

QUESTIONNAIRE SURVEYS IN MEDIA RESEARCH

Elirea Bornman

LEARNING OUTCOMES

At the end of this chapter, you should be able to:

- discern whether a questionnaire survey is appropriate to investigate a particular media-related research problem or issue, or not;
- make an informed choice of sampling design appropriate for investigating a particular media-related problem or issue;
- draw various probability and nonprobability samples;
- make an informed choice of a type of survey appropriate for investigating a particular media-related problem or issue;
- design a questionnaire to investigate a particular media-related problem or issue;
- critically evaluate questionnaire survey studies;
- conduct a questionnaire survey of your own on a limited scale.



THIS CHAPTER

This chapter gives an overview of the theory and practices associated with questionnaire surveys as a quantitative methodology, with specific reference to the ways in which surveys are used in media research. An application of how questionnaire surveys are used in audience research is illustrated by a case study. The most important topics dealt with in the chapter are the following:

- historical development of the use of questionnaire surveys in the social sciences and other applied disciplines;
- research topics appropriate for questionnaire surveys;
- steps in questionnaire surveys;
- probability and nonprobability sampling;
- types of surveys;
- designing of questionnaires;
- reliability and validity as yardsticks of quality for questionnaire surveys;
- sources of error in questionnaire surveys.

10.1 INTRODUCTION

Most of us are accustomed to questionnaire surveys. At some or other time each of us has been confronted by a fieldworker or a letter (with a questionnaire attached to it) with a request to answer some questions. A questionnaire could be included in your study material for a course in media studies. One of the questions could read as follows: 'What is your opinion of your textbook on media audiences? Do you think it is (a) an excellent book; (b) a good book; (c) an average book; (d) not such a good book; or (e) a poor book?'

You would probably not be ruffled by such a request, as you have encountered questionnaires before. The reason is that the questionnaire survey is one of the most widely used research methodologies in the social sciences as well as in the marketing industries and other applied fields (Babbie 1990; Neuman 2006). However, the popularity of surveys might be misleading. It might appear to a layperson that surveys are a quick and easy option to obtain information about a topic. A poorly conducted survey could, however, easily yield misleading or worthless results. Good surveys are complex endeavours that require a lot of thought, effort and dedication. In this chapter, you will learn more about the ingredients of good surveys as well as the strengths and limitations of questionnaire surveys as a research methodology.



10.2 BRIEF HISTORICAL OVERVIEW OF SURVEY RESEARCH

The survey is a very old technique. It can be traced back to ancient forms of the census. The aim of a census is usually to count people and/or to obtain information on the characteristics of the entire population of a particular territory (Babbie 1990; Neuman 2006).

In the Old Testament, Moses instructed Eleazar, the son of Aaron, to take a census of all the people of Israel. King David also conducted a census of all the people in his kingdom. Jesus was born when Joseph and Mary went to their ancestral town to be counted. Ancient Egyptian leaders conducted censuses to assist them in administering their domains. The *Domesday Book* represents a famous census of England conducted from 1085 to 1086 by William the Conqueror. Particularly in early censuses, the aim was to assess the potential income or property available for taxation or the number of young men available for military service. The power of ancient leaders was also established by the number of their subjects.

Over time, surveys also acquired other political functions. Babbie (2001:238) gives the example of a little-known survey that was conducted in 1880. A German socialist mailed approximately 25,000 questionnaires to French workers. The rather lengthy questionnaire contained intricate questions such as the following: 'Does your employer or his representative resort to trickery in order to defraud you of a part of your earnings?' The survey researcher in this case was none other than the socialist theorist, Karl Marx.

The development of representative democracy as a form of government has added to the importance of both censuses and surveys (Babbie 1990; Neuman 2006). Nowadays, censuses are, among others, used to assign the number of parliamentary representatives for a particular group or region. Census bureaus have also made important contributions to the development of various aspects of surveying. Opinion polls have furthermore become part and parcel of modern elections. Today, very few democratic elections take place without research organisations conducting polls in an attempt to predict the outcome. Commercial research organisations have furthermore played an important role in expanding the use of surveys to marketing research (especially during periods when no elections are conducted).

The refinement of questionnaire surveys as a respectable 'scientific'

survey research



research methodology in the social sciences started at universities in the United States and the United Kingdom. Babbie (1990) mentions two individuals in this regard, namely Samuel A. Stouffer and Paul Lazersfeld. Stouffer did pioneering work by employing survey methods to investigate social problems such as the effects of the Depression in the US and the status of black Americans during the 1930s. Paul Lazersfeld, who came from a European intellectual background, applied rigorous survey techniques in studying phenomena such as communications, leadership and economic behaviour. The use of questionnaire surveys in academia was furthermore stimulated by the quest of social scientists to model themselves after natural scientists and to become more professional, objective and nonpolitical.

World War II served as a further important incentive for the development and refinement of survey research techniques (Neuman 2006). The war also promoted cooperation between academic social researchers and marketing practitioners who joined forces in the war effort by studying morale, enemy propaganda, production capacity, and so forth. In doing so, they learnt from each other and gained experience in conducting large-scale surveys. Academic researchers helped practitioners to appreciate precise measurement, sampling and advanced methods of statistical analysis. Practitioners, on the other hand, introduced academic researchers to the practical side of organising and conducting the fieldwork in large-scale surveys. Many of these researchers joined universities after the war, where they promoted the use of survey methods in social research. Despite the fact that some university researchers were initially sceptical of a research technique that had been used predominantly by politicians and the marketing industry, survey research has constantly grown in use within academia.

Nowadays, survey research is widely used in both academic research as well as applied areas (Neuman 2006). Researchers in many social science disciplines (for example, communication, education, sociology, political science and social psychology) conduct surveys to expand the knowledge base within their respective fields. Quantitative survey research has furthermore become a major industry. Many applied fields rely heavily on surveys. Governments all over the world regularly conduct surveys to inform policy decisions. Surveys have, in particular, become indispensable in media industries, where broadcasting, print and other industries make use of surveys to obtain information about



media audiences (see Chapter 12). Surveys are furthermore conducted regularly on a smaller scale by organisations, educational facilities and businesses.

Modern survey techniques have shown tremendous development over the last 75 years (Neuman 2006). Technological development, such as computers, have facilitated survey research by making data capturing, data storage and statistical analysis easier and quicker. In recent decades, researchers have started to develop theories and to conduct research into aspects of the survey process itself, such as the communication–interaction processes associated with survey interviews, the effectiveness of visual clues, the impact of question wording and ordering, and the reasons for respondent cooperation and refusal. As such, this kind of metatheoretical research to enhance the validity of surveys has become an exciting field of scientific enquiry in itself.

10.3 WHAT IS A SURVEY?

Survey research has developed within the positivist paradigm of social research. The data produced by a survey are inherently statistical in nature. Robert Groves (1996:389) calls surveys 'quantitative beasts'. Although it is possible to include open-ended questions in a questionnaire, which can yield data of a more qualitative nature (see section 10.8.1), survey research is predominantly a quantitative methodology and the data are reported in the form of tables, graphs and statistics such as frequencies, means, standard deviations, t and F values, correlation coefficients, and so forth. In most cases, even the responses on open-ended questions are coded in such a way that the results can be reported in the form of statistics.

A further important characteristic of survey research is that it is essentially a self-report methodology (Neuman 2006). The term 'self-report' refers to the fact that we ask people – called respondents – questions when conducting a survey. Respondents are requested to provide information regarding themselves and/or to describe their own behaviour, attitudes, opinions, and so forth.

When studying surveys as a methodology, you need to keep in mind that survey research is just one of a range of research methodologies available to the social researcher (Babbie 1990). Whereas the questionnaire survey is a versatile methodology that can be applied in a variety of contexts to

I social research



investigate a multitude of topics, it is not appropriate to many research topics, and might not be the best approach to study some of the topics to which it is sometimes applied.

10.4 RESEARCH TOPICS APPROPRIATE FOR QUESTIONNAIRE SURVEYS

research topics

As surveys involve self-reporting, it is an appropriate methodology to investigate research questions dealing with topics on which we can ask questions which people will be able to answer. Questions can be asked on the following issues:

- Characteristics We can ask people questions on their demographic characteristics, such as their age, gender, race, language, educational qualifications, income, marital status, and so forth.
- Behaviour or behavioural intentions People can answer questions on what they did in the past, what they usually do and/or what they intend to do: did you watch Generations last night? Do you listen to the radio in the mornings? To which radio stations do you listen? How many hours do you watch television during weekends? Do you subscribe to satellite television? Do you intend to subscribe to satellite television in the near future?
- Self-classification People can be asked to classify themselves into various groups: into which social class would you categorise your family? Do you consider yourself to be a heavy television viewer, moderate television viewer, or light television viewer? Would you call yourself a soap opera addict?
- *Preferences* People can voice their preferences: which radio programmes do you prefer? If you have a choice, would you rather watch a rugby game or a soccer match?
- Attitudes/beliefs/opinions Surveys are, par excellence, appropriate
 to question people about their attitudes, beliefs and/or opinions on
 almost any topic: what kind of job do you think the SABC Board is
 doing? Do you think that the media in South Africa are really free?
 What is the most important factor threatening press freedom in
 Africa?
- Expectations People can answer questions on what they expect to happen: who do you think will become the new chairman of the SABC Council? Do you think that press freedom will improve, stay the same or deteriorate in Africa?



 Knowledge or awareness – People can be asked questions to establish their knowledge or awareness of certain issues: do we have commercial television stations in South Africa? If yes, name them. Which soap operas are broadcast on the television channels of the SABC?

When conducting survey research, it is important to remember that respondents can only provide reliable and valid answers on questions pertaining to themselves. Questionnaire surveys are therefore a very personal research method. Respondents cannot answer questions on behalf of other people. In audience research, parents are sometimes requested to provide information on the television viewing behaviour of their children (see Chapter 12). However, research has proven the results of such studies to be less reliable as parents tend to underestimate their children's television viewing (Webster, Phalen & Lichty 2006). Sometimes a representative is requested to complete a questionnaire on behalf of the organisation or group. In such a case, only specific information pertaining to the characteristics of the organisation or group can be answered. The representative cannot, for example, provide information on the attitudes or opinions of individual employees.

Questionnaire surveys are also not a good option to provide answers on 'why' questions, that is, research questions dealing with the reasons for particular trends or phenomena (Neuman 2006). It is, for example, not possible to establish by means of survey research the reasons why many South African learners watch television during school hours. A particular learner can, however, be asked about his or her own behaviour, that is, whether he or she watches television during school hours. People can also be asked for their opinions regarding such behaviour or what they think are the reasons for a particular trend or phenomenon. Researchers can also ask respondents what they think are the reasons for the lack of discipline in South African schools. They can also ask respondents whether they think that television plays a role. The responses will, however, represent nothing more than opinions, namely respondents' subjective understanding of the issue at hand. Few respondents will have the expertise to be fully aware of the complexity of factors that shape a particular phenomenon, such as undisciplined behaviour among children.

Survey research is also not a good option to investigate 'how' questions. The results of survey research can assist policy-makers in identifying



potential problem areas with regard to a particular phenomenon and could thus form the basis for deriving strategies to address problem areas. However, the results of a survey can only provide information on how respondents behave and/or what they think or feel should be done, that is, information on behaviour, perceptions, opinions and/or attitudes. Few respondents would, however, have the insight into the complexity of a particular phenomenon to make informed proposals to address problem areas.

In media research, surveys are also inappropriate where questions regarding the contents of media products are investigated. Surveys can, for example, not answer questions on the foreign and/or local contents of television schedules or the violent contents of particular programmes. Again, survey research can only provide information on respondents' behaviour (for example, whether they watch particular programmes or not), opinions and/or attitudes on such phenomena.

10.5 STEPS IN SURVEY RESEARCH

research steps

The steps in survey research are discussed here in order to indicate the complex nature of the decisions that have to be taken at each stage. The steps are, however, not necessarily in chronological order. The following basic steps are followed in most questionnaire surveys (Neuman 2006; Van Vuuren, Maree & De Beer 1998):

• Formulate the research question or problem, sub-problems and/ or hypotheses - In survey research, a deductive approach is usually followed (Neuman 2006). This means that the researcher begins with a research question and/or problem and ends with empirical measurement and data analysis. Research questions usually follow from an idea that requires research. Especially in academic environments, the idea could represent a personal preference or interest of a researcher. Experienced researchers in a particular field usually have such a proliferation of research ideas that they cannot research them all. In the end, the curiosity and creativity of the researcher play an important role in identifying research ideas. In institutional and marketing research environments, the research problem is usually predetermined by the needs of an institution or particular clients. In audience measurement, for example, the research questions are determined by the needs of the media and marketing industries. Whatever the case, the research question is the



rudder that steers the boat and determines the nature of the journey. Every single step in a survey is determined by the aim of a study, as represented by the research question or problem, sub-problems and/ or hypotheses. This means that, in the absence of absolute clarity on the aim of a survey study, the researcher(s) will fumble in making the crucial decisions involved with each of the other steps.

- Answer the question: is a survey a viable and/or the most appropriate methodology to investigate the research problem? Again, at this stage the researcher must seriously consider whether the research question indeed addresses one or more of the aspects that can be investigated by means of survey research as discussed in section 10.4. If not, it should be considered whether an alternative methodology would not be more appropriate.
- Study the literature Now the researcher has turned his or her research idea into an acceptable research question, problem, subproblem(s) and/or hypotheses. The researcher has also established that survey research is indeed a viable method to provide answers on these. What now? How does one start? Before starting with the real business of conducting a survey, the wise and sensible researcher will have a look at what other researchers have done in the field. This is done by doing a thorough review of the available literature. Reading of the available literature should be both extensive and intensive. Students often say that they could find nothing in the literature relating to their study. That usually means that they did not search hard enough, did not make use of modern technology available for literature searches or did not use the available facilities correctly. Even in the case of experienced researchers, it is wise to seek the assistance of a subject librarian. The survey researcher, in particular, should search for articles and/or reports that describe other survey studies that have been done on the same topic or related topics. How did these researchers approach the topic(s)? Which questions did they ask, and how did they ask questions on relevant issues? In doing so, the researcher will get a good idea on what has worked for other researchers and what has not.
- Do some homework and investigative research Apart from studying the literature, it is also a good idea to talk with experts and/or experienced researchers in the field. There are some tricks of the trade that are not written in textbooks, articles or reports. Many researchers also conduct investigative research, usually of a qualitative nature

429





(see Chapter 11). Interviews can be conducted with people who can offer insight into the topic and/or problem under investigation. Focus group interviews with people representative of the population could furthermore help the survey researcher to get a grasp on the problem area or the variety of issues at stake. Such research will help the researcher to set relevant questions in the questionnaire (Bailey 1987).

- *Identify the target population and accessible populations* The nature of the population is another factor determining most of the other steps and decisions in the process of survey research. The nature of the population should also be taken into account in the construction of the questionnaire and the formulation of questions. It is therefore imperative to identify and describe the target and accessible populations clearly and unambiguously (see section 10.6.1).
- Decide on the type of questionnaire survey to be employed, and when, where and how the survey will be conducted The aim of the study and the nature of the population should be considered when decisions on the nature of the survey are made. In addition, all kinds of practical considerations, such as the time factor and the availability of financial and other resources, should be taken into account in order to ensure the success of the survey. Various types of surveys as well as their strengths and limitations are discussed in section 10.7.
- Decide on the sample size, the type of sampling and the procedure of sampling to be employed Sampling is pertinent to questionnaire surveys. Both the reliability and validity of a survey depend largely on the type and procedure of sampling employed, and on the size of the sample. The survey researcher should therefore have a thorough knowledge of both the techniques of survey research, and the theory of sampling. Once again, in the case of complex populations, the researcher should not hesitate to obtain the service of sampling experts (usually statisticians who specialise in sampling). Sampling is further discussed in section 10.6.
- *Draw the sample* After deciding on the sample size, and the type and procedure of sampling, the procedure should be applied meticulously in every step of drawing the sample. No shortcuts should be taken; otherwise both the reliability and validity of the survey could suffer (see section 10.9).
- *Construct the questionnaire* The quality and usefulness of the data obtained will depend largely on the quality of the questionnaire. The



questionnaire is the survey researcher's methodological instrument, just as the scalpel is the instrument of the surgeon. It is imperative that this instrument should be sharp and effective. The researcher should therefore carefully consider the issues to be covered in the questionnaire, the wording and format of each question, the logical order of questions, and so forth. More information on questionnaire construction is provided in section 10.8.

- *Pilot-test the questionnaire; obtain the opinions of experts As already* indicated, no questionnaire researcher should ever depend solely on his or her own insight, knowledge and/or expertise when constructing a questionnaire. Other experts on questionnaire construction and experts within the field of investigation should be consulted. Even after the insight and advice of such experts has been obtained, the questionnaire should be pilot-tested, and adjusted according to the insights yielded by pilot-testing, before the final layout and printing are done. No changes can be made to a questionnaire once it has been printed or copied. Care is therefore necessary to ensure that data obtained by means of the final questionnaire will indeed meet the expectations of the researcher and other stakeholders. One of the techniques of pilot-testing is to ask a small number of people who are representative of the population to complete the draft questionnaire. After having done so, the researcher conducts interviews with the respondents in order to establish whether they experienced any problems with the questionnaire, whether there were any instructions, questions, words and/or phrases that they did not understand, and so forth. In doing so, potential problems with the questionnaire can be identified and rectified before the printing of the final questionnaire.
- Do the layout of the questionnaire The technical layout of a questionnaire should be carefully done, as the electronic capturing of the data depends on the correct layout. Researchers are advised to consult with the persons who will be responsible for data capturing in order to ensure that the layout meets their requirements.
- Train interviewers and conduct the survey Once the questionnaire has been finalised and printed or copied, the researcher reaches the stage where the actual survey can be conducted. If the type of survey requires the involvement of interviewers, thorough training of interviewers is imperative to ensure that the interviewers will follow the sampling instructions meticulously and conduct the interviews according to the requirements of the researcher.



- Code open-ended questions Once all the questionnaires have been completed and returned, open-ended questions have to be coded in order to allow for the electronic capturing of the data. (Closed or structured questions are usually presented in coded format in the questionnaire.) Principles similar to those for the thematic coding of text in content analysis are followed (see Cooper & Schindler (2001:424–34) for guidelines on the coding of questions).
- Capture the data in electronic format After the coding of openended questions, the data can be captured electronically. Most statistical packages nowadays have a program that enables researchers to capture the data themselves. However, mistakes in the capturing of data imply faulty data, from which wrong conclusions could be drawn. The capturing of data should therefore be meticulously controlled in order to ensure that each respondent's answers are captured correctly. If fairly large samples are involved, it is usually too time-consuming for the individual researcher to capture the data. In such cases, researchers make use of professional data-capturing services.
- Do the statistical analysis of the data This is the stage to which all survey researchers look forward! Finally, the researcher will be able to see what answers the data yield on the research question(s). As already mentioned, questionnaire data are essentially statistical in nature. Statistical analysis enables the researcher to summarise the data and draw relations between variables. Nowadays, a number of user-friendly statistical packages, such as the Statistical Analysis System (SAS) and Statistical Software for the Social Sciences (SPSS), are available. Analysing the data themselves is usually an immensely fulfilling experience for researchers who have been involved in the project from the beginning. However, the user-friendliness of statistical packages can be misleading. Researchers should have a sound background in the underlying assumptions and premises of the statistical techniques that they use. If they do not have the knowledge or expertise to do the data analysis, a statistician should be consulted.
- Interpret the statistics and draw conclusions Whereas the statistical analysis of the data can be done by a statistician, it is the responsibility and privilege of the researcher (sometimes with the help of a statistician) to interpret the data and draw conclusions in the light of the initial aim of the study.



• Disseminate the findings – The final stage of the research venture is perhaps the most important. All the trouble taken with the previous steps will be in vain if the research findings are not appropriately disseminated and applied in decision-making. The form in which the findings will be reported will depend on the aims of the study and the stakeholders involved. Academic research is usually reported in the form of dissertations, theses, conference papers and articles published in academic journals. In commercial settings, concise executive summaries of the most important findings are often preferred. Oral presentations, which allow the researchers to interact with groups of stakeholders on the findings, represent another popular form of dissemination. A researcher should always be ready to adapt the reporting of findings according to the needs of particular stakeholders.

Again, it needs to be emphasised that the steps as discussed in this section do not necessarily need to be followed in the same order. Although it is, for example, essential to develop a questionnaire before there can be any question of datagathering, the sample-drawing procedure and the development of the questionnaire could quite feasibly proceed simultaneously. After the pilot-testing of the questionnaire, the researcher will probably go back to the questionnaire development phase to address problems that have been identified in the pilot-testing. And this does not mean – nor is it advisable – that the researcher should take full responsibility for each phase. Good questionnaire surveys are usually a team effort rather than a solo flight. Few researchers have the expertise and capacity to conduct every step in survey research successfully on their own. A survey researcher should therefore not hesitate to seek the advice and assistance of experts when needed.

10.6 SAMPLING

If you go for a blood test, the nurse will take only a sample of your blood. As your blood is exactly the same throughout your body, the sample will be sufficient to provide information on the state of your health. More or less the same principle applies to sampling in questionnaire surveys. If, for example, you want to study the radio listening behaviour of the South African public, it is hardly possible to interview every member of the public. You will therefore have to take a so-called sample of the South African public.

sampling



The main motivations for sampling in survey research are time and cost. Theoretically, it might be possible to interview every member of the sector of society that you are interested in. For example, in the population censuses that are conducted at regular intervals in South Africa, the underlying principle is to obtain information about every person who sleeps within the borders of South Africa on a particular night. Unfortunately, surveys where you interview every member of the population are in most cases too expensive and time-consuming and just not practically possible. Survey researchers therefore make use of sampling methods to select a subgroup of the population, the members of which are then interviewed. Sampling as practised in survey research is, however, more complex than taking a blood test. Unlike the blood in your body, the members of most societies are not the same throughout, but tend to differ from one another in a variety of characteristics. Societal variety and heterogeneity are indeed some of the most difficult problems that sampling practitioners have to deal with.

10.6.1 Important concepts in sampling

Before we discuss methods of sampling in greater detail, you should acquaint yourself with the vocabulary used when sampling methods are discussed (Babbie 2001; Neuman 2006).

Unit of analysis

units of analysis

A sample is drawn from a larger pool of cases or elements. The sampling elements are called units of analysis. In media research, the units of analysis are usually individuals as most media researchers are interested in people's media behaviour, media consumption or reactions towards the media. The units of analysis could, however, also be organisations. A researcher can, for example, be interested in how various newspapers deal with race-related issues. Here, the focus of interest is not the opinions of individual journalists that work for a particular newspaper, but the official policy of the newspaper itself. The questionnaire will therefore be sent to the offices of newspapers. The questionnaire can be completed by the editor or any other person with the expertise to answer the questions. Only a single questionnaire will be completed for each newspaper. The unit of analysis will therefore be newspapers. In content analysis, the units of analysis are often newspaper articles, television programmes, policy documents, television schedules, and so forth.



Population or universe

The larger pool, from which a sample is drawn, is called the population. The term 'universe' is often used interchangeably with 'population'. The term 'population' as used in sampling should not be confused with the general way the term is used in the layman's world. In sampling, it is imperative to define the population very precisely. In the case of the All Media and Products Survey (AMPS) conducted by the South African Advertising Research Foundation (SAARF), the population is all members of the South African public of 16 years or older (SAARF sa). Three aspects are usually specified in the definition of the population:

- The units of analysis in the example of AMPS, the units of analysis are individual members of the society.
- Geographic location in AMPS, only people residing in South Africa are included in the population.
- Temporal or other boundaries in AMPS, the population is limited to individuals of 16 years or older, that is, persons born before a specific date.

In delineating the population of a study, vague terms such as 'adults' or 'teenagers' should be precisely defined by specifying age limits. When a term such as 'university students' is used, it should, for example, be indicated whether both full-time and part-time students, undergraduate and postgraduate students, university students and/or students of technical colleges, and so forth, are considered. In short, there should be no ambiguity or uncertainty about who is included in the population (or universe) and who is excluded.

Target and accessible population

Apart from the terms 'sample' and 'population', a distinction is sometimes drawn between the target and the accessible population. Whereas the *target population* refers to the total pool or section of society that we are interested in, the *accessible population* refers to that section to which we have access in order to draw a sample. Let us take the example of the researcher who wants to study the opinions of the South African public on the SABC. The target population in this case is all adult members of the South African public of 16 years or older. However, because of time and cost constraints, the researcher decides to draw a sample from the readily available address list of South Africans who pay television licences. In this case, the accessible

population or universe

target and accessible population



population is all the members of the public whose names appear on this address list.

The definitions of the target and accessible populations have serious implications for the generalisation of the results of a study and therefore the impact and applicability of the results (Du Plooy 2001). For example, if you draw a sample only from the names on the address list of people who pay television licences, you will only be able to generalise the findings to the people on the list.

Sampling frame

sampling frame

A sampling frame is the list, or quasi-list, that closely approximates all the elements or units of analysis from the population (Neuman 2006). If a sample is, for example, drawn from the address list of people who pay television licences, this list of addresses represents the sampling frame. The sampling frame plays a cardinal role in drawing probability samples. Examples of sampling frames are telephone directories, tax records, lists of the students registered at a particular university, the list of employees of an organisation, and so forth. Most techniques in probability sampling depend on the availability of a sampling frame. It is furthermore only possible to draw generalisations about the population represented by the sampling frame. A good sampling frame is therefore vital to good sampling. A mismatch between the conceptually defined population and the sampling frame can give rise to invalid sampling and sampling error.

Sampling ratio and sampling interval

sampling ratio and interval The sampling ratio is the sample size to the size of the population, that is the number of names on the sampling frame (Neuman 2006). For example, if the population consists of 50,000 people and a researcher draws a sample of 1000, the sampling ratio is 1:50. Thus the sampling ratio is the number of elements in the population divided by the number of elements in the sample. The sampling interval represents the standard distance between the elements selected for a sample. Given a sampling ratio of 1:50, every 50th element in the sampling frame will be selected. The sampling interval is therefore 50. The sampling ratio and sampling interval play an important role in probability sampling techniques such as systematic sampling.



Sample size

A common worry among questionnaire researchers is how large a sample should be in order to conduct a reasonably good survey (Nardi 2003). There is no simple answer to this question. Some textbooks offer formulae to calculate the desirable sample size. It is, however, the experience of this author that such formulae are seldom used in practice. Decisions on sample size are rather determined by practical considerations such as financial, manpower and time constraints. The type of survey is also an important factor. As internet, postal and self-administered surveys are relatively simple and cost-effective, it is possible to involve relatively large samples. Face-to-face interviews are, however, expensive and labour-intensive endeavours, and financial and manpower considerations will play an important role in decisions on sample size (see section 10.7).

There are a number of considerations to keep in mind. The first is rather simple: the larger the sample, the better. Sampling error is bigger when generalising from a small sample to the population (see section 10.10). According to Kerlinger (1986:119), probability sampling techniques require that the sample should be sufficiently large to give the principles of randomisation 'a chance to "work" (see section 10.6.2). Larger samples also have greater statistical power in the testing of hypotheses. However, if nonprobability sampling techniques are used, the sample size does not matter as much (see section 10.6.3). A large sample size will not correct the bias inherent to nonprobability sampling. Regardless of the sample size, generalisation to the population will not be possible (Nardi 2003).

The nature of the population that is being studied should also be kept in mind (Nardi 2003). When a population is relatively homogeneous, a smaller sample can be acceptable. Also, if it is expected that all members of the population think or behave more or less the same, that is, when there is not large variation in the phenomena being studied, a smaller sample could sufficiently reflect the population. However, larger samples are required for heterogeneous populations (for example, the South African population) and where large variations in the phenomena being studied are expected.

Another important consideration is the kind of statistical analyses you are planning to do. A smaller sample size can restrict the choice of

sample size



statistical analyses that can be conducted. If you are planning to compare subgroups, such as gender, age or racial groups, the sample size should be big enough to ensure a sufficient representation of these subgroups. A minimum sample size is also required to conduct more sophisticated statistical analyses such as factor analyses. In order to prevent later disappointment, it is therefore recommended that a statistician should be consulted in making decisions on sample size.

10.6.2 Probability sampling

In survey research, we distinguish between two types of sampling, namely probability and non-probability sampling (Babbie 2001; Neuman 2006). The terms 'random' and 'non-random' sampling are also used. The term 'random' refers to a mathematical process that yields a mathematical random result, that is, an outcome with no distinctive pattern. In a true random sampling process, each element or unit of analysis in the population has an equal chance of being selected for the sample.

random and non-random sampling In most cases, random or probability samples require more work and effort than non-random samples (or non-probability samples). Some of the most tiresome tasks are to acquire or compile a sampling frame and to reach each of the elements selected for the sample to complete the questionnaire. However, probability samples are in most cases superior to and preferred above non-probability samples. Firstly, probability sampling avoids a researcher's or fieldworker's conscious or unconscious bias in selecting respondents. It is consequently more likely to generate a sample that will be truly representative of the population. Probability sampling furthermore enables the researcher to use powerful statistical techniques in the analysis of the data and to generalise the findings to the population. Several options are available for drawing a probability sample.

Census-type sampling

census-type sampling When the pool from which elements are drawn is relatively small, the researcher can consider including all the elements in the study. Such studies are called census-type studies (Kotzé 2004). For example, if a researcher investigates the internet-usage behaviour of the employees of a small organisation, all employees can be asked to complete the questionnaire. (Of course, it will seldom be possible to involve each and every employee. Some might be sick or on leave during the period when



the study is conducted, while others might simply refuse to participate.) As a rule of thumb, a census-type study should be considered when the population has 500 or less elements. In the case of larger populations, a census-type study can be considered when the manpower and financial resources are available and one of the easier methods of interviewing, such as self-administration, is employed.

Simple random sampling

Simple random sampling is exactly what the term indicates. It represents the simplest way of drawing a probability sample (Babbie 2001; Neuman 2006). For this form of probability sampling an accurate sampling frame needs to be obtained, whereafter elements listed in the sampling frame are selected according to a mathematically random procedure. The simplest procedure is to throw all the names into a hat or other container, shuffle the names and draw elements for the sample one by one. More advanced techniques require that each name in the sampling frame is numbered. If there are 100 names on the list, 100 numbers are required. A table of random numbers is then used to draw the sample. Tables of random numbers and exact descriptions of the sample drawing procedure can be found in most general research and statistical handbooks (see the list of recommended books). If the sampling frame is available in a form that can be read by a computer program (for example a computer database, a DVD or CD or any other form of computer disc), a simple random sample can be drawn by means of a computer program. In practice, the computer program numbers all the elements in the sampling frame, generates its own list of random numbers and prints out the list of elements selected. Although simple random sampling is the purest form, on which all other forms of probability sampling are based, it is often not practical to implement.

simple random sampling

Systematic sampling

Instead of relying on random numbers, researchers often use systematic sampling when a list of elements (a sampling frame) is available. In systematic sampling, the sampling interval serves as a basis for selecting elements (Babbie 2001; Neuman 2006). Every kth element in the total list is selected. For example, if the list contains 10,000 elements and a sample of 1000 is required, every 10th element would be chosen. The sampling will start with any random number between 1 and 10. In order

systematic sampling



to avoid any human bias, it is important to select the first element at random. In order to select a random starting point between 1 and k, the person responsible for drawing the sample often closes his or her eyes and points (with closed eyes) to a number between 1 and k. The element having that number is included in the sample and from there on every kth element is selected.

In most cases, simple random sampling and probability sampling will yield similar results. However, systematic sampling is sometimes not desirable where there is a cyclical pattern in the way elements are listed in the sampling frame. Babbie (2001) quotes in this regard the example of a classic study conducted during World War II, in which every 10th soldier was selected from unit rosters. The rosters were, however, arranged squad by squad, with the sergeants first, then corporals and then privates. Each squad had 10 members. The end result was that the systematic sample contained only sergeants. If another random starting point was used, the sample could – for the same reasons – have contained no sergeants at all, which is also not desirable. When a systematic sample is considered, the sampling frame should therefore be examined carefully. If the elements are arranged in a specific order, the researcher should ask the question whether the order would bias the sample in any way.

However, Babbie (2001) regards systematic sampling not only as more convenient but also as superior to simple random sampling in most cases. Systematic sampling is furthermore often used in combination with other forms of sampling – see, in this regard, the discussions of stratified sampling and multi-stage cluster sampling.

Stratified sampling

stratified sampling Stratified sampling is a different but effective approach to probability sampling (Neuman 2006; Rosnow & Rosenthal 1996). The researcher firstly divides the population into subpopulations (also called substrata) on the basis of supplementary information on the population. After dividing the population into relevant strata, a random sample is drawn from each stratum – usually by means of systematic sampling. In doing so, the researcher controls the relative representation of each stratum in the final sample instead of allowing random processes to do it. The result is a more representative sample of the population than would normally be yielded by means of simple random sampling. However, prior knowledge of the population is required.



A simple example is a population of 10,000 that consists of 55% males and 45% females. If a sample of 1000 is required, a proportion of 550 (55%) will be drawn randomly from the male elements and 450 (45%) from the female elements. If a simple random sample is drawn, the composition of the final sample could deviate from the true gender ratio in the population.

Stratified sampling is, among other things, used when the researcher has reason to believe that a certain group (or stratum) of interest could differ from others with regard to the topic under investigation. The researcher could, for example, believe that the media behaviour of rural people differs from that of urban people. Stratified sampling is also used when certain strata are relatively small in number in comparison with others. When drawing a simple random sample, the smaller strata could be 'missed' altogether by the random procedures. For example, a researcher investigates the media attitudes of students of the Department of Communication Science at Unisa. However, as master's and doctoral students form only a very small portion of the total number of students, it could happen that no students from this stratum could be included in the final sample if simple random sampling is applied. Stratification could ensure a fair representation of this group.

Multi-stage cluster sampling

Multi-stage cluster sampling is a complex form of sampling usually employed for diverse and geographically dispersed populations where no sampling frame is available (Neuman 2006; Pitout & Du Plooy 2001). This form of sampling is mostly employed in large-scale surveys of a whole country or a particular region. As the term indicates, the sampling process involves various steps or stages.

Cluster sampling closely resembles stratified sampling in the sense that the population is divided into 'blocks' or 'clusters' in the first stages. The difference, however, is that the clusters are treated temporarily similar to elements in a population. A sampling frame of clusters – and not of individual elements – is therefore required. In national or regional surveys, the existing official division of the country into census enumerator areas is often employed for clustering purposes.

In a second stage, the clusters – which are still treated similar to individual elements – could be stratified according to relevant substrata.

multi-stage cluster sampling



In South Africa, clusters are usually stratified according to the various provinces as well as rural and urban areas. Racial group membership could furthermore serve as a basis for stratification. Once stratification of the clusters has been done, a certain number of clusters is drawn randomly – usually by means of systematic sampling – for each substratum, similar to the process of stratified sampling.

In the next step or stage, a number of households is systematically drawn within each of the clusters selected for the sample. The exact number of households to be drawn will depend on a number of factors such as the sample size, the number of households that can be comfortably completed by a single fieldworker or a team of fieldworkers during the survey period, and so forth. In the end, the households can be seen as smaller clusters within the larger clusters usually represented by census enumerator areas.

The last stage involves selection of one of the household members within each selected household. The question here is who to choose. The instinct of a fieldworker could be to select the first person he or she encounters, such as the person that opens the door. This could, however, lead to sample bias as certain household members are less likely to be at home during certain hours. Working people are, for example, not likely to be at home during daytime. A random procedure for withinhousehold sampling therefore needs to be followed. There are various techniques available to do so. In most South African surveys a selection grid is used. Household members that fulfil the sample requirements are listed from the eldest to the youngest (or the other way round) and a number is allocated to each one. A specific number (representing a household member) is then selected on the basis of the number allocated to the household according to the order in which households have been selected or visited in the larger cluster. Other techniques of simple random sampling can also be employed.

Cluster sampling is less expensive than simple random sampling for large and geographically dispersed populations. The reason is that the individual elements within a cluster are physically close to each other. If a simple random sample is drawn and the elements are widely dispersed over a large geographical area, it would be very expensive to travel to each selected respondent. Multi-stage cluster sampling is, however, less accurate than the other techniques of probability sampling. Sampling



error can occur in any of the different stages. It is nevertheless the sampling method used in most large-scale surveys. A complex method of multi-stage cluster sampling is, for example, followed in the AMPS studies conducted by SAARF (see section 10.11 – SAARF sa).

Overall, random sampling involves much more work than the techniques for non-probability sampling. A probability sample is, however, usually preferable above a non-probability sample because it enhances the reliability and validity of the findings (see section 10.9), it enables the researcher to use powerful statistical techniques in the analysis of the data and to generalise the findings to the population. The best possible sample for a particular study is usually well worth the time, cost and effort.

10.6.3 Non-probability sampling

Despite the advantages of probability sampling, some research studies are conducted in circumstances that do not allow for random selection of elements. In some studies probability sampling is also inappropriate. Suppose you want to study the information needs and media usage patterns of migrant workers. Probability sampling will not be possible as no list of migrant workers exists. Such cases call for non-probability sampling.

Quota sampling

Quota sampling is often used in audience and marketing research (Babbie 2001; Mytton 1999). Of all the non-probability sampling techniques, quota sampling shows the closest resemblance to probability sampling as it is modelled on the principles of stratified sampling. It is, however, not a true probability sampling technique. Quota sampling is used when interviewers need to look for fixed numbers of respondents of different categories or types. The variables involved are typically gender, age, housing type, social class, housing type or ethnicity. Quotas (the number of respondents to be interviewed for each category) are calculated for each category of each variable that are regarded as relevant for the particular study. One interviewer may, for example, be instructed to interview 10 male respondents in the age group 18 to 25 years that fall in the high socioeconomic group. Another may be given the task to interview 10 female respondents of the same age and socioeconomic group.

quota sampling



The sampling process usually starts with a matrix or table of variables important to the study, as well as information on the distribution of the population with regard to these variables. If a researcher wants to study the cellphone-usage behaviour of teenagers from 12 to 18 years old, there are various characteristics that could be important, such as socioeconomic background, age and gender. On the basis of the population proportions of each cell in the matrix representing these variables, a quota of the sample (a decision on the size of the sample should be made beforehand) is allocated to each cell of the matrix. That means that the researcher will need to know, for example, what percentage of the population is 12 years old, male and belongs to the higher socioeconomic group. Information on proportional distributions is necessary for all possible combinations of the variables at stake, that is, for all cells in the matrix.

Quota samples generally take less time and are cheaper than probability samples (Mytton 1999). Interviewers in surveys using quota sampling also complete many more interviews per day than when they need to seek specific individuals according to a random sampling plan. For these reasons, quota sampling is used in many commercial and marketing surveys. Interviewers are often seen with clipboards in shopping malls searching for respondents that fulfil their quota requirements. Often, when a respondent does not meet these requirements, the interview is politely ended.

Despite its resemblance to stratified sampling, there are several problems inherent to quota sampling. It is firstly, as already mentioned, difficult to obtain information on relative population proportions. It is therefore difficult to compile quota frames (the matrix with proportions for different cells) that represent a population correctly. The selection of sample elements could furthermore be biased even if correct information on cell proportions can be obtained. Fieldworkers instructed to find respondents fulfilling a complex set of characteristics could introduce bias in several ways. They could select and interview respondents that are easily available, such as family members, friends, people in their neighbourhood, church, and so forth. Fieldworkers will also tend to avoid people in places difficult to reach (for example, people living in high-rise buildings) and those living in remote areas. If those people that are more accessible differ with regard to the topic of the study from those living in difficult-to-reach places, the findings could be biased.



Interviewers can furthermore refrain from approaching people who look unfriendly or seem too busy. A further source of bias is the places where selections are made. Quota sampling done in shopping malls, for example, will be biased towards people who tend to spend more time in these malls.

There are, however, techniques that can be used to reduce the bias associated with quota sampling. The researcher can, for example, vary the places where respondents are sought. Some people can be selected at homes and others at places of work. In recent years, quota sampling has often been combined with probability sampling techniques such as cluster or systematic sampling. Fieldworkers are, for example, instructed to systematically visit homes in particular selected residential areas where they should attempt to find respondents fulfilling their quota requirements. The problem, however, is that it is difficult to estimate the representativeness of such hybrid sampling designs. At the same time, the logic of quota sampling can be applied to other research designs. When studying an organisation, representatives of different levels of management as well as employees can be interviewed, both men and women, and so forth.

Random-digit dialing (RDD)

The telephone has become an important tool in audience and marketing research. Random-digit dialing (RDD) is a sampling technique used for projects where the general public is interviewed by means of telephone interviews (see section 10.7.6 – Mytton 1999; Neuman 2006). Although telephone directories can serve as sampling frames, a number of groups are missed when telephone directories are used: people without telephones; people who have recently moved; people with unlisted numbers; people in hostels; people who only have cellphones; and so forth. People without telephones are perhaps the most important group in this regard because they represent the poor, the uneducated, the transient. These people would be difficult to reach in any telephone survey. However, in most developing countries ownership of mobile phones is more widespread than landline telephones.

When using RDD, the focus is on telephone numbers and not on directories. In South Africa, landline telephone numbers have three components: a three-digit area code (municipal area); a three-digit residential area code; and a four-digit number. Similar to landline

random-digit dialing (RDD)



numbers, cellphone numbers also have 10 digits. A computer program is used to randomly generate telephone numbers: landline numbers, cellphone numbers or both. A problem is that the program can select any number. This means that some selected numbers would be out of service, disconnected, payphones or numbers of businesses. Until an interviewer calls a particular number, it is not possible to know whether it really is what the researcher wants, namely the number of an individual or household. This can mean that a lot of time and money is spent on making calls to numbers that do not yield any interviews. Some research organisations make use of a computer program that dials the numbers automatically. However, it is still necessary for a human to listen to find out which numbers are individual or residential numbers.

As mentioned already, the sampling element in RDD is telephone numbers and not individuals. As in the case of multi-stage cluster sampling, interviewing in all cases the person who answers the telephone would give rise to sample bias. A secondary stage of within-household sampling is consequently necessary to select the person to be interviewed.

When landline numbers are drawn, it is possible to stratify the sample according to municipal areas and/or residential areas. Stratification is, however, generally not possible in the case of cellphone numbers. The most important reason why RDD cannot be regarded as a true probability sampling technique is, however, the fact that both landline and cellphone penetration is often low among the poor and in far-off rural areas – particularly in developing countries. That means that some degree of sample bias is almost inevitable.

Haphazard, accidental or convenience sampling: reliance on available subjects

Many survey studies rely on available subjects, such as stopping people visiting a shopping centre, at street corners or in other locations, interviewing members of clubs or other organisations, and so forth. University lecturers, for example, frequently request students enrolled in their classes to complete questionnaires (Babbie 2001; Neuman 2006).

The ease and inexpensiveness of such selection techniques explains their popularity. However, these techniques do not allow any control over the representativeness of the sample. It is consequently not possible to

convenience sampling



generalise the findings to the population and to establish the general value of the data. Available subjects could be selected to pre-test questionnaires. Although studies based on available subjects could yield valuable insights, care should be taken not to over-generalise the findings.

Purposive or judgmental sampling

Purposive or judgemental sampling implies that elements are selected on the basis of knowledge of the population and the aims of the study (Babbie 2001). For example, a researcher wants to study the role of opinion leaders in a society. It would be difficult to identify all opinion leaders in order to compile a sampling frame. However, the researcher could decide to focus on the most visible leaders voicing opinions during community meetings and through the media. In doing so, sufficient data for the purposes of the study could be collected. Some researchers might also be particularly interested in deviant cases that do not fit regular patterns of attitudes and behaviour. A researcher might, for example, be interested to study the reasons why some South African children watch television during school hours (see Chapter 12). In order to provide in-depth information on the phenomenon, it would be necessary to search for and select children involved in this form of deviant behaviour. Selecting such deviant cases represents a form of purposive sampling.

purposive sampling

Snowball sampling

Snowball sampling is appropriate when the members of a special population are difficult to locate: for example, homeless people, migrant workers or undocumented immigrants (Babbie 2001). In snowball sampling, the researcher first interviews those individuals that he or she is able to locate. These members are then requested to provide information on or to locate other members of the population that they happen to know. The term 'snowball' refers to the process of accumulating respondents as each located respondent suggests other potential respondents. The representativeness of such samples is, however, also questionable and the technique should only be used in exploratory studies.

snowball sampling

Volunteer sampling

Some researchers call for volunteers to participate in their research (Rosnow & Rosenthal 1996). For example, ethical guidelines at

volunteer sampling

447



universities sometimes prohibit lecturers from pressurising students to participate in research projects, and they therefore make an appeal on students to volunteer to participate in a study.

In recent years, researchers have come to know a great deal about the differences between typical volunteers and non-volunteers in research studies. Rosnow and Rosenthal (1996:204) name the following:

- Volunteers tend to have higher levels of education than non-volunteers.
- Volunteers tend to have a higher social status than non-volunteers.
- Volunteers tend to be more intelligent than non-volunteers.
- Volunteers tend to have a higher need for social approval than non-volunteers.
- Volunteers tend to be more arousal-seeking than is the case with non-volunteers.
- Volunteers tend to be more unconventional than non-volunteers.
- Women are more likely to volunteer than men.
- Volunteers tend to be less authoritarian than non-volunteers.

The differences between typical volunteers and non-volunteers imply that sampling bias is virtually inevitable when making use of volunteers. Apart from the general differences between volunteers and non-volunteers, it is furthermore difficult for a researcher to establish whether there are important differences between the two groups with regard to the topic of a study. For example, if a researcher makes use of volunteers to investigate students' attitudes towards soap operas, the volunteers might be predominantly soap-opera addicts, while students not really interested in soap operas could refrain from participating. If that is indeed the case, the use of volunteers can give rise to biased findings and conclusions. Of all the non-probability sampling techniques, volunteer sampling is probably the least desirable. It should be avoided if possible.

In conclusion, it can be said that non-probability sampling has its uses for certain research projects within particular contexts (Babbie 2001). However, researchers should at all times be aware of and acknowledge the shortcomings of non-probability sampling techniques especially with regard to the accurate and precise representation of the population.



10.7 TYPES OF QUESTIONNAIRE SURVEYS

Since survey research is used for a wide variety of purposes in a multitude of contexts, various types of interviewing may be used. Each of the interviewing types has particular advantages and disadvantages. The choice of a type of survey will depend on several factors, such as the aim of the study, the geographical distribution and characteristics of the population, the financial, human, technological and other resources available, and so forth. A sound knowledge of the advantages and disadvantages of each type is a prerequisite for making informed choices.

10.7.1 Self-administered surveys

In this type of survey, questionnaires are usually handed directly to respondents. The respondents have to read the instructions and complete the questionnaires by themselves – usually in their own time and at their convenience. They can also check personal records if necessary (Babbie 2001; Neuman 2006). The advantages of self-administered surveys are the following:

- Self-administered surveys are by far the cheapest and easiest way of conducting surveys. As self-administered surveys are also relatively simple to organise, they can be conducted by a single researcher. These considerations are usually important for poorly funded students and university lecturers.
- The response rates may also be high for well-educated populations who have a strong interest in the topic of a survey.
- Respondents might furthermore be more willing to respond to questions on sensitive issues in self-administered surveys than when they are interviewed.

Self-administered surveys hold, however, several important disadvantages:

Self-administered surveys are not really appropriate for populations where relatively large numbers of people are illiterate or semiliterate. In order to complete a questionnaire, a person needs to be able to read and comprehend both the instructions and questions. Open-ended questions furthermore require a person to express his or thoughts in words and to be able to write these down. If a person has not reached a sufficient level of literacy, his or her responses could be unreliable.

selfadministered surveys



- Self-administered surveys are not a good option for surveys with complex questionnaires.
- Where questionnaires are completed in private, the researcher has
 no control over the conditions under which questionnaires are
 completed, and does not know who really completed the questionnaire.
 A respondent could, for example, complete the questionnaire in a
 noisy bar where fellow drinkers may also render their inputs on the
 questions asked. A respondent can also assume a bogus identity in
 the questionnaire.
- Nobody is available to clarify questions or to provide assistance when needed.
- Poor questionnaire completion, where some questions are not answered, is another serious problem. Poorly completed questionnaires are regarded as a source of non-response error. Where completed questionnaires are picked up individually by the researcher or a fieldworker, the questionnaires can, however, be checked for full completion before they are received back.

10.7.2 Postal surveys

postal surveys

Postal surveys constitute a specific form of a self-administered survey. The questionnaires are posted to the respondents, who have to complete the questionnaires and return them to the researcher. A stamped and addressed envelope is usually included for posting the completed questionnaire back to the organisers (Babbie 2001; Neuman 2006; Van Vuuren & Maree 1999).

Postal surveys hold the same advantages as other forms of self-administered surveys. Particular advantages to postal surveys are the following:

- It is probably the cheapest form of surveying. For the price of an envelope and two stamps, a questionnaire can be sent to any address in the country.
- Questionnaires can be distributed over a wide geographical area.
- A postal survey is particularly convenient where an address list of the population is available.
- Postal surveys furthermore offer anonymity and avoid interviewer bias.

However, apart from the disadvantages associated with self-administered surveys in general (see section 10.7.1), a major problem with postal