How Many Should I Ask?:
Determining a Sample Size

There are many different factors that come into play when deciding how many people need to be asked a question or administered a survey. The first, and perhaps most obvious question being: Do I want to be able to measure statistically significant change? Other issues to examine include the type of sampling being used, the form in which the resulting data will be presented, the precision with which one wants to measure change, and the level of error one is willing to except.

ways to decide who will take part in an evaluation. Sometimes, everyone is included in an evaluation. This is called a *census*. However, a census can be very costly and time-consuming, especially when the population is large. Often, evaluations are based on a smaller group within the population: a *sample*. It is not easy to decide how large a sample should be or who should be in it.

Sampling methods depend on the type of data we need (qualitative or quantitative), the method of evaluation (focus groups, household survey, clinical measurement), and characteristics of the population involved. Since sample observations are used to tell us things about the whole group of people we are interested in, samples should be **representative**. That is, the characteristics of the sample should be similar to the population as a whole.

Defining the Population

First, it is important to clearly define the population that the sample group will represent. The population may be all children under five, women, the elderly, or an entire refugee camp. The list of all members of this population is often called a **sampling frame**. The individual members of this sampling frame, be they children, women, refugees or households, are often called **sampling units**, because they are eligible to be selected into the sample group.

Each sampling method described below is simply a way of choosing a smaller number (sample) of individuals or households from the larger, defined population.

Convenience Sampling

The only reason people are selected for a convenience sample is because they are easily available. Asking about the income of the first 10 households on the road in a village is an example of a convenience sample. There is no way we can be sure that their answers represent the community. It could be that the wealthiest people live in that area, or the poorest. This method is low cost, but scientifically weak.

Purposive Sampling

Purposive sampling involves choosing individuals from a population based on certain characteristics. For example, an evaluation team may choose to interview two mothers from low-income households, two mothers from higher-income households, a street vendor, a health worker and an elder because they will each have specific insights into the food security of a community. This method is low-medium cost. It is best suited to subjective-qualitative measurement (focus groups, interviews) where we are interested in the perspectives and experience of particular groups or individuals.

Random Sampling

Random sampling uses statistical probabilities to select a sample. Evaluators can decide how many people will be in the sample but the selection of each individual is left to random chance. Random sampling methods are most commonly used with quantitative data collection. Because sampling units are chosen at random, there is less chance of *bias* – incorrect results due to the exclusion of certain groups or types of people (e.g. surveying only people who live close to the centre of town, those who are in hospital, or those with children in school will not give results that reflect the wider community). There are several types of random sampling.

Simple Random Sampling gives each person the same chance of being chosen. Imagine a bag that contains one bean for each person in a village. Some beans are red and some beans are white. Each person picks a bean. Only those who pick a red bean are included in the sample. Each person has the same chance of being selected (equal to the number of red beans divided by the number of white beans). Other examples of simple random sampling include selecting households on a walk through the village, skipping a certain number of homes each time; writing names on pieces of paper and pulling them from a bag; or using a computer to generate random numbers that identify participants.

Cluster Sampling uses simple random sampling in more than one stage. Clusters may be households, villages, counties, hospitals or any other sub-group of individual sampling units. In the first round, a sample of clusters is chosen. For example: a subset of villages chosen from all the villages in a particular region. Once the clusters have been chosen, we can evaluate all individuals within the cluster or choose a random sample within each cluster. Cluster sampling is often used in large or remote populations.

> **A sample is a smaller group of individuals** *(children, women, refugees, households etc)* **that represents a larger population with similar characteristics.**
>
> **Choosing a sample can save time and money.**

Sound sampling methods reduce the risk of drawing mistaken conclusions from any survey. You will find more information on sample size calculation and selection methods in the resources listed below.

Resources

FANTA Project Sampling Guide. *Title II Indicator Guides, 1999.*

World Vision Canada, Design and Implementation of Nutrition Surveys, an excerpt from the MICAH Guide.

Both are available at: www.foodaidmanagement.org/mne3.htm

For more information on sampling or any other planning, monitoring or evaluation issues, contact the Canadian Foodgrains Bank or email: ***cfgb@foodgrainsbank.ca***