AUSTRALIAN BUREAU OF STATISTICS
VICTORIA

CATALOGUE NUMBER 1201.2

MARCH 1985

INFORMATION PAPER

GUIDELINES FOR THE DEVELOPMENT OF COMMUNITY SURVEYS, VICTORIA

CONTENTS

1. INTRODUCTION

Surveys are a major source of information for government, business, and community organisations. The increasing need for data to assist in planning and in the allocation of resources has resulted in large numbers of surveys being conducted. However, the growing popularity of surveys has not always been matched by a general recognition of the number of technical problems involved in their conduct, with the result that survey data are often used for purposes inappropriate to the context in which they have been collected.

While it is possible to use a variety of techniques to conduct a survey it is important that the methods used meet the survey's objectives. It is also important that the users of the survey's results are aware of the limitations imposed on the data by the particular methods used to collect them, so that they are used appropriately.

Any observation or investigation of the facts about a situation may be called 'a survey'. However, a 'sample survey' is conducted when the purpose is to collect information from a small group of people, a sample, in order to make inferences about the larger group from which it was drawn. Whilst there are many situations where inferences about the larger group are not required to be made, this paper is only concerned with surveys from which the results can be generalised. In particular, it is concerned with sample surveys in communities such as local government areas or postcode areas.

The aim of this paper is to alert the reader to some of the more important areas of concern in the development and assessment of community-based sample surveys. It can be used by researchers who are contemplating the development and conduct of such surveys and also by policy analysts and decision-makers who commission, use, and make decisions with the assistance of such surveys. It is not intended to provide a detailed expose on survey development and operations but rather an overview of the major areas that need to be addressed by both community-based survey practitioners and users of the data from such surveys. Designing and conducting a sample survey is relatively expensive, and professional advice will help to ensure that usable results are obtained by the most efficient means. These notes then, are not to be used as a manual to survey development but rather as a guide when planning and assessing sample surveys.

ERLE BOURKE
Deputy Commonwealth Statistician

2. BASIC CONSIDERATIONS FOR COMMUNITY BASED SAMPLE SURVEYS

The major objective of community-based sample surveys is to make inferences about particular groups of interest - either the community as a whole, or subgroups of it. In order to achieve this objective it is important that the sample studied be representative of the group of interest and that each person's chance of selection in the sample is known. For example, a sample of people chosen at a shopping centre or library is not representative of the resident population in a community because only certain types of people are likely to visit such places. Those people who never (or rarely) visit the shopping centre or library have no (or little) chance of selection in the sample, so that the results of a survey obtained in this way cannot be generalised to the resident population. It is therefore recommended that in order to be able to generalise results, a community survey be based on a person's place of residence. This methodology is referred to as a household survey.

The extent to which the results can be used to make inferences about the particular group of interest is known as the accuracy of the survey. In practice it is rarely possible to conduct a survey which strictly adheres to the principles of sample survey methodology, because of the resource constraints imposed.

It is therefore important that the survey taker is aware of the consequences of the particular methods used and the resultant sources of error, so that the overall accuracy of the data is the best achievable. It is also important that the researcher can anticipate when the resource constraints are so severe that the accuracy of the data will be unacceptable for the use for which they are intended.

3. STRATEGIC ISSUES IN SURVEY DEVELOPMENT

This section covers some of the more common strategic aspects of the development of small surveys, taking account of issues discussed in the previous sections. Similar problems are encountered in all surveys although every situation requires a slightly different approach. Therefore it is advisable to seek consultation from experts to ensure that the results will be as valid as possible within the project constraints.

Survey development is an iterative process - each step depends on others so that a decision about one aspect of the survey will influence other decisions and activities. Thus, although the planning stages are presented here in an ordered sequence, it is necessary throughout the development stages to reassess steps already specified.

3.1 Defining objectives

A survey is usually commissioned to provide some information about the frequency of occurrence of particular events, to investigate needs or opinions, and to assist in the determination of priorities in planning and decision making. The first and most crucial step in the development process is to identify in specific terms the objectives of the survey and to determine exactly what information is required from the survey. Questions such as

- What data is required?
- Who needs to be asked to provide the data (ie. who to sample)?
- Who will use the data?
- How will it be used?

are necessary in defining realistic and useful objectives.

3.2 Supplementary information

It is important to check whether work on the same topic has been done previously. Assessment of techniques and questionnaires from similar studies is useful and can reduce development time and problems.

If supplementary information (eg. census data, surveys, administrative information, etc.) is available and useful it is wise to ensure that as far as possible, concepts, data items, classifications etc., to be used in this survey are as comparable as possible with those used in the supplementary information so that as much information as possible can be used.

3.3 Determining the accuracy required and expected

The required level of accuracy for each of the specified objectives needs to be determined. Consideration of both sampling and non-sampling errors is involved (refer Section 4).

Specification of sampling errors requires nominating the precision necessary in the major variables to ensure that the results are sufficiently reliable to be used in decision-making. This can often be regarded as the precision required in the evidence or data to enable a choice to be made between possible responses to the problem under investigation. For example, does a survey which is investigating the demand for a pre-school and which estimates that somewhere between 50 and 150 children in an area are in need of this facility provide sufficient justification? That is, while the number could be as high as 150, it could also be as low as 50. It may be necessary to be more precise, for example, to restrict the range of the result to $\pm$ 20.

At the same time, consideration of non-sampling errors ensures that balanced attention is given to all sources of error to achieve an overall accuracy in the data sufficient for it to be useful in the context for which it is being collected. In the above example of the pre-school survey, the description of the term 'pre-school' needs to be provided in sufficient detail to enable respondents to give a meaningful opinion. The accuracy gained by including a specific detail can be balanced against the complexity of providing that detail by considering the effect on the total error in the survey e.g. will responses to possible fees of $8 per day or $10 per day be sufficiently different to enable a choice to be made between these alternatives, considering the other errors in the survey?

3.4 Defining the population

The group to which the survey is to relate is defined as the target population. The list, index, set of maps, or other records from which the sample is to be selected is called the sampling frame. Ideally the sampling frame should separately identify each unit in the target population without duplication. However, this is often not possible. Methods are available to adjust for some known deficiencies in the sampling frame.

A deficient sampling frame is a source of non-sampling error and the extent of the deficiency should be considered relative to all other sources of error in the project (refer Section 4) as well as to the objectives of the survey.

There are several alternative sampling frames for community surveys eg. lists of dwellings, lists of ratepayers, electoral rolls, telephone directories. Choice depends on the extent of bias likely to result from each sampling frame and the selected method of collection.

3.5 Determining the required sample size

The sample size required for a survey depends on several factors :

- the required accuracy of the results;
- the level of detail required in the results ie. the extent of disaggregation and cross-classification;
- the variability of the characteristic being measured;
- the proportion of the population estimated to have the characteristic being measured.

Surprisingly, the required sample size has little to do with the size of the population being studied except for very small populations. Hence it is not usually relevant to express the sample size as a percentage of the population but rather as an absolute number. It should be noted that one of the main factors in determining sample size is the size of the sub-groups being analysed, rather than the total sample size of the survey. It is necessary to have sufficient sample size within each subgroup analysed to ensure the reliability of the data at the level of interest.

A good approach is to determine which characteristics need to be measured and how they will need to be disagreggated and cross-classified. A blank shell of each table (cross-classification) can be drawn and the level of detail (or extent of disaggreggation) required for the characteristics of interest determined. Clearly the greater the number of cells required in a table, the larger the necessary sample size (and the greater the cost).

The actual number of selections required must be also adjusted to allow for the expected level of non-response. It is, however extremely important to recognise that the characteristics of non-respondents are likely to be different from those of respondents. Therefore boosting initial sample numbers to ensure the number of final responses are sufficient will not give totally reliable results and could produce quite misleading results.

In practice, the sample size actually used is a compromise between the size determined from the above factors and the resources available. However, if there are insufficient resources to satisfy the objectives then the objectives will need to be modified and, if necessary, consideration be given to cancelling the survey. A case study may satisfy more modest objectives and provide sufficient information. A case study may also be used to supplement a survey by providing more detailed information on a subset of the sample. However, it is not possible to make inferences about the target population from a case study and if this is a major requirement there is possibly no value in undertaking the exercise.

3.6 The method of collection

The main methods of collection from household surveys are mail, telephone, delivered questionnaires, and personal interview at the household.

3.6.1 Mail collection

Although they are cheap, mail collections have several disadvantages which usually render them unsuitable as a method of collection for a community survey. Some of these are :

- mail collections generally have low response rates (around 20-40 per cent);
- reply-paid envelopes need to be supplied to improve the response rate;
- the questionnaire must be extremely simple in order to reduce errors caused by misunderstanding or confusion;
- people with limited English or literacy will tend to be non-respondents.

Response rates can be improved by initiating personal follow up of non respondents but very little can be done to increase the level of complexity of the survey.

3.6.2 Telephone collection

The use of telephones for the conduct of surveys has become increasingly popular in recent years, principally because of cost considerations. However, the use of telephone interviewing techniques does have some problems.

The group of people that do not own or do not have access to a telephone are not represented. This group could be quite different in a number of important characteristics from those accessible by telephone. The likely effect on the data of omitting this group requires evaluation for particular topics. The exclusion of non-owners of phones may be critical for example, in a survey of users of welfare services.

Identification of the target population in a sampling frame is difficult. Selection of telephone numbers in a particular defined area is not easy. The method of allocating telephone numbers, the ordered nature of telephone directories, and the presence of business listings present further problems.

Questionnaires used for telephone interviewing need to be developed to achieve reliable and valid results for this particular method of collecting statistics, and interviewers must be trained in the techniques required for successful telephone interviewing.

Again, it is important that the proportion of selected people who respond to the survey be sufficient to provide a reasonably representative coverage of the population being studied.

In summary, telephone interviewing may be a satisfactory and cost effective alternative to face-to-face interviewing in some situations, while in other situations it may lead to results which are not representative of the population of interest.

3.6.3 Delivered questionnaire

As a more expensive alternative to a mail collection, the questionnaire can be delivered and/or collected by an interviewer. This method provides improved response rates compared with a mail collection and is particularly effective when information is required from several members of a household, or where records need to be checked by respondents.

Disadvantages include the cost, the questionnaire can still only be simple, and response rates may still not be adequate.

3.6.4 Personal interviewing

Personal interviewing at the household is generally the preferred method of collection because many of the non-sampling errors can be better controlled. However, it is expensive. Within the household, interviewing can be directed at each person, at any responsible adult, or at nominated individuals (one or more), depending on the content of the survey.

The main advantages of personal interviewing are in the improved quality of the data because non-response is minimised, more complicated procedures and questionnaires can be used and the trained interviewer can make observations (eg. dwelling structure).

The method can also be used in conjunction with other methods, for example to interview non-respondents from a mail survey.

3.7 The questionnaire

Questionnaire design is an extremely important phase of survey development. It requires time, patience, research, and testing to develop a useful questionnaire.

The questionnaire needs to

- identify the person(s) to be interviewed (eg. any person in each selected household, any male aged 18-24 in each selected household); and
- eliminate those not to be interviewed (eg. visitors not required).

As a general rule, there should be a separate question (or series of questions) to obtain information on each data item. It is necessary to clearly distinguish between the 'need to know' and 'nice to know' data. Avoid including the 'nice to know' unless the added cost of collecting them is small and justified.

Each question must be clear, concise, unambiguous, logical, free from jargon, and be relevant to the previously defined objectives. Any questions or topics which do not directly address the topics of interest should be dropped.

3.8 Pilot testing

Some form of pilot testing is essential to test the questionnaire and collection procedures. Pilot tests can also give an indication of incidence and response rates. This can help to refine the sample size and collection procedures. Even for a small survey, a small pilot test in a neighbouring area will assist in identifying problems before the survey commences. A pilot test may even indicate the futility of attempting a particular collection, for example, because the event or characteristic to be measured occurs infrequently, or because concepts cannot be unambiguously translated into questions, or because of people's unwillingness to answer questions on particular topics.

3.9 Training and supervision

If the method of collection is interviewing (personal or telephone), all interviewers need to be trained to ensure uniformity of collection procedures and to assist interviewers to become familiar with the questionnaire. Interviewers should also be monitored once interviewing has started to minimise errors and misunderstandings.

For similar reasons, in mail surveys someone must be trained to handle enquiries from respondents and no matter what collection method is used, people who will edit and process the survey returns also need to be trained.

3.10 Choice of processing method

The mode of processing survey responses must be decided at an early stage so that the questionnaire can be designed to facilitate processing. The costs of processing will also need to be considered in the overall budget and hence are a consideration at the time of determining sample size.

For small surveys manual tabulation is often the simplest and most efficient means of processing, especially if the results required do not involve detailed or complex cross-tabulations or data manipulation. Computer processing often requires too many overheads for a small one-off survey.

4. SOURCES OF ERROR IN A SAMPLE SURVEY

To assess the accuracy of the data (refer Section 2), it is necessary to consider the size of the various errors which are associated with the collection of that data. Errors in sample surveys are usually categorised into sampling and non-sampling errors.

Sampling errors are those errors caused by only surveying a sample of the group(s) of interest to make inferences about the total group(s) of interest. The size of sampling errors is usually expressed as a standard error or confidence interval and is a measure of the precision of the data. These measures can be estimated using mathematical formulae if the sample is selected and collected according to the principles outlined in this document.

Non-sampling errors are those errors which can occur during the conduct of a collection, regardless of whether a sample or complete enumeration of the group(s) of interest is undertaken. Some common sources of non-sampling errors are :

- failure to properly identify the group(s) of interest from which to select the sample;
- non-response (failure to obtain complete responses from initially selected units);
- poor questionnaire design;
- non-uniformity in interviewing standards;
- time biases (eg. seasonal influences, memory errors);
- processing errors.

Insufficient control of non-sampling errors can completely invalidate a survey's results.

While non-response is only one source of non-sampling error, it is usually an issue of great concern in community surveys because the reduction of non-response is a major cost. All collections can be expected to have some non-response. However, if the collection is a survey in which inferences are to be made about a specific group of interest it is important that non-response is minimised because respondents are generally different from non-respondents. The extent to which they are likely to be different in the characteristics being measured is important in assessing the validity of making inferences about the whole group of interest from the characteristics of respondents. Users of survey data, therefore, need to be aware of response rates so that potential biases from non-response can be considered when using the data. Survey takers on the other hand must ensure that the highest possible response rates are achieved.

It is emphasised however, that balanced attention to sources of both sampling and non-sampling errors is essential to minimise total errors.

5. RESULTS AND REPORTING

5.1 Editing and verification

All completed questionnaires need to be edited to check for validity and consistency of responses. Time spent on editing before or during processing is well spent as it reduces the number of inconsistencies that are detected and require resolution at the analysis and report stages.

Supplementary data from other sources or surveys can be used if available to assist in identifying possible omissions and errors, to verify the survey results, as well as to enable comparison with other available statistics (refer Section 4.2). Any errors or biases detected can be examined for their possible impact on the survey results. Often some effects of non-response become apparent at this stage when inconsistencies are identified.

5.2 Making inferences about the target group

Providing adequate selection and collection procedures have been followed, each selected unit in the sample is representative of other non-selected units in the population. The expansion of sample results to obtain population estimates is done by multiplying the results of sample units by the number of units they represent in the population. This is termed **weighting**. The results can then be tabulated according to the cross-classifications previously described (refer Section 14.5).

5.3 Presenting and assessing the results

Generally the results of the survey should be presented in a written report. To enable users to assess the adequacy of the survey results the following should be given :

- scope, ie. the information content and target population (eg. community needs of all persons in area);
- coverage, ie. details of the sample including any deficiencies (eg. a sample of 500 private households in an area was chosen and personal interviews were conducted with all persons aged 15 years and over. Adult household members responded on behalf of persons aged less than 15 years. Persons living in non-private dwellings (hotels, motels, nursing homes, etc.) were excluded);
- definitions of (at least) major data items;
- sampling error analysis. It is necessary to present estimates of sampling errors as indicated in Section 4;
- an explanation of how to use and interpret the sampling errors given;
- any potential sources of non-sampling error and their likely impact;
- a copy of the questionnaire.

Users of survey data should assess the accuracy and relevance of the data before applying the results to specific issues. Failure to do so may lead to decisions based on inapplicable or unreliable data.

6. TENDERING OUT SAMPLE SURVEYS

6.1 The tendering process

In calling for tenders, the survey organiser should be able to give some indication of the following:

- population with which the study is concerned;
- information required and the time period to which the information is to refer. Such things as the output, tabulation, and accuracy required should be detailed as much as possible;
- general methodology to be employed, covering
  - method of enumeration and measurement
  - type of questionnaire
  - sampling methods
  - any special requirements concerning operational procedures, interviews, etc;
- timing of the survey operations and development and presentation of results and reports;
- analysis and reporting required;
- costs ie. letting the consultant know how much money there is to spend.

The informational needs of the survey should also be presented, specifically to limit the boundaries of survey design and to control the common tendency to make one survey do more than what it is originally designed for.

Once the consultants are clear on the aims of the survey and the survey organiser's requirements, they should be able to submit detailed proposals on the following points which should be requested by the survey organiser:

- development of the survey - pilot testing, questionnaire development, development of survey methodology.
- sample design - this may be brief and subject to revision, but should deal with sample size and anticipated errors, selection method, stratification, coverage, estimation, handling of non-response and non-contact;
- survey operation, dealing with the type of questionnaire and enumeration, interviewer recruitment, training and supervision, monitoring and checking of the operations, evaluation and discussion of non-sampling errors;
- processing, analysis and reporting;
- costing and timing;
- the extent and nature of intended sub-contracting.

With such proposals, a rational choice of consultants can be made, choosing the consultant most likely to fulfil the aims at the least cost or best fulfil the aim for a given cost (the Public Service Board criteria is 'value for money').

For anyone seriously intending to tender out a survey to a private consultant it would be wise, at this and the contract stage, to be conversant with the following two documents:

- 'Market Research Society Code of Professional Behaviour' available from the Market Research Society of Australia; and
- 'Survey Guidelines' published by the New South Wales Privacy Committee to aid in the conduct of surveys and research.

6.2 The contract

Once it has been decided which consultant to hire, a contract should be drawn up. Guidelines for drawing up a contract are contained in the Australian Public Service Board guideline 'Engagement of Consultants and Contracts for Service', PSB May 1978. The contract should be designed to protect the rights of the survey organiser and the consultant and define the responsibilities of one with respect to the other.

The contract should cover the following points:

- the population to be surveyed;
- the aims of the survey;
- the information required regarding the population and the time period to which the information is to refer;
- sample design and survey operations;
- development stages; skirmishes, pilot tests, etc;
- timing of the different phases of the surveys;
- analysis and reporting to be performed;
- costing;
- the rights of the survey organiser to inspect the work at any stage;
- what phases of the survey need specific approval before they can be carried out;
- procedures for the termination of the contract by the consultant and the survey organiser;
- payment procedures and the handling of rising costs;
- ownership of the data as collected and after processing;
- publicity and confidentiality;
- final report, estimates, accuracy of estimates, analysis, and comment required.

It should be clear in the contract how changes in the specifications by either the consultant or the survey organiser are to be handled.

The contract should not be so detailed that the natural development of the survey is hindered. However, the interests of the survey organiser must be safeguarded. This can perhaps be achieved by making various parts of the survey subject to the survey organiser's approval.

7. THE ABS CONSULTATION SERVICE

The ABS has a small Statistical Services section in each of its Offices, their function being to provide, as resources permit, a free specialist service to Federal, State and Local government departments and authorities, research and teaching institutions, and other non-profit organisations. This specialist service involves the provision of advice and assistance on a wide range of mathematical and statistical problems, including technical and practical advice on sample survey development, design and conduct. As such, this service complements the many other specialist statistical services which the ABS offers.

Many requests for the assistance of the consultation group concern the development of surveys to provide information to identify community needs for planning purposes. These surveys usually cover small geographic areas such as suburbs, postcode areas, and local government areas and assistance is provided by advising on questionnaire design, methods of selecting a sample, collection and processing procedures.

Those who qualify for assistance are encouraged to contact

The Supervisor
Statistical Services Section
Australian Bureau of Statistics
Commonwealth Bank building
2 Elizabeth Street — Box 2796Y GPO
MELBOURNE VIC 3000 — MELBOURNE VIC 3001
Telephone: (03) 652 6185

8. CONCLUSION

As indicated in Section 1, survey taking is a complex process. This paper is not meant as a manual to enable researchers to develop and conduct their own surveys without any assistance. Rather it is hoped that these notes will indicate the complexities involved in developing and conducting a survey by identifying a range of important issues. The list of issues is by no means complete and there has not been an attempt to propose solutions to identified problems. The notes are intended to promote a scientific approach and to increase both intending survey-takers' and survey data users' awareness of the complexity of surveys of communities. For those organisations who qualify for assistance, the ABS' consultation service (refer Section 7) is an approach which can be taken to improve the survey's chance of success.