouped in categories.

The median is a measure that divides the distribution into two equal parts. The median is computed by locating the middle observation. For an odd number of cases, the middle item is calculated by adding 1 to N (the number of cases) and dividing the sum by 2. For example, Table 4 gives the membership figures for 9 and 10 farmers' clubs in the project area arranged by their size. Here the median is $(9+1)/2 = 5$, and since the number of members in this club is 38, it is the median membership for the farmers' clubs. Suppose there was also another club with a membership of 57, making an even number of cases, or 10 cases, the median will be a value between 38 and 42, that is, 40. The main advantage of the median is that it is not affected by extreme values. For example, even if the first club has only five members, it will remain the same.

Table 4. Membership of Farmers' Clubs

Club No.	Membership
1	25
2	30
3	35
4	35
5	38
6	42
7	45
8	50
9	56
10	57

The most widely used measure of the central tendency is arithmetic average or mean (symbolized \bar{x}). It is calculated by adding all the values of the distribution and dividing the sum by the total number of cases. For instance, in computing the mean for the membership of farmers' clubs as shown in Table 4

$$\begin{aligned}\bar{x} &= \frac{25+30+35+35+38+42+45+50+56}{9} \\ &= 39.6\end{aligned}$$

Thus the average membership of a farmers' club is around 40.

6.2.3 Range, Variance, and Standard Deviation

The measures of central tendency described above give an indication about the most representative value of the distribution. But they do not indicate the extent of the dispersion of the value. For example, two regions may have the same mean for the size of agricultural holdings, but in one region the land might be equitably distributed; all the farmers have more or less the same amount of land, whereas in the other region a small minority might own most of the land. Several measures of dispersion are designed to give information about the dispersion or variation in the values of a distribution. They are range, variance, and standard deviation.

The most simple, although not as useful, measure is the range. It measures the difference between the highest and the lowest values of the distribution and is computed by subtracting the latter from the former. Because the range depends on the two extreme scores, it is a very unstable measure.

Unlike the range, variance (s^2) takes into consideration the values of all the items in a distribution. It is computed by summing up the squared deviation from the mean and then by dividing the sum by the total number.

$$\text{Variance} = \frac{\sum (x_i - \bar{x})^2}{N}$$

Table 5 shows a distribution of seven cases and the calculation of its mean and variance.

Table 5. Mean and Variance for a Distribution

Serial No (1)	x (2)	$x-\bar{x}$ (3)	$(x-\bar{x})^2$ (4)	x^2 (5)
1	3	-4	16	9
2	3	-4	16	9
3	5	-2	4	25
4	6	-1	1	36
5	7	0	0	49
6	9	2	4	81
7	16	9	81	256
Total	49	0	122	465

Note: Mean (\bar{x}) = $\frac{49}{7} = 7$

Variance (s^2) = $\frac{(x-\bar{x})^2}{N}$

= $\frac{122}{7} = 17.4$

The steps involved in calculating the variance (s^2) are as follows. First, the arithmetic average is computed. In this case, the mean is $49/7 = 7$. The second step is to calculate the difference between the value in each category and the mean. This difference is shown in the third column of Table 5. The third step involves squaring the deviation from the mean as shown in column 4. Finally, the squares from the deviations are summed up and divided by the number of cases, which comes to 17.4 in this particular case.

The formula given above involves several computations that can be avoided, by employing a simpler formula that yields a close approximation to the true variance. The most common approximation formula is;

$$s^2 = \frac{\sum fx^2}{N} - \left(\frac{\sum fx}{N} \right)^2$$

The formula given above can be illustrated with reference to Table 5. Column 5 in the table gives x^2 (squares of the raw scores of observations). When the simpler formula is applied to the data in Table 5, we find:

$$\begin{aligned} \text{Variance} &= \frac{465}{7} - \left(\frac{49}{7}\right)^2 \\ &= 17.4 \end{aligned}$$

The variance expresses the average dispersion not in the original units of measurements but in squared units. This problem is solved by taking the square root of variance, which is called the standard deviation. Thus in the example $s = 4.2$.

6.2.4 "Pearson's" Coefficient of Correlation

The maximum value of r or the "coefficient of correlation" or "association" is 1, which can be both positive and negative. If two variables are positively associated, it means that available data suggest that an increase in the first variable is associated with an increase in the second. The negative correlation suggests that the association is inverse: When the one increases, the other decreases. Thus a value of r of $-.75$ is the same as of $+.75$ as far as the strength of the association is concerned. In everyday usage, an r of $.8$ and above is considered a high coefficient, an r around $.5$ is considered moderate, and an r of $.3$ and below is considered a low coefficient.

There are several formulas for computing r . Probably, the most simple is the following:

$$r = \frac{N\sum xy - (\sum x)(\sum y)}{\sqrt{[N\sum x^2 - (\sum x)^2][N\sum y^2 - (\sum y)^2]}}$$

A simple example will illustrate the above formula. Suppose there are scores for 10 farmers who were given two tests, one for general knowledge and the other for their knowledge of the recent agricultural innovations that a project is trying to promote. We want to know if the two are associated. Table 6 lists these scores.

Table 6. r Between General Knowledge Scores and Agricultural Innovation Scores

No.	General Knowledge Scores (x)	Agricultural Innovation Scores (y)	(x) ²	(y) ²	xy
1	20	12	400	144	240
2	18	16	324	256	288
3	16	10	256	100	160
4	15	14	225	196	210
5	14	12	196	144	168
6	12	10	144	100	120
7	12	9	144	81	108
8	10	8	100	64	80
9	8	7	64	49	56
10	5	2	25	4	10
10	130	100	1,878	1,138	1,440

When the figures from Table 6 are entered in the correlation formula, the following results:

$$r = \frac{10(1440) - [(130)(100)]}{\sqrt{[(10 \times 1878) - (130)^2] [(10 \times 1138) - (100)^2]}} = \frac{1400}{1610} = .87$$

The r of .87 is indeed a very strong correlation. It suggests that the general knowledge and the knowledge of agricultural innovations are closely related. The farmers who are more knowledgeable about general matters are also well informed about agricultural innovations.

To compute r for grouped data, the midpoint for the intervals is taken and multiplied with the frequencies.

6.2.5 Other Statistical Measures

In addition to the measures discussed above, several others, particularly chi square, simple and multivariate regression, "analysis of variance" and t-tests, can also be used to analyze data. However, lack of space does not permit discussion of them here. Readers are advised to consult any standard textbook on statistics for this purpose.

6.3 Presenting Data

6.3.1 Constructing Tables

Survey data are invariably presented in the form of tables. There are some simple rules for tabular presentation that help the reader to readily absorb the composition of the data set and appreciate without further analysis the most obvious patterns and relationships. These are given below.

- Limit the number of rows and columns. Numerous columns and rows can confuse the reader. What is appropriate for the professional journal is not necessarily permissible in the context of project and program analysis.
- Use clear, self-explanatory column and row headings. Liberal use of differential spacing is necessary to highlight comparisons.
- Use clear and unambiguous class intervals in frequency distributions.
- Transform the data into percentages and indexes as appropriate. Use averages, standard deviations, and so on to summarize the array of data.
- Present the data in logical order. One commonly used order is from most frequent to least frequent, although showing data the other way around may also be appropriate.
- Provide a title that summarizes the purpose and content of the table.
- If there is any additional information that can help in interpreting data, for example, statistical significance, mention it as a note at the end of the table.

- Summarize in the text the highlights revealed by the table and the conclusions that can be drawn from it.

6.3.2 Presenting Graphics

Graphic representation of the data can be very useful in communicating findings. It helps in dramatizing a point without deceiving the reader. If an investigator has access to a computer, relevant graphs can be prepared very quickly.

One of the simplest graphs is a pie diagram that shows the proportions of the whole in different categories. Suppose an investigator has conducted a survey of 90 participants in a microenterprise project to discover, as one of the survey objectives, the economic status of the project participants. The government is especially interested in determining whether the people below poverty line (e.g., earning less than \$1,000 a year) are adequately represented in the project. The tabulation of the survey responses generates the following frequencies:

Range of Household Income	Number of Respondents
0 - \$1,000	5
\$1,001 - \$2,000	10
\$2,001 - \$3,000	25
\$3,001 - \$4,000	40
\$4,001 - Above	10

The data given above can be presented in a pie diagram as shown in Figure 1.

The key to an accurate pie diagram is to draw it to scale. The slices should be restricted to 5 or 6, otherwise the diagram becomes too cluttered.

Bar graphs are also useful for presenting survey data. Unlike the pie diagram, bar graphs can provide an overview of many kinds of information at one glance. Suppose an investigator wants to compare the utilization of technical assistance by men and women entrepreneurs for a project. The sample of 80 was equally divided between men and women; the resulting data are given in Table 7.

Figure 1

Socio-Economic Status of the Participants

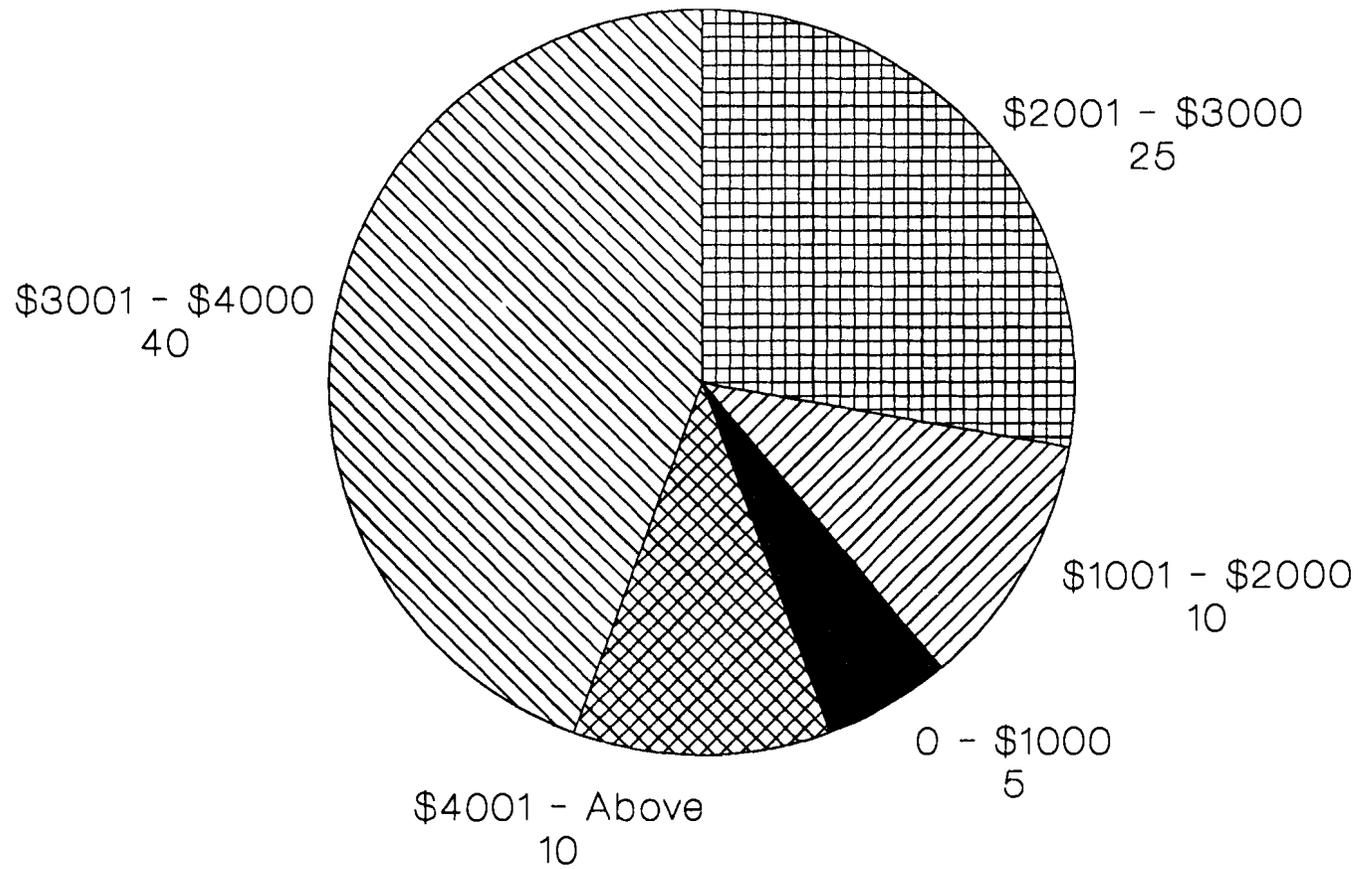


Table 7. Utilization of Technical Assistance by Gender

	Setting Up Business	Obtaining Loans	Purchasing Equipments and Inputs	Marketing
Women (40)	30	20	15	35
Men (40)	20	30	28	20
Total 80	50	50	43	55

The data can be presented in the form of a bar graph (see Figure 2).

Finally, line graphs can also be used to present data. Line graphs are most suitable for presenting time series data, although they can also be used to show frequency distributions and relationship between two variables. All graphs should have a title, scale (when appropriate), and key to define lines, values, and symbols.

6.3.3 Preparing the Report

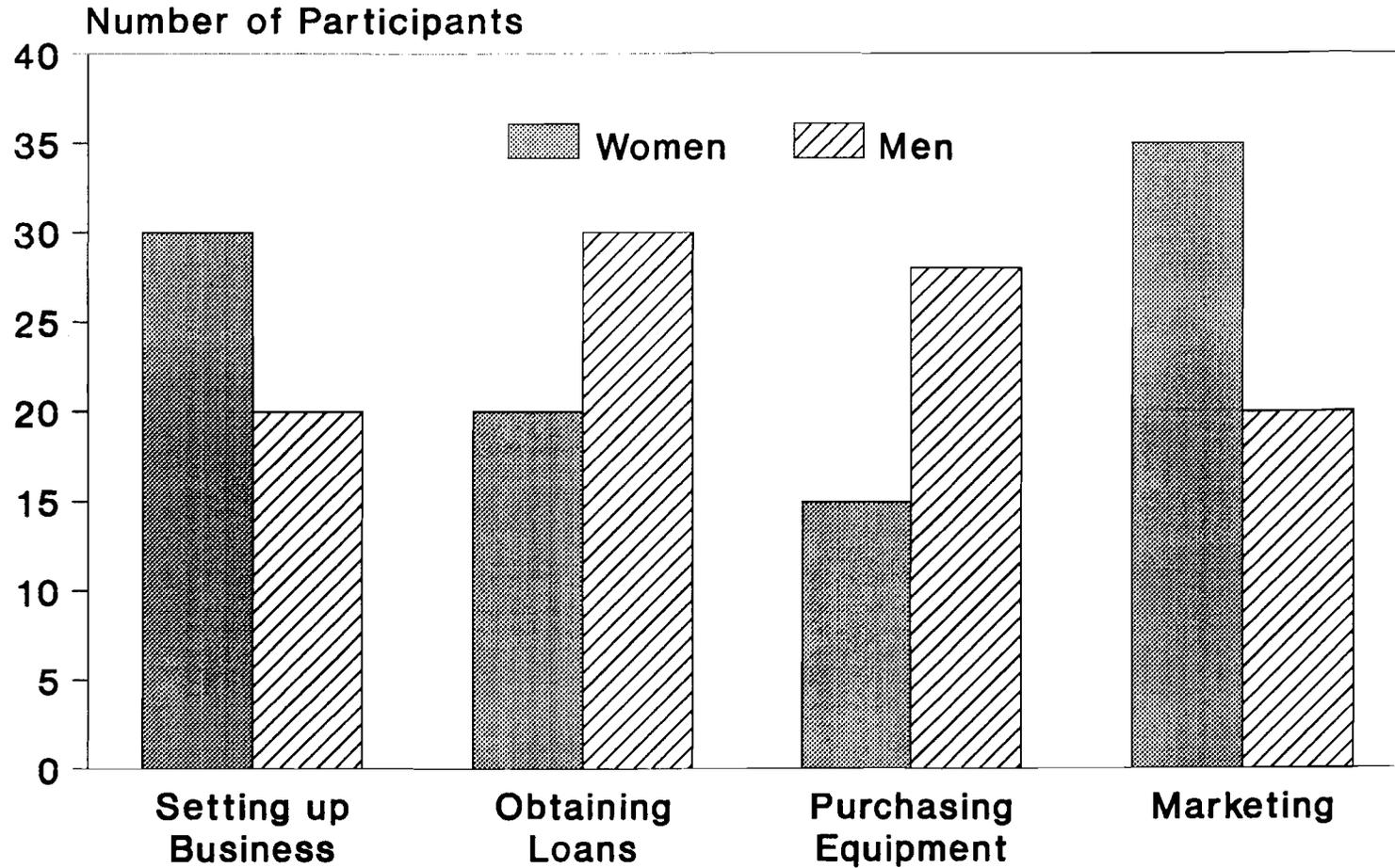
Finally a word about the preparation of the survey report. A typical academic report often follows the following format:

- a. The purpose and scope of the survey
- b. Conceptual framework
- c. Research methodology
- d. Summary of data
- e. Findings and their implications
- f. Summary
- g. Appendixes

This format, while appropriate for an academic milieu, is not the most useful in the contexts of project and program assessments. While all the items mentioned should be covered in a report, their sequence should be different. The investigator should begin with the summary and then move directly to the findings and their implications. Such an arrangement is preferable because the decision-makers who commission a mini survey are more interested in the findings and recommendations than in the

Figure 2

Utilization of Technical Assistance by Gender



methodology, sampling strategies, or the underlying conceptual framework. The customary presentation of the findings at the end of the report contributes little to stimulating the interest of decision-makers. The sections on conceptual framework, research methodology, and data may follow the chapter on the findings and implications. Even a better option is to put them in appendixes. A table of contents is always helpful to indicate the total coverage of the report and to guide the reader with an interest in the technical aspects.

A major limitation of many survey reports is that they often contain an elaborate discussion of the underlying concepts, sampling strategies, and procedures the investigator used to design the questionnaire. What is still worse, they include numerous tables that are unnecessary. The preferred approach is to cover each section, but as succinctly as possible. Investigators should emphasize the meaning and implications of data analysis for the projects and programs, not the methodological aspects of their research.

REFERENCES

- Barton, A.J. 1958. "Asking the Embarrassing Question." Public Opinion Quarterly. (22):67-68.
- Casely, Dennis J., and Krishna Kumar. 1988. The Collection, Analysis, and Use of Monitoring and Evaluation Data. Baltimore, Maryland and London: Johns Hopkins University Press for the World Bank.
- Chambers, Robert. 1981. "Rapid Rural Appraisal: Rationale and Repertoire." Public Administration and Development. 1(2):95-106.
- Finsterbusch, Kurt. 1976a. "Demonstrating the Value of Mini Surveys in Social Research." Sociological Methods and Research. 5(1).
- Finsterbusch, Kurt. 1976b. "Mini Surveys: An Underemployed Research Tool." Social Science Research. 5(1).
- Scott, Christopher. 1987. "Sampling for Monitoring and Evaluation." World Bank Working Paper. Washington, D.C.: World Bank.
- Sudman, Seymour, and Norman M. Bradburn. 1988. Asking Questions. San Francisco and London: Jossey-Bass Publishers.

List of other reports from the Program Design and Evaluation Methodology series

Manager's Guide to Data Collection, November 1979, Program Design and Evaluation Methodology Report No. 1 (PN-AAH-434).

Selection and Justification Procedures for Rural Roads Improvement Projects, January 1984, Program Design and Evaluation Methodology Report No. 2 (PN-AAL-032).

Selecting Data Collection Methods and Preparing Contractor Scopes of Work, August 1985, Program Design and Evaluation Methodology Report No. 3 (PN-AAL-057).

Evaluation Guidelines for Nonproject Assistance: Commodity Import Programs (CIPs) and CIP-Like Activities, August 1985, Program Design and Evaluation Methodology Report No. 4 (PN-AAL-058).

An Approach to Evaluating the Impact of A.I.D. Projects, March 1986, Program Design and Evaluation Methodology Report No. 5 (PN-AAL-059).

A Manual for Evaluating Small-Scale Enterprise Development Projects, November 1985, Program Design and Evaluation Methodology Report No. 6 (PN-AAL-065).

A.I.D. Evaluation Handbook, April 1987, Program Design and Evaluation Methodology Report No. 7 (PN-AAL-086).

Conducting Group Interviews in Developing Countries, April 1987, Program Design and Evaluation Methodology Report No. 8 (PN-AAL-088).

Guidelines for Data Collection, Monitoring, and Evaluation Plans for A.I.D.-Assisted Projects, April 1987, Program Design and Evaluation Methodology Report No. 9 (PN-AAL-089).

Rapid, Low-Cost Data Collection Methods for A.I.D., December 1987, Program Design and Evaluation Methodology Report No. 10 (PN-AAL-100).

Methodologies for Assessing the Impact of Agricultural and Rural Development Projects: A Dialogue, June 1989, Program Design and Evaluation Methodology Report No. 11 (PN-AAX-218).

Indicators for Measuring Changes in Income, Food Availability and Consumption, and the Natural Resource Base, September 1989, Program Design and Evaluation Methodology Report No. 12 (PN-AAX-223).

Conducting Key Informant Interviews in Developing Countries, December 1989, Program Design and Evaluation Methodology Report No. 13 (PM-AAX-226).

Preliminary Indicators for Measuring Changes in the Natural Resource Base, February 1990, A.I.D. Program Design and Evaluation Methodology Report No. 14 (PN-AAX-233)

