Republic of Botswana

GUIDELINES FOR DESIGNING AND IMPLEMENTING HOUSEHOLD SURVEYS IN BOTSWANA

February 2010

CSO
CENTRAL STATISTICS OFFICE

GUIDELINES FOR DESIGNING AND IMPLEMENTING HOUSEHOLD SURVEYS IN BOTSWANA

Published by
Central Statistics Office
Private Bag 0024, Gaborone
Telephone: 3671300 Fax: 3952201
E-mail: csobots@gov.bw
Website: www.cso.gov.bw

Contact Unit: Surveys
Telephone: 3671300

February 2010

PREFACE

This survey methodological monograph entitled "G**uidelines for designing and implementing household surveys in Botswana"** prepared by CSO is designed to assist various Government Departments and Researchers in the formulation and implementation of small scale and ad-hoc surveys.

The CSO has been requested on various occasions to provide comments and guidance in the design and execution of surveys focusing on a variety of demographic and socio-economic issues. The CSO continues to provide such technical service. However, it is noted that in some cases the survey instruments and the scope of documents forwarded to the CSO for comments have either been limited or the statistical activities envisaged often follow undesirable sequence of operations. Considerable variations in methods and procedures have also been noted from one survey to another. Therefore, in order to ensure the standardization of methods and also provide statistical guidelines, the CSO has prepared a check-list of survey activities.

It will be appreciated that this guideline refers largely to the more quantitative and statistical surveys rather than the sociological, anthropological and attitudinal studies.

The contribution of the Research and Survey Methods team comprising of Dr. V. K. Dwivedi and Mrs. Thapelo Sediadie is highly appreciated.

A. N. Majelantle,
Government Statistician
February 2010

CONTENTS

1. BACKGROUND

The Central Statistics Office is the principal data collecting, processing and disseminating agency responsible for coordinating, monitoring and supervising the National Statistical System. It thus has the statutory mandate to produce and provide the Government, the Private Sectors, NGOs, Parastatal Organisations, International Organisations, the Civil Society and the general public with statistical information for decision-making, policy formulation and planning purposes. The statutory mandate also includes the responsibility of providing advisory and technical service to all users on statistical matters.

Being a perennial source of data for policy formulation, monitoring and evaluation, household surveys occupied a central role in the Central Statistics Office activity programme over the years. Household surveys, together with censuses, constitute the major sources of data in the country and thus enable CSO to provide socio-economic and demographic data on a continuous basis.

Household surveys became a programmed feature of the Central Statistics Office (CSO) work programme in 1983. The institution of a programme of household surveys was a national response to the United Nations global household survey programme whose aim was to develop sustainable household survey capability in developing countries. Prior to this development, household surveys in the CSO were conducted on an ad-hoc basis, with little or no conscious effort for capability and capacity building.

Household Surveys Unit was established within the general philosophy of the African Household Survey Capability Programme (AHSCP). The Unit, started operations on a continuous basis in 1983, as opposed to before when household surveys were planned and conducted on an *ad hoc* basis. Household Surveys Conducted During Inter Censual Period (1981-2010) is given in Appendix-I

The CSO has been requested on various occasions to provide guidance in the design and execution of surveys focusing on a variety of demographic and socio-economic issues. In order to ensure the standardization of survey methodology and also provide statistical guidelines, the CSO has prepared a check-list of survey activities.

2. INTRODUCTION

The planning of the survey is an extremely important task, since the quality of the survey results depends considerably on the preparation made before the survey is conducted. The amount of planning needed varies greatly with the type of material available and the nature of the information sought.

It is recognized that the design and execution of household surveys involves a wide range of technical and operational activities including:

i. Budget Estimate & Resource Mobilization,
ii. The formulation of the objectives of the survey,
iii. The specification of the output, the selection of variables,
iv. The design of questionnaires,
v. Pre-testing and final designing of questionnaires and the preparation of manuals and instructions,
vi. Gazetting of the questionnaires,

vii. The design of sample, the recruitment and training of personnel,
viii. Tabulation plan,
ix. Recruitment and training of field staff,
x. The collection of data,
xi. The supervision of data collection,
xii. Quality control,
xiii. The processing of data (manual editing & coding, online editing)
xiv. Tabulation and production of preliminary tables
xv. The analysis and interpretation of results
xvi. Report writing, and
xvii. Dissemination of the results (printing report, website, CD rom, dissemination seminar, etc)

It is generally not possible to give rules of thumb regarding the actions to be taken at the different stages of the survey, since these would depend very much on (i) the data requirements, (ii) available resources and (iii) the operational conditions of individual surveys.

Flow chart of surveys activities for the design and execution of household surveys is given in **Appendix-I.**

The important aspects requiring attention at planning and operating stages of the sample survey are discussed in sequence in the following sections.

3. BUDGETING

Preparation of a preliminary budget estimate is a priority activity that should be planned and executed at an early stage. The draft budget will be based on the assumption about (i) the number of household (or other elements of the population as the case may be) to be covered, and (ii) the time needed to interview a household.

The major items for budget estimate include:

i. Salaries (depends upon the number of field staff (i) Enumerators (ii) Supervisors), [Estimation of number of enumerators required to complete the enumeration of total number of households in the survey period is given in **Appendix-I]**
ii. Transport cost,
iii. Training,
iv. Equipments and supplies,
v. Fuel, maintenance and insurance,
vi. Printing of questionnaires and manuals
vii. Data collection
viii. Data processing software
ix. Publicity cost
x. Consulting fee (data analysis and report writing)
xi. Dissemination seminar
xii. Miscellaneous

This will be an estimate since the details of some costs will be unknown and as such will be examined or revised as the survey activities progress. Once the budget is drafted, the funds required

should be found.

Remark: (a) A good way to start a budget is to look at the budget of similar surveys already done in the country.(b) Contingency costs should be added.

For the brief discussion on the cost components and the suggested format, refer **Appendix-III.**

4. OBJECTIVES OF THE SURVEY

The first task when designing a household survey is to lay down its objectives and put them down in writing. The objectives should be precise, for example; it is not enough to say that the survey is intended to find out about the living conditions of old people. One should exactly define what is meant by the living conditions and what is meant by old person. Failure to think of the survey objectives fully and precisely must inevitably undermine its ultimate value. Once the objectives are settled, the plan of the survey is directed to achieving them with the required accuracy and within the given resources and time period.

The objectives of the survey should define (i) the population the survey intends to cover, (ii) data relevant to the purpose of the survey to be collected, and (iii) Geographic coverage.

5. JUSTIFICATION

The justification of the survey should be established in close consultation with planners and the major data user (stakeholders) of the survey.

i. If it is relevant to collect the existing basic informative data related to survey.

ii. Review existing data and literature with respect to

- Data availability (gaps and omission if any);
- The extent to which existing data has been utilized and its major constraints.

iii. The extent to which the Output of the survey overcomes the above constraints.

iv. Relationship with other studies.

6. VARIABLES OF STUDY

Plan the Subject- Matter Content of the survey, for instance

1. Select the variables (topics) to be studied by the survey.
2. Consider the concepts to be measured, observed or reported and their theoretical basis if any. The concepts should be as simple and unambiguous as possible and meaningful and understandable to the population being studied.
3. Consider the level of detail and depth of information required.

7. FREQUENCY OF DATA COLLECTION

- Determine the frequency of data collection by sampling units (village, households, etc.). From a study conducted in one season it is not possible to make generalization about events taking place throughout the years (if seasonal variation exists). For example, to minimize any seasonality variation the data collection for Household Income and Expenditure Survey (HIES) and Botswana Labour Force Survey are for 12 months period.

- Determine the method of interview and data collection (whether personal visits or mail questionnaire etc.).
- Determine the general approach to be used in covering the population i.e. whether de-jure1 or de-facto2 approach to be used)

8. OUTPUT DESIGN

Specify the output of the survey.

8.1 Outline the Tabulation Plan:

(a) List the tables required (keeping publications in mind!).

(b) Indicate the table formats.

(c) Give the main classificatory variables.

For example in most of the national household surveys the main classificatory variables considered are

i. –**Relationship:**
ii. –**Sex**: Male and Female
iii. –**Age-groups**: Age groups at interval of 5 years each (viz. <1, 1-4, 5-9, 10-14,…)
iv. –**Place of Residence**: (i) Cities/Towns, (ii) Urban Villages and (iii) Rural
v. –**Education**: (i) Pre-Primary, (ii) Primary (ii) Secondary, and (iii) Tertiary
vi. **–Training**:
vii. **–Economic Activities:** As per 2001 Census Categories
viii. –**Marital Status**: (i) Never Married, (ii) Married (iii) Living Together (iv) Separated (v) Divorced, and (vi) Widowed
ix. **–Geographical Area:** Administrative District/Sub-District
x. **–Parents Survival**
xi. **–Place of birth**
xii. **–Language**
xiii. **–Religion**

Extra classificatory variables may be added as and when needed.

8.2 Determine the Levels of Estimation3 i.e. give list of tables by:

-Stratum
-Cities
-District
-Region
-Country

8.3 Specify the types of tabulation and statistical analysis intended, i.e. whether

-Totals
-Measures of location (means, median, mode, etc.)
-Measures of dispersion (variance, standard error, coefficient of variation, range)
-Frequency distributions
-Associations (correlation, regression); multivariate analysis

1 De-jure: Under a de-jure approach, persons are counted in accordance with their usual place of residence.
2 De-facto: Under a de-facto approach persons are counted at the place they are staying at time of enumeration.
3 The level of estimation is also essential for the determination of the sample size and its allocation.

-Trend and time-series analysis etc.
-Multivariate analysis

9. QUESTIONNAIRE DESIGN

The information collected in an interview survey ultimately depends on what questions are asked and how they are asked. One of the most difficult problems in survey preparation is the translation of complex ideas and concepts into usable questions for survey purposes. As is often the case in translating from one language to another, something is frequently lost in the process.

There is a great deal of experience in questionnaire design and most of the specific issues and problems are well known. Nevertheless, questionnaire design is more of an art than a science. One has to try to optimise the requirements of three categories of audiences (i.e. respondents, the enumerators, and the data processors).

The way in which a questionnaire is designed will have a crucial influence on the accuracy of the responses obtained.

10. PRETEST AND PILOTING

No household survey questionnaire should be finalized without being tried out on a small number of households. Pretest involves administering a questionnaire to selected few households or trying out selected sections of the questionnaire.

Piloting is a comprehensive field test of a draft questionnaire which covers more households. The households should belong to several households that represent the population of interest. Example; for surveys which intends to cover both urban and rural, the pilot should be done in different parts of the country so that it covers people very much similar like those to be sampled.

Pretests and pilot test serve many purposes but the main ones are given below;

(i) To check the adequacy of the questionnaire. This is probably the most valuable function of the pilot and pretest survey.

(ii) To test the editing and tabulation programs.

(iii) To check the efficiency of the survey procedures i.e. manuals, instructions

(iv) To mark the time to complete a questionnaire, thus determine data collection period.

All members of the survey team should be involved in the pilot testing process. Usually pretest is conducted by a team which designed the questionnaire and the field supervisors. After the pilot testing or pretest, the survey team should hold meetings to report their findings and revise the questionnaires and survey procedures.

11. SAMPLE DESIGN

A sample design is a set of rules or procedures that specify how a sample (subset of the population) is to be selected. The design of samples is always subject to administrative constraints and the type of auxiliary information available. The considerations involved for sample designs are:

(i) Define the sample domain i.e. area and scope of the sample.

(ii) Define the sampling units (villages, Enumeration Areas, localities, Lolwapa/ Compound, Dwellings, Holdings, Households, and Individuals). It depends on the kind of study undertaken.

(iii) Determine the sampling procedures*.

*****Remark:*** *A variety of sampling procedures exists (including simple random sampling, cluster sampling, stratified sampling, area sampling, multistage sampling, etc.) which can be considered in planning surveys. For example in the western region of Botswana, where the population is sparse and traveling conditions are difficult, some degree of clustering is reasonable.*

(iv) Determine the sample size.

(v) Construct the Sampling Frame.

(vi) Allocate the sample by sampling units.

(vii) Outline the Estimation Procedures for the estimate required and for the computation of sampling errors (the estimation procedures are linked with the sampling procedures adopted).

12. WORK-PLAN

Drawing up a realistic plan of all the survey activities from the planning stage until the report preparation plays an important role in close supervision of survey activities to survey manager.

12.1 Survey Activities

(i) List the and determine their timings

(ii) Schedule of preparatory activities

(iii) Appointment of staff (if any)

(iv) Training of field personnel

(v) Training of office personnel

(vi) Schedule of data collection

(vii) Schedule of data processing

(viii) Schedule of report preparation and dissemination of survey results.

If some activities are delayed because of the reasons beyond control, it will affect other activities too. Thus it is recommended that work-plan should be updated from time to time and remedial measures to be noted for their implementation.

The format of the surveys activities and their timing is depicted in **Appendix-IV.**

12.2 Manpower

Establish the minimum manpower4 requirements for the survey

4 No matter how limited in size and scope, a good household survey would require minimum 'nuclear staff' of (i) survey manage (ii) subject matter specialist (iii) sampling statistician (iii) data processing specialist (iii) assistants in office and field work

Manpower requirements should include the number of:

(i) Professional and technical staff

(ii) Field personnel (a) enumerators (b) supervisors (c) drivers

(iii) Data processing personnel (a) editors (b) coders (c) data entry operators

12.3 Administrative Structure

Establish the administrative structure of the survey

(i) Organizational Chart

  (a) Linking the survey programme to a specific Government department or agency

  (b) Linking the various snits (sections) created for the design and execution of the survey; viz (i) Survey operational section, and (ii) analysis and planning section

(ii) Manpower allocation

(iii) Job description and line of responsibilities

13. RECRUITMENT AND TRAINING OF FIELD STAFF, CODERS AND DATA OPERATORS

The quality of household survey depends to a significant extent on the quality of the field staff as well as coders and data entry staff. Training for this group should be adequate. Usually the decisions on whether to employ temporary staff or permanent staff depend on the periodicity of the survey. If it is a continuous survey; permanent field staffs are usually considered e.g. annual agric surveys. Surveys which are conducted after a number of years; say 5 or 10 years employ some temporary staff.

The period of training will depend on numerous factors including the number of trainees, complexity of the survey as well as the work experience of the trainees and length of the questionnaires. Usually those who gained experience for the same survey will not require more time for training, whereas those who are new will need a thorough training so that the data obtained is of good quality.

It is important to train more than required enumerators and data entry staff and administer some kind of a test to trainees as well as practical interviews involving real household. The results of the test can then be used to select enumerators, coders and supervisors.

Remark: The quality of training will affect the quality of the survey and ultimately the quality of data collection.

- Conduct Pilot Exercise: This is a obligatory particularly for relatively larger surveys.

(Pilot surveys are recommended as standard practice for pre-testing the sample design, questionnaire and execution of the study.)

14. FIELD WORK PERIOD OR SURVEY PERIOD

The actual work of going out to the areas being sampled and interviewing the sampled households is typically referred to as the fieldwork. Fieldwork should start as soon as possible after training in order to minimize any forgetting of what was learned in the training.

All survey materials such as questionnaires, forms, tents etc. should be made available on time so

that interviewers are not hampered by lack of materials.

A detailed plan must be drawn up that matches the household that has been selected with the survey team. The plan needs to be realistic and it should be drawn based on the past experience. The field work should follow the work plan drawn. It is important to prepare a time table indicating a sequence and estimated duration of various operations of the survey and also to keep track of the sequence and summary returns of the data collection process.

Field staffs are usually organized in teams led by supervisors and each team is assigned a portion of the sample to cover. The supervisor is responsible for ensuring that households in his/her assigned portion are interviewed. Adequate transportation should be provided for each field team.

15. DATA COLLECTION

1. Consider the Technical Aspects of data collection
   - Prepare instructions and manuals for:

     -Sample selection,

     -Completing the questionnaires (for enumerators)

     -Supervision of field work (for supervisors)

     -Ttraining of field personnel (including demonstration and testing materials).
   - Prepare procedures for the supervision of data collection and quality control: The supervisor should check sample selection, spot checking of enumerators, re-interview a small sub sample of respondents for cross checking etc.

2. Consider the Operational Aspect of data collection.
   - Establish the criterion for the selection of field personnel (for enumerators, for supervisors, for selected field personnel).
   - Plan for the training of field personnel (arrange for training materials, duration of training, etc.).
   - Consider organization

     -Distribute the sample by geographical locations

     -Allocate enumerators

     -Allocate supervisors (ratio of 1:2 to 1:3 is desirable in most cases

     -Organize logistics-transport etc.

3. Consider the Publicity and Public Relations aspect of data collection, for instance

   -Addressing Kgotla meetings

   -Addressing schools

   -Radio and Television programmes; News papers; Pamphlets and Posters in public places

16. INTERVIEWERS' SUPERVISION

The quality of work done by enumerators is of crucial importance to any household survey and interviewer by supervision is often a key to good quality on the part of the interviewing staff.

Assuring quality is not an easy task since some interviewers may simply not be able to do the work and others may not put forth their full effort. A key to maintaining quality work is an effective system of field work supervision. The general management of the survey team lies upon supervisors.

Some of the responsibilities of a supervisor are:

- General management of the team.
- Create a good atmosphere within team members for better management and work relations.
- Identification of EA's and the selected households.
- Use the checklist provided for completed questionnaires on errors and omissions.
- Ensure that all procedures are followed when interviewing.

17. QUALITY CONTROL

Controlling the quality of the survey is important. Many activities such as training and supervision are an integral part of quality control, however there are some specific activities usually referred to as quality control which relates to controlling the quality of the interviewer's work. This is where by survey team members visits the field team members for field editing, validation, observation of interviews and addressing any field work problem.

Quality Control Measures at Data Collection Stage: The major sources of errors in data collection are usually those related to the completion of interviews. There are number of operational controls at the disposal of survey manager (both preventive measures as well as corrective methods) that could be used in order to minimize these errors that would arise during the process of survey operations.

1. Errors associated with incomplete listings or incorrect listings of households within a village EA could be prevented by rechecking the supervisors in the field and correcting the records by adding/deleting sampling units (household) to/from the list. Careful attention has to be paid to this survey operation since it could be an important source of error which could upset the random selection of the sample as well as the sampling weight.

2. Non-interviews5 of all kinds also pose problems. The seriousness of this source of error basically depends on the proportion of the sample not completed and also on the extent to which the household not covered differ from those included in the sample in-terms of the subject under investigation. These problems could be handled by repeated call-backs by enumerator/supervisor, for example in all national household surveys three call backs are attempted and considered adequate provided they were made on strategic dates and at strategic times of the day when members of the household are expected to be home.

3. Because of any reasons, the blank items (i.e. missing data) can distort the survey results6. Preventive measures usually include (i) the tabulation of missing data by enumerators, (ii) training of enumerators, and (iii) the revision of the questionnaire format. Corrective measures could include checking back with the respondents by the supervisors.

Therefore, in-order to improve the quality of the survey data it would be desirable to concentrate one's energy and resources on minimizing errors in (i) household listing, (ii) the rates of non-

5 Non-interviews usually cover occupied units for which interviews should have been obtained but were not due to several reasons viz. (i) Not at home, (ii) refusal by respondents

6 The problem of missing cases would particularly create problems in multi-variate analysis.

interview of households, (iii) the rates of non-response of items in the questionnaires, (iv) respondent error, and (v) enumerators error in asking the questions and recording information etc. at the data collection stage rather than relying on the adjustment of the data at a later stage. These adjustments, in any case, would entail the use of assumptions and might not hold true.

Remark: After completion of quality control exercise a report on the findings should be compiled for future use.

18. FIELD WORK PROBLEMS AND THE POSSIBLE SOLUTIONS

Refusals

Some people of the public who have been selected for interview may refuse to be interviewed. The reasons that they give may be that they do not have time or that the enumerators invade their privacy. Some of them may answer the household questionnaire but refuse to answer some questions in the individual questionnaire. Attempts to interview them through revisits sometimes do fail.

The conduct of the field staff may also contribute to refusals. Field staff are expected to conduct themselves in the manner that reflects a descent behavior that is worthy of a government official. This covers such things as the way one dresses when they visit households to conduct interview and where for instance, an enumerator involves themselves in a dispute at a social places.

Refusals may be minimized through publicity of the survey through the media, use of cover letters that explains the purpose of the enumerator's visit.

Terminations

A number of enumerators may terminate a contract for various reasons. This becomes a critical problem especially for surveys such as Household Income and Expenditure Survey (HIES) which takes twelve months as it demands proper trained staff. To overcome this problem, more than required enumerators and supervisors should be recruited and trained.

Transport constraints

Reliable transportation is crucial to the work of the survey teams. Shortage of vehicles and continuous vehicle breakdown are usually encountered during the field work period. If transport problems are unattended they will result in delaying work progress. Each team should have a dependable transportation so that it can move from one area to another. Emergency transportation must also be planned in case of breakdowns.

Budgetary constraints

Even though a survey is budgeted for, financial shortage may be encountered and may resulting in shortage of equipments as well as late payments for field staff. Therefore it is important to draw a realistic budget and it should be based on the past experience of a similar survey.

Enumeration area maps

The enumeration maps used in the survey may not be up to date such that some dwelling reflected in them may not exist during the survey period. For example; during Household Income and Expenditure Survey, some demolished dwellings were on the map but not on the ground. In some cases the floods may swipe away some dwelling and this will result in fewer households been listed during the listing exercise.

19. DATA ENTRY AND PROCESSING

A crucial task is to enter the data and put them into a form that is amenable to data analysis. Most data entry is now performed using computers with data entry software i.e CSPRO, IMPS. The software should be designed to check the logical consistency of the data.

The data from the field should arrive as at a central location as soon as possible so that it can be checked for any serious problems.

19.1 Data processing

Data processing is a set of activities aimed at converting the survey data from its raw state as the output from data collection to a cleaned and corrected state that is can be used in analysis, presentation or dissemination. The data may be changed by a number of operations which are intended to improve their accuracy. The data may be checked, compared corrected keyed and scanned coded tabulated and so on until the survey manager is satisfied that the results are 'fit for use'.

Data processing operations may be quite prone to human error when performed manually. By reducing reliance on manual labor, automation reduces the types of errors in the data caused by manual processing but may also introduce other types of errors that are specific to the technology used..

19.2 Data entry /capture

Data entry is the process where the questionnaire data are converted into computer-readable form. Data can be entered manually using keying equipment or automatically by using scanning or optical recognition devices. *Most data entry is now performed using computers with data entry software i.e CSPRO, SAS.*

19.3 Data editing

This is a set of methodologies used for verifying that data capture responses are plausible and if not correcting the data. *The data from the field should….*Editing rules can be developed for each variable or a combination of variables. The editing rules specify likely variable values or likely values for combinations of variables often as acceptable value intervals. In other cases, values are inserted or changed by means of deducing the correct value based on other information on the record. Most editing is performed automatically by specially designed computer software.

19.1 Coding

This is a procedure for classifying open-ended responses into pre-defined categories that are identified by numeric or alpha numeric code numbers. For example; for the occupation categories to be consistent across surveys by different organizations, a standard occupational classification (SOC) system is used.

19.4 Weighting

During the file preparation stage, a number of activities are involved. First weights have to be computed for the sample units. Weights are applied to data to compensate for the fact that sample units may be selected at unequal probability rates and have different probability of responding to the survey and that some population elements may not be included in the list or frame used. The main aim of weighting is to reduce bias in survey estimates by making each respondent represent a different fraction of the target population.

There are 3 forms of weights;

The base weights that account for the unequal probability of selecting elements from the sampling frame. The second form of weights is too adjust the base weights of the respondents for sampled elements that did not respond to the survey. Lastly poststratification may be used to make the sample estimates conform to the known population estimates

20. DATA ANALYSIS

All data is collected for purposes of analysis, so it is hardly necessary to point out that the final activity after data collection is the analysis. Data analysis usually takes longer time than planned but the findings based on the data are likely to be accurate and more useful if the analysts consult the survey team.

21. SURVEY REPORT AND DISSEMINATION

Generally two kinds of reports may be presented: either a general report giving a description of the survey for the use of those who are primarily interested in the results or a technical report giving details of sample design, computational procedures, accuracy and allied aspects.

The producers of statistical data should review data dissemination media, consulting representatives of users' groups to evaluate the demand for data and adapting their products and services to users' needs, thus fostering and contributing to the utilization of statistical data by academic and governmental institutions, NGOs and other users.

22. REFERENCES

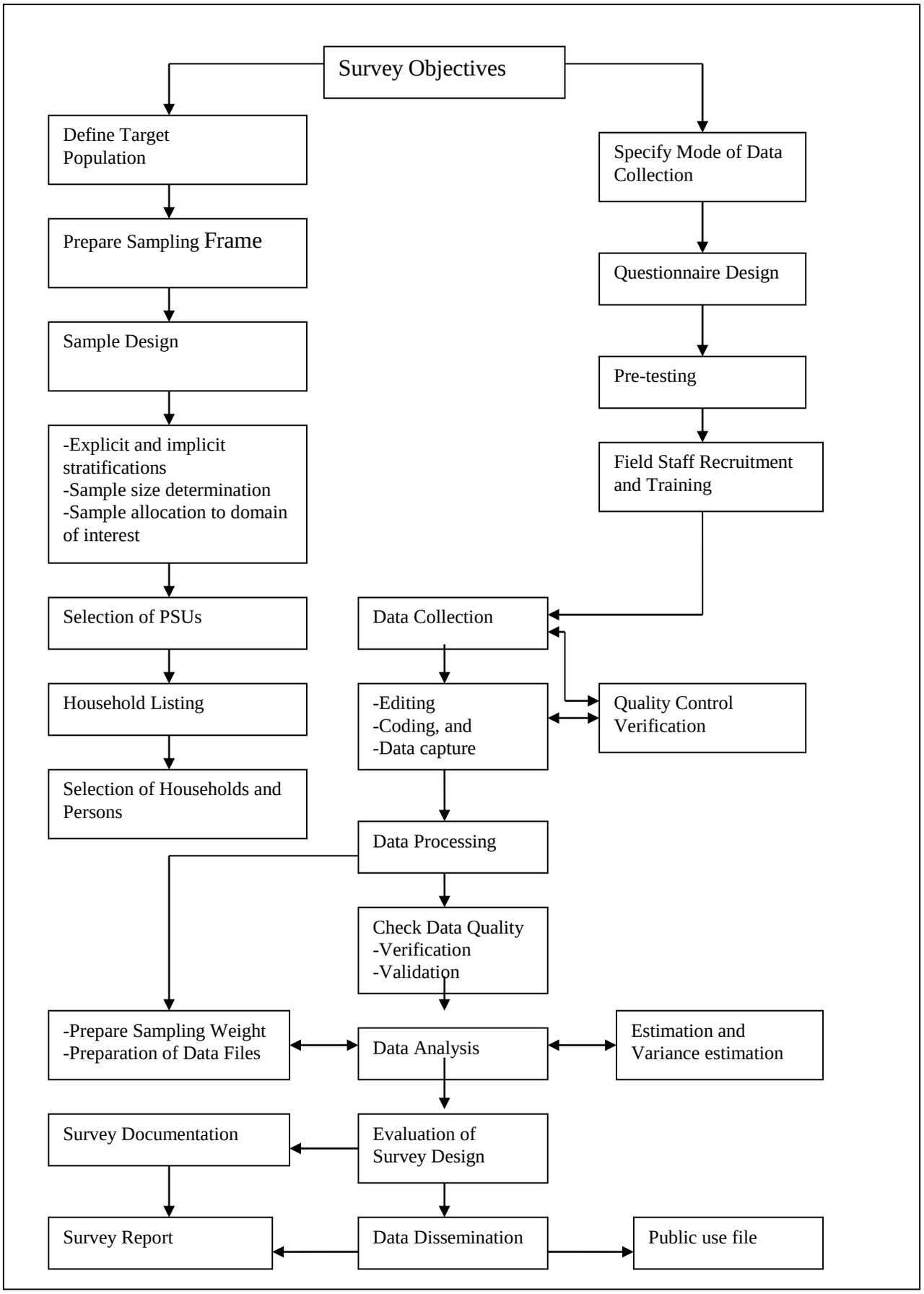
1. Cochran, W. G. (1977), Sampling Techniques, 3rd edition, Wiley, New York
2. Kish, L. (1965). Survey Sampling. Wiley, New York
3. Singh D. and Chaudhary, F.S. (1986). The Theory and Analysis of Sample Survey DesignWiley Eastern, New Delhi.
4. United Nations (2005): Households Sample Surveys in Developing and Transition Countries. United Nations Statistics Division (UNSD) Publication. (ST/ES/STAT/SER.F/96). New York. This publication can be downloaded free of cost from internet http://unstats.un.org/unsd/HHsurveys

23. APPENDIX-I: HOUSEHOLD SURVEYS CONDUCTED DURING INTER CENSUAL PERIOD

	Census 1981-1990	**Census 1991-2000**	**Census 2001-2010**
ECONOMIC SURVEYS			
Agricultural Census	1982	1993	2004
Household Income and Expenditure Survey (HIES)	1985/86	1993/94	2002/3
Informal Sector Survey (ISS)		1999	2007
Labour Force Survey (LFS)	1984	1995/6	2005/6
Botswana CWI Survey (BCWIS)			2009/10
SOCIAL SURVEYS			
Botswana Multiple Indicator		2000	
Botswana Literacy Survey (BLS)	-	1993	2003
Botswana AIDS Impact Survey (BAIS)			2001; 2004; 2008
Botswana Demographic Survey (BDS)	1987	1997	2006
Botswana Family Health Survey (BFHS)	1984 and 1988	1996	2007**

** Also contains the Botswana Multiple Indicator Survey (MIS) modules.

24. APPEDIX-II: FLOW CHART: SURVEY PROCESS



25. APPENDIX-III: ESTIMATION OF NUMBER OF ENUMERATORS

The detailed calculation process for the estimation of number of enumerators required to complete the survey in specified period is given below. The example is given below pertains to 2008 Botswana Aids Impact Survey-III. This estimation helps in budgeting and procurement of survey resources (tents, transport etc.).

CALCULATION PROCESS	
ITEMS	**Value**
(A) Listing Time	
(1) Number of days that a team of 2 enumerators lists all the households in **one** EA (past surveys experience)	1 day
(2) Thus one EA will be listed (listing of all households) by one enumerator in (2x1)	2 days
(3) Number of selected EAs to be listed	459
(4) =(2)x(3) Number of days taken by one enumerator for listing of total selected EAs	918 days
(5) Number of working days in a month	22 days
(6)=(4)/(5) Number of months taken by one enumerator for listing of total selected EAs	41.7 Months
(B) Expected Time Spent per HH per Interviewer	
(7) Traveling time to and from and within the EA per household (past surveys experience)	25 Minutes
(8) Time to administer the Household questionnaire (pre-test estimate)	45 Minutes
(9) Time to administer the Individual questionnaire per respondent (pre-test estimate)	120 Minutes
(10) Estimating the average number of qualified individual respondents (10-64 Yrs)7 per household	2.58 individual per household
(11)=(9)x(10) Total time to administer the Individual questionnaires per HH	310 Minutes
(12) Time to administer the Individual questionnaire per respondent-DBS (pre-test estimate)	15 Minutes
(13) Estimating the average number of qualified individual respondent (≥18 Months)-for DBS8 (pre-test estimate)	3.08 individual per household
(14)=(12)x(13) Total time to administer the Individual questionnaires per HH -DBS	46 Minutes
(15)=(7)+(8)+(11)+(14) The total expected time spent per Household per interviewer	426 Minutes
(C) Expected Number of Households to be Interviewed per Interviewer per Month	
(16) Usual working hours per day for the interviewers = 8 hours	480 Minutes
(17)=(16)/(15) Number of HH to be completed per interviewer per day	1.1 households
(18)=(17)x(5) Thus the expected Number of Households to be Interviewed per Interviewer per month	22.5 households
(19) To be conservative side: Number of Households to be Interviewed per Interviewer per month	22 households
(D) Expected Number of Months required by One interviewer to complete the survey field work all by him/herself	
Two basic settlements type strata are used in the PSU sample (EAs) and they have the following EAs allocation:	
(20) Number of EAs selected in Cities/Towns and Urban Villages	320 EAs
(21) Number of EAs selected in Rural Areas	139 EAs
(22) Number of Household to be selected per EA in Cities/Towns and Urban Villages	15 Households/EA
(23) Number of Household to be selected per EA in in Rural Areas	25 Households/EA

7Qualified individual respondent (10-64 Yrs)= Household size x Proportion of 10-64 yrs respondents x Individual response rate = 4.1 x0.7 x 0.9 = 2.58

8 Qualified individual respondent (≥18 Months)-for DBS= Household size x Proportion of respondent (≥18 Months) x Individual response rate = 4.1*0.94*0.8= 3.08

CALCULATION PROCESS	
ITEMS	**Value**
(24)=(19)/[(22) x 0.9] The estimated EAs that can be completed per month by one enumerator in Cities/Towns and Urban Villages = (adjusted for 90% household response rate)	1.6 EA
(25)=(19)/[(23) x 0.9] The estimated EAs that can be completed per month by one enumerator in Rural Areas = (adjusted for 90% household response rate)	1 EA
(26) Using the information (20) and (21) in (24) and (25), the number of months by one enumerator to complete the survey in each settlement is given as	
(27) =(20)/(24) Cities/Towns and Urban Villages	196.4 months
(28)=(21)/(25) Rural Areas	142.2 months
(29)=(27)+(28) Months for one enumerator to complete the survey fieldwork all by him/herself	338.6 months
(30) At this point the assumption is that an interviewer will find all the respondents at home for every single Household visited. We know that not to be the case. Households sometimes have to be revisited latter the same day or on following days or week ends. To compensate for this, let an additional time cost to be added is	30 percent
(31) Total number of increased months= (29)+ 30% of (29)	440.1 months
[E] Expected Number of interviewers (excluding drives and supervisors)	
(32)= (6)+(31) Total time to do listing and the survey itself for **one enumerator** would therefore be	481.8 months
(33) The survey is to be completed in	2.75 months
(34) Percent resignations and absence of enumerators and supervisors	15 percent
(35)=(adjustment for 34 i.e. 1.15)x(32)/(33) Number of enumerators required to complete the survey	201 enumerators
(36) Number of enumerators per team (may vary from survey to survey)	4
(37)=(35)/(36) Number of teams	50

26. APPENDIX-IV: FORMAT FOR SURVEY COSTS

The provision of sufficient funds for (i) the employment, training and travel costs of survey personnel, (ii) purchase of equipments and supplies etc. is a pre-requisite for the design and execution of any survey. The salary and allowance of survey personnel would depend on the number of personnel, salary/allowance rate and duration of employment. The supply of stationary, printing materials, medicines, camping equipments is needed by all surveys. However, the expenditure on non-expendable items such as survey equipment (measuring and weighing scale), personal computers would depend on the nature of the survey. The cost components for the design and execution of household surveys and the suggested format is given as follows:

DESCRIPTION	**Estimated Period**	**Quantity**	**Unit Cost**	**Total**
1	**2**	**3**	**4**	**5**
1. Personnel Emoluments & allowances				
1.1 Salaries:				
Temporary Enumerators, team supervisors, coders and editors				
1.2 Overtime:				
Enumerators, supervisors and drivers				
1.3 Commuted Allowances				
Enumerators, supervisors, drivers				
1.4 Subsistence, Training and Meal Allowances				
Quality control trips, training allowance, meal allowance				
Sub-Total-1				
2. Charges for Hired Vehicles				
Main survey				
(i)Pool Vehicle Hire Rate-Enumeration				
(ii) Pool Vehicle Hire Rate-Quality control				
(iii)Fuel Cost				
(iv) Pool Driver hire				
(v) Mileage				
Sub-Total-2				
3. Training to Field Staff				
(I) Meals charges (person meals)				
(ii) Conference Hall				
(iii) Photocopying-Questionnaire for training				
(iv) Photocopying-100 Enumerators Manual-100 pages each				
Sub-Total-3				
4. Stationary and Camping Materials				
4.1 Materials/Supplies				
(I) Scientific Calculators				
(ii) Service Kit				
(iii) Pumps				
(iv) Paper files				

DESCRIPTION	Estimated Period	Quantity	Unit Cost	Total
(v) Blue Pens				
(vi) Green Pens				
(vii) Pencils				
(viii) Eraser				
(ix) Clip boards				
(x) Enumerators Bag				
(xi) Photo copying Papers (A4)-Reams for pre-test and training (Questionnaires and manuals)				
4.2 Camping equipment etc				
Sub-Total-4				
5. Procurements of PCs, and Printers				
(i) Micro computers				
(ii) Printer				
Sub-Total-5				
6. Other: Questionnaire Printing for Main survey				
Questionnaire Printing				
Sub-Total-6				
7. Consultancy: Analysis of Survey data and report writing				
Consultancy-Data analysis and report writing (may be a team of 4 experts)				
Sub-Total-7				
8. Dissemination Seminar on the survey results				
Conference facilities etc.				
Printing/ photocopying charges for 275 reports (includes papers, printing and binding charges)				
Sub-Total-8				
9. Printing of Final Survey Report				
Sub-Total-9				
10. Advertising and publicity of survey				
Sub-Total-10				
Total				
Contingencies				
VAT				
GRAND. TOTAL				

27. APPENDIX-V: SURVEYS ACTIVITIES AND THEIR TIMING

In order to ensure the smooth execution of the survey, the expected duration9 as well as the timing10 of the various survey activities has to be determined in advance.

The format for the timing of the survey activities is given as follows:

Activity Number	Survey Activities	Expected Duration	Target Dates		Action
			From	To	
A	Planning and Preparatory Activities				
1	Initial overall planning – (adoption of benchmark objectives, selection of variables etc.) Identification of Reference Group (RG) & Technical Working Group (TWG) members.				
i.	Data capture Systems Planning				
ii.	Preparation of Edit Specifications				
iii.	Drafting of Survey questionnaires				
iv.	Letter to Reference Group				
v.	First Technical Group Meeting				
vi.	Sample Design/frame				
vii.	Reference group meeting				
viii.	Preparation of instructions & manuals				
ix.	First Stage Selection of Sample (Enumeration Area).				
x.	Photocopy of Questionnaires for Pre-test				
2.	Pre-Test Activities				
i.	**Process of vehicle Procurement and drivers (Pre-test)**				
ii.	**Training of interviewers for pre-test**				
iii.	Pre-test exercise in the field				
iv.	Revision & Finalization of Questionnaires and Manuals				
v.	*Process of Gazette Notification*				
3.	**Booking of Training Facilities**				
4.	Questionnaire & Manuals				
i.	Finalization Edit specs				
ii.	Preparation of Editing and Coding Manuals				

9 The expected duration of survey activities should be realistic based on past experiences and the existing conditions. When faced with uncertain circumstances an attempt should be made to allow for sufficient time, for example the process of hiring a consultant (if necessary) takes more expected duration than stipulated in the time frame.

10 As much as possible, the timing should indicate the starting and finishing dates for the execution of a given survey activity; otherwise the week and month numbers would suffice.

Activity Number	Survey Activities	Expected Duration	Target Dates		Action
			From	To	
iii.	Printing of questionnaires, Enumerators & Supervisors Manual				
iv.	**Advertising Posts (field staff)**				
5.A	**Identification of training Personnel**				
B	Field Operations				
i.	Process of recruitment of Field staff Personnel for availability by time of training.				
ii.	Procurement of vehicles and Drivers (Main survey)				
iii.	Training of survey field staffs. (One week supervisor two weeks enumerators)				
iv.	Publicity				
v.	Scheduling of field operations. Organizing field survey personnel.				
vi.	**Data collection.**				
vii.	Quality Control Checks				
C	**Data Processing**				
i.	Tabulation Plan & Formats				
ii.	Identification of a consultants (Tendering if out sourcing)				
iii.	Training of Editors and Coders				
iv.	Office coding and Editing of survey questionnaires.				
v.	Finalization Systems Development & Testing				
vi.	Data capture.				
vii.	Validation and errors corrections.				
viii.	Evaluation of the data				
ix.	Tabulation.				
D	**Report Preparation and Dissemination**.				
i.	Formation of Analysis team (DU)				
ii.	Data review and analysis.				
iii.	Preparation of Stats Brief of survey results				
iv.	Preparation of draft report.				
v.	Circulation of report to Editors for comments.				
vi.	Discussion on report by reference group.				
vii.	Dissemination seminar on survey results. (Including preparations)				
viii.	Finalization of report.				

28. APPENDIX-VI: DESIGN OF QUESTIONNAIRES

(a) QUESTIONNAIRE ADMINISTRATION

There are two ways in which a questionnaire may be administered.

i. It can be given directly to the respondent to fill in,
ii. An enumerator may read the questions to the respondents and complete it on their behalf.

Sometimes a kind of in-between arrangement is used, where an enumerators calls on a respondent, who should have already completed the questionnaire, to give assistance should it be required. Population Censuses in developing countries often use this in-between method.

Questionnaire for self-completion are usually sent out by post or else handed out at a convenient distribution point – for example a bus company might distribute questionnaires about its services to passengers whilst they are seated on the buses.

The two basic problems with self-fill questionnaire are that the respondent might not understand the questions or, even if he/she does, he/she might not feel inclined to answer. However the advantages of the self fill approach are that they are cheaper, quicker and less demanding upon manpower.

(b) TYPE OF QUESTIONNS

There are two basic types of questions: (i) **open** and (ii) **closed** (or **pre-coded**).

Open questions are those for which the respondent is to give his or her response in the provided blank space. These questions are more difficult to deal with at the data processing stage, especially as they require a coder to **classify** the response. Another disadvantage is that open questions require **more time** to complete than do closed.

Closed questions are supplied with a set of alternative answers which the respondent is expected to select the answer from. A simple example is the question, 'What is your sex? For which 'Male' and 'Female' are the choices.

The great advantage of closed question is that they are very easy to code afterwards. However, the disadvantage is that it is all too easy to choose a response **at random**, without even understanding the question.

Most of the survey questionnaires are of **mixture type** i.e. contains open and closed questions.

(c) THE DESIGN OF QUESTIONNAIRES

The way in which a questionnaire is designed will have a crucial influence on the <u>accuracy</u> of the responses obtained. The most important things to remember when designing a questionnaire are to try to keep it <u>simple</u> and <u>short.</u>

These and other general issues of design are now listed.

I. Reference Period of the Questions
An important type of timing consideration is the appropriate reference period for the questions. For example, personal and demographic characteristics such as age, marital status, educational attainment, etc. will usually relate to the date of interview. For certain subjects it relates to previous day (food consumption), previous week (employment activity) and even a yearly reference (entrepreneurial income).

II. Simplicity
The questionnaire should avoid unnecessary complications, such as technical words, which may make it hard to follow and be understood.

There are situations where complex questionnaires, with unusual structure and/or are appearances, are unavoidable due to the nature of the investigation. The complexity should never be more than is absolutely necessary.

III. Shortness
Short questionnaires are better because respondents will tend to become restless and bored when the questioning goes on for too long.

Short questionnaires also, of course, use up less paper: this is a real issue when the sample size is large. Shortness does not, however, mean brevity in this instance. Individual questions must be long enough to explain the issue thoroughly and not considered merely saving the space.

IV. Layout
The questionnaire should be attractive and easy to read in order to tempt the respondent to complete it.

The most successful questionnaire layout maintains a logical flow, with one question leading to the next wherever possible, and separate section for topics of different types.

With a good layout it will be difficult to accidentally omit a question, both when the questionnaire is being completed and when it is processed after being returned to survey headquarters.

V. Language
The respondent's natural language should be used to avoid misunderstanding and non-response. This may require several versions of the questionnaire and careful selection and distribution of enumerators.

VI. Consistent Wording
Where possible, a consistent choice of words and phrases should be used.

For instance, if the answer choices for questions 1-6 are Yes/No then question 7 should not change to something else simply for the sake of change.

The respondent will answer more accurately if able to settle into a pattern. However, fatigue will set in if the format remains unchanged for more than a few pages; a complete change of style is then advisable.

VII. Printing
Good quality printing promotes attractiveness and lends authority to the questionnaire. Small size print should be avoided to prevent non-response because of reading difficulty.

VIII. Clear Rules: The rules for answering questions should be stated fully and clearly, especially when the questionnaire is to be self-filled.

For example, with optional choice answers we should specify exactly how we expect the chosen answer to be indicated, e.g. ring the response, tick the accompanying box, etc.

We should also make it clear whenever more than one response may be given.
For instance, the questionnaire may offer a list of possible reasons why a businessman has not shown an operating profit in the past year. The businessman completing the form might be allowed to select several answers; for example, he might both blame his problems on lack of skilled manpower and unfair taxation policies.

IX. Leave Adequate Space
Unless enough space is left for answer to be given to open questions, we can expect many answers to be insubstantial or cramped and consequently illegible.
Response should be encouraged to continue on additional sheets if necessary.

X. Supply Units of Measurements
Where measurements, such as the amount of rainfall per day, are to be recorded the units to be used should be stated on the questionnaire.

XI. Give Instructions for Question Skipping
Often it is only a particular types of respondents that is expected to answer a given question, or a set of questions.

For example, a question on gender may be followed with instructions such as,
'If male go to question 17'
(Where question 17 is the next question, which is not aimed solely at women).

(d) THE CONTENTS OF QUESTIONNAIRE

It is obvious that any specific investigation will require a specific set of questions. Reference to earlier questionnaires on the same subject, if any exist, may be useful guide to content, but should always be treated critically: the earlier investigation may have made errors or may have had critically different circumstances.

For instance, there are several 'model' questionnaires produced by international organization may not be suitable for use in Botswana, where customs and terminology differs.

Questionnaires should always be tried out during a pilot survey before it is finalised. The questionnaire compilers should incorporate all the suggestions and comments in the questionnaire. He/she should keep in mind all conceivable types of respondent, such as all ethinic groups, all levels of education attainment and both sexes.

There are several general points about questionnaire content which will now be itemised. In each

case the aim is to increase the probability that,

i. An answer will be obtained, and
ii. It is an appropriate one for the question that the investigator wished to convey.

I. No Ambiguity

The question should have a unique meaning that is understood by all respondents. This implies that complex term should be avoided.

Sometimes very simple terms that are in common use should also be avoided, because they do not have a unique meaning. Examples are 'nice' and 'household'.

There is also ambiguity in simple objectives such as 'large' and 'heavy'. The question, 'Is your business large?' is useless, except to discover how many businessmen *think* that they own a large business. A factual question such as, 'How many people does your business employ?' would be preferable.

III. Ability to Answer

All respondents should be able to answer each question accurately. Hence there is no purpose in asking questions to which the respondent must guess an answer.

Inaccurate reporting will arise when too long a recall period is specific in a question such as,

'How many times have you visited Gaborone town centre in the past year?'

Or

'How many bottles of pop did you drink in the last month?

Few people would be able to answer such questions accurately. To make matters worse, most respondents are unwilling to admit that they are unable to answer a question, even when the question itself is at fault, because they do not wish to appear ignorant. This can make it difficult to identify problem questions, for instance at the pilot stage.

IV. Relevance

Long questionnaires are to be avoided; this implies that all questions should be relevant.

Even short questionnaires have been known to contain questions that are irrelevant because their answers cannot possibly be put to any useful purpose or, more specifically, the questions fail to address the aims of the survey.

Often very general questions, such as,

'Do you think this is a good shop?

Fall into this category. Questions on particular issues, such as the quality of goods, availability of goods etc. in this example, have a real purpose, since they can identify problem areas.

If only the general question is asked, the 'Yes' answer may mean an overall approval with no indication of whether there are still one or two drawbacks. The 'No' answer simply tells us there is at least one problem, but not what it is.

V. Avoid Leading Questions

Those who do respond to questionnaires are often eager to do a good job; they may see a need to get the answer 'right' or to please the enumerator (if applicable). Consequently they will tend to agree with any proposition made to them through a leading question.

For examples,

'Do you agree with the thirty million other Africans that Coca-Cola is the world's best drink?
(Leading to answer, 'yes')

Sometimes leading questions can provide hilarious results due to so called prestige-error. As mentioned above, relatively few people like to appear ignorant so that they are likely to be several affirmative answers to the question,

'Have you ever read the newsmagazine, 'New Citizen'?
Even when such a magazine has never existed!

VI. No Distress
Questions that are likely to cause upset or embarrassment are best avoided.

VII. No Calculations
Questions involving calculations should be avoided. For example, 'How much do you earn each week?' should not be asked of employees who are paid monthly.

(e) CODING
The responses to questionnaires are usually initially summarized, before analysis, as a set of codes on a database. Codes are easily assigned to closed questions, and are best laid out on the questionnaire itself, for ease and accuracy when the database is being compiled (e.g. during keying-in of data).

29. APPENDIX-VII: BASIC CONCEPTS

1 POPULATION

The population is the set of people or entities to which findings are to be generalized. The population must be defined explicitly before a sample is taken. Care must be taken not to generalize beyond the population. Doing so is a common error in social science writing. It is essential to define the population in terms of:

- **Content** refers to the definition of the type and characteristics of the elements which comprise the population; e.g. list of establishment, list of households,
- **Extent** refers to geographic boundaries as they relate to coverage; and e.g. list of establishments, list of Households in Gaborone.
- **Time** would refer to the time period to which the population refers. e.g. list of establishments, list households in Gaborone in 2001 Census.

2 TARGET POPULATION

The target population is the population (of elements of course), which we want to investigate by means of a sample survey (i.e. population which required to meet the survey objectives). For example population 12 years and above, children below 5 years etc.

3 SURVEY POPULATION

The survey/study population is the population actually covered by the survey, or better still, the population we have access to by means of the sampling frame. Ideally the target and survey population will be the same but for practical reasons they may not be identical. The survey population is usually a sub-set of the target population. When the two populations are not identical then the results of the sample should be generalized to the survey population. Generalizing the sample results to the target population may only be done with some degree of caution.

Examples

1. Many national surveys in the Botswana would ideally include hospitals, hotels, prisons, army barracks and other institutions. However, the severe problems involved in collecting responses from such persons frequently lead to their exclusion from target population. The advantage of starting with the ideal target population is that the exclusions are explicitly identified, thus enabling the magnitude and consequences of the restrictions to be assessed.

2. In HIES survey, one of the main objectives was to measure average income and expenditures of all households in Botswana (Target population) but the survey population comprised of those households residing in private dwellings. Thus in this instance generalizing the sample results to all households in Botswana should be done with some degree of caution.

- *Additions; I suggest we highlight on the sampling frame and define it under basic concepts.*

4 SAMPLING FRAME

For using sampling methods in the collection of data, it is essential to have a *frame* of all the

sampling units11 belonging to the population to be studied with their proper identification particulars, and such a frame is termed as *sampling frame*. This may be a list of units with their identification particulars or the maps showing the boundaries of the sampling units. The list should contain all the units of the population under consideration (target population). This indicates that sampling frame is the basic requirement to execute the sampling method.

TIPS AND GUIDELINES ON SAMPLING FRAMES

- Assess different possible frames at the planning stage for their suitability and quality.
- Avoid using multiple frames
- Use the same frame for surveys with the same target population
- Incorporate procedures to eliminate duplication and to update for births, deaths, and out of scope units in order to improve and maintain the level of quality of the frame.
- Implement map checks for area frames.

11 *Sampling units*: These are units into which the population is divided for the purposes of taking the sample. It forms the basis of a sampling procedure and as such should be clearly defined. When sampling people in a city, the sampling unit might be an individual member of a family or a group of elements or a household. For agricultural surveys, a sampling unit might be a field, a farm.

5. Enumeration Areas: An Enumeration Area (EA) is the smallest geographic unit, which represents an average workload for an enumerator over a specified period. The average size of an EA is approximately 120-150 malwapa. An EA may be a whole locality (this is the case of a small village which is an EA by itself), a part of a locality (this is the case of a bigger village which has been divided into more than one EA) or a group of localities (this is the case of cattle posts, lands areas or freehold farms).on the selection stage.

Insert the EA pictures from Urban and rural

6. Household: A household consists of one or more persons, related or unrelated, living together "under the same roof" in the same lolwapa, eating together "from the same pot" and/or making common provision for food and other living arrangements.