

Unit 21

Survey Design

Contents

- 21.1 Introduction
- 21.2 Preliminary Considerations
- 21.3 Stages / Phases in Survey Research
- 21.4 Formulation of Research Question
- 21.5 Survey Research Designs
- 21.6 Sampling Design
- 21.7 Conclusion

Learning Objectives



It is expected that after reading Unit 12 you would be able to

- ❖ Work out (a) the purpose of enquiry, (b) the unit on which the survey is to focus and (c) the availability of resources
- ❖ Divide the survey research into stages or phases
- ❖ Formulate the kinds of questions to be asked during the survey
- ❖ Decide the type of research design to employ
- ❖ Choose a sampling plan for survey.

21.1 Introduction

Design elements are crucial in understanding the complexity of undertaking a survey. The survey method involves several aspects, which are crucial for its execution in a systematic and scientific manner. Unit 21 provides a description of the basic dimensions of survey in social research. The topics discussed in this unit are: preliminary considerations for undertaking survey research, different stages/phases in conducting a survey, formulation of research questions, types of research design, and details of sampling design.

21.2 Preliminary Considerations

There are three preliminary considerations before any researcher decides to undertake survey as a method of social investigation. The considerations are

- i) Purpose of enquiry,
- ii) Population on which survey is to focus, and
- iii) Resource availability.

Each of these considerations is essential for survey research.

The first prerequisite for undertaking survey research is clear and well-defined objectives, purposes or research questions. The researcher must be clear that the survey method is the best method to study those objectives (purposes or research questions). If there is a clear idea

about the objectives of a research study, it leads to the choice of the appropriate survey as a method of investigation.

The second prerequisite is specifying the population of study. This influences the decisions about sampling and resources. The accessibility of the population is an important element in developing survey design. It judges the duration of study, the different layers of people to be included and the size of the research staff as well as the sample.

The third important prerequisite in designing a survey is to take stock of the resources availability for study. Resources may be in the form of financial costs, manpower requirement and the total time within which the survey has to be completed. Surveys are usually labour intensive, the single largest expenditure being the fieldwork costs which include the interviewing time, travel time, and transport costs, etc. There are additional costs involved in survey research, namely, training and supervision of interviewers, questionnaire construction, pilot testing the questionnaire, printing, posting, coding, computer programming, etc. Thus, the researcher must estimate the availability of resources before starting the survey research.

21.3 Stages / Phases in Survey Research

Survey research calls for a systematic and comprehensive collection of information about the attitudes, beliefs and behaviour of people. In a survey, the researcher follows a deductive approach. He begins with a theoretical or applied research problem and ends with empirical measurement and data analysis. Once a researcher decides that survey is the best and appropriate method to undertake his research, then elaborate steps are taken in order to make the application of the method more scientific and systematic. Further, there is no unanimity among the researchers on the actual number of steps or their ordering. However, here we deal with a few major steps in three phases/stages that cover the preparation, planning and execution of survey research (See Figure 21.1).

Stages / Phases in Survey Research		
Design and Planning Phase	Data Collection Phase	Data Analysis and Reporting
1. Research Question Formulation	1. Locating the Respondents	1. Editing
2. Decide Survey Research Design	2. Accessing the Respondents and Field Settings	2. Coding
3. Draw the Sample	3. Supervision and Monitoring of Main Survey	3. Tabulation and Analyses
4. Data Gathering Techniques		4. Write-up
5. Questionnaire Construction		
6. Pilot Survey		
7. Finalising the Survey Instrument		
8. Training of the Investigators		

Figure 21.1 Major Steps in Three Stages of Survey Research

The first phase covers design and planning for a survey. This is a very elaborate and long phase in which the details of conducting a survey are worked out. It is crucial as it sets the ground for the actual survey.

The important steps in this phase are formulating a research question, deciding the survey research design, working out the sampling procedure, deciding the data gathering techniques, construction of a questionnaire, conducting a pilot survey, modifying the survey instrument, training field investigators/interviewers.

The second phase involves the execution of the main survey for data collection. It includes locating and accessing the respondents and field settings, conducting the survey, monitoring and supervision of the field personnel.

In the third and final phase, the researcher edits the data collected, prepares codebook, and decides the tabulation for data analysis. Finally, the researcher prepares the Write-up/Report of the study.

The listing of steps in a sequential form may give the impression that once a step is covered there is no turning back. However, to the contrary, it is often necessary to alter decisions made at an earlier point in the light of developments during the later stages of research. For instance, the problems encountered during the execution stage may call for modification either of the sampling design or of the type of instrument for collection or of the questionnaire. Further, it is also a mistake to assume that these steps can be followed one at a time without any regard for subsequent steps. In reality, it is often important to keep in mind the other steps while working out the details of one step. For instance, it is necessary to keep in mind the intended data analysis procedures while constructing a questionnaire.

21.4 Formulation of the Research Question

The first step in a survey design is the formulation of a research question. Research question here refers to the major goal problem or objective of the study. In fact, the formulation of a question should precede the selection of survey method for data collection. Though survey has a wide range of applications in the social sciences, it may at times turn out that the other methods of data collection (such as in-depth interviewing or participant observation, etc.) may be more appropriate or suitable to the problem chosen. Thus, it is necessary to think of the appropriateness of the survey method for investigating a particular research problem.

Further, it may be kept in mind that there is a need to formulate a research question at the beginning of research. But, this emphasis does not mean that the initial research question is the final one (see Box 21.1). Often research questions arise unexpectedly in the middle of a conversation with informants. These need not be ignored. Their addition or modification of questions decided earlier is always a useful exercise.

Box 21. Refine Your Research Questions

According to David de Vaus (2002), it is important to know what we are looking for, it is a mistake to let the initial focus blind us to other unanticipated questions which may perhaps be more interesting, important or manageable than the initial one. Research questions can be refined and new issues emerge in the course of survey research.

How does one arrive at a research question? Often, the researcher's curiosity about a particular social phenomenon guides him to choose a particular research question. However, the researcher can formulate his research question in more than one way. First, the starting point could be a theory or a model. The choice of theoretical framework or model will have implications for the kind of questions that are to be formulated and the kind of data that are to be collected. Second, significant social policy implications may guide researchers in formulating their research problem. In other words, the problem chosen should be relevant to the issues of social improvement, social change, or social action. Third, the goal of producing social criticism may lead some researchers to formulate research questions suitable for that objective. Some research questions may meet several or some of these criteria, but it is rare that a research question will meet all of them.

Further, it is important that the survey's general purpose be translated into a specific central problem/aim/objective. It is not enough to say, for example, 'I am interested in getting some answers about inequality'. One should be clear about the kinds of questions one wants to answer. What answers to what questions? Does he want to know the extent of inequality, its distribution, its causes, its effects, or what? What sort of inequality one is interested in - social, political, economic, etc? Over what period? Where?

The process is of narrowing down from a general area of interest to a level that is specific enough for concrete empirical research to be carried out. After specifying the primary objective of the survey, the next step involves identification and itemising of the subsidiary topics that relate to the central purpose. For example, in a study of student attitudes to the quality of undergraduate education in a college, the subsidiary issues may include the quality of teachers, the infrastructural facilities, the admission procedures, the type of courses offered, the content of the courses, the interaction between teachers and students, the facilities for remedial study, the facilities and scope for extra-curricular activities, etc. Once the subsidiary items have been identified, the researcher looks for the specific information required relating to each of those items. For example, with respect to the type of courses required, detailed information is needed on the duration of courses, the status of the courses (credited or non-credited), orientation of the courses (theoretical or practical), etc.

A thorough review of the existing literature (books and journals) on the

topic is helpful while narrowing the specific topic, identifying the subsidiary topics, itemising the subsidiary topics and spelling out the kind of information required on these topics for data collection. A review of literature is helpful because it reveals to us the scope and coverage of research that has already been conducted and helps us in identifying important gaps so that an attempt could be made to fill those gaps through survey research.

There are other ways too that can help a researcher in formulating his research question. One is the exploratory study, which is a tentative and relatively unstructured investigation of a few people who are similar in many respects to those we intend to study as part of the main investigation. For instance, if we intend to work on the quality of undergraduate education in a college, one may have to visit an undergraduate college, interview a few students, and identify the specific research questions for the study. The most important goal of such an exploratory study is that it stimulates the thinking of the researcher and helps him identify specific issues that may be worthy of inclusion in the study. Sometimes, it may help us in understanding the way the concepts are understood differently by different people. For instance, in our earlier example, different students may view differently the notion of 'quality' in undergraduate education. For some, it may mean good teachers/teaching. For others, it may be good infrastructure or a range of relevant courses of study, etc.

Another aspect of the research question formulation is that of consultation with those who are knowledgeable about the topic and who have done related research. An encounter with these experts can help researchers in anticipating the pitfalls of the proposed research study.

Reflection and Action 21.1

Suppose you have to carry out a survey of waste disposal methods, and you need to formulate research questions. Work out the major goal problem or objective of your study. Write down answers to the following questions on a sheet of paper.

Questions

- ❖ In the light of the objective of your study, do you find the survey method a useful and appropriate way of data collection? Provide specific reasons of accepting or rejecting survey method as a method of data collection.
- ❖ How have you arrived at each of the research questions? State clearly the choice of your theoretical model.
- ❖ Have you looked at the social policy with regard to waste disposal? Elaborate the implications of the policy in terms of it guiding your research problem.
- ❖ Is your study going to generate social criticism of an issue of importance for the general public? If yes, you need to work out how your study is going to lead other researchers to formulate further research in the same area.
- ❖ Have you transformed the general objective of your study into a specific central problem? It is important that you do this exercise at this stage and state the narrowing down of the problem from general to specific level for concrete research to be carried out.

- ❖ Have you consulted the experts in the area of your research to find out what to expect in your field of inquiry and what are the likely pitfalls to anticipate and to be careful about?

Write your answers to above questions with the help of the relevant text in Section 21.4.

Types of research questions

There is no simple way to define research questions. However, according to David de Vaus, we can delineate different kinds of questions and provide guidelines to help focus research -- descriptive research questions or explanatory research questions or both.

a) Descriptive research questions

It is difficult to focus a descriptive question but five questions can help.

- ❖ What is the time-frame of our interest?
- ❖ What is the geographic location of our interest?
- ❖ patterns for sub-groups?
- ❖ What aspect of the topic are we interested in?
- ❖ How abstract is our interest?

b) Explanatory research question

The first step in formulating explanatory research question is to decide whether we are looking for causes or consequences. We must list possible causes and consequences and then collect relevant data. There are a number of ways of coming up with such a list, namely, previous research, the facts, our own hunches, an interaction with the informants.

Four questions might help formulating explanatory research question.

- ❖ What am I trying to explain?
- ❖ What are the possible causes?
- ❖ Which causes will I explore?
- ❖ What are the possible mechanisms?

Thus, there are varieties of issues that need to be kept in mind while formulating the research question that proposes to undertake survey as a method of data collection. It actually triggers the mind of the researcher as how to operationalise the research.

21.5 Survey Research Designs

After the formulation of the research question, we must decide the type of design to employ. Research designs for surveys are mainly of two types, namely, experimental and descriptive. Experimental designs use environmental arrangements and rely on two or more groups of a participants or observations. When a randomly constituted group of a hundred children is compared three times, the survey design is called experimental. The descriptive design produces information on groups and phenomena that already exist. No new groups are created. Descriptive designs are also called 'observational' or 'explanatory' designs.

Experimental designs

According to Arlene Fink (1995), experimental designs are characterised by comparing two or more groups, at least one of which is experimental. The other is a control (or comparison) group. An experimental groups is given a new or untested, innovative programme, intervention, or treatment. The control is given an alternative (e.g. the traditional programme or no programme at all). A group is any collective unit. Sometimes, the unit is made up of individuals with a common experience, such as men who have participated in a war or those who have had undergone treatment at a hospital or those who have attended a particular school. At other times, the unit is naturally occurring: a classroom, business, hospital, or a prison.

There are different types of controls. Firstly, there are concurrent controls in which participants are not randomly assigned to groups. 'Concurrent' means that each group is assembled at the same time. For example, when 10 of 20 schools are randomly assigned to an experimental group while, at the same time, 10 are assigned to a control group, you have a randomised trial or true experiment. Secondly, there are concurrent controls in which participants are not randomly assigned to groups. These are called non-randomised controlled trials, quasi-experiments, or non-equivalent controls. Thirdly, in 'self-controls', a group is surveyed at two different times. These require pre-measure, post-measures, and are called 'longitudinal' or 'before-after' designs. Fourthly, 'Historical Controls' which make use of data collected for participants in other surveys. Finally, 'combinations' of all consist of concurrent controls with or without pre or post-measures. Of all the experimental designs, longitudinal designs are used quite often in social research.

Longitudinal designs

In a longitudinal[®] design, data are systematically collected over a period of months or years in such a way that it is possible to observe trends in attitudes or behaviours over the specified period. Consider the following example. If a researcher is interested in studying the changes in the spending behaviour of the white-collar employees, she or he will have to conduct the study at different points of time in a year or periodically. A study of this kind of a problem involves longitudinal (or time-series) designs.

Trend studies and panel studies

There are two types of longitudinal designs: Panel Study and Trend Study. In a panel study, the same respondents (people) are interviewed two or more times. In a trend study, two or more different samples of people (respondents) are drawn at different times from the same population (for examples see Box 21.2). You may have noticed that the television and newspaper media draw samples from time to time to study trends.

Box 21.2 Examples of Panel Study and Trend Study

Suppose we wish to understand the changing support for the leading political parties in the general elections to Parliament, two months before the election. For this study, we draw a sample of adults in the voting age group and 60 percent support, let us say, Political Party I. One month before the election, we re-interview the people who were a part of our sample. Perhaps, 55 percent of them now support Political Party I. This would be a panel study. On the other hand, if we draw a new sample of voters, one month before, and find that 55 percent of the new voters preferred Political Party I, this would be a trend study.

Trend studies provide information on net changes and panel studies provide information on both net and gross changes. Suppose at Time 1, sixty out of every 100 people say they would vote for Political Party A. Between Time 1 and Time 2, 10 of the Political Party A voters might change their opinion and vote for Political Party B and five of the Political Party B might vote for Political Party A. Thus, between Time 1 and Time 2, many people have changed their vote (opinion) and this is a gross change. The net result of these individual changes is that at Time 2, 55 of every 100 people would vote for Political Party A. Thus, if we use a panel design we can count the number of people in our sample who change their opinion between Time 1 and Time 2. If we use a trend design, we can describe the net change in the population, but we cannot count the number of individual change. If the net change is 5 percent (from 60 percent to 55 percent supporting Political Party A), the gross change might be anywhere between 5 percent and 85 percent. Trend studies describe how the distribution of a variable is changing for the population studied. Panel studies describe how individual members of a population are changing (see Box 21.3 for an example of panel design).

Box 21.3 An Example of Panel Design

Lazarsfeld, et.al, find a classic example of panel design in "The People's Choice". (1944). The researchers interviewed a panel of 600 residents of Erie County, Ohio (USA), once a month between May and November of 1940 with respect to how they intended to vote in the presidential election of 1940. Of particular interest to these investigators was the process by which panel members decided to vote the way they eventually did.

There are both advantages and disadvantages to panel designs. Most of the advantages relate to being able to interview the same respondents repeatedly over a period of months or years to see if any changes occur in their thinking and attitudes. If we find shifts in attitudes or reported behaviour, we are in a better position to argue that there has been a real change in the population.

Another advantage of panel design is that they permit much more information to be collected about each respondent than is feasible through

other designs. Panel designs also avoid a heavy dependence on the memory of respondents for information about his time covered by the panel. However, they may still rely on memory for questions about the respondent's past.

The primary disadvantage of panel study is sample mortality; that is, there will be a loss of panel members owing to lack of cooperation, death, or change of residence, etc. Another disadvantage is that it takes a long time to collect data and hence the costs involved may be more. Further, repeated surveys on the same population may cause respondents to become overly sensitized to their role as "selective" respondents.

Descriptive designs

There are two types of descriptive designs, namely, cross-sectional designs and cohort designs.

Cross-sectional designs

In a cross-sectional[®] design, the data is collected at one point in time. This is the most frequently used and most fundamental design in social survey research and is sometimes called 'single-shot survey'. This design is most appropriate for making inferences about the characteristics of the population from which the sample is drawn, and inferences about the relationships between variables at that point in time at which data is collected. The interviewing for such studies is sometimes carried out in less than a week, but more typically requires a few weeks and in some cases a few months. For example, a survey of current people's choices of a political party is a cross-sectional survey.

A researcher might use the data either to describe a sample on one variable or a number of variables or to demonstrate the association between these variables. Consider the following example. A researcher is interested in exploring the attitudes of the parents on the education of their children. He selects a sample of households with school going children in a community (Rural/Urban) and interviews the parents in those select households. After the interviews are completed, the researcher analyses the data and draws conclusions. The research design involved here is a cross-sectional design.

One of the advantages of this design is that a researcher can classify a sample into many quite different sub-groups to explore the separate dimensions of the research topic. A researcher may also use this design in causal analysis. Interestingly, by asking questions about the past it is possible to study certain longitudinal (overtime) social phenomena such as social mobility using a cross-sectional design.

However, the cross-sectional designs have two important limitations: (1) it is difficult to establish the time order of the variables, (2) it is difficult to analyse change over time. To illustrate the first point, suppose in a study, we find that the organisational affiliation is associated with political efficacy, feelings that one can influence the political system. The question

arises as to whether our respondents first joined organisations and later came to feel efficacious because of their experiences in that organisation, or whether they felt efficacious first, which led them to join a particular organisation. Alternatively, whether both causal sequences are operating, either for different people or for the same people. If we gathered data at one point in time, it is difficult to know when differences on each variable developed. Since causes come before effects, it is difficult to know in a cross-sectional design which variable should be taken as independent, causal agent.

A second limitation of the cross-sectional design is that it is difficult to use it for analysing changes over time. Consider the analysis of voting behaviour. At any given point during an election, we can ask respondents how they would vote 'if the elections were held today'. However, voting decisions might change during a campaign and different voters may exhibit different patterns of change. Some know their choice from the beginning and never vary; some switch back and forth between candidates; some are undecided until Election Day. Moreover, people who decide early in a campaign and never waver may have different characteristics and may experience different pressures than people who decide late. In sum, data collected at one time make it difficult to understand social processes that occur over time.

ii) Cohort designs

These forward-looking, or prospective, designs provide data about changes in a specific population. Suppose a survey of the aspirations of athletes participating in the 1996 Olympics is given in 1996, 2000, and 2004. This is a Cohort design and the Cohort is 1996 Olympics.

Cohort designs can also be retrospective, or look back over time (a historical Cohort) if the events being studied actually occurred before the onset of the survey. For example, suppose a group of persons was diagnosed ten years ago with disease X. If you survey their medical records to the present time, you are using a retrospective Cohort.

Reflection and Action 21.2

In the case of your study of waste disposal methods, you would have formulated research questions and then also refined them. For making clear the type of research questions and the type of research design you have formulated and adopted, answer the following questions on a sheet of paper.

Questions

- ❖ Are the questions formulated by you a mix of different types mentioned in the text or of one specific type only?
- ❖ What type of survey design have you selected for the survey in your research on waste disposal methods?

Write in detail about the type of design you have selected and also give reasons why you have selected this one and not any of the other designs mentioned in the text.

21.6 Sampling Design

After survey design has been selected, a sampling plan must be chosen. Many details about sampling have been discussed earlier in the units of Block 4 and Block 5, and therefore I am including only those that are new.

Why sampling?

Because the researchers may not include the whole of the population in survey research for reasons of costs and manpower, a part of the population is selected for understanding the characteristics of a population. Sampling, therefore, is a process of systematically selecting cases for inclusion in a research project. A researcher gets a set of cases (or a sample) that is more manageable and cost effective to work with than the pool of all cases. For example, it would be much less costly and time consuming to measure variables on 200 than 20,000 people. The researcher is not just interested in a small subset of cases. Instead, he would like to generalise about the entire pool. If well done, sampling lets the researcher measure variables on the smaller set of cases but generalise results accurately to all cases. For instance, if sampling is well conducted, a researcher can measure variables with, say, 2000 cases, and generalise to 200 million. Survey researchers argue strongly that there would only be 2 - 4 percent variation if all the 200 million cases were used for the study.

The next question is how it is possible to generalize from a handful of cases. It is not based on any magic, but on logical statistical reasoning that has been tested repeatedly with empirical evidence. Moreover, a researcher cannot use just any sample for the purposes of generalization. There are well laid down sampling procedures that require rigorous exercises in order to increase the level of precision in a study.

Basic terms in sampling

i) Sampling element

A sampling element is the unit of analysis (or case) in a population. It can be a person, a group, an organisation, a written document, or symbolic message, or even a social action (a divorce, a fight, etc.) that is being measured.

ii) Population

It is the large pool of elements from which a sample is drawn. Sometimes, the term 'universe' is used interchangeably with 'population'. To define the universe or population, a researcher specifies the unit being sampled, the geographical location, and the temporal boundaries of populations. The term 'target population' refers to the specific pool of cases the researcher wants to study.

iii) Sampling ratio

The ratio of the size of the sample to the size of the target population is the 'sampling ratio'. For example, a college has 2000 students, and a researcher draws a sample of 200 from it. Thus the sampling ratio is $200/2000 = 0.1$, or 10 percent.

iv) Sampling frame

A population is an abstract concept. Except for specific small populations, one can never truly freeze a population to measure it. For example, in a city at any given moment, some people are dying, some are traveling outside the city, some have gone on temporary migration to other cities, and some are taking birth. The researcher must decide exactly who to count. Should he count a city resident who happens to be on a holiday or outside the city when the time is fixed for the study? Therefore, the notion of 'population' is abstract and it exists in the mind, but is impossible to pinpoint concretely. Since it is an abstract concept, except for small populations, a researcher needs to estimate the population. Thus, it requires an operational definition.

A researcher operationalises a population by developing a specific list that closely approximates all the elements in the population. This list is a 'sampling frame'. The researcher may choose from many types of sampling frames: telephone directories, driving licences, ration cards, membership of a club, students' registration in a university, etc. A good sampling frame is crucial to good sampling. A mismatch between the sampling frame and the conceptually defined population can be a major source of error. For instance, if you select the telephone directory as a sampling frame, it constitutes only 5-10 percent of the population in a city, say, Delhi. The directory does not list those who do not have a telephone connection. It also does not update quickly the frequent shifts of residences and changes in telephone numbers in a city.

v) Parameter and statistic

Any characteristic of a population is a 'population parameter'. For instance, students from the science stream of children in the age group 6-11 years, etc. It is the true characteristic of the population. Parameters are determined when all elements in a population are measured. The parameter is never known with absolute accuracy for large populations, so researchers must estimate it as based on samples.

Researchers use information from the sample, called a 'statistic', to estimate population parameters.

Sampling errors

If many samples are taken from the same population, it is unlikely that they will all have characteristics identical either with each other or with the population. In brief, there will be what is called 'sampling error'. Sampling error is not necessarily the result of mistakes made in the sampling procedure. Rather, variations occur due to the chance selection of different individuals.

Types of sampling designs

Sampling designs are of two types. The first is called probability-sampling design and the second is non-probability sampling design. In Unit 15 of Block 5 our focus was on how to carry out calculations while here we are discussing each type of sampling design in terms of its essential character

and procedure. All the same you are likely to find some repetition of subject matter which has been retained to reinforce the learning of the same.

i) **Probability sampling design:** Probability sampling provides a statistical basis for stating that a sample is representative of the 'target population'. In probability sample, every element in the population has a known chance of being included in the sample. That means, every member of the target population has a non-zero probability of being included in the study/sample. This allows for estimates of the accuracy of sample findings in approximating what we would find out if we had conducted a census of the total population.

ii) **Non-probability sampling design:** In non-probability sampling designs, we do not know whether an element of the population has an equal chance of being selected. Its probability of selection cannot be determined, as is that with probability sampling where each element has a 50 percent chance of being selected and 50 percent chance of not being included in the sample. The non-probability samples are drawn based on judgment regarding the characteristics of the target population and the needs of survey. With non-probability sampling, some members of the eligible target population have a chance of being chosen and others do not. Thus, the statistical estimates of precision cannot be made with this sample. The non-probability sampling designs are preferred when there is no possibility of probability sampling. Whenever feasible, probability-sampling designs are preferred.

Probability sampling procedures

i) Simple random sampling

In simple random sampling, each member of the population under study has an equal chance of being selected. The method involves selecting at random from a list of the population (a sampling frame) the required number of subjects for the sample. Because of the probability and chance, the sample should contain subjects with characteristics similar to the population as a whole. For example, some old, some young, some tall, some short, some rich, some poor, etc. One essential requisite for this kind of sampling is that a complete list of the population (sampling frame) is needed.

The biggest advantage of simple random sampling is that a researcher can get an unbiased sample without much technical difficulty. For instance, once a member (or element) is selected, he (or it) is not eligible for a second chance and is not returned to the pool. This is what makes a simple random sample relatively unbiased. The typical ways of selecting a simple random sample are mainly through a lottery or through a table of random numbers or now through computer generated random numbers. The lottery method is adopted for smaller populations or sampling frame. For bigger sampling frames, the computer-generated numbers are selected.

It may be kept in mind that simple random sampling may not guarantee a perfect representation of the population. In other words, it may be wise to say that most random samples are close to the population most of the time but may not perfectly match the entire population.

ii) Systematic sampling

This is a modified version of simple random sampling. It involves selecting cases/elements from a population list in a systematic rather than random fashion. Here, the researcher calculates a 'sampling interval' rather than using a list of random numbers. The interval becomes his quasi-random selection method. Thus, in a systematic sample, every 'n'th member has a chance to be included in the sample. For instance, if we are to select 10 cases out of a total of 100, every tenth will have a chance to be selected. In this case, 'n' is any number between 1 and 10. Thus, the starting point in the systematic sampling is chosen at random.

In most cases, a simple random sample and a systematic sample yield virtually equivalent results. One important situation in which systematic sampling cannot be substituted for simple random sampling occurs when the elements in a sample are organised in a cycle or pattern. For example, if sampling is organised of married couples with the male first and female second, such a pattern gives the researcher an unrepresentative sample if a systematic sample is used. His systematic sample can be non-representative and include only wives or only husbands, depending upon the manner in which cases are organised or patterned.

iii) Stratified sampling

Stratified sampling involves dividing the population into homogenous groups (sub-groups or strata) each group containing subjects with similar characteristics. For example, in the earlier example, Group A might contain only males and Group B only females. After dividing the population into strata or sub-groups, the researcher draws either a simple random sample or systematic sample or both from each of the sub-groups.

How does one decide on sub-groups? The strata or sub-groups are chosen because evidence is available that they are related to the outcome. The justification for the selection of the strata can come from literature and expert opinion.

In stratified sampling, the researcher controls the relative size of each stratum, rather than letting random processes control it. This guarantees representation of different strata within a sample. However, one condition is that the stratified sampling procedures produce samples that are more representative of the population than simple random sample or systematic sample if the stratum information is accurate.

iv) Cluster sampling

The most widely employed probability sample design in survey research is cluster sampling. It addresses two problems - first is the lack of good sampling frame, and second the costs involved in reaching a sampled element or a case. For example, there is no single list of undergraduates

in colleges of a city. Even if one gets an accurate sampling frame, it would cost too much to reach many of the undergraduates as the colleges are spread out geographically in the city. In this case, instead of using a single sampling frame, researchers use a sampling design that involves 'clusters'. In this case, the cluster will be the college.

A cluster is a naturally occurring unit (e.g., a school, which has many classrooms, students, teachers; a city with zones, namely, East, West, South, Central, North, etc., states, etc.). The clusters are selected randomly and all members of the selected clusters are included in the sample or simple random or systematic or stratified samples are taken out of each cluster. Cluster sampling is used in large surveys. It differs from stratified sampling in that with cluster sampling one starts with a naturally occurring constituency. The researcher selects from among the clusters and either surveys all members of the selection or randomly selects from among them. The resulting sample may not be representatives of areas not covered by the cluster, nor does one cluster necessarily represent another.

v) Stage sampling

Stage sampling, also known as multi-stage sampling, is an extension of cluster sampling. It involves selection of a sample in several stages. That is, taking samples from samples. Suppose we want to survey children's academic performance in schools from a large city. One type of stage sampling might be to select a number of schools; at random, and from each of these schools select a number of classes and select children from within these classes. Another type of stage sampling could be to select one school, in terms of either the geographic region of the school in the community/city or in terms of the type of school (public, private aided, private unaided, etc.) and take their simple random sampling, systematic sampling, or stratified sampling or even mixture of all these.

Non-probability sampling procedures

i) Accidental or convenience sampling

This is also called 'man-on-the-street' survey and it involves choosing the individuals/cases that are readily available on the street, at a market place, at a school, or at a cinema theatre, etc., until the required sample size is obtained. This kind of survey can produce ineffective, highly unrepresentative samples and is not recommended. When a researcher haphazardly selects a sample that is convenient, he can easily get a sample that seriously misrepresents the population. Such samples are cheap and quick, although biases and errors are there in plenty. An example of this kind of survey is the kind of interviews television programmes conduct on the street.

ii) Quota sampling

Quota sampling has often been described as the non-probability equivalent of stratified sampling. In quota sampling, the researcher divides the population into sub-groups or categories such as men and women, reserved

castes and non-reserved castes, arts and commerce, younger and older, etc., and then decides the proportion of individual in each category or sub-group. Thus, the number of respondents is fixed in various categories of the sample.

Quota sampling is an improvement because the researcher can ensure that some population differences are in the sample. However, it is not to argue that it is fully representative. For instance, once quota categories are selected in quota sampling, the actual number of individuals in each category is selected based on the convenience or accidental sampling. Further, the researcher bias might also creep into the selection of actual cases of individuals who are included in the study.

iii) Purposive or judgmental sampling

In purposive sampling, researchers handpick the cases to be included in the sample based on their judgment of their typicality. That means, the judgment of the researcher is used in selecting cases with a specific purpose in mind.

Purposive sampling is appropriate in the following three situations. First, a researcher uses it to select unique cases that are especially informative. Second, a researcher may use purposive sampling to select members of a difficult-to-reach, specialised population. Third, the purposive sampling is preferred when a researcher wants to identify particular types of cases for in-depth investigation.

iv) Dimensional sampling

This is a further refinement of quota sampling. It involves identifying various factors of interest in a population and obtaining at least one respondent of every combination of those factors.

v) Snowball sampling

It is also called a network or chain referral or reputational sampling and it is a method for identifying and sampling (or selecting) the cases in a network. Snowball sampling is based on an analogy to a snowball which begins small but becomes larger as it rolls on wet snow and picks up additional snow. Friendship networks are the most important sources of this kind of sampling. Further, professional associations, chat groups, etc. may also provide bases for such sampling.

Sample size

A question often asked is of the size of the sample. How large should it be? The best answer could be 'it depends'. That means there is no clear-cut answer to this question. However, the literature on survey sampling does give a general advice to prospective survey researchers. For some people, a sample of thirty is the bare minimum if the researcher plans to some form of statistical analysis of the data. However, the size of the sample depends on the plans of the researcher and how accurate the sample has to be for his purposes, and on population characteristics. It may be stated here that a large sample alone does not guarantee a

representative sample. Further, when we increase the sample size, we increase costs. Larger samples mean increased costs. A researcher may keep in mind all these considerations while judging the size of the sample.

In Section 15.3 of Unit 15, you have already learnt how to calculate sample size, when estimating mean and when sampling for proportion.

Reflection and Action 21.3

Discuss in a note of about five hundred words the sampling type and design of the survey you need to carry out in your research on waste disposal methods. While writing the note, take help of the relevant text in the unit.

21.7 Conclusion

This unit has attempted to discuss the initial concerns of the survey researcher. It spelled out the essential prerequisites for undertaking a survey and the steps to be followed in survey research in order to make it more systematic. The unit provided a detailed account of research question formulation, survey research designs, and the sampling designs. Thus, it tried to equip the student with the basic terms and concepts and the process with suitable examples.

Further Reading

Aldridge, Alan. And Levine, Ken. 2001. *Surveying the Social World - Principles and Practice in Survey Research*. Open University Press: Buckingham

Fink, Arlene. 1995. *The Survey Hand Book*. Sage Publications: Thousand Oaks

Cohen, Louis and Manion, Lawrence. 1994. *Research Methods in Education*. Routledge: London and New York

Moser, C. A. and G. Kalton 1973. *Survey Methods in Social Investigation*. The English Language Book Society: London

de Vaus, David. 2002. *Surveys in Social Research*. Routledge: London