Assessing Limitations and Uses of Convenience Samples: A Guide for Graduate Students

S. David Kriska[1], Marcia M. Sass[2], and Mark C. Fulcomer[3]
[1] Walden University, 3758 Surrey Hill Place, Upper Arlington, OH 43220.
[2] UMDNJ (Rutgers University). 48 Trainor Circle, Bordentown, NJ 08505
[3] Restat Systems, Inc., 48 Trainor Circle, Bordentown, NJ 08505

Abstract
Graduate students in applied areas such as psychology and public health often collect data for their dissertation projects and are typically faced with constraints that result in their samples being viewed as mere convenience samples. In describing their sampling plan and its limitations, students often struggle to evaluate the quality of their sampling procedures. This poster takes the view that all samples of human participants are convenience samples to some degree, if for no other reason than ethical considerations make participation voluntary and financial limits make pure random sampling exorbitantly expensive. According to this view, psychology and other applied fields are disciplines built on convenience samples. In spite of the use of convenience samples, applied statistics and data analysis procedures are useful in making advances in applied research. Because some convenience samples may be better than others, this poster session will examine factors and issues in sample selection. The aim is initiate discussion that will result in a framework that graduate students can use to address how to generalize their results.

Key Words: Sampling, sampling frames, graduate research, non-random samples, external validity

1. The Graduate Student Sampling Problem

Graduate students conducting empirical quantitative research face the task of selecting a sample to provide data that will help answer the research question they have identified. In their course work, they have studied simple random sampling and also know about more complex sampling strategies. Importantly, they will also have been warned about the danger of using convenience samples, such as judgment, snowball, or any non-probability sample. They also have studied research design and are aware of threats to internal and external validity. They also know of the importance of addressing expected effect sizes, statistical power, sample size, alpha, and the relationship among these statistical constructs as discussed by Austin, Boyle, and Lualhati (1998). Nevertheless, when faced with actually having to collect data, their sampling plans appear to be plagued by being overly convenient.

Hays (1973) emphasizes that researchers should not take lightly the assumption of random sampling that is a premise for many statistical procedures. Indeed, researchers reporting statistical inference results act as if their sample was selected randomly even though often it is not truly random. How can the students take seriously Hays's advice to realize that they are acting as if they have obtained random samples from some well defined population (Hays, 1973, p. 292) when their sampling plan is a low-cost, voluntary research study?

1.1 Do Convenience Samples Have Value?

First, we believe that non-random or convenience samples can have value, but are not as valuable as probability-based samples. As support, we note the work of two preeminent statisticians. Deming (1966) extols the virtues of probability samples and does say that judgment samples are not amenable to statistical analysis (p. 11), but his use of the term "statistical analysis" appears focused on examining standard errors associated with parameter estimates in population surveys. He goes on to comment that judgment samples are useful for examining biases and also notes that judgment samples can deliver useful results even though understanding them is difficult. Deming most importantly notes the importance of understanding the strengths and weakness of judgment samples (p. 12).

While Tukey (1977) notes that the development of inferential statistics is one of the great intellectual achievements of the 20th century, his text, *Exploratory Data Analysis*, examines ways to study batches of data and he specifically defines a batch as not a random sample. A batch is a set of data that has relevance to a research question. Review of Tukey's work shows there can be value in analyzing quantitative data even when the data set is not a truly random sample of a population.

1.2 Research on Non-random Samples

Even though the literature in many fields contains numerous applications of statistical inference based on a probability structure that is often best described as simple random sampling, the research is actually based on sampling plans that capitalize on convenience. No doubt, non-random samples are common. In an acknowledgement of the pervasiveness of non-random sampling, McCready (2006) proposes an alternative interpretation of sampling error. Rather than thinking of the standard error as the standard deviation of a theoretical statistic whose values might arise through repeated sampling, he chooses to think of the standard error as the standard deviation of the statistic in repeated non-random sampling. In another acknowledgement of non-random sampling, Oleson and Arkin (2006) raise the question of how well do sample participants represent the population the researcher claims they do. They note that all research is flawed and researchers need to be most concerned about the big deficiencies and errors.

Researchers are clearly aware of their use of samples that are not simply random. For example, Bhutta (2012) discusses snowball sampling using Facebook. The student researcher will see Facebook as providing an opportunity to obtain a large sample of data inexpensively, but not without raising questions of possible bias. Heckathorn (1997, 2002) and Salganik and Heckathorrn (2004) address respondent driven sampling, a form of snowball sampling, as a means of identifying individuals in hidden populations, e.g., people with AIDS. Sampling from populations in which individuals are hard to identify can justify clever approaches, but the role of the simple random sampling model in analyzing data is unclear . In another example, Kittleson (2003) writes about suggestions for using the Web to collect data. By sampling users of the Web, some individuals in the general population are excluded from appearing in the sample. Again, question are raised about using the simple random sampling model in data analysis.

Perhaps one of them of the most widely admitted uses of biased samples is in the area of employment test validation. There has been much discussion over many years of a range restriction bias in the in the applied psychology literature, e.g., Barrett, Phillips, and Alexander (1981) and Sackett and Yang (2000). The range restriction is a common issue and several models are used to estimate a correlation in the unrestricted population based on the sample correlation obtained in the restricted sample.

In sum, we believe there is no doubt that non-random samples are collected and analyzed with models that incorporate simple random sampling. We believe the researcher needs to address the potential bias.

2. Sampling Perspective: Suggestions for Students

In making suggestions for students, we first propose stepping back and noting how knowledge is accumulated. Then we make some specific suggestions for students to consider when designing their study. Finally, we asked participants at these ASA meetings to offer comments and suggestions. Their comments are included below.

2.1 Accumulating Knowledge

To see the value of convenience samples, we examine how knowledge is accumulated. Coombs, Dawes, and Tversky (1970), see Figure 1, summarize the model building process in scientific investigation. Understanding the real world is the focus of science. Because the world is too complex to explain completely, scientists develop abstractions and formulate relatively simple models of the real world. To test the validity of the model, the researcher derives a prediction using the model and compares the prediction to empirical data where the data come from the real world. When the data and prediction do not agree, the researcher goes back to the model and modifies it. Coombs et al. note that models can only be rejected; models are not proved by the data. Hence, models are always in need of improvement.

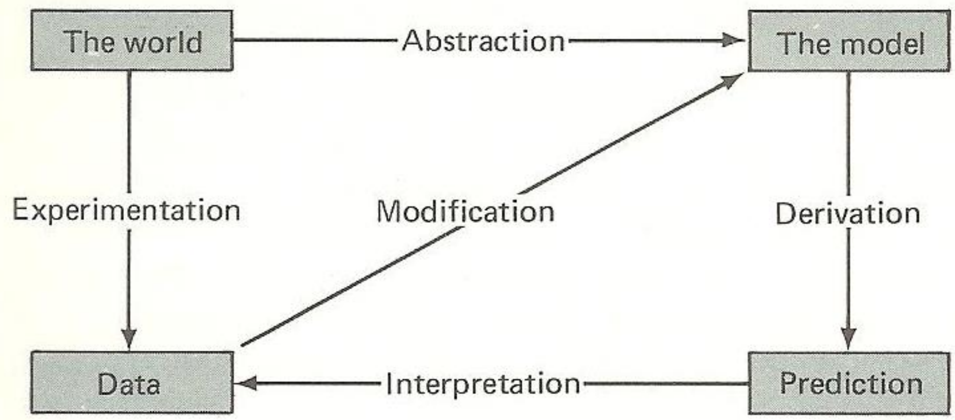

Fig. 1.1 Schematic illustration of a scientific investigation.

Reprinted with permission from Prentice-Hall. Based on a figure in Coombs, Raiffa, and Thrall, 1954 that is now in the public domain.

Figure 1: Scientific Investigation (From Coombs, Dawes, and Tversky (1970).

For the student researcher, the Coombs, et al. (1970) model can be a useful reminder that the thesis or dissertation does not need to be a grandiose advancement in current knowledge. While an aim may be to generalize to all of humanity, a more reasonable approach may be to generalize to one subset and then through additional research to a second subset, and then to another. As the accumulation of data lends support to the model, the viability of the model is enhanced. As data lead to the rejection of the model, it is modified and subjected to further testing. The basic point is that the accumulation of knowledge can be accomplished by taking small steps and testing models in a variety of situations.

2.2 Conducting the Research

While admitting that virtually all sampling plans contain an element of convenience, but at the same time, trying to obtain as valuable of a sample as possible, what can the student do? To begin, Bracht and Glass (1968) suggest distinguishing between the target population to which generalizations are to be made and the accessible population from which data will be drawn. Others may call the accessible sample a sampling frame, e.g., Hahn & Meeker, 1993. The student needs to recognize the risk associated with making inferences from the sample to the accessible population and then to the target population.

In addition, we think students should consider using a good secondary data source as the basis for their study. If the data collection was carefully performed and the variables in the data file allow research questions to be answered, the approach can be quite cost-effective for the student. One potential issue that must be addressed, however, is about the relevancy of the data to the research questions. Moreover, accessing a secondary data set does not absolve the student from the responsibility of evaluating data quality. There is a good chance, though, that the secondary data set was formed with more resources than the student has available.

For students who must collect data, here are our suggestions:

1. Be specific in identifying and describing the target population and accessible population or sampling frame. Students should describe these two populations in detail, and should be clinically objective in describing how the target may differ from the accessible population.

2. Once the target and accessible populations have been identified, the student needs to describe how elements in the accessible population, or sampling frame, are selected. In practice, these procedures include recruitment methods, screening procedures, IRB withdrawal options, compensation amounts, and other steps that help determine who ends up in the sample. Snowball, cluster, quota, and other methods may be involved. These non-probability sampling methods are less desirable than probability samples, but sometimes unavoidable. As noted at the *Joint Meetings* in Montréal, the recruitment of friends to be participants in one's study is an especially weak method. In practice, self selection will often be a part of research involving human subjects. Any relevant, potential biases that may be attributed to sample selection should be noted.

3. We also suggest the student review relevant literature keeping an eye on the state-of-art for relevant field and report on these current practices. For example, in organizational psychology, studies are published based on data collected within a single organization or physical facility and statements are made that are presumed to be applicable beyond the sampling frame. We should also note that we believe most researchers are willing to generalize beyond the sampling frame, but much subjectivity is involved in determining how far beyond is reasonable. The bottom line is that student researchers should describe the leap from the accessible population to the target population.

4. Related to reviewing sampling literature on the content of the study, the student should also review statistical literature related to sample selection. As noted above, there is literature on snowball, cluster, and quota sampling along with other analytic models such as those addressing range restrictions.

5. The student should build in randomness whenever possible. If an experiment is being conducted, random assignment to treatments is imperative. If a quasi-experiment is being performed involving intact groups, students should attempt to randomly select from the each group to the extent practical.

6. In analyzing data, the student needs to compare the sample to the target population and other research. For example, report demographic statistics for the student's sample and compare these results to census or other research data. For other variables, make comparisons to other

researchers' results for studies in which the same variables have been used.

7. In writing the results and conclusion chapters, the student should take a conditional event perspective. Students should view the results section as a description of the data given the model, and the conclusion as the implications of the study given the data. When a null hypothesis is rejected, the researcher is saying the data are unlikely given the model of no effect. In drawing a conclusion about the target population, the inferences become more difficult. Potential sampling biases work to minimize the generalizability of sample results, and the student needs to carefully evaluate the possibility of sampling bias and evaluate the potential limitations. On the other hand, the student should realize that a result in one sampling frame may be important and suggest the need for further researcher in other sampling frames. The researcher should use the Coombs et al. (1970) model and note that as results are replicated in a variety of settings, knowledge is accumulated.

3. Conclusions

We see a gap in what students are taught about sampling in their courses and the practice of sampling when doing graduate research. As noted in their class work, random sampling is a powerful procedure that allows researchers to draw conclusions about populations using only sample data. In practice, however, true random samples are rare and to some extent, all samples are convenience samples. When working with humans, for example, ethical procedures allow participants to withdraw from a study resulting in the exclusion of these individuals. Clearly, selecting a random sample is not a trivial task.

We believe the advice of Bracht and Glass (1968) is still relevant. They divide external validity into two categories, population validity and ecological validity. While ecological validity is concerned with challenges related to moving from the experimental environment to the real world, population validity is focused on moving from the sample to the experimentally accessible population and then to the target population. They emphasize the importance of having thorough knowledge of characteristics of both populations. To the extent the two populations are not identical, the results of a study may be limited. They go on to discus the interaction of personological variables and treatments. Their example is one in which a programmed text for English grammar was better than classroom instruction for high ability students, but the reverse result was obtained for low ability students. They note the importance of careful describing both the accessible and target populations and provide interesting examples of the two not matching.

In sum, our suggestions to students are practical. We begin by suggesting the student keep the big picture in mind. In some instances, students may be able to used a secondary data source to answer their research questions. We also suggest a model building perspective. Understanding model building, testing, and revising may allow the student to work with a subset of the total population. Knowledge can be gained when a model is tested in one subpopulation and then subsequently extended by working with another subpopulation. Most importantly, though, we encourage students to be clinically critical and complete in examining their own work. Knowing the literature and state-of-the-art practice may provide some comfort.

References

Austin, J. T., Boyle, K. A., & Lualhati, J. C. (1998), "Statistical conclusion validity for organizational science researchers: A review," *Organizational Research. Methods,* 1:2, 164-208, doi: 10.1177/109442819812002

Barrett, G. V., Phillips, J. S., & Alexander, R. A. (1981), "Concurrent and predictive validity designs: A critical reanalysis," *Journal Of Applied Psychology*," 66:1, 1-6. doi:10.1037/0021-9010.66.1.1

Bhutta, C. B., (2012), "Not by the book: Facebook as a sampling frame," *Sociological Methods & Research 41*, 57-88. doi: 10.1177/0049124112440795

Bracht , G. H. & Glass. G. V. (1968), "The external validity of experiments," *American Educational Research Journal,* 5: 4, 437-474, http://www.jstor.org/stable/116199

Coombs, C. H., Dawes, R. M., & Tversky, A. (1970), *Mathematical Psychology: An Elementary Introduction*, Englewood Cliffs, NJ: Prentice-Hall.

Deming, W. E. (1966), *Some Theory of Sampling,* New York: Dover.

Hahn, G. H. & Meeker, W. Q. (1993), "Assumptions for statistical inference," *The American Statistician,* 47: 1, 1-11. doi.org/10.1080/00031305.1993.10475924

Hays, W. L. (1973), *Statistics for the Social Science*, 2nd *ed.*, New York: Holt, Rinehart and Winston.

Heckathorn, D.D. (1997), "Respondent driven sampling: A new approach to the study of hidden populations," *Social Problems,* 44: 1, 174 199, http://www.jstor.org/stable/3096941

Heckathorn, D.D. (2002), "Respondent-driven sampling II: Valid population estimates from chain-referral samples of hidden populations," *Soc*ial *Problems,* 49: 1, 11 34, http://www.jstor.org/stable/10.1525/sp.2002.49.1.11

Kittleson, M. J. (2003), "Suggestions for using the web to collect data, "*American Journal of Health Behavior,* 27, 170-72.

McCready, W. (2006), "Applying sampling procedures," In F. Leong, & J. Austin (Eds.), *The psychology research handbook: A guide for graduate students and research assistants.* (2nd ed., pp. 147-161), Thousand Oaks, CA: SAGE Publications, Inc., doi: 10.4135/9781412976626.n10

Oleson, K., & Arkin, R. (2006), "Reviewing and evaluating a research article," In F. Leong, & J. Austin (Eds.), *The psychology research handbook: A guide for graduate students and research assistants.* (2nd ed., pp. 59-75), Thousand Oaks, CA: SAGE Publications, Inc. doi: 10.4135/9781412976626.n4

Sackett, P. R., & Yang, H. (2000), "Correction for range restriction: An expanded typology," *Journal Of Applied Psychology,* 85:1, 112-118, doi:10.1037/0021-9010.85.1.112

Salganik, M. J. and Heckathorn, D. D. (2004), "Sampling and estimation in hidden populations using respondent-driven sampling," *Sociological Methodology,* 34, 193-239. http://links.jstor.org

Tukey, J. W. (1977), *Exploratory Data Analysis*. Reading, MA: Addison-Wesley.