Chapter 2

Survey Research: What is a representative sample?

A representative sample is a group that closely matches the characteristics of its population as a whole. In other words, the sample is a fairly accurate reflection of the population from which the sample is drawn. For example, If the population is 100 students in a class and I want to learn things about them by taking a sample. If 50 students are female in the class, the half of my sample should be female. If 15% of the class is Hispanic, then 15% of my sample should be Hispanic. If20% of my class is rich, then 20% of my sample should be rich. Getting a sample that is representative is enhanced with random sampling – which means that my sample is drawn from the population in an unbiased way. I.e. everyone in the population has a chance of being randomly selected for my sample. This selection process does not guarantee a representative sample, but one is more likely as the size of the sample increases. For example if I randomly select only 10 students from the class, the chance that 5 are female isn't so high. But if I have a sample of 30, the probability that half are female is increased.

If a sample isn't representative, any conclusions I draw from the sample, may not apply to the population. For example assume I want to measure attitudes toward gay marriage of all Americans. Assume 5% of all Americans live in San Francisco. If I have a sample with half of them from San Francisco, my conclusions won't generalize to the population.

Survey Research: What is volunteer bias?

When we construct a sample, we can't force them to participate. Volunteer bias happens when the sample is different from the population in at least one important way. In this case, the sample all volunteered to participate and so they might be different from all the others that refused or didn't volunteer. Volunteer bias is more likely when there is a low return rate. For example, I mail questionnaires out to 1,000 people but only get responses from 20. Those 20 are probably different in some important ways from all the others who refused or didn't return the questionnaire. Some research areas are more susceptible to volunteer bias. Those areas that are not controversial are less susceptible. But areas like asking people to report on illegal activity or sexual behaviors are susceptible. Do you think that people who volunteer to talk about their sex lives have different sex lives than people who refused to talk about their sex lives? Absolutely.

What is a confounding variable in an experiment?

Recall that the logic of the experiment (presented in lecture) is that you start with 2 groups exactly equal (on all important variables) and then do just one thing to make them different. If when you measure at the end, they are now different, it must be because of what you did. You just found the cause. But what happens when you do sloppy work and the groups are different on 2 things instead of just one? Suppose you have one group drink 5 beers and another drink just 1 because you want to see if alcohol impairs memory. But imagine that the group that drinks 5 beers also got together while you were away and they all smoked some pot. The other group didn't get high. Now when you measure memory and if you find that the 5 beer group had worse memory, you aren't sure if it is because of the alcohol difference between the groups or the pot smoking difference. You have allowed pot smoking to become a confounding variable – because it also differs between the groups in addition to the IV (dose of alcohol).