

Brooks Morgan

PTC 604 - Communication Theory and Research

7 October 2008

This document describes the sampling plan for my PTC 604 research proposal. In order to provide context for my proposal, the document has been divided into three sections:

- Sampling Overview – provides an introduction to samples and explains the various methods for determining samples (probability sampling and nonprobability sampling)
- Sampling for Usability Testing – provides a brief description of usability testing, including information about how sample sizes are determined for traditional usability testing and whether or not sample sizes are relevant for usability testing
- PTC 604 Proposal – synthesizes the information in the first two sections into a sample plan for my PTC 604 research proposal

### **Sampling Overview**

In terms of social research, sampling is the process of selecting individuals or groups from a population of interest. The sample is studied with the intent to generalize the results back to the population from which the sample was chosen (Trochim). The ability to study a subset of a population, gather valid data, draw conclusions from the data and then extend the conclusions to the entire population is the cornerstone of social research. Financial and time constraints typically ensure that studying an entire population is not feasible. For example, if researchers are interested in studying a

particular aspect of single, white males between the ages of 40-50 who live in rural areas of the United States, the research team will have to study a representative sample of the population because gathering data on the entire population of single, white, American males between 40-50 who live in rural areas is not a realistic possibility.

The initial steps of a research project often include a sampling plan, which ensures that a selected sample is a valid representation of the population being studied. If research is conducted on an invalid or biased sample, the research data is suspect and the ability to generalize the conclusions is compromised. The sample plan defines the sampling frame, which lists the accessible population from which the research will draw a sample. The sample plan then narrows the sampling frame to a precise sample, which indicates the group of people who will participate in the study. Samples may be selected via two methods: probability and nonprobability. In a probability sample, or unbiased sample, each individual has exactly the same chance as every other individual of being selected (Bernard, 147). Probability sampling relies on one of many random selection processes:

- *simple random sampling* assigns a number to each member of the population being studied and then relies on a computer's random number generator or a random number table to select a sample from the list of assigned numbers.
- *systemic random sampling* is often used for populations whose size precludes simple random sampling (e.g. when a population is too large to assign a number to each member). Instead, systemic random sampling defines a sampling interval,  $N$ , enters the sampling frame at a randomly selected location and includes every  $N$ th person in the sample.

- *stratified random sampling* is used to ensure that samples include specific subpopulations. The population is divided (“stratified”) based on key dependent variables (e.g. age, sex, marital status) and then a random sample is taken from each subpopulation.
- *cluster sampling* is primarily used for efficiency purposes. Cluster sampling narrows the sampling field down from large, heterogeneous chunks to small, homogeneous ones that are relatively easy to sample directly (Bernard, 154).

For research that attempts to draw conclusions from a sample and generalize the results back to the population from which the sample was chosen, probability sampling is the best choice. However, social research may include circumstances where it is not practical to do random sampling. For example, some research requires input from expert informants instead of randomly selected samples. In situations such as these, one of the following nonprobability sampling techniques may be used:

- *quota sampling* selects samples non-randomly based on a fixed quota (e.g. the parameters of samples are initially established and data collection on subjects continues until the sample parameters are satisfied).
- *purposive sampling* requires the researcher to seek specific, pre-defined groups for the research. Purposive sampling is similar to quota sampling, but does not include a predefined number of subjects for samples.
- *convenience sampling* uses whomever is most convenient for the study (e.g. “person on the street” interviews)
- *snowball sampling* begins by identifying someone who meets the criteria for inclusion in a study. Researchers collect data from the initial person

and then ask them to recommend others who they may know who also meet the criteria.

After selecting an appropriate sampling method, researchers must consider an appropriate sample size. Volumes of information are dedicated to properly determining sample sizes for social research projects, but two factors consistently determine whether or not a sample is representative:

- *sample accuracy* – verifies a sample is truly unbiased by ensuring that every member of a population has an equal chance of being selected
- *sample precision* – increased sizes of unbiased samples result in more precise data

### **Sampling for Usability Testing**

Usability testing and social science research are both “empirical” methods, that is, they both focus on observations of actual behaviors (Redish, 35). However, the similarities must be cautiously considered, as the two fields are vastly different in many ways. The major goal of social science research is to test whether or not some phenomenon exists; the test must be done with a sample large enough to detect if the phenomenon is present. The goal of a usability test is to uncover problems with a document or product (Redish, 36). In research studies, researchers go to elaborate lengths to ensure that they have a random sample of participants from some relevant population. In usability testing, a purposive sample often meets the needs of the test. However, much research has gone into determining appropriate sample sizes for usability tests.

The earliest use of a mathematical formula for justifying a sample size in usability studies was by J.R. Lewis in 1982. Lewis, recognizing the need to balance project

concerns with the questionable return-on-investment in usability studies, proposed the following formula to estimate usability test sample size (n):

$$\text{Proportion of unique problems found} = 1 - (1 - p)^n$$

where:

p= mean problem discovery rate computed across test participants

In 1991, R.A. Virzi critically examined the work done by Lewis. Virzi's work made three claims regarding sample size for usability studies (Virzi, 459):

- observing four or five participants allows testers to discover 80% of a product's usability problems (as long as  $p$  for a study is in the approximate range of .30 to .40).
- observing additional participants reveals fewer and fewer new usability problems
- observers detect the more severe usability problems with the first few participants

### **PTC 604 Proposal**

My PTC 604 research proposal will not be a traditional social research study; instead, my project will be a documentation usability study in the tradition of Janice Redish.

Several months ago, my manager asked me to lead the documentation effort for a particular product. At the time, the product was in the initial planning phase. In the design phase of product development, a team of 30 software engineers will collaboratively author ~20 specification documents (some of which are for internal-only audiences while others are customer-oriented). I was asked to develop two documentation deliverables to support the product design team:

- a customized set of FrameMaker templates that software engineering team members will use to write the specification documents.
- a FrameMaker tutorial that provides introductory lessons on how to use FrameMaker and how to use the customized templates (existing variables, conditional text tags, tables, figures etc.)

In lieu of observed usability testing, I propose to distribute a Likert-type survey to all 30 engineers. The survey will include questions that are designed to measure the usability/effectiveness of each chapter in the tutorial. I plan to investigate the following construct:

*“The FrameMaker tutorial content was correct and well-organized; the tutorial also taught me how to properly use certain features in the FrameMaker templates, such as conditional text, variables, and tables/figures.”*

The team of 30 engineers will be my sample, which is nonprobability purposive in this case. Additionally, I may conduct interviews with select representatives from the software development team to gather qualitative information about the usefulness of the tutorial.

## Works Cited

Bernard, H. Russell. *Social Research Methods: Qualitative and Quantitative Approaches*. Thousand Oaks: Sage Publications, Inc, 2000. Print.

Dumas, Joseph S., and Janice Redish. *A Practical Guide to Usability Testing*. Bristol, UK: Intellect Ltd., 1999. Print.

Virzi, R.A. "Refining the test phase of usability evaluation: How many subjects is enough?" *Human Factors* 34 (1992): 457-468. Print.

Note:

<I also need to include a reference to the Trochim site, but I'm not sure how to cite Web sites only in the appropriate format>