Module

3
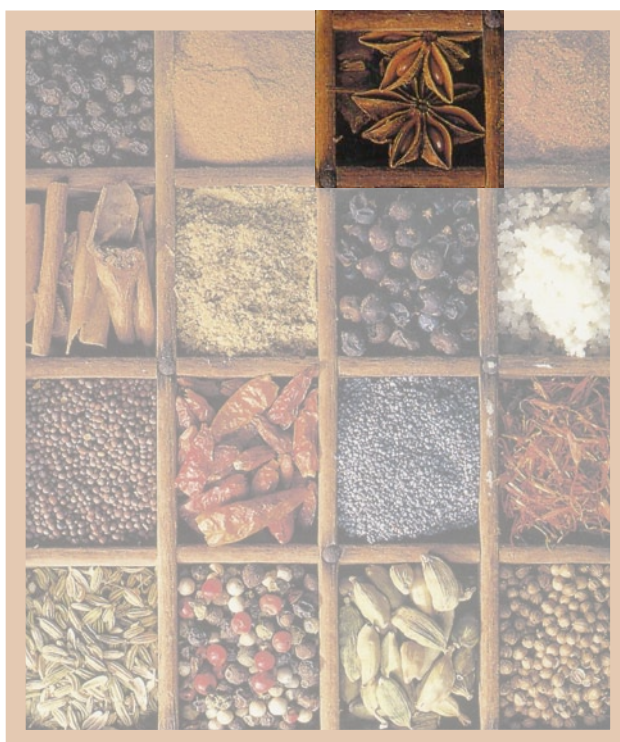
Kenneth N. Ross

Sample design for educational survey research

Quantitative research methods in educational planning

These modules were prepared by IIEP staff and consultants to be used in training workshops presented for the National Research Coordinators who are responsible for the educational policy research programme conducted by the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ).

The publication is available from the following two Internet Websites:
http://www.sacmeq.org and **http://www.unesco.org/iiep**.

Content

Basic concepts of sample design for educational survey research

Sampling in educational research is generally conducted in order to permit the detailed study of part, rather than the whole, of a population. The information derived from the resulting sample is customarily employed to develop useful generalizations about the population. These generalizations may be in the form of estimates of one or more characteristics associated with the population, or they may be concerned with estimates of the strength of relationships between characteristics within the population.

Provided that scientific sampling procedures are used, the selection of a sample often provides many advantages compared with a complete coverage of the population. For example, reduced costs associated with gathering and analyzing the data, reduced requirements for trained personnel to conduct the fieldwork, improved speed in most aspects of data summarization and reporting, and greater accuracy due to the possibility of more intense supervision of fieldwork and data preparation operations.

The social science research situations in which sampling is used may be divided into the following three broad categories: **experiments** – in which the introduction of treatment variables occurs according to a pre-arranged experimental design and all extraneous variables are either controlled or randomized; **surveys** – in which all members of a defined target population have a known

non-zero probability of selection into the sample; and **investigations** – in which data are collected without either the randomization of experiments or the probability sampling of surveys.

Experiments are strong with respect to internal validity because they are concerned with the question of whether a true measure of the effect of a treatment variable has been obtained for the subjects in the experiment. Surveys, on the other hand, are strong with respect to external validity because they are concerned with the question of whether the findings obtained for the subjects in the survey may be generalized to a wider population. Investigations are weak with respect to both internal and external validity and their use is due mainly to convenience or low cost.

Populations :
desired, defined, and excluded

In any educational research study it is important to have a precise description of the population of elements (persons, organizations, objects, etc.) that is to form the focus of the study. In most studies this population will be a finite one that consists of elements which conform to some designated set of specifications. These specifications provide clear guidance as to which elements are to be included in the population and which are to be excluded.

In order to prepare a suitable description of a population it is essential to distinguish between the population for which the results are ideally required, the desired target population, and the population which is actually studied, the defined target population. An ideal situation, in which the researcher had complete control over the research environment, would lead to both of these populations containing the same elements. However, in most studies, some differences arise due, for example, to (a) noncoverage:

the population description may accidently omit some elements because the researcher has no knowledge of their existence, (b) lack of resources: the researcher may intentionally exclude some elements from the population description because the costs of their inclusion in data gathering operations would be prohibitive, or (c) an ageing population description: the population description may have been prepared at an earlier date and therefore it includes some elements which have ceased to exist.

The defined target population provides an operational definition which may be used to guide the construction of a list of population elements, or sampling frame, from which the sample may be drawn. The elements that are excluded from the desired target population in order to form the defined target population are referred to as the excluded population.

Sampling frames

The selection of a sample from a defined target population requires the construction of a sampling frame. The sampling frame is commonly prepared in the form of a physical list of population elements – although it may also consist of rather unusual listings, such as directories or maps, which display less obvious linkages between individual list entries and population elements. A well-constructed sampling frame allows the researcher to 'take hold' of the defined target population without the need to worry about contamination of the listing with incorrect entries or entries which represent elements associated with the excluded population.

Generally the sampling frame incorporates a great deal more structure than one would expect to find in a simple list of elements. For example, in a series of large-scale studies of Reading Literacy carried out in 30 countries during 1991 (Ross, 1991), sampling frames were constructed which listed schools according to a number

of stratification variables: size (number of students), program (for example, comprehensive or selective), region (for example, urban or rural), and sex composition (single sex or coeducational). The use of these stratification variables in the construction of sampling frames was due, in part, to the need to present research results for sample data that had been drawn from particular strata within the sampling frame.

Representativeness

The notion of 'representativeness' is a frequently used, and often misunderstood, notion in social science research. A sample is often described as being representative if certain percentage frequency distributions of element characteristics within the sample data are similar to corresponding distributions within the whole population.

The population characteristics selected for these comparisons are referred to as 'marker variables'. These variables are usually selected from among those demographic variables that are readily available for both population and sample. Unfortunately, there are no objective rules for deciding which variables should be nominated as marker variables. Further, there are no agreed benchmarks for assessing the degree of similarity required between percentage frequency distributions for a sample to be judged as 'representative of the population'.

It is important to note that a high degree of representativeness in a set of sample data refers specifically to the marker variables selected for analysis. It does not refer to other variables assessed by the sample data and therefore does not necessarily guarantee that the sample data will provide accurate estimates for all element characteristics. The assessment of the accuracy of sample data can only be discussed meaningfully with reference to the value of the mean square error, calculated separately, for particular sample estimates (Ross, 1978).

The most popular marker variables in the field of education have commonly been demographic factors associated with students (sex, age, socio-economic status, etc.) and schools (type of school, school location, school size, etc.). For example, in a series of educational research studies carried out in the United States during the early 1970's, Wolf (1977) selected the following marker variables: sex of student, father's occupation, father's education, and mother's education. These variables were selected because their percentage frequency distributions could be obtained for the population from tabulations prepared by the Bureau of the Census.

Probability samples and non-probability samples

The use of samples in educational research is usually followed by the calculation of sample estimates with the aim of either (a) estimating the values of population parameters from sample statistics, or (b) testing statistical hypotheses about population parameters. These two aims require that the researcher has some knowledge of the accuracy of the values of sample statistics as estimates of the relevant population parameters. The accuracy of these estimates may generally be derived from statistical theory – provided that probability sampling has been employed. Probability sampling requires that each member of the defined target population has a known, and non-zero, chance of being selected into the sample.

In contrast, the stability of sample estimates based on non-probability sampling cannot be discovered from the internal evidence of a single sample. That is, it is not possible to determine whether a non-probability sample is likely to provide very accurate or very inaccurate estimates of population parameters. Consequently, these types of samples are not appropriate for

dealing objectively with issues concerning either the estimation of population parameters or the testing of hypotheses.

The use of non-probability samples is sometimes carried out with the (usually implied) justification that estimates derived from the sample may be linked to some hypothetical universe of elements rather than to a real population. This justification may lead to research results which are not meaningful if the gap between the hypothetical universe and any relevant real population is too large.

In some circumstances, a well-planned probability sample design can be turned accidentally into a non-probability sample design if subjective judgement is exercised at any stage during the execution of the sample design. Some researchers fall into this trap through a lack of control of field operations at the final stage of a multi-stage sample design.

The most common example of this in educational settings occurs when the researcher goes to great lengths in drawing a probability sample of schools, and then leaves it to the initiative of teaching staff in the sampled schools to select a 'random sample' of students or classes.

Types of non-probability samples

There are three main types of non-probability samples: judgement, convenience, and quota samples. These approaches to sampling result in the elements in the target population having an unknown chance of being selected into the sample. It is always wise to treat research results arising from these types of sample design as suggesting statistical characteristics about the population – rather than as providing population estimates with specifiable confidence limits.

1. Judgement sampling

The process of judgement, or purposive, sampling is based on the assumption that the researcher is able to select elements which represent a 'typical sample' from the appropriate target population. The quality of samples selected by using this approach depends on the accuracy of subjective interpretations of what constitutes a typical sample.

It is extremely difficult to obtain meaningful results from a judgement sample because no two experts will agree upon the exact composition of a typical sample. Therefore, in the absence of an external criterion, there is no way in which in the research results obtained from one judgement sample can be judged as being more accurate than the research results obtained from another.

2. Convenience sampling

A sample of convenience is the terminology used to describe a sample in which elements have been selected from the target population on the basis of their accessibility or convenience to the researcher.

Convenience samples are sometimes referred to as 'accidental samples' for the reason that elements may be drawn into the sample simply because they just happen to be situated, spatially or administratively, near to where the researcher is conducting the data collection.

The main assumption associated with convenience sampling is that the members of the target population are homogeneous. That is, that there would be no difference in the research results obtained from a random sample, a nearby sample, a co-operative sample, or a sample gathered in some inaccessible part of the population.

As for judgement sampling, there is no way in which the researcher may check the precision of one sample of convenience against another. Indeed the critics of this approach argue that, for many research situations, readily accessible elements within the target population will differ significantly from less accessible elements. They therefore conclude that the use of convenience sampling is likely to introduce a substantial degree of bias into sample estimates of population parameters.

3. Quota sampling

Quota sampling is a frequently used type of non-probability sampling. It is sometimes misleadingly referred to as 'representative sampling' because numbers of elements are drawn from various target population strata in proportion to the size of these strata.

While quota sampling places fairly tight restrictions on the number of sample elements per stratum, there is often little or no control exercised over the procedures used to select elements within these strata. For example, either judgement or convenience sampling may be used in any or all of the strata. Therefore, the superficial appearance of accuracy associated with proportionate representation of strata should be considered in the light that there is no way of checking either the accuracy of estimates obtained for any one stratum, or the accuracy of estimates obtained by combining individual stratum estimates.

Types of probability samples

There are many ways in which a probability sample may be drawn from a population. The method that is most commonly described in textbooks is simple random sampling. This method is rarely used in practical social research situations because

(a) the selection and measurement of individual population elements is often too expensive, and (b) certain complexities may be introduced intentionally into the sample design in order to address more appropriately the objectives and administrative constraints associated with the research. The complexities most often employed in educational research include the use of stratification techniques, cluster sampling, and multiple stages of selection.

1. Simple random sampling

The selection of a simple random sample is usually carried out according to a set of mechanical instructions which guarantees the random nature of the selection procedure.

For example, Kish (1965) provides the following operational definition in order to describe procedures for the selection of a simple random sample of elements without replacement from a finite population of elements:

From a table of random digits select with equal probability n different selection numbers, corresponding to n of the N listing numbers of the population elements. The n listings selected from the list, on which each of the N population elements is represented separately by exactly one listing, must identify uniquely n different elements. (p. 36)

Simple random sampling, as described in this definition, results in an equal probability of selection for all elements in the population. This characteristic, called 'epsem sampling' (equal probability of selection method), is not restricted solely to this type of sample design. Equal probability of selection can result from either the use of equal probabilities of selection throughout the sampling process, or from the use varying probabilities that compensate for each other through several stages of multistage sampling. Epsem sampling is widely applied in educational research because it usually leads to self-weighting samples in which the simple arithmetic mean obtained from the sample data is an unbiased estimate of the population mean.

2. Stratified sampling

The technique of stratification is often employed in the preparation of sample designs because it generally provides increased accuracy in sample estimates without leading to substantial increases in costs. Stratification does not imply any departure from probability sampling – it simply requires that the population be divided into subpopulations called strata and that probability sampling be conducted independently within each stratum. The sample estimates of population parameters are then obtained by combining information from each stratum.

In some studies, stratification is used for reasons other than obtaining gains in sampling accuracy. For example, strata may be formed in order to employ different sample designs within strata, or because the subpopulations defined by the strata are designated as separate 'domains of study' (Kish, 1987, p. 34).

Variables used to stratify populations in education generally describe demographic aspects concerning schools (for example, location, size, and program) and students (for example, age, sex, grade level, and socio-economic status).

Stratified sampling may result in either proportionate or disproportionate sample designs. In a proportionate stratified sample design the number of observations in the total sample is allocated among the strata of the population in proportion to the relative number of elements in each stratum of the population.

That is, a stratum containing a given percentage of the elements in the population would be represented by the same percentage of the total number of sample elements. In situations where the elements are selected with equal probability within strata, this type of sample design results in epsem sampling and therefore 'self-weighted' estimates of population parameters.

In contrast, a disproportionate stratified sample design is associated with the use of different probabilities of selection, or sampling fractions, within the various population strata. This can sometimes occur when the sample is designed to achieve greater overall accuracy than proportionate stratification by using 'optimum allocation' (Kish,1965:92). More commonly disproportionate sampling is used in order to ensure that the accuracy of sample estimates obtained for stratum parameters is sufficiently high to be able to make meaningful comparisons between strata.

The sample estimates derived from a disproportionate sample design are generally prepared with the assistance of 'weighting factors'. These factors, represented either by the inverse of the selection probabilities or by a set of numbers proportional to them, are employed in order to prevent inequalities in selection probabilities from causing the introduction of bias into sample estimates of population parameters. The reciprocals of the selection probalities, sometimes called 'raising factors', refer to the number of elements in the population represented by a sample element (Ross, 1978).

In the field of educational research, the weighting factors are often calculated so as to ensure that the sum of the weighting factors over all elements in the sample is equal to the sample size. This ensures that the readers of research reports are not confused by differences between actual and weighted sample sizes.

3.   Cluster sampling

A population of elements can usually be thought of as a hierarchy of different sized groups or 'clusters' of sampling elements. These groups may vary in size and nature. For example, a population of school students may be grouped into a number of classrooms, or it may be grouped into a number of schools. A sample of students may then be selected from this population by selecting clusters of students

as classroom groups or school groups rather than individually as would occur when using a simple random sample design.

The use of cluster sampling in educational research is sometimes undertaken as an alternative to simple random sampling in order to reduce research costs for a given sample size. For example, a cluster sample consisting of the selection of 10 classes – each containing around 20 students – would generally lead to smaller data collection costs compared with a simple random sample of 200 students. The reduced costs occur because the simple random sample may require the researcher to collect data in as many as 200 schools.

Cluster sampling does not prevent the application of probability sampling techniques. This may be demonstrated by examining several ways in which cluster samples may be drawn from a population of students. Consider the hypothetical population, described in Figure 1, of twenty-four students distributed among six classrooms (with four students per class) and three schools (with two classes per school).

A simple random sample of four students drawn without replacement from this population would result in an epsem sample with each element having a probability of selection, p, equal to 1/6 (Kish. 1965:40). A range of cluster samples, listed below with their associated p values, may also be drawn in a manner which results in epsem samples.

- Randomly select one class, then include all students in this class in the sample. (p = 1/6 x 4/4 = 1/6).

- Randomly select two classes, then select a random sample of two students from within these classes. (p = 2/6 x 2/4 = 1/6).

- Randomly select two schools, then select a random sample of one class from within each of these schools, then select a random sample of two students from within these classes. (p = 2/3 x 1/2 x 2/4 = 1/6).

Figure 1.   Hypothetical Population of 24 students

Schools	School 1		School 2		School 3	
	/ \		/ \		/ \	
Classes	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6
	| | | |	| | | |	| | | |	| | | |	| | | |	| | | |
Students	a b c d	e f g h	i j k l	m n o p	q r s t	u v w x

The accuracy of estimates obtained from probability samples

The degree of accuracy associated with a sample estimate derived from any one probability sample may be judged by the difference between the estimate and the value of the population parameter which is being estimated. In most situations the value of the population parameter is not known and therefore the actual accuracy of an individual sample estimate cannot be calculated in absolute terms. Instead, through a knowledge of the behaviour of estimates derived from all possible samples which can be drawn from the population by using the same sample design, it is possible to estimate the probable accuracy of the obtained sample estimate.

1.   Mean square error

Consider a probability sample of n elements which is used to calculate the sample mean as an estimate of the population mean. If an infinite set of samples of size n were drawn independently from

this population and the sample mean calculated for each sample, the average of the resulting sampling distribution of sample means would be referred to as the expected value.

The accuracy of the sample mean as an estimator of the population parameter may be summarized in terms of the mean square error (MSE). The MSE is defined as the average of the squares of the deviations of all possible sample estimates from the value being estimated (Hansen et al, 1953).

$$MSE(\bar{x}) = E(\bar{x} - \bar{X})^2$$

$$= E(\bar{x} - E(\bar{x}))^2 + E(\bar{x} - \bar{X})^2$$

$$= \text{VARIANCE OF } \bar{x} + (\text{BIAS OF } \bar{x})^2$$

WHERE, $\bar{x}$ = THE SAMPLE MEAN,

$\bar{X}$ = THE POPULATION MEAN,

$E(\bar{x})$ = THE EXPECTED VALUE OF $\bar{x}$,

$MSE(\bar{x})$ = MEAN SQUARE ERR OF $\bar{x}$

A sample design is unbiased if the expected value of the sample mean is equal to the population mean. It is important to remember that 'bias' is not a property of a single sample, but of the entire sampling distribution, and that it belongs neither to the selection nor the estimation procedure alone, but to both jointly.

For most well-designed samples in educational research the sampling bias is usually very small – tending towards zero with increasing sample size. The accuracy of sample estimates is therefore generally assessed in terms of the magnitude of the variance term in the above equation.

2. The accuracy of individual sample estimates

In educational settings the researcher is usually dealing with a single sample of data and not with all possible samples from a population. The variance of sample estimates as a measure of sampling accuracy cannot therefore be calculated exactly. Fortunately, for many probability sample designs, statistical theory may be used to derive formulae which provide estimates of the variance based on the internal evidence of a single sample of data.

For a simple random sample of n elements drawn without replacement from a population of N elements, the variance of the sample mean may be estimated from a single sample of data by using the following formula:

where s2 is the usual sample estimate of the variance of the element values in the population (Kish, 1965:41).

$$var(\bar{x}) \; = \; \frac{N - n}{N} \; x \; \frac{s^2}{n}$$

For sufficiently large values of N, the value of the finite population correction, (N - n)/N, tends toward unity. The variance of the sample mean in this situation may be estimated to be equal to s2/n.

The sampling distribution of the sample mean is approximately normally distributed for many educational sampling situations. The approximation improves with increased sample size even though the distribution of elements in the parent population may be far from normal. This characteristic of sampling distributions is known as the Central Limit Theorem and it occurs not only for the sample mean but also for most estimators commonly used to describe survey research results (Kish, 1965).

From a knowledge of the properties of the normal distribution we know that we can be "68 per cent confident" that the population mean lies within the range specified by:

(the sample mean ± one standard error of the sample mean)

where the standard error of the sample mean is equal to the square root of the variance of the sample mean. Similarly the we can be "95 percent confident" that the population mean lies within the range specified by:

(sample mean ± two standard errors of the sample mean)

While the above discussion has concentrated mostly on sample means derived from simple random samples, the same approach may be used to set up confidence limits for many other population values derived from various types of sample designs. For example, confidence limits may be calculated for complex statistics such as correlation coefficients, regression coefficients, multiple correlation coefficients, etc. (Ross, 1978).

3.　Comparison of the accuracy of probability samples

The accuracy of probability samples is usually compared by considering the variances associated with a particular sample estimate for a given sample size. This comparison has, in recent years, been based on the recommendation put forward by Kish (1965) that the simple random sample design should be used as a standard for quantifying the accuracy of a variety of probability sample designs which incorporate such complexities as stratification and clustering. Kish (1965:162) introduced the term 'deff' (design effect) to describe the ratio of the variance of the sample mean for a complex sample design, denoted c, to the variance a simple random sample, denoted srs, of the same size.

$$deff \quad = \quad \frac{var(\overline{x}_c)}{var(\overline{x}_{srs})}$$

The potential for arriving at false conclusions in educational research by using incorrect sampling error calculations has been demonstrated in a study carried out by Ross (1976). This study showed that it was highly misleading to assume that sample size was, in itself, an adequate indicator of the sampling accuracy associated with complex sample designs. For example, Ross (1976: 40) demonstrated that a two-stage cluster sample of 150 students (that was selected for a study conducted in Australia by randomly selecting 6 classes followed by the random selection of 25 students within these classes) had the same sampling accuracy for sample means as would a simple random sample of 20 students.

Sample design for two-stage cluster samples

The two-stage cluster sample is probably the most commonly used sample design in educational research. This design is generally employed by selecting either schools or classes at the first stage of sampling, followed by the selection of either clusters of students within schools or clusters of students within classes at the second stage. In many studies the two-stage cluster design is preferred because this design offers an opportunity for the researcher to conduct analyses at more than one level of data aggregation. For example, the selection of students within classes at the second stage of sampling would, provided there were sufficient numbers of classes and numbers of students selected within classes, permit analyses to be carried out at (a) the between-student level (by

using data describing individual students), (b) the between-class level (by using data based on class mean scores), or (c) both levels simultaneously using 'multilevel analysis' procedures.

In the following discussion some properties of two-stage cluster samples have been described by using the simplifying assumption that the clusters in the population are of equal size. This assumption permits the use of 'sample design tables' (see *Appendix 2*) which provide an excellent aid for designing two-stage cluster samples with pre-specified levels of sampling accuracy.

1. The design effect for two-stage cluster samples

The value of the design effect (Kish, 1965:257) for a two-stage cluster sample design depends, for a given cluster size, on the value of the coefficient of intraclass correlation.

$$deff \;=\; \frac{var(\overline{X}_c)}{var(\overline{X}_{srs})} \;=\; 1 + (b - 1)\,roh$$

where b is the size of the selected clusters, and roh is the coefficient of intraclass correlation

The coefficient of intraclass correlation, often referred to as roh, provides a measure of the degree of homogeneity within clusters. In educational settings the notion of homogeneity within clusters may be observed in the tendency of student characteristics to more homogeneous within schools, or classes, than would be the case if students were assigned to schools, or classes, at random. This homogeneity may be due to common selective factors (for example,

residential zoning of schools), or to joint exposure to the same external influences (for example, teachers and school programs), or to mutual interaction (for example, peer group pressure), or to some combination of these.

2. The effective sample size for two-stage cluster samples

The "effective sample size" (Kish, 1965:259) for a given two-stage cluster sample is equal to the size of the simple random sample which has a level of sampling accuracy, as measured by the variance of the sample mean, which is equal to the sampling accuracy of the given two-stage cluster sample. A little algebra may be used to demonstrate that the actual sample size, n_c, and the effective sample size, n^*, for a two-stage cluster sample are related to the design effect associated with that sample in the following manner (Ross, 1978:137-138).

$$n_c \ = \ n^* \times deff$$

From previous discussion, we may replace deff in this formula by an expression which is a function of the cluster size and the coefficient of intraclass correlation.

$$n_c \ = \ n^* \times (1 + (b - 1)\ roh)$$

For example, consider a two-stage cluster sample based on a sample of 10 schools followed by the selection of 20 students per school. In addition, consider a student characteristic (for example, a test score or attitude scale score) for which the value of the coefficient of intraclass correlation is equal to 0.1. This value of roh would

be typical for clusters of students selected randomly from within secondary schools in Australia (Ross, 1983). In this situation, the above formula simplifies to the following expression.

$$200 \quad = \quad n^* \text{ x } (1 + (20 \quad 1) \, 0.1)$$

Solving this equation for n* gives a value of 69 for the value of the effective sample size. That is, given the value of 0.1 for roh, a two-stage cluster sample of size 200 that is selected by sampling 10 schools followed by sampling 20 students per school would have sampling accuracy equivalent to a simple random sample of 69 students.

For a given population of students, the value of roh tends to be higher for clusters based on classes rather than clusters based on schools. Ross (1978) has obtained values of roh as high as 0.5 for mathematics test scores based on classes within Australian secondary schools.

Sample design tables for two-stage cluster samples

Sample design tables are often prepared for well-designed research studies in which it is intended to employ two-stage cluster sampling. These tables present a range of sample design options – each designed to have a pre-specified level of sampling accuracy. A hypothetical example has been presented in the following discussion in order to illustrate the sequence of calculations and decisions involved in the preparation of these tables.

Consider an educational research study in which test items are administered to a two-stage cluster sample of students with the aim

of estimating the percentage of students in the population that are able to obtain correct answers. In addition, assume that a sampling accuracy constraint has been placed on the design of the study so as to ensure that the sample estimate of the percentage of students providing the correct answer, p, will provide p ± 5 per cent as 95 per cent confidence limits for the value of the percentage in the population.

For reasonably large samples it is possible to assume normality of the distribution of sample estimates (Kish, 1965:13-14) and therefore confidence limits of p ± 5 per cent are approximately equivalent to an error range of plus or minus two standard errors of p. Consequently, the error constraint placed on the study means that one standard error of p needs to be less than or equal to 2.5 per cent.

Consider a simple random sample of n^* students selected from this population in order to calculate values of p. Statistical theory may be employed to show that, for large populations, the variance of the sample estimate of p as an estimate of the population value may be calculated by using the following formula (Kish, 1965:46).

$$var\ (p)\ =\ \frac{p\ (100\ -\ p)}{n^*\ -\ 1}$$

The maximum value of p(100 - p) occurs for p = 50. Therefore, in order to ensure that we could satisfy the error constraints described above, the following inequality would need to be valid.

$$(2.5)^2\ \geq\ \frac{50\ (100\ -\ 50)}{n^*\ -\ 1}$$

That is, the size of the simple random sample, n^*, would have to be greater than, or equal to, about 400 students in order to obtain 95 per cent confidence limits of $p \pm 5$ per cent.

Now consider the size of a two-stage cluster sample design which would provide equivalent sampling accuracy to a simple random sample of 400 students. The design of this cluster sample would require knowledge of the numbers of primary sampling units (for example, schools or classes) and the numbers of secondary sampling units (students) which would be required.

From previous discussion, the relationship between the size of the cluster sample, n_c, which has the same accuracy as a simple random sample of size $n^* = 400$ may be written in the following fashion. This expression is often described as a 'planning equation' because it may be used to explore sample design options for two-stage cluster samples.

$$n_c \; = \; 400 \; \times \; (1 + (b - 1)\, roh)$$

The value of n_c is equal to the product of the number of primary sampling units, a, and the number of secondary sampling units selected from each primary sampling unit, b. Substituting for n_c in this formula, and then transposing provides the following expression for a in terms of b and roh.

$$a \; = \; \frac{400}{b} \; \times \; (1 + (b - 1)\, roh)$$

As an example, consider roh = 0.1, and b = 20. Then we have the following value for a.

$$a \;=\; \frac{400}{20} \;\times\; (1 + (20 - 1)\,0.1)$$

$$=\; 58$$

That is, for roh = 0.1, a two-stage cluster sample of 1160 students (consisting of the selection of 58 primary sampling units followed by the selection of clusters of 20 students) would have sampling accuracy equivalent to a simple random sample of 400 students.

In *Table 1.1* the planning equation has been employed to list sets of values for a, b, deff, and n_c which describe a group of two-stage cluster sample designs that have sampling accuracy equivalent to a simple random sample of 400 students. Three sets of sample designs have been listed in the table – corresponding to roh values of 0.1, 0.2, and 0.4. In a study of school systems in ten developed countries, Ross (1983: 54) has shown that values of roh in this range are typical for achievement test scores obtained from clusters of students within schools.

The most striking feature of *Table 1.1* is the rapidly diminishing effect that increasing b, the cluster size, has on a, the number of clusters that are to be selected. This is particularly noticeable when the cluster size reaches 10 to 15 students.

Table 1   **Sample design table for two-stage cluster samples with sampling accuracy equal to a simple random sample of 400**

Cluster Size b	roh = 0.1			roh = 0.2			roh = 0.4		
	deff	n_c	a	deff	n_c	a	deff	n_c	a
1	1.0	400	400	1.0	400	400	1.0	400	400
2	1.1	440	220	1.2	480	240	1.4	560	280
5	1.4	560	112	1.8	720	144	2.6	1040	208
10	1.9	760	76	2.8	1120	112	4.6	1840	184
15	2.4	960	64	3.8	1530	102	6.6	2640	176
20	2.9	1160	58	4.8	1920	96	8.6	3440	172
30	3.9	1560	52	6.8	2730	91	12.6	5040	168
40	4.9	1960	49	8.8	3520	88	16.6	6640	166
50	5.9	2400	48	10.8	4350	87	20.6	8250	165

Consider, for example, two sample designs applied in a situation where a value of roh = 0.4 may be assumed: (a) a total sample of 2640 students obtained by selecting 15 students per cluster from 176 clusters, and (b) a total sample of 8250 students obtained by selecting 50 students per cluster from 165 clusters. From *Table 1.1*, it may be observed that both of these sample designs have sampling accuracy equivalent to a simple random sample of 400 students. These two sample designs have equivalent sampling accuracy, however there is a striking difference between the each design in terms of total sample size. Further, the magnitude of this difference is not reflected proportionately in the difference between the number of clusters selected.

This result illustrates an important point for the planning of educational research studies that seek to make stable estimates of student population characteristics. This is that the sampling accuracy levels of two-stage cluster sample designs, for cluster sizes of 10 or more, tend to be greatly influenced by small changes in the number of clusters that are selected at the first stage of sampling, and relatively less influenced by small changes in the size of the selected clusters.

The main use of sample design tables like the one presented in *Table 1.1* is to permit the researcher to choose, for a given value of roh, one sample design from among a list of equally accurate sample design options. The final choice between equally accurate options is usually guided by cost factors, or data analysis strategies, or a combination of both of these.

For example, the cost of collecting data by using 'group administration' of tests, questionnaires, etc. often depends more on the number of selected schools than on the number of students surveyed within each selected school. This occurs because the use of this methodology usually leads to higher marginal costs associated with travel to many schools, compared with the marginal costs of increasing the number of students surveyed within each selected school. In contrast, the cost of collecting data by using 'individual administration' of one-to-one tests, interviews, etc., normally depends more on the total sample size than on the number of selected schools.

The choice of a sample design option may also depend upon the data analysis strategies that are being employed in the research. For example, analyses may be planned at both the between-student and between-school levels of analysis. In order to conduct analyses at the between-school level, data obtained from individual students may need to be aggregated to obtain files consisting of school records based on student mean scores. This type of analysis

requires that sufficient students be selected per school so as to ensure that stable estimates are able to be made for individual schools. At the same time, it requires that sufficient schools are available so as to ensure that meaningful results may be obtained for estimates of population characteristics.

How to draw a national sample of students: a hypothetical example for 'country x' $\quad$ 2

In this section a hypothetical example has been presented in order to illustrate the 12 main steps involved in selecting a national sample of schools and students for 'Country X'. The numbers of schools and students in the example have been made very small so that all calculations and tabulations can be carried out by hand. In the next section of this module a 'real' example has been presented in which the required calculations and tabulations must be carried out by computer.

The desired target population for the example consists of 1100 students attending 20 schools located in two administrative regions of Country X. The sample design requires that five schools be selected across four strata with probability proportional to the number of students in the defined target population. Within the five selected schools a simple random sample of 20 students is to be selected. The 12 steps required to complete the sample design cover the following five areas.

- **Specification of Sample (Steps 1 to 4)**

The development and listing of 'specifications' for the sample design – including target population, sample size, sampling stages, stratification, sample allocation, cluster size, and pseudoschool

rules. In most research situations the specifications are prepared following a detailed investigation of the availability of information that is suitable for the construction of sampling frames, and an evaluation of the nature and scope of the data that are to be collected.

- **Stratification and Sample Allocation (Steps 5 to 7)**

A description of the stratification variables and an analysis of the breakdown of the defined target population across strata. The breakdown is used with the sample design specifications to guide the allocation of the sample across strata.

- **Marker Variables (Step 8)**

This is an 'optional' area. It provides information that can be used to make comparisons between the population and sample. However, it has no real impact on the scientific quality of the final sample.

- **Construction of Sampling Frame (Steps 9 and 10)**

The construction of a separate list of schools for each stratum. Each school is listed in association with the number of students in the defined target population. Schools that are smaller than a specified size are linked to a similar (and nearby) school to form a 'pseudoschool'.

- **Sample Selection (Steps 11 and 12)**

The selection of the sample of schools within strata (by using a 'lottery tickets' approach), and the selection of a sample of students within schools (by using a table of random numbers).

EXERCISE A

The reader should work through the 12 steps of the sample design. Where tabulations are presented, the figures in each cell should be verified by hand calculation using the listing of schools that has been provided in step 3. After working through all steps, the following questions should be addressed:

1. What is the probability of a student in the first school of the first stratum being selected into the sample? Is this probability the same for students in other schools in the first stratum? Is it the same for students in other strata?

2. How would the sampling of students be undertaken for a pseudoschool that was selected at the first stage of sampling?

3. What other variables might be used to stratify the defined target population? Are any of these 'better' than the ones used in the example?

4. What would *Table 3* and *Table 4* look like if the stratification variables were 'Region' and 'School Type'?

Step 1
List the basic characteristics of the sample design

1. **Desired target population**: Grade Six students in Country X.

2. **Stratification variables**: 'Region' and 'School size'.

3. **Sampling stages**: First Stage: Schools selected within strata with probability proportional to size. Second Stage: Students selected within schools by using simple random sampling.

4. **Minimum cluster size**: Fixed cluster of twenty students per school.

5. **Number of schools to be selected**: A total of five schools.

6. **Allocation of the sample across strata**: Proportionate allocation across strata. That is, the size of the sample within a stratum should be proportional to the total size of the stratum.

7. **Pseudoschools**: Each school in the defined target population with less than 20 students in the defined target population is to be combined with another similar (and nearby) school to form a 'pseudoschool'.

8. **Selection equation**

$$Probability \quad = \quad a_h \times \left(\frac{N_{hi}}{N_h} \right) \times \left(\frac{n_{hi}}{n_h} \right)$$

$$= \quad \frac{a_h \times n_{hi}}{N_h}$$

where a_h = the number of schools selected in stratum h,
$\quad N_h$ = the total number of students in stratum h,
$\quad N_{hi}$ = the total number of students in school i in stratum h, and
$\quad n_{hi}$ = the number of students selected from school i.

Step 2
Prepare brief and accurate written descriptions of the desired target population, the defined target population, and the excluded population

1. **Desired Target Population**: "Grade Six students in Country X".

2. **Defined Target Population**: "All students at the Grade Six level who are attending registered Government or Non-government primary schools in all administrative regions of Country X".

3. **Excluded Population**: "All students at the Grade Six level attending special schools for the handicapped in Country X".

Step 3
Locate a listing of all primary schools in Country_X that includes the following information for each school that has students in the desired target population

1. Description of listing

a. *School name*. The official (unique) name of the school or a suitable school identification number.

b. *Region*. The major state, province, or administrative region in which the school is located.

c. *District*. The district within an administrative region that is normally supervised by a District Education Officer.

d. *School type*. The name of the authority that administers the school. For this example, a simple classification into 'Government' and 'Non-government' schools. (Many studies use more detailed subgroups.)

e. *School location*. The location of the school in terms of urbanisation. For this example, 'Urban' or 'Rural' schools.

f. *School size*. The size of the total school enrolment. For this example, a simple classification into 'Large' and 'Small' schools. (Many studies use more detailed subgroups and some use the 'exact' enrolment.)

g. *School program*. A description of school program that is suitable for identifying schools in the 'excluded population'. For this example, 'Regular' and 'Special' schools.

h. *Target population*. The number of students within the school who are members of the desired target population. For this example, the total enrolment of students at the Grade Six level.

2. Contents of Listing

School_A	Reg_1	Dist_1	Govt	Urban	Large	Regular	100
School_B	Reg_1	Dist_1	Govt	Urban	Large	Regular	60
School_C	Reg_1	Dist_1	Non-Govt	Urban	Small	Regular	60
School_D	Reg_1	Dist_2	Govt	Urban	Large	Special	60
School_E	Reg_1	Dist_2	Non-Govt	Urban	Large	Regular	70
School_F	Reg_1	Dist_3	Govt	Rural	Small	Regular	40
School_G	Reg_1	Dist_3	Govt	Rural	Large	Regular	80
School_H	Reg_1	Dist_4	Non-Govt	Urban	Small	Regular	50
School_I	Reg_1	Dist_4	Non-Govt	Urban	Small	Regular	50
School_J	Reg_1	Dist_4	Govt	Urban	Large	Regular	90
School_K	Reg_2	Dist_5	Govt	Urban	Large	Regular	60
School_L	Reg_2	Dist_5	Non-Govt	Urban	Large	Regular	40
School_M	Reg_2	Dist_6	Govt	Urban	Small	Regular	10
School_N	Reg_2	Dist_6	Govt	Urban	Small	Regular	80
School_O	Reg_2	Dist_6	Non-Govt	Urban	Large	Regular	50
School_P	Reg_2	Dist_7	Govt	Rural	Small	Regular	30
School_Q	Reg_2	Dist_7	Govt	Rural	Large	Regular	50
School_R	Reg_2	Dist_8	Govt	Urban	Large	Special	40
School_S	Reg_2	Dist_8	Non-Govt	Urban	Small	Regular	50
School_T	Reg_2	Dist_9	Non-Govt	Urban	Small	Regular	30

Step 4
Use the listing of schools in the desired target population to prepare a tabular description of the desired target population, the defined target population, and the excluded population

There are 1100 students attending 20 schools in the desired target population. Two of these schools are 'special schools' and are therefore their 100 students are assigned to the excluded population – leaving 1000 students in 18 schools as the defined target population.

Table 2   **The desired, defined and excluded populations**

Desired		Defined		Excluded	
Schools	Students	Schools	Students	Schools	Students
20	1100	18	1000	2	100

Step 5
Select the stratification variables

The stratification variables are 'Region' (which has two categories: 'Region_1' and 'Region_2') and 'School Size' (which has two categories: 'Large' and 'Small' schools). These two variables combine to form the following four strata.

Stratum_1: Region_1 Large Schools

Stratum_2: Region_1 Small Schools

Stratum_3: Region_2 Large Schools

Stratum_4: Region_2 Small Schools

Step 6
Apply the stratification variables to the desired, defined, and excluded population

Table 3   **Schools and students in the desired, defined and excluded populations listed by the four strata**

Stratum	Desired		Defined		Excluded	
	Schools	Students	Schools	Students	Schools	Students
Stratum_1	6	460	5	400	1	60
Stratum_2	4	200	4	200	0	0
Stratum_3	5	240	4	200	1	40
Stratum_4	5	200	5	200	0	0
Country X	**20**	**1100**	**18**	**1000**	**2**	**100**

Step 7
Etablish the allocation of the sample across strata

The sample specifications in Step 1 require five schools to be selected in a manner that provides a proportionate allocation of the sample across strata. Since a fixed-size cluster of 20 students is to be drawn from each selected school, the number of schools to be selected from each stratum must be proportional to the total stratum size.

Proportionate Number of Schools. The proportionate number of schools required from the first stratum is 5 x (400/1000) = 2 schools. The other three strata have 200 students each and so the proportionate number of schools for each is 5 x (200/1000) = 1 school.

Table 4 **The sample allocation for the defined target population**

Stratum	Population of students		Sample allocation			
			Schools		Students	
	Number	Percent	Exact	Rounded	Number	Percent
Stratum_1	400	40.0	2.0	2	40	40.0
Stratum_2	200	20.0	1.0	1	20	20.0
Stratum_3	200	20.0	1.0	1	20	20.0
Stratum_4	200	20.0	1.0	1	20	20.0

Step 8 (*an optional step*)
Prepare tabular displays of students and schools in the defined target population – broken down by certain variables that may be of interest for the construction of 'marker variables' after the data have been collected

This is an optional step in the sample design because it has no impact upon the scientific quality of the final sample. The main justification for this step is that it may provide useful information (sometimes referred to as 'marker variables') that can be used to compare several characteristics of the final sample and the defined target population.

Table 5   **Schools with students in the defined target population listed by categories of school size, school location, and school type**

Stratum	School size		School location		School type		Stratum
	Large	Small	Rural	Urban	Govt	Non-Govt	
Stratum_1	5	0	1	4	4	1	5
Stratum_2	0	4	1	3	1	3	4
Stratum_3	4	0	1	3	2	2	4
Stratum_4	0	5	1	4	3	2	5
Country X	**9**	**9**	**4**	**14**	**10**	**8**	**18**

Table 6   **Students in the defined target population listed by categories of school size, school location, and school type**

Stratum	School size		School location		School type		Stratum
	Large	Small	Rural	Urban	Govt	Non-Govt	
Stratum_1	400	0	80	320	330	70	400
Stratum_2	0	200	40	160	40	160	200
Stratum_3	200	0	50	150	110	90	200
Stratum_4	0	200	30	170	120	80	200
Country X	**600**	**400**	**200**	**800**	**600**	**400**	**1000**

Step 9
For schools with students in the defined target population, prepare a separate list of schools for each stratum with 'pseudoschools' identified by a bracket ([ )

The sample design specifications in Step 1 require that 20 students be drawn for each selected school. However, one school in the defined target population, School_M, has only 10 students. This school is therefore combined with a similar (and nearby) school, School_N, to form a 'pseudoschool'.

The pseudoschool is treated as a single school of 90 students during the process of selecting the sample. If a pseudoschool happens to be selected into the sample, it is treated as a single school for the purposes of selecting a subsample of 20 students.

Note that School_D and School_R do not appear on the list because they are members of the excluded population.

Stratum 1. Region_1 Large schools

School_A	Reg_1	Dist_1	Govt	Urban	Large	Regular	100
School_B	Reg_1	Dist_1	Govt	Urban	Large	Regular	60
School_E	Reg_1	Dist_2	Non-Govt	Urban	Large	Regular	70
School_G	Reg_1	Dist_3	Govt	Rural	Large	Regular	80
School_J	Reg_1	Dist_4	Govt	Urban	Large	Regular	90

Stratum 2. Region_1 Small schools

School_C	Reg_1	Dist_1	Non-Govt	Urban	Small	Regular	60
School_F	Reg_1	Dist_3	Govt	Rural	Small	Regular	40
School_H	Reg_1	Dist_4	Non-Govt	Urban	Small	Regular	50
School_I	Reg_1	Dist_4	Non-Govt	Urban	Small	Regular	50

Stratum 3. Region_2 Large schools

School_K	Reg_2	Dist_5	Govt	Urban	Large	Regular	60
School_L	Reg_2	Dist_5	Non-Govt	Urban	Large	Regular	40
School_O	Reg_2	Dist_6	Non-Govt	Urban	Large	Regular	50
School_Q	Reg_2	Dist_7	Govt	Rural	Large	Regular	50

Stratum 4. Region_2 Small schools

School_M	Reg_2	Dist_6	Govt	Urban	Small	Regular	10
School_N	Reg_2	Dist_6	Govt	Urban	Small	Regular	80
School_P	Reg_2	Dist_7	Govt	Rural	Small	Regular	30
School_S	Reg_2	Dist_8	Non-Govt	Urban	Small	Regular	50
School_T	Reg_2	Dist_9	Non-Govt	Urban	Small	Regular	30

Step 10
For schools with students in the defined target population, assign 'lottery tickets' such that each school receives a number of tickets that is equal to the number of students in the defined target population

Note that the pseudoschool made up from School_M and School_N has tickets numbered 1 to 90 because these two schools are treated as a single school for the purposes of sample selection.

Stratum 1.   Region_1 Large schools

								Lottery tickets
School_A	Reg_1	Dist_1	Govt	Urban	Large	Regular	100	1-100
School_B	Reg_1	Dist_1	Govt	Urban	Large	Regular	60	101-160
School_E	Reg_1	Dist_2	Non-Govt	Urban	Large	Regular	70	161-230
School_G	Reg_1	Dist_3	Govt	Rural	Large	Regular	80	231-310
School_J	Reg_1	Dist_4	Govt	Urban	Large	Regular	90	311-400

Stratum 2.   Region_1 Small schools

								Lottery tickets
School_C	Reg_1	Dist_1	Non-Govt	Urban	Small	Regular	60	1-60
School_F	Reg_1	Dist_3	Govt	Rural	Small	Regular	40	61-100
School_H	Reg_1	Dist_4	Non-Govt	Urban	Small	Regular	50	101-150
School_I	Reg_1	Dist_4	Non-Govt	Urban	Small	Regular	50	151-200

Stratum 3. Regional_ 2 Large schools

								Lottery tickets
School_K	Reg_2	Dist_5	Govt	Urban	Large	Regular	60	1-60
School_L	Reg_2	Dist_5	Non-Govt	Urban	Large	Regular	40	61-100
School_O	Reg_2	Dist_6	Non-Govt	Urban	Large	Regular	50	101-150
School_Q	Reg_2	Dist_7	Govt	Rural	Large	Regular	50	151-200

Stratum 4. Regional_2 Small schools

								Lottery tickets
School_M	Reg_2	Dist_6	Govt	Urban	Small	Regular	10	1-90
School_N	Reg_2	Dist_6	Govt	Urban	Small	Regular	80	
School_P	Reg_2	Dist_7	Govt	Rural	Small	Regular	30	91-120
School_S	Reg_2	Dist_8	Non-Govt	Urban	Small	Regular	50	121-170
School_T	Reg_2	Dist_9	Non-Govt	Urban	Small	Regular	30	171-200

Step 11
Select the sample of schools

- **Selection of two schools for stratum 1**

In the first stratum two schools must be selected with probability proportional to the number of students in the defined target population. This is achieved by allocating a number of 'lottery tickets' to each school so that the first school on the list, School_A, with 100 students in the defined target population receives lottery tickets that are numbered from 1 to 100. The second school on the list, School_B, with 60 students receives lottery tickets that are numbered from 101 to 160, and so on – until the final school in the first stratum receives tickets numbered 311 to 400.

Since two schools must be selected there is a need to identify two 'winning tickets'. The ratio of the total number of tickets to the number of winning tickets is known as the 'sampling interval'. For the first stratum the sampling interval is equal to 400/2 = 200. That, is each ticket in the first stratum should have a 1 in 200 chance of being drawn as a winning ticket. Note that within the other three strata, one winning ticket must be selected from a total of 200 tickets – which gives the same value for the sampling interval (200/1 = 200).

The 'winning tickets' for the first stratum are drawn by using a 'random start – constant interval' procedure whereby a random number in the interval of 1 to 200 is selected as the first winning ticket and the second ticket is selected by adding an increment of 200. Assuming a random start of 105, the winning ticket numbers would be 105 and 305. This results in the selection of School_B (which holds tickets 101 to 160) and School_G (which holds tickets 231 to 310). The probability of selection is proportional to the number of tickets held and therefore each of these schools is selected with probability proportional to the number of students in the defined target population.

- **Selection of one school each for the other three strata**

Only one winning ticket is required for the other strata and therefore one random number must be drawn between 1 and 200 for each stratum. Assuming these numbers are 65, 98, and 176, the selected schools are School_F, School_L, and School_T.

- **Sample of Schools**

Stratum 1.   Region_1 Large schools

								Lottery tickets
School_B	Reg_1	Dist_1	Govt	Urban	Large	Regular	60	105
School_G	Reg_1	Dist_3	Govt	Rural	Large	Regular	80	305
School_F	Reg_1	Dist_3	Govt	Rural	Small	Regular	40	65
School_L	Reg_2	Dist_5	Non-Govt	Urban	Large	Regular	40	98
School_T	Reg_2	Dist_9	Non-Govt	Urban	Small	Regular	30	176

Step 12
Use a table of random numbers to select a simple random sample of 20 students in each selected school

Within a selected school a table of random numbers is used to identify students from a sequentially numbered roll of students in the defined target population. The application of this procedure has been described in detail by Ross and Postlethwaite (1991). A summary of the procedure has been presented below.

1. Obtain Grade Six register(s) of attendance

These registers are obtained for all students in the defined target population – in this case Grade Six students. In multiple session schools, both morning and afternoon registers are obtained.

2. Place a sequential number beside the name of each Grade Six student

For example, consider a school with one 'shift' and a total of 48 students in Grade Six. Commence by placing the number '1' beside the first student on the Register; then place the number '2' beside the second student on the Register; ...etc...;finally, place the number '48' beside the last student on the Register.

As another example, consider a school with 42 students in the morning 'shift' and 48 students in the afternoon session of Grade Six. Commence by placing the number '1' beside the first student on the Morning Register; ...etc...; then place a '42' beside the last student on the Morning Register; then place a '43' beside the first student on the Afternoon register; ...etc...;finally place a '90' beside the last student on the afternoon register.

3. **Locate the appropriate set of selection numbers**

In *Appendix 6* sets of "random number tables" have been listed for a variety of school sizes. (Note that only the sets relevant for school sizes in the range 21 to 100 have been presented in this Appendix.) For example, if a school has 48 students in Grade Six, the appropriate set of selection numbers is listed under the 'R48' heading. Similarly, if a school has 90 Grade Six students then the appropriate set of selection numbers is listed under the 'R90' heading.

4. **Use the appropriate set of random numbers**

After locating the appropriate set of random numbers, use the first random number to locate the Grade Six student with the same sequential number on the Register. Then use the second random number to locate the Grade Six student with the same sequential number on the Register. Continue with this process until the complete sample of students has been selected.

For example, in Step 11 the first school selected in Stratum_1 is School_B which has 60 students in the defined target population. Within this school the students selected would be those with the following sequential numbers: 1, 5, 6, 7, 9 ..., 52, and 54. (These numbers are obtained from the set of random numbers labelled 'R60' in *Appendix 6*).
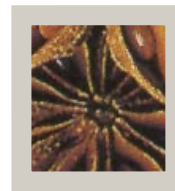
3 How to draw a national sample of schools and students: a *real* example for Zimbabwe

In *Section Two* of this module a hypothetical example was presented in order to provide a simple illustration of the 12 main steps involved in selecting a national sample of schools and students. In this section a 'real' example has been presented by applying these 12 steps to the selection of a national sample of schools and students in Zimbabwe. For most of these steps, an 'exercise' has been given which requires the use of a computer to carry out the kinds of calculations and tabulations that were possible to conduct by hand for the hypothetical example. The example draws upon computer-based techniques that were developed for a large-scale study of the quality of education at Grade Six level in Zimbabwe (Ross and Postlethwaite, 1991). The study was undertaken in 1991 by the International Institute for Educational Planning (IIEP, UNESCO) and the Zimbabwe Ministry of Education and Culture.

In order to work through the exercises in this section it will be necessary to employ the SAMDEM software (Sylla et al., 2004) in association with a sampling frame that was constructed from the 1991 'school forms' survey of primary schools in Zimbabwe. The database (described in Step 3 of this section, and entitled 'Zimbabwe.dat') was developed in association with George Moyo, Patrick Pfukani, and Saul Murimba from the Policy and Planning Unit of the Zimbabwe Ministry of Education and Culture.

The 12 steps described in this section have been used in many countries to select national samples of schools and students (Ross, 1991). However, it should be emphasized that the general ideas are applicable to samples that might be selected for smaller studies such as surveys of an administrative region or educational district.

Step 1
List the basic characteristics of the sample design

1. **Desired Target Population**: Grade Six students in Zimbabwe.

2. **Stratification Variables**: The following two stratification variables were used.

    a. *'Region':* This variable referred to the nine major administrative regions of Zimbabwe and took the following nine values: 'Harare', 'Manica' (Manicaland), 'Mascen' (Mashonaland Central), 'Masest' (Mashonaland East), 'Maswst' (Mashonaland West), 'Matnor' (Matabeleland North), 'Matsou' (Matabeleland South), 'Maving' (Masvingo), 'Midlnd' (Midlands).

    b. *'School size':* This variable referred to the size of the school in terms of total school enrolment and took the following two values: 'L' (Large) and 'S' (Small).

3. **Sampling Stages**: First Stage: Schools selected within strata with probability proportional to size. Second Stage: Students selected within schools by using simple random sampling.

4. **Minimum Cluster Size**: A fixed cluster of twenty students per school was accepted because this represented a manageable cluster size for data collection requirements within any single school.

5. **Number of Schools to be Selected**: The number of schools required for this example was governed by the requirement that final sample should have an effective sample size for the main criterion variables of at least 400 students. That is, the final sample was required to have sampling accuracy that was equivalent to, or better than, a simple random sample of 400 students.

a. Sample design tables

The general sample design framework adopted for the example consisted of a stratified two-stage cluster sample design. This permitted the use of **sample design tables** (Ross, 1987) to provide estimates of the number of schools and students required to obtain a sample with an effective sample size of 400. In order to use the sample design tables, it is necessary to know: the **minimum cluster size** (the minimum number of students within a school that will be selected for participation in the data collection), and the **coefficient of intraclass correlation** (a measure of the tendency of student characteristics to be more homogeneous within schools than would be the case if students were assigned to schools at random).

From above, it was known that the minimum cluster size adopted for the example was 20 students. Also, from a previous study of Reading Achievement conducted in Zimbabwe (Ross and Postlethwaite, 1991) it was found that the value of roh was around 0.3.

In *Appendix 2* a set of sample design tables has been presented for various values of the minimum cluster size, and various values of the co-efficient of intraclass correlation. The construction of these tables has been described by Ross (1987). In each table, 'a' has been used to describe the number of schools, 'b' has been used to describe the

minimum cluster size, and 'n' has been used to describe the total sample size.

The rows of *Table 4* that correspond to a minimum cluster size of one refer to the effective sample size. That is, they describe the size of a simple random sample which has equivalent accuracy. Therefore, the pairs of figures in the fourth and fifth columns in the table all refer to sample designs which have equivalent accuracy to a simple random sample of size 400. The second and third columns refer to an effective sample size of 1600, and the final two pairs of columns refer to effective sample sizes of 178 and 100, respectively.

The most important columns of figures in *Appendix 2* for this example are the fourth and fifth columns because they list a variety of two-stage samples that would result in an effective sample size of 400.

To illustrate, consider the intersection of the fourth and fifth columns of figures with the third row of figures in the first page of *Appendix 2*. The pair of values a=112 and n=560 indicate that if roh is equal to 0.1 and the minimum cluster size, b, is equal to 5, then the two-stage cluster sample design required to meet the required sampling standard would be 5 students selected from each of 112 schools – which would result in a total sample size of 560 students.

The effect of a different value of roh, for the same minimum cluster size, may be examined by considering the corresponding rows of the table for roh=0.2, 0.3, etc. For example, in the case where roh=0.3, a total sample size of 880 students obtained by selecting 5 students from each of 176 schools would be needed to meet the required sampling standard.

b. The number of schools required for this example

In this study the value selected for the minimum cluster size was 20, and the estimated value of the co-efficient of intra-class correlation was 0.3. From the sample design tables in Appendix 8, it may be seen that, in order to obtain a two-stage cluster sample with an effective sample size of 400, it is necessary to select a sample of 134 schools – which results in a total sample size of 2680 students.

6. **Allocation of the sample across strata**: Proportionate allocation across strata. That is, the size of the sample within a stratum should be proportional to the total size of the stratum.

7. **Pseudoschools**: Each school with less than 20 students in the defined target population is to be combined with another similar (and nearby) school to form a 'pseudoschool'.

8. **Selection Equation**

$$Probability \quad = \quad a_h \times \left(\frac{N_{hi}}{N_h} \right) \times \left(\frac{n_{hi}}{n_h} \right)$$

$$= \quad \frac{a_h \times n_{hi}}{N_h}$$

where a_h = the number of schools selected in stratum h,
N_h = the total number of students in stratum h,
N_{hi} = the total number of students in school i in stratum h, and
n_{hi} = the number of students selected from school i.

Step 2
Prepare brief and accurate written descriptions of the desired target population, the defined target population, and the excluded population

1. **Desired target population**: "Grade Six students in Zimbabwe".

2. Defined target population: "All students at the Grade Six level in September 1991 who are attending registered Government or Non-government primary schools that offer the 'regular' curriculum, that contain at least 10 Grade Six students, and that are located in all nine administrative regions of Zimbabwe".

3. Excluded Population: "All students at the Grade Six level in September 1991 who are attending (a) special schools for the handicapped in Zimbabwe (these schools do not offer the 'regular' curriculum), or (ii) primary schools in Zimbabwe with less than 10 Grade Six students (these schools are smaller than the specified minimum size)".

Step 3
Locate a listing of all primary schools in Zimbabwe that includes the following information for each school that has students in the desired target population

---

EXERCISE B

Use the SAMDEM software to list the following groups of schools in the Zimbabwe.dat file:

1. the first ten schools in the file,

2. schools that contain students in the 'excluded population' (special and small schools) and,

3. schools that must be combined to form `pseudoschools'.

Description of Listing (Stored in 'Zimababwe.dat' File)

- *Schlname* (Columns 1-15). This variable describes the official name given to each the school by the Zimbabwe Ministry of Education and Culture.

- *Region* (Columns 17-22). This variable describes the administrative regions of Zimbabwe. The variable takes nine values: 'Harare', 'Manica' (Manicaland), 'Mascen' (Mashonaland Central), 'Masest' (Mashonaland East), 'Maswst' (Mashonaland West), 'Matnor' (Matabeleland North), 'Matsou' (Matabeleland South), 'Maving' (Masvingo), 'Midlnd' (Midlands).

- *District* (Columns 24-29). This variable describes the administrative district within each administrative region of Zimbabwe. The variable takes 55 values ranging from 'Harare' (which is congruent with the region of Harare) to 'Zvisha' (which is located in the region of Midlands).

- *Schltype* (Columns 31-33). This variable describes schools according to the authority responsible for school administration. The variable takes two values: 'Gov' (Government Schools), and 'Non' (Non-government schools administered by District Councils, Rural Councils, Missions, Farm School Councils, Trust Councils, etc.).

- *Schllocn* (Column 35). This variable describes the location of the school in terms of the degree of urbanization of the surrounding community. The variable takes two values 'R' (rural location) and 'U' (urban location).

Contents of Listing (First Ten Schools in 'Zimbabwe.dat' File)

Alpha Brick	Harare	Harare	Non	R	L	Reg	86
Royden	Harare	Harare	Non	R	L	Reg	34
St Marnocks	Harare	Harare	Non	R	L	Reg	125
Chizungu	Harare	Harare	Non	R	L	Reg	302
Nyabira	Harare	Harare	Non	R	L	Reg	102
Kubatana	Harare	Harare	Non	R	L	Reg	43
Kintyre	Harare	Harare	Non	R	L	Reg	61
Gwebi	Harare	Harare	Non	R	L	Reg	90
Henderson	Harare	Harare	Non	R	L	Reg	36
Beatrice	Harare	Harare	Non	R	L	Reg	24

- *Schlsize* (Column 37). This variable describes each school in terms of the size of its enrolment at the Grade Six level. The variable takes two values: 'S' (a 'small' school) and 'L' (a 'large' school).

- *Schlprog* (Columns 39-41). This variable differentiates between school programs according to whether schools are regular primary schools or special schools designed for handicapped students. The variable takes two values 'Reg' (regular schools) and 'Spe' (special schools).

- *Enrolg6* (Columns 43-45). This variable describes the Grade Six enrolment for each Zimbabwe primary school in 1991.

Step 4
Use the listing of schools in the desired target population to prepare a tabular description of the desired target population, the defined target population, and the excluded population

---

EXERCISE C

Use the SAMDEM software to prepare frequency distributions for the desired, defined, and excluded populations. Then complete Table 3.1.

Table 7   **Schools and students in the desired, defined and excluded populations (Zimbabwe Grade Six 1991)**

Desired		Defined		Excluded			
				Special Schl		Small Stdt	
Schools	Students	Schools	Students	Schl	Stdt	Schl	Stdt
4487	283781	-	-	-	-	65	393

Step 5
Select the stratification variables

EXERCISE D

Use the names of the stratification variables presented in Step 1 to complete the following list.

Stratum_1: Harare Region Large Schools
Stratum_2: Harare Region Small Schools
Stratum_3: Manica Region Large Schools
Stratum_4: Manica Region Small Schools
. . . .
. . . .
. . . .
. . . .
. . . .
. . . .
. . . .
. . . .
. . . .
. . . .
. . . .
. . . .
Stratum_17: Midlnd Region Large Schools
Stratum_18: Midlnd Region Small Schools

Step 6
Apply the stratification variables to the desired, defined, and excluded population

> **EXERCISE E**
>
> *Use the SAMDEM software to prepare frequency distributions for the desired, defined, and excluded populations – separately for each stratum. Then complete Table 3.2*

Table 8   **The desired, defined, and excluded populations (Zimbabwe Grade Six 1991)**

No	Stratum		Desired		Defined		Excluded			
							Special		Very-small	
	Region	Sch. size	Sch.	Stu.	Sch.	Stu.	Sch.	Stu.	Sch.	Stu.
01	Harare	Large	140	25054	140	25054	0	0	0	0
02	Harare	Small	82	3367	72	3298	8	61	2	8
03	Manica	Large	-	-	-	-	-	-	-	-
04	Manica	Small	-	-	-	-	-	-	-	-
05	Mascen	Large	-	-	-	-	-	-	-	-
06	Mascen	Small	-	-	-	-	-	-	-	-
07	Masest	Large	-	-	-	-	-	-	-	-
08	Masest	Small	-	-	-	-	-	-	-	-
09	Maswst	Large	-	-	-	-	-	-	-	-
10	Maswst	Small	-	-	-	-	-	-	-	-
11	Matnor	Large	-	-	-	-	-	-	-	-
12	Matnor	Small	-	-	-	-	-	-	-	-
13	Matsou	Large	-	-	-	-	-	-	-	-
14	Matsou	Small	-	-	-	-	-	-	-	-
15	Maving	Large	-	-	-	-	-	-	-	-
16	Maving	Small	-	-	-	-	-	-	-	-
17	Midlnd	Large	-	-	-	-	-	-	-	-
18	Midlnd	Small	-	-	-	-	-	-	-	-
Zimbabwe			4487	283781	4399		23	275	65	393

Step 7
Establish the allocation of the sample across strata

EXERCISE F

Use the SAMDEM software to establish a proportionate allocation of the sample across the strata. Then complete Table 3.3.

Table 9 **The sample allocation for the defined population (Zimbabwe Grade Six 1991)**

No	Stratum		Defined		Sample				
					Schools		Students		
	Region	Sch. size	Number	Pct.	Exact	Planned	Number	Pct.	
01	Harare	Large	25054	8.8	11.9	12	240	8.9	
02	Harare	Small	3298	1.2	1.6	2	40	1.5	
03	Manica	Large	-	-	-	-	-	-	
04	Manica	Small	-	-	-	-	-	-	
05	Mascen	Large	-	-	-	-	-	-	
06	Mascen	Small	-	-	-	-	-	-	
07	Masest	Large	-	-	-	-	-	-	
08	Masest	Small	-	-	-	-	-	-	
09	Maswst	Large	-	-	-	-	-	-	
10	Maswst	Small	-	-	-	-	-	-	
11	Matnor	Large	-	-	-	-	-	-	
12	Matnor	Small	-	-	-	-	-	-	
13	Matsou	Large	-	-	-	-	-	-	
14	Matsou	Small	-	-	-	-	-	-	
15	Maving	Large	-	-	-	-	-	-	
16	Maving	Small	-	-	-	-	-	-	
17	Midlnd	Large	-	-	-	-	-	-	
18	Midlnd	Small	-	-	-	-	-	-	
	Zimbabwe Total		283113	100.0	134.0	135	2700	100.0	

Step 8 (an optional step)
Prepare tabular displays of students and schools in the defined target population – broken down by School Size, Location, and Type

---

EXERCISE G

Use the SAMDEM software to prepare 'marker variable' information. Complete Tables 3.4 and 3.5.

Table 10 **Schools with students in the defined target population listed by school size, school location, and school type (Zimbabwe Grade Six 1991)**

No	Stratum		School size		School location		School type		Stratum Total
	Region	**Sch. size**	**Large**	**Small**	**Rural**	**Urban**	**Govt**	**Non-Govt**	
01	Harare	Large	140	0	8	132	99	41	140
02	Harare	Small	0	72	26	46	17	55	72
03	Manica	Large	-	-	-	-	-	-	-
04	Manica	Small	-	-	-	-	-	-	-
05	Mascen	Large	-	-	-	-	-	-	-
06	Mascen	Small	-	-	-	-	-	-	-
07	Masest	Large	-	-	-	-	-	-	-
08	Masest	Small	-	-	-	-	-	-	-
09	Maswst	Large	-	-	-	-	-	-	-
10	Maswst	Small	-	-	-	-	-	-	-
11	Matnor	Large	-	-	-	-	-	-	-
12	Matnor	Small	-	-	-	-	-	-	-
13	Matsou	Large	-	-	-	-	-	-	-
14	Matsou	Small	-	-	-	-	-	-	-
15	Maving	Large	-	-	-	-	-	-	-
16	Maving	Small	-	-	-	-	-	-	-
17	Midlnd	Large	-	-	-	-	-	-	-
18	Midlnd	Small	-	-	-	-	-	-	-
	Zimbabwe Total		1142	3257	4036	363	270	4129	4399

Table 11 **Students in the defined target population listed by school size, school location, and school type (Zimbabwe Grade Six 1991)**

No	Stratum		School size		School location		School type		Stratum Total
	Region	Sch. size	Large	Small	Rural	Urban	Govt	Non-Govt	
01	Harare	Large	25054	0	1175	23879	17798	7256	25054
02	Harare	Small	0	3298	1062	2236	913	2385	3298
03	Manica	Large	-	-	-	-	-	-	-
04	Manica	Small	-	-	-	-	-	-	-
05	Mascen	Large	-	-	-	-	-	-	-
06	Mascen	Small	-	-	-	-	-	-	-
07	Masest	Large	-	-	-	-	-	-	-
08	Masest	Small	-	-	-	-	-	-	-
09	Maswst	Large	-	-	-	-	-	-	-
10	Maswst	Small	-	-	-	-	-	-	-
11	Matnor	Large	-	-	-	-	-	-	-
12	Matnor	Small	-	-	-	-	-	-	-
13	Matsou	Large	-	-	-	-	-	-	-
14	Matsou	Small	-	-	-	-	-	-	-
15	Maving	Large	-	-	-	-	-	-	-
16	Maving	Small	-	-	-	-	-	-	-
17	Midlnd	Large	-	-	-	-	-	-	-
18	Midlnd	Small	-	-	-	-	-	-	-
Zimbabwe Total			139642	143471	239428	43685	35763	247350	283113

Step 9
For schools with students in the defined target population, prepare a separate list of schools for each stratum with 'pseudoschools' identified with a bracket ([ )

EXERCISE H

Use the SAMDEM software to list the schools having students in the defined target population and located in 'stratum 2: Harare region small schools!'.

Step 10
For the defined target population, assign 'lottery tickets' such that each school receives a number of tickets that is equal to the number of students in the defined target population

---

EXERCISE I

Use the SAMDEM software to assign 'lottery tickets' for schools having students in the defined target population and located in 'stratum 2: Harare region small schools'. repeat this exercise for one other stratum.

Step 11
Select the sample of schools.

> **EXERCISE J**
>
> *Use the SAMDEM software to to select the sample of schools from schools having students in the defined target population and located in 'stratum 2: Harare region small schools'. repeat this exercise for one other stratum.*

Step 12
Use a table of random numbers to select a simple random sample of 20 students in each school.

> **EXERCISE K**
>
> *use the random number tables (see Appendix 1) to identify student selection numbers for the schools that were selected from 'stratum 2: Harare region small schools'. repeat this exercise for one other stratum.*
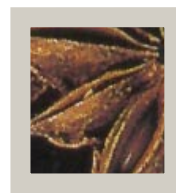
4 The estimation of sampling errors

The computational formulae required to estimate the variance of descriptive statistics, such as sample means, are widely available for simple random sample designs and some sample designs that incorporate complexities such as stratification and cluster sampling. However, in the case of more complex analytical statistics, such as correlation coefficients and regression coefficients, the required formulae are not readily available for sample designs which depart from the model of simple random sampling. These formulae are either enormously complicated or, ultimately, they prove resistant to mathematical analysis (Frankel, 1971).

In the absence of suitable formulae, a variety of empirical techniques have emerged in recent years which provide «approximate variances that appear satisfactory for practical purposes» (Kish, 1978:20). The most frequently applied empirical techniques may be divided into two broad categories : Subsample Replication and Taylor's Series Approximation.

In Subsample Replication a total sample of data is used to construct two or more subsamples and then a distribution of parameter estimates is generated by using each subsample. The subsample results are analyzed to obtain an estimate of the parameter, as well as a confidence assessment for that estimate (Finifter, 1972). The main approaches in using this technique have been Independent Replication (Deming, 1960), Jackknifing (Tukey, 1958), and Balanced Repeated Replication (McCarthy, 1966).

Alternatively, Taylor's Series Approximation may be used to provide a more 'direct' method of variance estimation than these three approaches. In the absence of an exact formula for the variance, the Taylor's Series is used to approximate a numerical value for the first few terms of a series expansion of the variance formula. A number of computer programs have been prepared in order to carry out the extensive numerical calculations required for this approach (Wolter, 1985, *Appendix*).

In the remainder of this section the Jackknife procedure has been described. This procedure offers the following two important benefits.

- **Greater Flexibility**
  The Jackknife may be applied to a wide variety of sample designs whereas (i) Balanced Repeated Replication is designed for application to sample designs that have precisely two primary sampling units per stratum, and (ii) Independent Replication requires a large number of selections per stratum so that a reasonably large number of independent replicated samples can be formed

- **Ease of Use**
  The Jackknife does not require specialized software systems whereas (i) Balanced Repeated Replication usually requires the prior establishment by computer of complex Hadamard matrices that are used to 'balance' the half samples of data that form replications of the original sample, and (ii) Taylor's Series methods require specific software routines to be available for each statistic under consideration.

The Jackknife Procedure

The development of the Jackknife procedure may be traced back to a method used by Quenouille (1956) to reduce the bias of estimates. Further refinement of the method (Mosteller and Tukey, 1968) led to its application in a range of social science situations where formulae are not readily available for the calculation of sampling errors.

The Jackknife procedure requires that an initial estimate, y_{all}, of a statistic, y, be made on the total sample of data. The total sample is then divided into k subgroups and y_i is computed, using the same functional form as y_{all}, but based on the reduced sample of data obtained by omitting the ith subgroup. Then k 'pseudovalues' y_i^* (i=1,2,...,k) can be defined – based on the k reduced samples:

$$y_i^* = ky_{all} - (k - I)y_i$$

Quenouille's estimator (also called the 'Jackknife value') is the mean of the k pseudovalues:

$$y^* = I/k \sum_{i=I}^{k} y_i^*$$

Quenouille's contribution was to show that, while y_{all} may have bias of order 1/n as an estimate of y, the Jackknife value, y^*, has bias of order $1/n^2$.

The variance of y^* may be estimated from the pseudovalues.

$$var(y^*) = \frac{var(y_i^*)}{k}$$

$$= \frac{1}{k(k-1)} \sum_{i=1}^{k} (y_i^* - y^*)^2$$

Tukey (1958) set forward the proposal that the pseudovalues could be treated as if they were approximately independent observations and that Student's t distribution could be applied to these estimates in order to construct confidence intervals for y^*. Later empirical work conducted by Frankel (1971) provided support for these proposals when the Jackknife technique was applied to complex sample designs and a variety of simple and complex statistics.

Substituting for y_i^* in the expression for var(y^*) permits the variance of y^* to be estimated from the k subsample estimates, y_i, and their mean – without the need to calculate pseudovalues.

$$var(y^*) = \frac{(k-1)}{k} \sum_{i=1}^{k} (y_i - \bar{y}_i)^2$$

Wolter (1985, p. 156) has shown that replacing y^* by y_{all} in the right hand side of the first expression for var(y^*) given above provides a conservative estimate of var(y^*) – the overestimate being equal to $(y_{all}-y^*)^2/(k-1)$.

In practice, these expressions for var(y*) have also been used to estimate the variance not only of Quenouille's estimator y*, but also of y_{all} (Wolter, 1985, pp. 155, 172).

Substituting y_{all} for y* provides the following estimator for the variance of y* (and also a conservative estimator for y_{all}).

$$var(y^*) = \frac{1}{k(k-1)} \sum_{i=1}^{k} (y_i^* - y_{all})$$

Wolter (1985, p.180) and Rust (1985) have presented an extension of these formulae for complex stratified sample designs in which there are k_h primary sampling units in the hth stratum (where h = 1,2,...,H). In this case, the formula for the variance of y_{all} employs y_{hi} to denote the estimator derived from the same functional form as y_{all} – calculated after deleting the ith primary sampling unit from the hth stratum.

$$var(y_{all}) = \sum_{h=1}^{H} \frac{(k_h - 1)}{k_h} \sum_{i=1}^{k_h} (y_{hi} - y_{all})^2$$

where K = Σkh is the total number of samples that are formed

In many educational research studies, schools are used as the primary sampling units, and therefore the K estimates of y_{hi} are obtained by removing one school at a time from the total sample and then applying the same functional form used to estimate y_{all} for each reduced sample. In studies where geographical areas are used

as the primary sampling units, followed by the selection of more than one school per area, the estimates of y_{hi} are based on reduced samples formed by omitting one geographical area at a time.
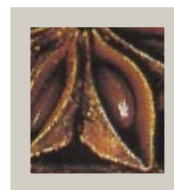
5 References

Brickell, J. L. (1974). "Nominated samples from public schools and statistical bias". *American Educational Research Journal*, <u>11</u>(4), 333-341.

Deming, W. E. (1960). *Sample design in business research*. New York: Wiley.

Finifter, B. M. (1972). "The generation of confidence: Evaluating research findings by random subsample replication". In H. L. Costner (Ed.), *Sociological Methodology*. San Francisco: Jossey-Bass.

Frankel, M. R. (1971). *Inference from survey samples*. Ann Arbor, Michigan: Institute for Social Research.

Haggard, E.A. (1958). *Intraclass correlation and the analysis of variance*. New York: Dryden.

Hansen, M. H., Hurwitz, W. N.; Madow, W. G. (1953). *Sample survey methods and theory* (Vols. 1 & 2). New York: Wiley.

International Association for the Evaluation of Educational Achievement (IEA) (1986). *Technical committee meeting: sample design and sample sizes*. Paper distributed at the Annual General Meeting of the IEA, Stockholm, Sweden.

Jessen, R. L. (1978). "Statistical survey techniques". New York: Wiley.

Kalton, G. (1983). "Introduction to survey sampling" (Sage Series in Quantitative Applications in the Social Sciences, No. 35). Beverly Hills, CA: Sage.

Kish, L. (1965). *Survey sampling*. New York: Wiley.

Kish, L. (1978). "On the future of survey sampling". In N. K. Namboordi (Ed.), Survey sampling and measurement. New York: Academic Press.

McCarthy, P. J. (1966). Replication: An approach to the analysis of data from complex surveys. Washington: United States National Centre for Health Statistics.

Mosteller, F. and Tukey, J. (1968). "Data Analysis Including Statistics". In G. Lindzey and E. Aronson (Eds.), The handbook of social psychology. (2nd ed.). Reading, Massachusetts: Addison-Wesley.

Quenouille, M. J. (1956). "Notes on Bias in Estimation". Biometrica, 43, 353-360.

Ross, K. N. (1976). Searching for uncertainty: Sampling errors in educational survey research. Hawthorn, Victoria: Australian Council for Educational Research.

Ross, K. N. (1978). Sample design for educational survey research (Monograph). Evaluation in Education, 2, 105-195.

Ross, K. N. (1987). Sample design. International Journal of Educational Research, 11(1), 57-75.

Ross K. N. (1991). Sampling Manual for the IEA International Study of Reading Literacy. Hamburg: IEA International Study of Reading Literacy Coordinating Centre.

Ross, K. N. and Postlethwaite, T. N. (1992). "Indicators of the Quality of Education: A National Study of Primary Schools in Zimbabwe". Paris: International Institute for Educational Planning.

Rust, K. (1985). Variance estimation for complex estimators in sample surveys. Journal of Official Statistics, 1(4), 381-397.

Sylla, K.; Saito, M. and Ross,K. (2004). SAMDEM: Sample Design Manager. Paris. International Institute for Educational Planning.

Tukey, J. W. (1958). Bias and confidence in not-quite large samples (Abstract). Annals of Mathematical Statistics, 29, 614.

Wolf, R. M. (1977). Achievement in America: National report of the United States for the International Educational Achievement Project. New York: Teachers College Press.

Wolter, K. M. (1985). Introduction to Variance Estimation. New York: Springer-Verlag.

Appendix 1

Random number tables for selecting a simple random sample of twenty students from groups of students of size 21 to 100

Size of Group

R21	R22	R23	R24	R25	R26	R27	R28	R29	R30
1	1	1	2	1	1	1	1	1	4
2	2	2	3	2	2	2	2	2	6
3	3	3	4	3	3	3	3	3	7
4	5	5	5	4	5	4	4	4	8
5	6	6	7	5	6	6	6	5	10
6	7	7	8	6	7	7	7	6	12
7	8	9	9	7	8	8	8	9	15
8	9	11	10	9	9	11	11	10	16
9	10	12	11	11	10	12	12	11	17
10	11	13	12	12	12	13	13	12	18
11	12	14	13	13	13	15	15	14	19
12	13	15	15	14	14	17	17	16	20
13	14	16	16	15	15	18	18	21	22
14	15	17	17	16	16	19	19	22	23
16	16	18	18	18	18	20	20	23	24
17	17	19	19	20	19	21	21	24	25
18	18	20	21	21	21	22	22	25	26
19	20	21	22	23	22	25	25	27	27
20	21	22	23	24	25	27	27	28	29
21	22	23	24	25	26	28	28	29	30

Size of Group

R31	R32	R33	R34	R35	R36	R37	R38	R39	R40
1	4	1	2	1	4	1	1	1	2
4	6	2	3	2	5	3	4	2	3
5	7	4	4	7	7	6	7	3	5
8	9	5	5	8	8	7	9	5	6
10	10	7	6	9	10	8	11	6	8
11	11	10	710	11	13	9	12	7	12
14	12	12	12	14	16	10	14	8	13
15	13	13	13	15	17	11	16	12	16
16	14	14	14	16	19	12	17	14	17
17	15	16	16	17	21	13	18	15	18
19	16	18	18	21	22	17	19	17	23
20	17	19	19	22	24	18	22	20	24
22	21	21	21	23	25	19	23	24	25
23	23	23	23	24	26	24	24	26	31
24	24	25	25	25	27	25	25	27	33
25	25	26	26	26	28	28	26	28	34
27	26	28	28	32	29	29	28	29	37
29	28	31	31	33	31	30	31	31	38
30	30	32	32	34	33	34	32	36	39
31	31	33	33	35	36	36	37	38	40

Size of Group

R41	R42	R43	R44	R45	R46	R47	R48	R49	R50
1	2	1	1	3	1	3	1	1	3
4	3	2	3	5	11	5	3	2	4
5	4	6	4	7	12	10	6	4	5
7	5	11	6	8	15	17	15	6	6
8	6	13	7	10	17	24	19	8	7
10	8	14	8	11	21	25	20	13	8
12	10	15	9	12	22	27	24	15	11
13	13	17	11	16	23	28	26	17	13
15	16	20	13	19	24	30	27	20	14
17	20	25	15	23	25	31	29	23	16
20	21	26	18	24	26	32	30	24	20
21	23	28	19	25	28	33	31	27	22
25	24	31	24	32	29	34	32	31	24
27	27	32	25	34	32	36	34	32	29
28	29	33	32	35	34	37	36	34	32
29	31	37	35	37	36	38	37	37	33
32	34	38	38	38	40	39	38	41	36
34	35	39	40	40	41	42	40	43	37
35	36	40	43	43	44	44	44	44	40
39	41	42	44	44	45	45	48	47	45

Size of Group

R51	R52	R53	R54	R55	R56	R57	R58	R59	R60
3	4	3	1	1	2	2	3	3	1
4	5	6	2	6	5	6	5	4	5
8	6	8	4	7	7	7	6	5	6
9	10	9	8	9	10	10	8	8	7
10	13	11	10	10	12	17	16	10	9
12	17	12	11	12	15	19	17	11	11
13	18	14	16	14	18	23	21	15	14
15	19	15	20	16	24	29	23	16	16
17	20	18	24	29	25	34	27	17	19
26	22	19	25	31	27	35	32	20	21
27	25	23	27	36	32	37	37	24	24
29	27	26	28	38	34	38	38	26	28
31	28	28	29	40	38	43	42	29	29
32	32	33	32	42	43	44	44	31	39
36	35	38	37	44	45	46	46	36	40
39	40	42	39	47	47	48	48	38	49
40	42	43	41	48	49	52	49	42	50
44	49	48	46	52	50	53	50	48	51
47	51	51	47	53	52	55	52	50	52
48	52	53	49	54	55	57	56	53	54

Size of Group

R61	R62	R63	R64	R65	R66	R67	R68	R69	R70
2	9	8	5	3	1	4	2	6	5
5	12	12	8	4	2	5	4	9	6
7	14	14	10	7	8	6	5	11	7
8	22	15	17	10	9	10	10	14	9
11	23	16	19	14	11	14	16	23	11
18	25	18	21	18	19	17	17	24	13
19	227	24	25	19	22	18	21	29	17
21	28	32	26	20	23	20	22	31	18
22	29	33	27	28	25	21	23	36	21
27	33	34	29	37	26	25	28	40	22
30	41	38	33	40	27	29	31	44	36
34	42	39	35	42	30	30	32	48	39
35	43	40	37	46	28	34	33	49	41
39	46	42	39	47	46	37	37	50	47
46	50	45	45	49	50	39	42	52	49
48	51	46	48	51	52	41	49	55	50
49	53	48	49	58	53	47	55	59	53
50	57	54	54	59	54	57	58	60	55
53	59	61	58	61	61	58	61	67	61
54	62	63	61	65	66	63	67	68	67

Size of Group

R71	R72	R73	R74	R75	R76	R77	R78	R79	R80
9	3	1	4	2	6	3	12	8	6
15	6	5	15	7	20	5	13	10	10
17	10	9	19	9	21	6	23	14	17
20	12	11	23	15	22	7	24	15	19
22	18	12	24	16	24	10	28	21	21
25	20	13	27	23	25	16	32	23	23
27	22	14	28	29	30	24	46	31	26
28	26	21	35	36	31	31	47	32	31
34	29	28	38	37	35	32	48	24	32
35	32	29	42	41	37	36	49	41	36
37	38	34	47	43	39	38	52	46	41
39	41	37	49	45	50	45	53	48	42
46	43	42	50	46	51	48	57	54	44
48	44	48	51	50	58	62	59	58	45
50	47	54	52	53	61	64	63	61	46
53	48	57	56	60	63	70	64	62	50
60	50	62	62	66	65	71	67	68	61
61	52	63	64	67	67	72	70	71	71
62	61	68	65	69	68	73	75	75	73
64	66	72	72	72	73	75	76	77	76

Size of Group

R81	R82	R83	R84	R85	R86	R87	R88	R89	R90
9	9	2	4	3	1	1	6	2	1
13	15	4	6	6	8	3	8	12	2
16	16	14	7	7	13	7	16	14	5
33	23	17	13	25	24	14	23	15	10
40	27	30	14	27	34	16	26	39	22
41	28	35	19	30	35	17	29	48	24
42	29	41	25	32	39	19	32	52	25
44	33	42	31	36	45	20	35	56	27
45	42	49	34	45	47	26	42	57	31
46	43	52	39	48	50	28	43	60	35
54	48	53	40	50	52	30	48	62	40
59	50	58	44	54	61	40	50	64	46
64	56	63	59	58	64	53	55	66	51
71	57	67	62	63	65	54	59	67	53
72	60	69	70	64	66	60	63	68	54
73	61	71	73	66	67	66	64	70	73
74	66	74	74	68	68	77	70	75	81
75	67	76	77	72	75	80	74	76	83
76	69	77	83	83	80	83	81	88	85
79	71	80	84	85	83	86	86	89	90

Size of Group

R91	R92	R93	R94	R95	R96	R97	R98	R99	R100
7	8	1	1	7	2	3	8	3	7
15	15	7	8	9	4	10	11	11	21
20	17	9	9	12	6	26	17	20	22
29	21	19	10	14	9	33	23	29	25
36	25	20	14	18	16	37	25	30	27
38	26	21	15	24	18	39	31	34	35
42	31	22	20	32	21	40	36	38	40
43	34	30	23	35	28	41	38	42	49
49	40	31	36	46	32	48	42	50	50
50	41	32	46	49	47	50	48	51	52
54	49	35	48	54	50	51	54	53	56
58	54	36	57	60	51	53	58	54	57
59	61	40	58	63	63	61	63	60	75
67	69	46	60	64	64	62	70	62	79
73	71	51	61	69	70	64	71	63	81
80	78	62	72	78	73	65	78	68	86
84	79	72	79	87	78	68	89	79	87
85	81	73	81	88	80	70	91	92	92
87	83	76	87	92	89	82	92	94	93
91	84	89	91	95	92	97	93	99	94

Appendix 2

Sample design tables (for roh values of 0.1 to 0.9)

Cluster size	95% confidence limits for means/percentages							
	±0.05s/±2.5%		±0.1s/±5.0%		±0.15s/±7.5%		±0.2s/±10.0%	
b	a	n	a	n	a	n	a	n
roh = 0.1								
1 (SRS)	1600	1600	400	400	178	178	100	100
2	880	1760	220	440	98	196	55	110
5	448	2240	112	560	50	250	28	140
10	304	3040	76	760	34	340	19	190
15	256	3840	64	960	29	435	16	240
20	232	4640	58	1160	26	520	15	300
30	208	6240	52	1560	24	720	13	390
40	196	7840	49	1960	22	880	13	520
50	189	9450	48	2400	21	1050	12	600
roh = 0.2								
1 (SRS)	1600	1600	400	400	178	178	100	100
2	960	1920	240	480	107	214	60	120
5	576	2880	144	720	65	325	36	180
10	448	4480	112	1120	50	500	28	280
15	406	6090	102	1530	46	690	26	390
20	384	7680	96	1920	43	860	24	480
30	363	10890	91	2730	41	1230	23	690
40	352	14080	88	3520	40	1600	22	880
50	346	17300	87	4350	39	1950	22	1100
roh = 0.3								
1 (SRS)	1600	1600	400	400	178	178	100	100
2	1040	2080	260	520	116	232	65	130
5	704	3520	176	880	79	395	44	220
10	592	5920	148	1480	66	660	37	370
15	555	8325	139	2085	62	930	35	252
20	536	10720	134	2680	60	1200	34	680
30	518	15540	130	3900	58	1740	33	990
40	508	20320	127	5080	57	2280	32	1280
50	503	25150	126	6300	56	2800	32	1600

Cluster size	95% confidence limits for means/percentages							
	±0.05s/±2.5%		±0.1s/±5.0%		±0.15s/±7.5%		±0.2s/±10.0%	
b	a	n	a	n	a	n	a	n
roh = 0.4								
1 (SRS)	1600	1600	400	400	178	178	100	100
2	1120	2240	280	560	125	250	70	140
5	832	4160	208	1040	93	465	52	260
10	736	7360	184	1840	82	820	46	460
15	704	10560	176	2640	79	1185	44	660
20	688	13760	172	3440	77	1540	43	860
30	672	20160	168	5040	75	2250	42	1260
40	664	26560	166	6640	74	2960	42	1680
50	660	33000	165	8250	74	3700	42	2100
roh = 0.5								
1 (SRS)	1600	1600	400	400	178	178	100	100
2	1200	2400	300	600	134	268	75	150
5	960	4800	240	1200	107	535	60	300
10	880	8800	220	2200	98	980	55	550
15	854	12810	214	3210	95	1425	54	810
20	840	16800	210	4200	94	1880	53	1060
30	827	24810	207	6210	92	2760	52	1560
40	820	32800	205	8200	92	3680	52	2080
50	816	40800	204	10200	91	4550	51	2550
roh = 0.6								
1 (SRS)	1600	1600	400	400	178	178	100	100
2	1280	2560	320	640	143	286	80	160
5	1088	5440	272	1360	122	610	68	340
10	1024	10240	256	2560	114	1140	64	640
15	1003	15045	251	3765	112	1680	63	945
20	992	19840	248	4960	111	2220	62	1240
30	982	29460	246	7380	110	3300	62	1860
40	976	39040	244	9760	109	4360	61	2440
50	973	48650	244	12200	109	5450	61	3050

Cluster size	95% confidence limits for means/percentages							
	±0.05s/±2.5%		±0.1s/±5.0%		±0.15s/±7.5%		±0.2s/±10.0%	
b	a	n	a	n	a	n	a	n
roh = 0.7								
1 (SRS)	1600	1600	400	400	178	178	100	100
2	1360	2720	340	680	152	304	85	170
5	1216	6080	304	1520	136	680	76	380
10	1168	11680	292	2920	130	1300	73	730
15	1152	17280	288	4320	129	1935	72	1080
20	1144	22880	286	5720	128	2560	72	1440
30	1136	34080	284	8520	127	3810	71	2130
40	1132	45280	283	11320	126	5040	71	2840
50	1130	56500	283	14150	126	6300	71	3550
roh = 0.8								
1 (SRS)	1600	1600	400	400	178	178	100	100
2	1440	2880	360	720	161	322	90	180
5	1344	6720	336	1680	150	750	84	420
10	1312	13120	328	3280	146	1460	82	820
15	1302	19530	326	4890	145	2175	82	1230
20	1296	25920	324	6480	145	2900	81	1620
30	1291	38730	323	9690	144	4320	81	2430
40	1288	51520	322	12880	144	5760	81	3240
50	1287	64350	322	16100	144	7200	81	4050
roh = 0.9								
1 (SRS)	1600	1600	400	400	178	178	100	100
2	1520	3040	380	760	170	340	95	190
5	1472	7360	368	1840	164	820	92	460
10	1456	14560	364	3640	162	1620	91	910
15	1451	21765	363	5445	162	2430	91	1365
20	1448	28960	362	7240	162	3240	91	1820
30	1446	43380	362	10860	161	4830	91	2730
40	1444	57760	361	14440	161	6440	91	3640
50	1444	72200	361	18050	161	8050	91	4550

Appendix 3

Estimation of the coefficient of intraclass correlation

The coefficient of intraclass correlation (roh) was developed earlier this century in connection with studies carried out to measure 'fraternal resemblance' – such as in the calculation of the correlation between the heights of brothers. To establish this correlation there was generally no reason for ordering the pairs of measurements obtained from any two brothers. Therefore the approach used was to calculate a product-moment correlation coefficient from a symmetrical table of measures consisting of two interchanged entries for each pair of brothers. In the days before computers, these calculations became extremely arduous when large numbers of brothers from large families were being studied. Some computationally simpler formulae for calculating estimates of this coefficient were eventually developed by several statisticians (Haggard, 1958). These formulae were based either on analysis of variance methods, or on the calculation of the variance of elements and the variance of the means of groups, or clusters, of elements.

In the field of educational survey research, the clusters of elements which define primary sampling units often refer to students that are grouped into schools. The value of roh then provides a measure of the tendency for students to be more similar within schools (on some given characteristic such as achievement test scores) than would be the case if students had been allocated at random among schools. The estimate of roh is prepared by using schools as 'experimental treatments' in calculating the between-clusters mean square (BCMS) and the within-clusters mean square (WCMS). The term b refers to the number of students per school.

$$estimated\ roh = \frac{BCMS - WCMS}{BCMS + (b - 1)\ WCMS}$$

If the numerator and denominator on the right hand side of the above expression are both divided by the value of WCMS, the estimated roh can be expressed in terms of the value of the F statistic and the value of b.

$$estimated\ roh = \frac{F - 1}{F + (b - 1)}$$

The application of the following alternative formula which is based upon variance estimates for elements (students) and cluster means (school means) has been discussed by Kish (1965: 176). The term sc2 is the variance of cluster means, and s2 is the variance of element values.

$$estimated\ roh = \frac{b\,s_c^{\,2} - s^2}{(b - 1)\,s^2}$$

Note that, in situations where the number of elements per cluster varies, the value of b is sometimes replaced by the value of the average cluster size.

Since 1992 UNESCO's International Institute for Educational Planning (IIEP) has been working with Ministries of Education in Southern and Eastern Africa in order to undertake integrated research and training activities that will expand opportunities for educational planners to gain the technical skills required for monitoring and evaluating the quality of basic education, and to generate information that can be used by decision-makers to plan and improve the quality of education. These activities have been conducted under the auspices of the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ).

Fifteen Ministries of Education are members of the SACMEQ Consortium: Botswana, Kenya, Lesotho, Malawi, Mauritius, Mozambique, Namibia, Seychelles, South Africa, Swaziland, Tanzania (Mainland), Tanzania (Zanzibar), Uganda, Zambia, and Zimbabwe.

SACMEQ is officially registered as an Intergovernmental Organization and is governed by the SACMEQ Assembly of Ministers of Education.

In 2004 SACMEQ was awarded the prestigious Jan Amos Comenius Medal in recognition of its "outstanding achievements in the field of educational research, capacity building, and innovation".

These modules were prepared by IIEP staff and consultants to be used in training workshops presented for the National Research Coordinators who are responsible for SACMEQ's educational policy research programme. All modules may be found on two Internet Websites: **http://www.sacmeq.org** and **http://www.unesco.org/iiep**.