



Commentary

Qualitative research sample design and sample size: Resolving and unresolved issues and inferential imperatives

Robert T. Trotter II*

Department of Anthropology, 575 East Pine Knoll Drive, Northern Arizona University, Flagstaff, AZ 86011, USA

ARTICLE INFO

Available online 16 July 2012

Keywords:

Qualitative sampling
Qualitative sample size
Qualitative sampling models
Qualitative methods

Introduction

Originally there was a wide gap between quantitative and qualitative research design and methodological discussions. Researchers can follow the original qual/quant debate by reviewing the methods and project design discussions from about 1990 to 2000 representing the epistemological and inferential imperatives of both camps (Johnson, 1990). However, qualitative research methodologies have evolved extensively in the recent past, producing scientifically defensible theory, sampling strategies, and analytical strategies. Today the qual/quant gap has narrowed to the point that there can be a seamless transition between qualitative and quantitative research designs, or, even more importantly, integrated mixed methods research designs that incorporate the strengths of both approaches and simultaneously reduce the limitations of each approach (Bernard, 2011; Creswell, 2009).

Comparison of approach

The original qual/quant debates centered on differences in assumptions and inferential orientations towards key methodological concepts including appropriate sample design to reduce inferential bias, and issues of reliability, validity, replicability, and generalizability. Each side felt they had the correct answer, possibly the only answer, and the other did not. Much of the debate revolved around the epistemological differences in the two approaches, summarized in Fig. 1. (CDC, 2004) The sampling debate and potential resolution for key

elements of that debate are succinctly summarized by Luborsky and Rubinstein (1995) who provide a comprehensive baseline discussion of the logical framework behind qualitative sampling in response to inappropriate quantitatively oriented critiques of qualitative design.

The bulk of the epistemological issues identified by Luborsky and Rubinstein (1995) has been progressively addressed and are resolving or resolved through current methodological approaches and empirical tests of theory. Most NIH funded researchers who utilize both qualitative and quantitative theory, sampling, methods, and analytical strategies design projects with correctly targeted methods, defensible sampling designs, and theory based analysis strategies for each type of research, enhancing the strengths of each while also reducing the weaknesses of either or both (NIH, 2000). On the other hand, there is a small group of disciplinary chauvinists (including recently trained scholars) who continue to oppose any research, regardless of sound theory and design that does not fit with their strongly held beliefs and the accepted wisdom of their disciplinary paradigm (either qualitative or quantitative). Consequently, one of the primary issues needing continuing resolution is a lingering us-versus-them cultural theme in programs on both ends of the qualitative–quantitative continuum which produce *ad hominem* critiques, rather than evidence based distinctions or matches of the appropriate paradigms, theoretical frameworks, and methodological strengths and weaknesses of each approach.

Qualitative research design sampling issues: lingering discussions

There are several lingering issues for determining scientifically validated sampling frameworks and sample sizes for qualitative research. One issue for qualitative sampling (and qualitative research in general), is the linguistic/definitional heritage of qualitative research. Classic qualitative methodology derives from parallel, seldom convergent, methodological developments in anthropology (Bernard, 2011) and qualitative sociology (Kendall, 2011) followed by subsequent contributions from such diverse disciplines as nursing (Speziale et al., 2011; Tuckett, 2004), education (Bogdan and Biklen, 2007), and public health (Andresen and DeFries Bouldin, 2010; Curry et al., 2006; Trotter et al., 2000). Each produced a profusion of discipline specific terminology that, to date, has not been standardized. Consequently, individuals engaged in qualitative research design (and grant writing) end up miscommunicating by naming the same process using different terminology, using the same terminology for different designs, or using discipline specific jargon that is not shared by other researchers. This confusion is being slowly resolved in areas where multidisciplinary approaches require a consistent terminology and definition of

* Fax: +1 928 523 9135.

E-mail address: Robert.trotter@nau.edu.

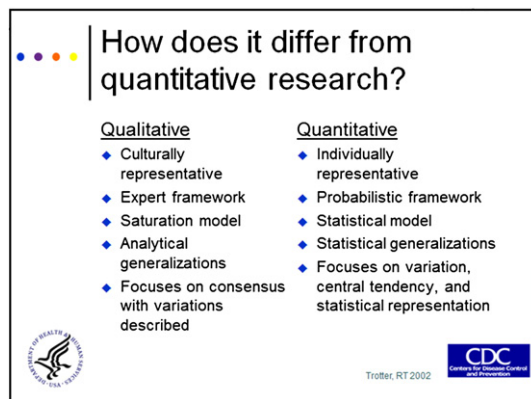


Fig. 1. The logic of qualitative research design.

approach. However, an interim, pragmatic, approach is to acknowledge the need for a thesaurus for the most common labels for qualitative sampling. Pending the creation of such a thesaurus, the following terms are the most commonly cited qualitative sampling strategies: qualitative sampling, ethnographic sampling, theoretical sampling, convenience sampling, snowball sampling, chain referral sampling, nominated sampling, nominated expert sampling, targeted sampling, network sampling, respondent driven sampling, and non-probability sampling. While space does not allow for a citation or definition of each, authors and grant writers ignore the profusion at their own peril.

A second partially unresolved issue for qualitative sampling is agreement on the ideal sample size. The gold standard for probabilistic research sample size is to conduct a power analysis (Murphy et al., 2009) taking into account the theoretical frame, population size and characteristics, and statistical analysis for a specific research project. In contrast, the ideal standard for qualitative sample size is to interview to redundancy (the process of sequentially conducting interviews until all concepts are repeated multiple times without new concepts or themes emerging) (Bernard, 2011), or to interview to saturation (a point at which all questions have been thoroughly explored in detail, no new concepts or themes emerge in subsequent interviews) (Schensul and LeCompte, 2010). The quantitative approach produces a specific number. The qualitative approach produces a process. The difference in approach continues to be debated, but there appear to be several frameworks that allow researchers to compatibly apply both standards correctly in mixed methods research (Creswell, 2009; Johnson, 1990).

A third slowly resolving controversy is tied to the condition that there are two distinct (but compatible) approaches to qualitative research that may require subtly different sampling approaches. One is an exploratory (emergent theory) approach that requires the researcher to empirically but inductively find the theory in the data and the other is a confirmatory approach that applies specific culture theory to a research topic. The emergent theory approach includes grounded theory design in sociology (Bryant and Charmaz, 2007) and classic ethnography in anthropology (Bernard, 2011; Schensul and LeCompte, 2010). The emergent theory approach primarily utilizes the saturation sampling model, since a valid, reliable and (culturally) generalizable sample size cannot be pre-determined given the need for a thorough exploration of an as yet unknown area of belief or behavior. The lack of clear sample size determination creates some on-going issues and debates in the methodological literature, particularly around the issue of operationally defining saturation or redundancy. Alternately, theoretically based qualitative research design (or qualitative social theory approach) allows for sampling and analysis to be linked to an existing theoretical framework such as cognitive theory, feminist theory, embodiment theory, constant comparative theory, etc. These theory based designs tend to favor nominated

expert sampling designs and can often be specifically enumerated since the total population (rather than a sample) of experts can often be interviewed and rarely exceeds 15 and 25 recognizable (consensus) experts for any community based study.

The fourth sample design issue that is resolving, evolving, but persistent is the condition that both the emergent theory and extant theory approaches rely on qualitative, rather than probabilistic, measures of reliability, replicability, and validity to assure that the data coverage is correct. (Bernard, 2011; Johnson, 1990; Morse et al., 2002; Schensul and LeCompte, 2010). Qualitative reliability can be addressed through a basic redundancy test. If the group, as a whole, is essentially giving the same answers to the questions asked, and saying it over and over, reliability is satisfied. An alternative qualitative reliability test is the process of triangulation. If the same information is acquired from multiple (unconnected) sources and multiple methods (e.g. interview, observation, survey, cognitive techniques) then both qualitative reliability and qualitative replicability have been achieved. The difference in inferential imperatives between these quality control measures and their quantitative counterparts is an on-going source of discussion.

The fifth sampling issue arises from the fact that the most common focus for qualitative research is pursuing in-depth descriptive studies of modal culture within well-defined communities. As a consequence, qualitative samples tend to be relatively small and consensus oriented, especially in comparison to probabilistically driven population samples based on power analysis. This condition creates concerns over the generalizability of the data. Qualitative sampling is designed to confirm consensus and systematic variation from consensus, while quantitative sampling is designed to identify the distribution of the variability of traits (or the lack of those traits) across a population. A very useful theoretical justification for small but representative qualitative sample sizes is provided by Weller and Romney (1988), and by Romney et al. (1986). Their consensus theory approach allows the size of populations represented in qualitative research to range from whole cultures and subcultures within a larger society, to special populations that need exploration in terms of specific beliefs, behavior, or relevance to the larger group, without having to increase the actual sample size, if appropriate sampling processes are followed.

Current models for qualitative sampling and sample size: resolving the unresolved

The ongoing discussions over the five issues described above have produced a small set empirically tested and epistemologically congruent models for qualitative sampling design and sample size. Nominated expert sampling is a classic ethnographic approach to exploring cultural and social meanings in various populations, communities, and cultural groups. The sample recruitment process is to identify all of the consensus experts (i.e. those nominated by multiple other individuals in a community) who have the most extensive expertise in a specific area of social or cultural knowledge. The ideal standard is to recruit the entire expert group to provide a saturation level of information about the targeted research topic. Since experts tend to agree about the vast majority of their subject area, and also provide virtually the whole explanation of the variability in expert views (they know each other's agreements and disagreements) this provides an in depth investigation of a topic that is qualitatively valid, reliable, and culturally generalizable (Romney et al., 1986). Intensive case-finding through geographical sampling is based on the concept that likeminded people tend to congregate. Consequently, there are many different social settings that are ideal for identifying qualitative samples and collecting data within subgroups of a larger population. This approach works particularly well for research on place bound behaviors (such as clinics, schools, bars, martial arts groups, sports events, etc.). The approach is to identify a set of known (or discoverable) locations where the target behaviors occur on a regular basis and to recruit research participants from those locations using either a probabilistic or purposeful recruitment

frame (Curtis et al., 2000). Referral sampling (snowball sampling is the most common label) starts with an index individual who is identified as having the key characteristics required by the research design, and asking that individual to nominate others with similar characteristics. The nominated individuals constitute a second wave of data collection. The same sample nomination questions are asked of the second wave in order to construct the third wave and so on. In many cases, this sample design utilizes multiple index individuals as starting points, in order to create large samples. One positive aspect of carefully designed snowball samples is that they can provide a framework for conducting both a power analysis of the sample and sample size and allow for estimation of population characteristics for hidden populations (Frank and Snijders, 1994). One of the most widely used variations on referral sampling is Heckathorn's (1997) respondent driven sampling (RDS). Group identification and network sampling strategies can be used to create sampling frameworks for studies of communities and cultural groups (Schensul et al., 1999). Social network samples focus on specific relationships (e.g. drug use, needle sharing, sexual partnerships) as well as their intensity, directionality, and frequency. A network sample is designed to describe a larger segment of a community or group that is tied together by some common relationship. This approach allows the researcher to make inferences about the type and quality of the relationships, about core versus peripheral participation in the group, about roles and statuses in the group, and about dynamic interactions within the group (Salganik and Heckathorn, 2004). Targeted sampling is a well substantiated sampling choice for mixed methods designs in hard to reach populations (i.e. when too many key parameters about the characteristics of a targeted population are unknown and as a consequence it is impossible to use a pre-existing sampling frame that is appropriate to sample the specific population). Targeted sampling (Robinson et al., 2006; Watters and Biernacki, 1989) can be used as a reasonable substitute for strict probabilistic sampling designs in situations where both qualitative and quantitative research is needed within the same project.

For the most part, qualitative sampling easily resolves most validity issues but has more epistemological difficulty with generalizability to large populations, especially highly diverse populations that are multi-modal on beliefs, values, knowledge, processes, etc. The sample designs and sample size considerations discussed above have evolved to meet appropriate standards and inferential imperatives of the concepts of reliability, generalizability, and replicability. There are still a number of epistemological questions that need further resolution, especially where qualitative and quantitative data are combined for analysis. Mixed methods projects and approaches can be extraordinarily problematic when both qualitative and quantitative approaches are misused and miss-design, and particularly powerful when correctly designed.

Conflict of interest statement

The authors declared that there are no conflict of interest.

References

- Andresen, E., DeFries Bouldin, E. (Eds.), 2010. *Public Health Foundations: Concepts and Practices*. John Wiley and Sons, Inc.
- Bernard, H.R., 2011. *Research Methods in Anthropology: Qualitative and Quantitative Approaches*, Fifth edition. Rowman Altamira, New York.
- Bogdan, R., Biklen, S. K., 2007. *Qualitative Research for Education: An Introduction to Theory and Methods*. Pearson/Allyn and Bacon.
- Bryant, A., Charmaz, K., 2007. *The SAGE Handbook of Grounded Theory*. Sage Publications.
- Centers for Disease Control, Prevention (CDC), 2004. *Qualitative Research September 2004 CDC/KEMR Training Workshop*, PowerPoint.
- Creswell, J.W., 2009. *Research Design: Qualitative, Quantitative, and Mixed Methods Approaches*, 3rd edition. Sage Publications.
- Curry, L., Shield, R.R., Wetle, T., 2006. *Improving Aging and Public Health Research: Qualitative and Mixed Methods*. American Public Health Association.
- Curtis, S., Gesler, W., Smith, G., Washburn, S., 2000. Approaches to sampling and case selection in qualitative research: examples in the geography of health. *Soc. Sci. Med.* 50, 1001–1014.
- Frank, O., Snijders, T., 1994. Estimating the size of hidden populations using snowball sampling. *J. Off. Stat.* 10, 53–67.
- Heckathorn, D.D., 1997. Respondent driven sampling: a new approach to the study of hidden populations. *Soc. Probl.* 2 (44), 174–199.
- Johnson, J.C., 1990. *Selecting Ethnographic Informants*. Sage Publications.
- Kendall, D., 2011. *Sociology in Our Times: The Essentials*, 8th edition. Wadsworth.
- Luborsky, M.R., Rubinstein, R.L., 1995. Sampling in qualitative research: rationale, issues, and methods. *Res. Aging* 17 (1), 89–114.
- Morse, J.M., Barrett, M., Mayan, M., Olson, K., Spiers, J., 2002. Verification strategies for establishing reliability and validity in qualitative research. *Int. J. Qual. Meth.* 1 (2), 1–19.
- Murphy, K.R., Myers, B., Wolach, A.H., 2009. *Statistical Power Analysis: A Simple and General Model for Traditional and Modern Hypothesis Tests*, 3rd edition. Taylor and Francis, Routledge.
- National Institutes of Health (NIH), 2000. *Qualitative Methods in Health Research: Opportunities and Considerations*. Application and Review Office of Behavioral and Social Sciences Research National Institutes of Health. Bethesda, MD.
- Robinson, W.T., Risser, J.M.H., McGoy, S., Becker, A.B., Rehman, H., Jefferson, M., Griffin, V., Wolverton, M., Tortu, S., 2006. Recruiting injection drug users: a three-site comparison of results and experiences with respondent-driven and targeted sampling procedures. *J. Urban Health* 83 (7).
- Romney, A.K., Weller, S.C., Batchelder, W.C., 1986. Culture as consensus: a theory of culture and informant accuracy. *Am. Anthropol.* 88 (2), 313–338.
- Salganik, M.J., Heckathorn, D.D., 2004. Sampling and estimation in hidden populations using respondent-driven sampling. *Socio. Meth.* 34 (1), 193–240.
- Schensul, J.J., LeCompte, M.D., 2010. *Designing and Conducting Ethnographic Research: An Introduction*. Altamira Press, Walnut Creek, CA.
- Schensul, J.J., LeCompte, M.D., Trotter II, R.T., Cromley, E.K., Singer, M., 1999. *Mapping Social Networks, Spatial Data, and Hidden Populations*. Ethnographers Toolkit. Altamira Press, Walnut Creek, CA.
- Speziale, H.S., Streubert, H.J., Carpenter, D.R., 2011. *Qualitative Research in Nursing: Advancing the Humanistic Imperative*, 5th edition. Lippincott, Williams and Wilson.
- Trotter II, R.T., Medina Mora, M.A., 2000. *Qualitative Methods. Guide to Drug Abuse Epidemiology, Mental Health and Substance Dependence Department, Noncommunicable Disease and Mental Health Cluster*. World Health Organization, Geneva, pp. 93–123.
- Tuckett, A.G., 2004. Qualitative research sampling: the very real complexities. *Nurse Res.* 12 (1), 47–61.
- Watters, J.K., Biernacki, P., 1989. Targeted sampling: options for the study of hidden populations social problems 36 (4), 416–430. <http://dx.doi.org/10.2307/800824> (Oct.).
- Weller, S.C., Romney, A.K., 1988. *Systematic data collection*. Qual. Res. Meth. 10 (Sage Publications).