BIBLIOGRAPHY

Board of Science and Technology for International Development (BOSTID), (1988), *Cutting Edge Technologies and Microcomputer Applications for Developing Countries,* London: Intermediate Technology Publications/Westview.

Bulmer, M. and Warwick, D.D. (eds) (1983), *Social Research in Developing Countries: Surveys and Censuses in the Third World,* London: Wiley.

Casley, D.J. and Lury, D.A. (1981), *Data Collection in Developing Countries,* Oxford: Clarendon Press.

Cochran, W.G. (1977), *Sampling Techniques,* London: Sage.

Coleman, G. (1982), *Labour Data Collection in African Traditional Agricultural Systems,* University of East Anglia, Development Studies Occasional Paper, Norwich: University of East Anglia.

Connell, J. and Lipton, M. (1977), *Assessing Village Labour Situations in Developing Countries,* Delhi, India: Oxford University Press.

Devine, F. (1982), *Basic Statistics,* London: Harrap.

Feuerstein, M.T. (1986), *Partners in Evaluation,* London: Macmillan.

Fink, A. and Kosecoff, J. (1985), *How to Conduct Surveys: A Step by Step Guide,* Beverley Hills, USA: Sage.

Gardner, G. (1978), *Social Surveys for Social Planners,* Milton Keynes: Open University Press.

Hoinville, G. and Jowell, R. (1978), *Survey Research Practice*, London: Heinemann Educational Books.

Idaikkador, N.M. (1979), *Agricultural Statistics : a Handbook for Developing Countries*, Oxford: Pergamon.

Lipton, M. and Moore, M. (1972), *The Methodology of Village Studies in Less Developed Countries*, Institute of Development Studies Discussion Paper No. 10, Brighton: University of Sussex.

Moser, C.A. and Kalton, G. (1979), *Survey Methods in Social Investigation*, London: Heinemann Educational Books.

Peil M. et al, (1982) *Social Science Research Methods : an African Handbook*, London: Hodder and Stoughton.

Pratt, B. (forthcoming), *Data Collection for Fieldworkers, Development Guidelines Series*, Oxford: Oxfam.

Rowntree, D. (1981), *Statistics without Tears: A Primer for Non-mathematicians*, London: Penguin.

Sethuraman, S.V. (ed) (1981), *The Urban Informal Sector in Developing Countries: Employment, Poverty and Environment*, Geneva, Switzerland: International Labour Organisation.

Simpson-Herbert, M. (1983), *Methods for Gathering Socio-Cultural Data for Water Supply and Sanitation Projects*, Technical Note No. 1, Washington, USA: IBRD/World Bank.

UNICEF (1978), *Keep Your Village Clean: a Handbook for Community Workers*, New York, USA: UNICEF.

World Health Organisation (1977), *Rapid Village Nutrition Survey Technique*, Geneva, Switzerland: WHO.

GLOSSARY

active (mode):
Sentences in the form subject-verb-object, for example 'He dug the hole'. Sentences in the form object-verb, for example 'The hole was dug', are termed passive.

answer list:
A standard range of possible answers which the interviewers read out to the respondent. Respondents choose which answer(s) they agree with.

attitude scale:
A list of attitude statements, arranged in order, which the interviewer reads out loud. Respondents select the statement which they agree with.

average:
A measure of the centre of a distribution.

axis:
Straight line measuring distance in a single dimension. Used of graphs.

back-translation:
A technique for checking the translation of questionnaires. One person translates the original into the required language. A second person translates the translation back into the original language.

base variable:
To convert a two-way table into percentage form, you must first choose a base variable. Within each category of the base variable, percentages must sum to 100.

bias:

Repeated, systematic error. Bias in sampling, for example, means ignoring or under-representing some parts of the target population.

call-back:

Repeat attempt to contact a sample household in a social survey.

case-study:

Study based on a small number of 'typical' examples. Results are not statistically reliable.

chain sample:

See **sample**.

closed question:

(*See also* **open question**). The full range of possible answers is listed in advance. The interviewer marks the answer given in each case.

cluster sample:

See **sample**.

code:

System for summarising information collected in surveys. Each possible answer is given a code (usually a number). Each person's answer is then categorised using the appropriate code for each answer given.

coding instructions:

A full listing of all possible codes, together with the answers they summarise, for each question.

coding sheet:

A sheet recording the coded answers from many respondents. The codes for each respondent are usually recorded along a single row (or column).

communication strategy:

A plan to communicate survey findings. This can include the use of reports, pamphlets, presentations, workshops, tape-slide shows, drama and newspaper or radio coverage.

computer:
Electronic device to store and analyse data.

confidence interval:
Range of values which you can confidently expect will contain the true target population value you are trying to estimate. The confidence limits are the lower and upper values in the confidence interval.

control (group):
A 'standard' group for use in comparisons. You can compare a project area with a 'control' village, for example; this is a village which is similar in many ways to the project area, but without the project of interest.

correlation:
The strength of relationship between two (or more) variables. Positive correlation means that one variable tends to increase together with another variable. Negative correlation means that one variable decreases as the other one increases. The strength of a correlation is measured using values between -1 and >1. A score of zero means there is no relationship.

counting sheet:
A sheet for counting the number of responses in each possible category. A tally is used to record each response in the appropriate position.

coverage rate:
The proportion of the selected sample successfully contacted during fieldwork (usually given as a percentage).

cross-tabulation:
See two-way table.

cumulative frequency:
(*See also distribution.*) A count of the number of distribution cases less than a given range of values. For example, a cumulative income distribution may count the number of values less than $50, the number less than $100, the number less than $150, and so on.

data:

Information. The term 'data' often describes information stored in numerical form. Hard data is precise, numerical information. Soft data is less precise, verbal information. Raw data is the name given to survey information before it has been processed and analysed. Missing data are values or responses which fieldworkers were unable to collect (or which were lost before analysis).

data base:

System for storing data in an organised way. The aim is to give future users easy access to available data.

data cards:

Cards on which responses from a single questionnaire are written in summary form.

data processing:

Converting raw data into a usable, meaningful form.

data sheets:

Large sheets of paper on which responses from many questionnaires are recorded in a summary form. Usually, responses from each questionnaire are recorded in a single row (or column).

debriefing:

Discussion with the fieldwork team when fieldwork is complete.

demography:

The study of populations, especially statistics of birth, fertility and death.

deviation:

(*See also* **standard deviation.**) The size of the difference between two numbers. A common measure is deviation from the mean.

distribution:

The full set of sample or population values. A frequency distribution is a table which shows how many times each particular value or item occurs in the sample data.

economic:
Relating to work, income, spending or business.

economic activity:
Work to earn money, whether formal or informal, large-scale or small-scale. This includes small-scale agriculture and work in the home to earn money (such as brewing or baking for sale).

editing:
'Cleaning up' raw data. Checking for mistakes, correcting them where possible, dealing with missing data and creating new variables, as appropriate.

eligible:
Suitable for inclusion in the sample: a member of the target population.

error:
Variation between sample data and the full target population. Error can be systematic (*see* **bias**), or random. Random error is a result of chance differences from sample to sample. Systematic error (bias) occurs where there is a fault in the way the sample is chosen.

evaluation:
Review to assess whether a project has met its objectives. To perform a thorough evaluation, project monitoring is essential.

event calendar:
List of well-known local, regional or national events used to help pin down dates of birth, or other personal events of interest.

experiment:
Carefully designed trial or test, controlling the effect of outside factors. Used in agriculture, for example, to test alternative seed or crop types in controlled conditions.

fertility:
Number of births for each woman. Usually expressed as number of births to a fixed number of women over a fixed time period; per 1000 women per year, for example.

filter:
Form design system to ensure a question, or group of questions, is only put to a particular sub-group in the sample. A preliminary question 'filters' out those who do not apply.

frequency distribution:
See distribution.

formal methods:
(*See also informal methods*.) Structured research methods, usually using structured interview forms with large samples.

genealogy-based sample:
See sample.

guessed mean:
See mean.

hard data:
See data.

horizontal axis:
See axis.

infant mortality:
Number of babies dying before their first birthday out of a fixed number of live births (usually 1000).

informal methods:
(*See also formal methods.*) Less structured, exploratory research methods, such as case-studies or participant-observation.

interquartile range:
A measure of spread. The difference between the upper and lower quartile (*see percentile*).

key-informant:
Person carefully chosen for interview because of their special knowledge of some aspect of the target population.

key-words:

Summary question on interview form. The fieldworker is free to decide on exact wording during the interview.

malnutrition:

See nutrition.

matched sample:

See sample.

mean:

A kind of average for interval variables. Total of the sample values divided by the number of values in the sample. You can use a guessed mean (a round number, close to the true mean) to simplify calculation of the standard deviation.

median:

A kind of average for interval variables. The middle value when the data are arranged in order of size. Where the set of data has an even number of values the median is the mean of the two middle values.

memory:

Power of a computer to store data and programmes.

missing data:

See data.

monitoring:

Regular, systematic information collection to assess progress in meeting project objectives.

non-contact:

Member of the sample which the fieldworkers are unable to contact.

non-random sampling:

See sample.

non-sampling error:

Error due to factors other than random variation from sample to sample. Non-contacts, refusals, interviewer error, poor form design

and errors in data processing all contribute to non-sampling error.

nutrition:
Feeding habits. Science of food, nourishment and health. Mal-nutrition is poor nutrition.

one-shot survey:
Survey at a single point in time, in contrast to a repeat survey.

one-way table:
See table.

open question:
(*See also closed question.*) A question to which there can be many different answers. The interviewer writes down the answer in full, using the respondent's own words.

over-sample:
Used in stratified sampling. Instead of fixing the sample size in each strata according to its share of the total population, a larger sample is drawn from those strata with a small total population.

panel survey:
Survey involving several repeat visits to the same fixed sample over a long time period. Useful to study or monitor change over time.

participant observation:
An informal research method. The fieldworker lives and works among the target population for a period of several weeks or months, completing a systematic programme of observation and interview.

percentage:
The proportion (or fraction) in a particular class or sub-group, multiplied by 100.

percentile:
The value in the sample below which a given percentage of the values lie: for example, the 60th percentile has 60% of the sample values below it. The median is the 50th percentile. The upper

quartile is the 75th percentile. The lower quartile is the 25th percentile.

personal computer:
A small, desk-top computer designed for use in a small office or business.

pilot survey:
Also known as a pre-test. A small survey, in advance of the main fieldwork, to test the form, sampling procedures and fieldwork management procedures.

pre-test:
See pilot survey.

primary sampling unit:
Used in two-stage sampling. Clusters selected in the first stage are called primary sampling units.

probe:
A neutral phrase to encourage respondents to expand on their answer.

programme:
A detailed list of instructions to control the operation of a computer, written in a special programming language. A programme-package contains a wide range of ready-made programmes for use by the non-specialist. Examples of programming languages are BASIC, FORTRAN and PASCAL.

purposive sample:
See sample.

questionnaire:
A standard form used by the fieldworker to record a survey interview.

quota sample:
See sample.

random error:
See error.

random sampling:
See sample.

random walk:
Fieldworkers follow a carefully pre-set route, interviewing households at regular intervals.

range:
Difference between the highest and lowest value in a set of data.

raw data:
See data.

refusal:
A member of the sample who refuses to be interviewed.

regression:
A statistical technique for studying the strength of relationship between two interval variables.

repeat survey:
Study involving repeated fieldwork over a period of time, rather than a one-shot survey.

respondent:
Person being interviewed.

response rate:
Proportion (or fraction) of the sample who are successfully interviewed.

robust:
Affected little by small changes in the sample values (usually applied to summary measures, such as averages).

rotating survey:

Repeat visits are made to a panel, but on each new visit a fraction of the panel is replaced.

sample:

A selection of units chosen to represent the target population. In random sampling, the method of selection is based on chance and all units in the target population have a known change of begin selected. In non-random sampling the method of selection is based at least partly on the judgement of the researcher or interviewer.

Types of random sampling include: the simple random sample (selection of units by chance in its purest, simplest form);the systematic sample (choose a random point on the list or map and select units spaced at regular intervals from then on); the stratified random sample (use existing information to divide the sample into sub-groups called strata, then select a random sample within each sub-group); the cluster sample (to save time and money, choose the sample from several randomly selected clusters or areas of concentration rather than from the full target population); the two-stage sample (choose a random sample of units from each randomly selected cluster).

Types of non-random sampling include: the purposive sample (selecting units which you believe to be 'typical'); the quota sample (fieldworkers look for an agreed number of people or units of various 'types'); the random walk (fieldworkers follow a pre-set route, interviewing households at regular intervals); the genealogy-based sample (sample entire families, including all close relatives, to get a cross-section of the community); the chain sample (if you are looking for rare or minority units you can sometimes use a first contact to identify further respondents); the matched sample (choose pairs of units which you judge to be 'alike' in some way, for a comparative study).

sampling frame:

A list or map of all the units in the target population.

simple random sample:

See *sample.*

skip:
> Instruction to fieldworker to by-pass one or more questions on the form. The by-passed questions are only asked to a particular sub-group in the sample.

social survey:
> Fieldworkers use a specially designed form to collect a standard set of information from members of the target population.

soft data:
> *See data.*

spread:
> Measures of spread tell you how spread out (dispersed) a set of values is. Spread is measured from the centre (middle) of a distribution.

square:
> To square a number, multiply it by itself.

square root:
> The square root of a number B is the number A which, when squared, will give B. For example, 7 is the square root of 49.

standard deviation:
> (*See also deviation.*) The mean of squared deviation from the mean is called the variance. The standard deviation is the square root of the variance.

statistical function:
> On an electronic calculator, a function is the power to complete a certain type of operation. Addition, multiplication, taking the square root and squaring are examples of possible functions. Statistical functions, included on some calculators, will add up values as you input them and compute their mean and standard deviation.

strata, stratify:
> *See sample.*

stratified random sample:

See sample.

structured interview survey:

(*See also unstructured interview survey.*) A type of social survey. The range of possible answers to each question is known in advance. Often, possible answers are listed on the form so that the interviewer simply marks the appropriate reply in each case.

sub-sample:

A sample within a sample. For example you can draw a random sub-sample of completed forms to check the quality of fieldwork.

systematic error:

See error and bias.

systematic sampling:

See sample. ·

table:

A table presents a division of the sample into sub-groups. A one-way table uses the classes of a single variable to define the sub-groups. A two-way table uses the classes of a second variable to further sub-divide each class in the one-way table. A three-way table is actually a series of tables with a separate two-way table for sample values falling into each class of a third variable.

tabulation plan:

List of tables you plan to prepare.

tally:

Method of counting sample values by hand. Before you begin, lay out a table framework with a large box for each sub-group in the table. Work through the forms, marking each response in the appropriate box with a tally - a dash or mark. For easy counting, write the tallies in groups of five like this: .

target population:

The full group of people or units of interest.

three-way table:
See table.

two stage sample:
See sample.

two-way table:
See table.

unstructured interview survey:
(*See also structured interview survey.*) A type of social survey. An informal interview, not structured by a standard list of questions. Fieldworkers are free to deal with the topics of interest in any order and to phrase the questions as they think best.

variable:
A single characteristic of the target population, such as household size, household income, spending priorities.

variance:
See standard deviation.

vertical axis:
See axis.

weight:
Where you have over-sampled some strata, you need to use weights when you combine results from separate strata to give an estimate for the full target population. Divide the total for each stratum by the sample size in that stratum, then multiply by the proportion (fraction) of the target population in that stratum before adding up the strata results.

work status:
Worker 'types' such as self-employed, work for a wage or salary, unpaid family worker, or unpaid non-family worker.