Selecting Samples to Examine in an Evaluation

) many evaluations, it is necessary to select a *sample* (some members) of a population from which to collect data in order to answer an evaluation question. The focus of the evaluation will determine whether the population of interest is human (individuals, households), physical areas (plots of land, districts) or things (samples of manufactured products drawn for quality control purposes).

The most basic choice evaluators make when deciding what approach to use in selecting a sample is whether to rely on judgment to select elements that will represent the population to be examined. The alternative to sampling methods that rely on judgment are lottery type methods that give, or attempt to give, every member of a population an equal chance of being selected.

Samples that rely on judgments about which members of a population to examine (*purposive samples*) are particularly useful when we want to learn about elements that might not be selected if a lottery type method was used. For example:

- Members of minority groups
- Diverse conditions, e.g., micro-climates, cultural differences
- Extreme cases of any sort, e.g., exceptional performers, worst case situations

In these situations, *criteria samples* (must fall in a certain category), *quota samples* (to ensure that rare case or minorities are adequately studied) and other kinds of deliberate selection procedures are the most appropriate choice.

Judgment-based or purposive samples may be exactly right for getting to the members of a population]at an evaluation team wants to interview, but such samples are never considered to be *representative* of .ie population as a whole.

A representative sample is generally defined as one that *has been selected in such a way that, as far as can be ascertained, the main characteristics of the sample match those of the parent population; i.e. the population from which the sample has been drawn.* In a representative sample, minorities, diverse conditions and extreme cases will show up no more frequently than they would in the population as a whole.

The most common way of selecting a sample that is, or attempts to be, representative in the sense that term is used here is a lottery or random sampling approach. How successful random sampling approaches are in representing a population depends in part upon whether the size of the population is known. Technically, a *random sample* is one in which each member of the population has an equal chance of being selected.

True random samples (*probability samples*) can only be drawn in situations where we have a list or count of all of the members of a population, whether that is a list of all of the students in a school, all of the villages in a country or all of the registered businesses in a city. Probability samples, in which every member of the population has a known, nonzero probability of being selected, are the only type of samples for which evaluators and other researchers can calculate the statistical likelihood that any specific member of a population would have been selected.

When the population from which a true random sample is to be selected is large and widely dispersed, evaluators sometimes use techniques for ensuring that the sample they draw is concentrated geographically, without doing any harm to their ability to calculate the statistical likelihood of any member of that population being selected. The best known of these methods is *cluster sampling* which, typically, involves using existing groupings of a population, such as provinces or districts as a first stage in a *multi-stage sampling* process. Since cluster sampling ensures that every province or district has an equal chance of being selected, statistical principles for drawing a representative sample are not violated. Thus, evaluators can randomly sample clusters first and then randomly sample individual units of a population from which to collect data inside the clusters that were drawn. Techniques of this sort help evaluator's address, in a cost-effective manner, questions that call for answers based on a *representative* sample.

Evaluators also face challenges when the size of the population from which it is important to try to draw a representative sample is unknown. In these situations, options that rely on chance rather than evaluator judgment to select which members of a population to examine are useful. A technique called *systematic sampling* often fills this need. In sequential sampling, every "nth" member of a population is selected. Using systematic sampling, an evaluation team might select every 5th person coming into a market on a busy day, or every 10th car traveling on a given road. The evaluator's intent in this type of situation is to be as representative and unbiased as possible, even if a true random sample is not possible. Systematic sampling, it should be noted, is not reserved for situations where the size of a population is unknown. This method is also frequently used when a long list of a known population exists and is most easily sampled by picking a starting point at random and selecting every "nth" item or name from the list.

Another sampling technique that evaluators often find useful, because it serves as a bridge between a desire to look at special populations and the need for representative samples, is *stratified sampling*. Strata are groups or subpopulations which, for sampling purposes, are treated as separate populations. For example, in certain countries one might know that few female doctors would be selected in a representative sample of all doctors. By treating female doctors and male doctors as separate strata or populations and drawing random samples from each, both representativeness and an interest in a special population (female doctors, ethnic minorities) would be satisfied.

When considering sampling methods, it is useful to recognize that methods that rely on judgment (*purposive samples*) and those that rely on random sampling procedures in an attempt to select an unbiased and representative sample, whether or not the size of the population is known, are all better choices that selection procedures that are likely to be biased in ways that the evaluator does not understand. Sampling techniques that fall in this category, and are sub-optimal for that reason, include:

- Convenience samples – choosing places that are easy to reach or close to the capital city or a main road.
- Volunteer samples – where motives may not be understood.
- "Snowball" samples – which are developed by asking one person who they would recommend be interviewed next. While snowball samples may be a good way to find rare cases, i.e., somebody else who happened to be present at a specific event, the chances are high that evaluators will not be fully aware of why certain names are suggested to them as interview candidates.