Learning Guide: Sampling Part 2: When Is a Sample “Good Enough”?

Topp, Barker & Degenhardt

1. Why is it difficult, or even perhaps impossible, to obtain a probability sample of a hidden population?

2. Why is it difficult, no matter what sampling approach is used, to ensure that the sample is sufficiently representative of the hidden population to permit generalization of conclusions to the theoretical population as a whole?

3. The authors compare their results from a non-probability sample described on p. 35 to those of a probability sample described on p. 34. What kind of probability sample was used?

4. The authors specifically chose two subsets from the probability sample. What were these subsets and why did they select them?

5. Note that these authors use the term “purposive” sample rather freely. It is not a purposive sample as most methodologists would define the term “purposive sample”. Which of the non-probability sample types did they take?

6. What screening criteria did they use to select participants? Note – they may be confusing using screening criteria with purposive sampling, a common practice that does not make a sample a purposive sample. In fact, screening criteria are used with many if not most probability samples.

7. The authors give descriptive statistics (should be familiar term to you) about the three samples, the two subgroups extracted from the total data base for the probability sample and the non-probability sample. They point out that these descriptive statistics are similar among all three samples. Why would the similarity in descriptive characteristics like age, income, and sex provide evidence that the non-probability sample may well be adequate for generalizing conclusions to the theoretical population?

8. The authors also compare the results about ecstasy use among three samples. They point to differences in the data about drug use collected as potential reasons for the observed differences. However, they also point to the similarities in the data about frequency and type of drug use. Why would the similarity in the data about drug use provide evidence that the non-probability sample may well be adequate for generalizing conclusions to the theoretical population?

9. The authors conclude with these statements. Explain the key points about sampling they are making in your own words.

“The results presented in this paper should provide some reassurance that inferences drawn from such research can reasonably be generalised to the entire population of users. Future related work could usefully examine the degree of external validity of population parameters estimated from ecstasy users sampled using purposive methods different to those described in this paper, or recruited in different cultural contexts. It will also be valuable to determine whether the extent of generalisability demonstrated in the
present results would be observed in comparisons of users of other illicit drugs recruited using distinct probability and non-probability sampling methods."

Summary Questions Base your answers on what you have learned about internal validity, external validity, and explanatory power and what you have learned about sampling.

1. This article deals with the problems of sampling “hidden” populations. Give another example of a population of interest to you that would be a hidden population. My example might be food insecure single mothers.

2. How can you use what these authors learned and concluded to help you evaluate the degree to which you can depend on the research results reported in journals when a non-probability sample was used, such as some study about the hidden population of interest to you?

3. What ideas did you get from the article that would help you conduct a good needs assessment for a hidden population?

Auerswald et al.

1. Why is it difficult, no matter what sampling approach is used, to ensure that the sample is sufficiently representative of a marginalized population to permit generalization of conclusions to the theoretical population as a whole?

2. What are some of the sampling barriers or limitations faced by researchers whose work involves marginalized population?

3. The authors of this article use the term venue-based sampling. This is like the intercept sampling that I will discuss in class – sampling based on going where the participants you want to recruit will be, like going to a NASCAR race to interview NASCAR fans. This article reports on a pilot study, which was followed by the “official” or final study, reported elsewhere. What were the authors’ three objectives for this pilot study related to sampling?

4. Look at the first paragraph in the methods section (p. 134). Explain how this first stage in their procedures is relevant to understanding the context in which your research will be conducted.

5. Three different non-probability sampling techniques were used to identify the youth participants in the interviews. What were they?

6. On p. 133, the authors say that: “The success of venue-based sampling depends heavily on the choice of initial venues. However, the methods for the choice of initial venues for recruitment have not been well established.” How many venues were identified in this pilot study? Why would the procedures used in this pilot study be superior to a researcher simply relying on venues used in other study or asking people like clinic workers to identify venues?

7. The authors also wanted to identify the specific venues that are used by different “crowds.” They report that there are several different crowds in this population and that the crowds do congregate at different venues. From a sampling perspective, why would the procedures used in this pilot study to associate “crowds” and “venues” improve the researcher’s ability to develop robust procedures for taking a probability sample later?
8. How can you apply the authors’ general conclusion to sampling in general? The authors conclude that: “Identifying high-risk youth for interventions through individual level screening can be labor-intensive and ineffective, particularly if such screening takes place in clinic settings that high-risk youth might not access. The identification of high-risk venues could prove a more effective way of accessing youth for sampling, testing, implementing interventions, and delivering services.”

**Summary questions** Base your answers on what you have learned about internal validity, external validity, and explanatory power and what you have learned about sampling.

1. This article deals with the problems of sampling marginalized populations. Give another example of a population of interest to you that would be a marginalized population. My example might the homeless.

2. How can you use what these authors learned and concluded to help you evaluate the degree to which you can depend on the research results reported in journals when a non-probability sample was used, such as some study about a marginalized population of interest to you?

3. What ideas did you get from the article that would help you conduct a good needs assessment or implement a program to serve a marginalized population?