Sampling and Experimental Designs

Sampling Methods:

Good Types of Sampling: attempts to balance out the variability in the population through random selection of sample participants.

Census: *all* of the population is used
Sample: a subset of the population is used.

Simple Random Sample (SRS): any one person or group of people have an equal chance of being selected (often referred to as a random sample). Gold standard.

Stratified Sample: "some of all"; Some people from all groups; like blocking in an experiment; homogeneous strata (groups) are formed and an SRS is conducted within each strata (group)

Cluster Sample: "all of some"; All people from some groups; an SRS is conduct to select a cluster (homogeneous-like group) and then the entire cluster is surveyed (*census*-like)

Systematic Sample: "every-n^{th} person is selected" by an algorithm (or system);

Bad Types of Sampling: "won't be representative of the entire population" or ways bias can be introduced into the results.

Convenience Sample: closest or first people encountered – convenient for the lazy

Voluntary Sample: radio call-ins, write-ins (letters to congress, editor, etc) – usually leans toward the extremes

Incomplete Frame: sampling leaves out people of the survey (like a telephone survey leaves out those without a phone)

Indimidation: When someone of authority interviews people (they are unlikely to tell the complete truth on a subject)

Nonresponse: small percentage of people respond to survey; generally representing the extremes (people strongly in favor or opposed)

Poorly Worded Questions: can also introduce bias into a survey by use of inflammatory words or statements to lead people into answering a particular way

Bias: a systematic (not necessarily deliberate) favoring/disfavoring of one group or another

Remember the dart board analogy: If you aim towards the center but don't hit it every time- that's variation!! If your aim is off and all your shots are going right – that's bias.

Observational Study: **data is collected from observing things, but no "treatment" is applied to the experimental units. Observational studies can lead to experiments (where treatments are applied) and they can establish correlations between variables but they *cannot establish cause and effect relationships* between variables.**

Experimental Design:

Three Key Design Criteria: Control, Randomization and Replication are all meant to *minimize the variability* encountered in the experiment.

Control: controlling what each subject is treated with

Randomization: assigning the treatments to subjects in a random manner

Replication: the number of repetitions that a treatment has within the experiment

Basic Experimental Design Types:

Completely Randomized: treatments assigned to all subjects randomly

Matched Pair: pairs of subjects are matched (before and after a treatment has been applied, left foot and right foot, husband and wife) and each part of the pair receives a different treatment

Randomized Blocked Design: subjects are blocked (grouped) according a common characteristic that may effect the response variable (like gender, breed, or age). Treatments are assigned randomly within blocks.

Must be able to identify the response variable (what are we measuring), the test subjects (what are the treatments applied to – may not be what we are measuring for a response variable), and what are the treatments being applied in every experiment.

A KEY part is the random assignment of subjects to treatments. If asked to detail a random assignment process, you must layout an algorithm in great detail. This must be detailed enough that a middle school student could follow the algorithm correctly. Random assignment algorithms should ensure that groups (or treatments) are approximately equal in size. Flipping coins or rolling dice generally will not ensure this because of the probabilistic nature of the process.

Blinding:

"double-blind": neither the person administering the subject or the subject knows which treatment is being applied

"single-blind": person receiving treatment doesn't know which treatment they are getting, but the person administering the treatment does know

***placebo*: a "sugar-pill" (in a drug test); acts as a control group to test the effectiveness of the other treatments against a baseline**

***control group*: used when the treatments need to be compared against an unknown baseline (like the baseline level of stress in a company). Often times the current drug or treatment is the baseline against which the new item is measured**