

# A Call for Qualitative Power Analyses\*

ANTHONY J. ONWUEGBUZIE<sup>1,\*</sup> and NANCY L. LEECH<sup>2</sup>

<sup>1</sup>*Department of Educational Measurement and Research, College of Education, University of South Florida, USA;* <sup>2</sup>*University of Colorado at Denver and Health Sciences Center*

**Abstract.** The purpose of this paper is to emphasize the importance of sampling and sample size considerations in *all* qualitative research. Such considerations would help qualitative researchers to select sample sizes and sampling designs that are most compatible with their research purposes. First, we discuss the importance of sampling in qualitative research. Next, we outline 24 designs for selecting a sample in qualitative research. We then discuss the importance of selecting a sample size that yields data that have a realistic chance of reaching data saturation, theoretical saturation, or informational redundancy. Based on the literature, we then provide sample size guidelines for several qualitative research designs. As such, we provide a framework for making sampling and sample size considerations in interpretive research.

## 1. Setting the Scene

According to The American Heritage College Dictionary (1993, p. 1206), a sample is “a portion, piece, or segment that is representative of a whole.” Also, sampling is “an act, process, or technique of selecting an appropriate sample” (p. 1206). Thus, sampling is a term that transcends research in general and research paradigms in particular. In quantitative research, there appears to be unequivocal agreement that sampling plays a very important role.

Given the importance placed on sampling within quantitative research, it is interesting that sampling has not been given the same prominence in qualitative research. Indeed, using the keywords “qualitative research” and “sampling,” as well as “qualitative research” and “sample size,” a review by the authors of the ERIC (i.e., Educational Resource Information Center) and PsycINFO databases yielded only three journal articles (i.e., Crowley, 1994/1995; Sandelowski, 1995; Jones, 2002) that discussed the issue

---

\*Author for correspondence: Anthony J. Onwuegbuzie, Department of Educational Measurement and Research, College of Education, University of South Florida, 4202 East Fowler Avenue, EDU 162, Tampa, FL, 33620-7750, USA. E-mail: tonyonwuegbuzie@aol.com

\*An earlier version of this article received the 2004 Southwest Educational Research Association (SERA) Outstanding Paper Award.

of sampling and/or sample size in qualitative research. Yet, as noted by Onwuegbuzie (2003), qualitative researchers make *inferences* from the sample of words to each respondent's truth space when conducting thematic analyses on data from interviews and focus groups. However, if the sample of words collected is not representative of each respondent's total truth space, then the sampling error associated with the researcher generalizing the sample of words will be large. Second, in situations where the purpose of the qualitative study is to generalize the findings, both the size of the sample and the sampling scheme used clearly are crucial. However, sampling should be a consideration in all qualitative inquiries, regardless of purpose of research. Unfortunately, this is not often the case; sample sizes are often selected in a seemingly arbitrary manner in many research studies and little or no rationale is provided for the sampling scheme used.

Therefore, the purpose of this paper is to advocate the use of qualitative power analyses. The current paper is divided into several sections. First, we discuss the importance of sampling in qualitative research. Second, we outline the 22 designs for selecting a sample in qualitative research identified by Onwuegbuzie and Leech (2005), comprising 5 probability sampling schemes and 17 non-probability sampling schemes. We identify two new non-probability sampling schemes that have not been discussed in the literature. We then discuss the importance of selecting a sample size that yields data that have a realistic chance of reaching data saturation (Morse, 1995; Flick, 1998), theoretical saturation (Strauss and Corbin, 1990) or informational redundancy (Lincoln and Guba, 1985). Based on the literature, we then provide sample size guidelines for several qualitative research designs. As such, we provide a framework for making sampling and sample size considerations in interpretive research.

## **2. Sample Size Considerations in Qualitative Research**

Generally, the goal in quantitative research is to obtain large, representative samples and to generalize findings to populations. This is not the case in qualitative research; where purposive sampling of one or a few cases (i.e., individuals, groups) is more appropriate (Onwuegbuzie and Daniel, 2003). Thus, qualitative researchers have not attached the same level of importance to sample size. The premise for this relative lack of focus on sample size appears sound. Specifically, qualitative researchers correctly contend that the goal of qualitative research typically is to obtain insights into particular educational, social, and familial processes and practices that prevail within a specific location (Connolly, 1998). In order to gain insights, qualitative researchers usually strive to extract meaning from their data.

According to Donmoyer (1990), the concept of generalization should be rejected for qualitative researchers in education who study people and the

meaning in their lives. While it is true that in many situations qualitative researchers are not interested in generalizing findings beyond the people they directly study, we contend that in virtually every qualitative study, one or more of the following generalizations nearly always are made: (a) from the sample of words to the voice; (b) from the sample of observations to the truth space; (c) from the words of key informants to the voice of the other sample members; (d) from the words of sample members to those of one or more individuals not selected for the study; or (e) from the observations of sample members to the experience of one or more individuals not selected for the study.

### 2.1. SAMPLE OF WORDS TO THE VOICE

As noted by Onwuegbuzie (2003), words that arise from a study participant serve as sample units of data that represent the total number of words existing from that sample member. Thus, for example, when a researcher interviews a person, the feelings, experiences, thoughts, and opinions expressed by that person can never represent the whole picture. Therefore, in order to capture the voice, it is crucial that the interviewer collects a sufficient number of words from the interviewee. As such, in interviews, the size of the sample, where the sample represents the data collected (e.g., words, body language), is an important consideration. Failure to collect an adequate sample of words would help to yield a crisis of representation and legitimation (Denzin and Lincoln, 2000), where a crisis of representation stems from an inability to capture lived experience and a crisis of legitimation stems from to inability to interpret and evaluate data. These crises are even more likely to permeate qualitative studies during group interviews (i.e., focus groups). This is because focus group interviewers are faced with three challenges that are not pertinent in individual interviews: (a) the interviewer must prevent one person or a subset of focus group interviewees from dominating the group by *sampling* their voice on a disproportionate amount of occasions and/or for a disproportionate length of time; (b) the interviewer must find a way to encourage recalcitrant interviewees to contribute appropriately to the session; and (c) the interviewer must capture the voice from all members of the focus group to ensure the most complete coverage of each underlying topic (Fontana and Frey, 2000).

Our contention that sample size considerations must be made when attempting to capture the voice is supported by the statements of Lincoln and Guba (1985); prolonged engagement and persistent observations help to increase the researcher's chances of understanding the underlying phenomena, events, or cases. Although not described by Lincoln and Guba as such, the concepts of prolonged engagement and persistent observations represent sampling concepts. Prolonged engagement and persistent observations should be the goal of every qualitative researcher conducting

interviews, focus groups, and the like. Consequently, researchers should pay attention to both the length and number of interviews and focus groups. Indeed, qualitative researchers should provide a rationale for decisions made regarding the length and number of interviews and focus groups as part of their discussions of the conceptual framework and research design/procedures used. Unfortunately, few qualitative researchers, if any, provide such a rationale (Onwuegbuzie and Leech, 2005).

It appears that many qualitative researchers only conduct one interview session per study participant, with these sessions often lasting for one hour or less (Onwuegbuzie and Leech, 2005). However, as noted by Lincoln and Guba (1985), there is little or no discussion in the literature as to whether 60 min or less of interview/focus group data typically is sufficient to obtain saturation. Therefore, a qualitative power analysis is needed. A qualitative power analysis could be conducted via qualitative metasyntheses and qualitative metasummaries. According to Sandelowski and Barroso (2003a), a “qualitative metasynthesis (also referred to as a qualitative meta-data-analysis, qualitative meta-analysis, and meta-ethnography) is a form of systematic review or integration of qualitative research findings in a target domain that are themselves interpretive syntheses of data, including phenomenologies, ethnographies, grounded theories, and other integrated and coherent descriptions or explanations of phenomena, events, or cases” (p. 227). A qualitative metasummary is “a form of systematic review or integration of qualitative findings in a target domain that are themselves topical or thematic summaries or surveys of data” (p. 227). These analytical techniques could be used to compare and contrast studies of the same phenomenon, event, or case with respect to the length and number of interviews/focus groups. Information from these meta-studies could help qualitative researchers to optimize the design of their studies. For example, interviews lasting for one hour should not be scheduled without careful consideration of the power of this length of time to capture the voice. Obviously, the unit sample size also will be a function of time, money, and resources.

Researchers conducting even the most naturalistic of studies (e.g., ethnographic studies) have to make some design decisions prior to the start of the inquiry. For example, ethnographic researchers must decide for how long the initial interview(s) should be. Decisions about subsequent interviews might have to wait until the study has begun; however, decisions regarding the first interview should be made at the design stage. Thus, a qualitative power analysis has great utility in qualitative research.

## 2.2. SAMPLE OF OBSERVATIONS TO THE TRUTH SPACE

In much the same way as when words are the units of analysis, taking observations from a setting involves sampling. The techniques of persistent

observations and prolonged engagement represent sampling issues. Consistent with this is the fact that theoretical sampling is an important strategy used by grounded theorists. Inadequate theoretical sampling leads to a formal theory not being developed or developed prematurely (Straus and Corbin, 1990; Charmaz, 2000). Therefore, grounded theorists should continuously assess how much data should be sampled until data saturation, theoretical saturation, or informational redundancy is reached.

### 2.3. FROM THE WORDS OF KEY INFORMANTS TO THE VOICE OF THE OTHER SAMPLE MEMBERS

It is not unusual for qualitative researchers to study a small sub-sample of individuals, called key informants, who are selected from the overall sample of research participants. These key informants often generate a large part of the researcher's data. The extent to which generalizing from the key informants to the other study participants is valid depends on two issues. First, the key informants must be representative of the study participants who were not selected as key informants. Second, the sample of words must be representative of each key informant's voice. Thus, the sampling units of particular interest are both the persons (i.e., key informants) and words. Researchers must make careful decisions about each of these sampling units early on, which may evolve (e.g., theoretical sampling) as the data are collected and analyzed. Failure to make optimal sampling decisions could lead to what is termed key informant bias (Maxwell, 1996). Unfortunately, key informant bias is a common occurrence because of the unrepresentativeness of key informants (Poggie, 1972; Pelto and Pelto, 1975; Maxwell, 1995). Thus, some researchers recommend that the key informants be randomly selected from the set of study participants to increase the likelihood that the key informants' words represent those of the other study members (Sankoff, 1971; Maxwell, 1996). In addition to selecting a representative sub-sample of key informants, as noted previously, qualitative researchers must attempt to capture the voice of each key informant by sampling a sufficient number of words.

### 2.4. FROM THE WORDS/OBSERVATIONS OF SAMPLE MEMBERS TO THOSE OF ONE OR MORE INDIVIDUALS NOT SELECTED FOR THE STUDY

When qualitative researchers seek to generalize words or observations from the sample to one or more individuals who were not selected for the study, it is important that they attempt to ensure that the words/observations of these study participants are adequately represented by sampling adequate numbers of units. In addition, because, in some cases, interpretations and theories emerging from the findings might be expected to be generalized to individu-

als or groups outside the sample, researchers should pay special attention to both the method of selecting and the size of the sample. Thus, what follows is a presentation of the array of sampling schemes from which qualitative researchers can choose, as well as a discussion of sample size considerations.

### 3. Sampling Schemes

As contended by Onwuegbuzie and Leech (2005), different methods for selecting samples in qualitative research are appropriate at different times. Moreover, these methodologists asserted that qualitative researchers not only have at their disposal sampling schemes that have been traditionally associated with the qualitative paradigm, they also have in their arsenals all the sampling designs that have been routinely associated with the quantitative paradigm. Indeed, after eliminating redundancies (i.e., convenience sampling, purposeful sampling, network sampling), Onwuegbuzie and Leech (2005) identified 22 sampling designs that they contended qualitative researchers (and quantitative researchers) have available for selection. Thus, below, we outline these 22 sampling schemes, and identify two additional sampling schemes that, to date, have not been discussed in the literature. All of these sampling designs fall into one of two categories: random sampling schemes and non-random sampling schemes.

#### 3.1. RANDOM (PROBABILITY) SAMPLING

Before deciding on whom to select for the inquiry and how to select the sample, qualitative researchers must decide what the objective of the study is (Onwuegbuzie et al., 2004). If the objective of the study is to generalize the interpretations to a population, then the researcher should attempt to select a sample that is both random and large (Creswell, 2002; Johnson and Christensen, 2004). In this situation, the qualitative researcher can select one of five random sampling schemes at a particular stage of the sampling process discussed below.

##### 3.1.1. *Simple Random Sampling*

In simple random sampling, participants are selected in such a way that every person in the population has the same probability of being selected for the study, and the selection of one individual does not affect selection of any other individual (i.e., independence).

##### 3.1.2. *Stratified Random Sampling*

Stratified random sampling represents a sampling scheme in which a population is divided into sub-populations such that members of each

sub-population are relatively homogeneous with respect to one or more characteristics and relatively heterogeneous from members of all other sub-groups with respect to this/these characteristic(s). In order to obtain a stratified random sample, the sampling frame is first divided into sub-populations, or *strata*. Next, a random sample is selected from each strata. The goal of stratified random sampling is to select a sample in such a way that the target subgroups are represented in the sample in the same proportion that they exist in the population.

### 3.1.3. *Cluster Random Sampling*

Cluster random sampling is a method of randomly selecting clusters of individuals instead of randomly selecting individuals one at a time.

### 3.1.4. *Systematic Random Sampling*

Systematic random sampling is a method of sampling in which individuals are selected from a list by choosing every  $k$ th sampling frame member, where  $k$  represents the population size divided by the desired sample size.

### 3.1.5. *Multi-stage Random Sampling*

Multi-stage random sampling involves selecting a sample in two or more stages because either the population is relatively large or its members cannot easily be identified. In multi-stage random sampling, the first stage often involves cluster sampling, whereas subsequent stages involve simple random sampling, stratified random sampling, cluster random sampling, and/or systematic random sampling.

## 3.2. NON-RANDOM SAMPLING

If the goal is not to generalize to a population but to obtain insights into a phenomenon, individuals, or events, as will typically be the case in qualitative research, then the researcher purposefully selects individuals, groups, and settings that maximize understanding of the phenomenon. As such, the most common method of sampling in qualitative research is purposeful sampling. Here, individuals, groups, and settings are considered for selection if they are “information rich” (Patton, 1990: 169).

Miles and Huberman (1994) have identified 16 strategies for purposive sampling. These strategies differ with respect to whether they are implemented before data collection has started or after data collection begins (Creswell, 2002). Also, the appropriateness of each strategy is dependent on the research objective, purpose, and question. The 16 strategies are outlined

below. In addition to these 16 strategies, there is quota sampling, another non-random type of sampling scheme.

### 3.2.1. *Maximum Variation Sampling*

According to Sandelowski (1995), maximum variation sampling is one of the most frequently used purposeful sampling techniques. In this method, a wide range of individuals, groups, or settings is purposively selected such that all or most types of individuals, groups, or settings are selected for the inquiry. In this way, the multiple perspectives of individuals can be presented that exemplify the complexity of the world (Creswell, 2002).

### 3.2.2. *Homogeneous sampling*

In stark contrast to maximum variation sampling, homogeneous sampling involves sampling individuals, groups, or settings because they all possess similar characteristics or attributes. Participants are selected for the study based on membership in a subgroup or unit that has specific characteristics. This sampling approach often is used to select focus groups (Onwuegbuzie et al., 2004).

### 3.2.3. *Critical Case Sampling*

In critical case sampling, individuals, groups, or settings are selected that bring to the fore the phenomenon of interest such that the researcher can learn more about the phenomenon than would have been learned without including these critical cases.

### 3.2.4. *Theory-based Sampling*

In theory-based sampling, individuals, groups, or settings are selected because they help the qualitative researcher to develop a theory. This sampling scheme also is used to expand a theory.

### 3.2.5. *Confirming and Disconfirming Cases Sampling*

This method of sampling often is used after data collection has commenced. The former (i.e., exploration) tends to improve interpretation of the findings, whereas the latter (i.e., confirmation) tends to assist in data validation (Miles and Huberman, 1994).

### 3.2.6. *Snowball Sampling*

This sampling approach, also known as network sampling, usually comes to the fore after data collection has begun. Snowball sampling involves



asking participants who have already been selected for the study to recruit other participants.

### 3.2.7. *Extreme Case Sampling*

In extreme case sampling, an outlying case or one that possesses one or more extreme characteristics is studied. The method is to select extreme cases and then to compare them.

### 3.2.8. *Typical Case Sampling*

In typical case sampling, the researcher studies an individual, group, or setting that is typical. The researcher should consult several experts in the field of study in order to obtain a consensus as to what example(s) is typical of the phenomenon and should, therefore, be studied (Johnson and Christensen, 2004).

### 3.2.9. *Intensity Sampling*

In intensity sampling, the researcher studies individuals, groups, or settings that experience the phenomenon intensely but not extremely.

### 3.2.10. *Politically Important Sampling*

In politically important sampling, the researcher selects pertinent informants who may need to be included/excluded because they connect with politically sensitive expected in the analysis (Miles and Huberman, 1994).

### 3.2.11. *Random Purposeful Sampling*

In random purposeful sample, the researcher chooses cases at random from the sampling frame consisting of a purposefully selected sample. That is, the researcher first obtains a list of individuals of interest for study using one of the 15 other methods of purposeful sampling, and then randomly selects a desired number of individuals from this list. Although not stated by Miles and Huberman (1994) or other qualitative methodologists, when selecting a random purposeful sample, the researcher can use any of the random sampling techniques discussed above. According to Miles and Huberman (1994), random purposeful sampling “adds credibility to sample when potential purposeful sample is too large” (p. 28).

### 3.2.12. *Stratified Purposeful Sampling*

Stratified purposeful sampling is similar to stratified random sampling. In order to obtain a stratified purposeful sample, the sampling frame is first

divided into strata; then a purposeful sample is selected from each strata. Such a sampling scheme can facilitate group comparisons (Miles and Huberman, 1994).

### 3.2.13. *Criterion Sampling*

In criterion sampling, individuals, groups, or settings are selected that meet criteria. According to Miles and Huberman (1994), this sampling technique typically is utilized for the purpose of quality assurance.

### 3.2.14. *Opportunistic Sampling*

In opportunistic sampling, the researcher capitalizes on opportunities during the data collection stage to select cases. These cases could represent typical, negative, critical, or extreme cases (Johnson and Christensen, 2004). Opportunistic sampling takes place after the study begins in order to take advantage of developing events. This form of sampling is particularly useful when the researcher is unable or unwilling to declare in advance of the inquiry every case that will be included in the investigation.

### 3.2.15. *Mixed Purposeful Sampling*

This method of sampling involves the mixing of more than one sampling strategy. For example, a researcher might begin by selecting two samples: one via extreme case sampling and the other via critical case sampling. The researcher could then compare the results emerging from both samples. Consequently, mixed purposeful sampling can help to triangulate data (Miles and Huberman, 1994).

### 3.2.16. *Convenience Sampling*

Convenience sampling techniques used by qualitative researchers involves selecting individuals or groups that happen to be available and are willing to participate at the time.

### 3.2.17. *Quota Sampling*

In quota sampling, the researchers decide on the specific characteristics and quotas of sample members to be selected. A main limitation of this method of sampling is that only those who are accessible at that time of selection have a chance of being selected.

In addition to the 16 types of purposive designs outlined by Miles and Huberman (1994), and quota sampling, we have identified two more

purposive sampling schemes, namely, what we call *multi-stage purposeful random sampling* and *multi-stage purposive sampling*.

### 3.2.18. *Multi-stage Purposeful Random Sampling*

Multi-stage random purposeful sampling involves selecting a sample in two or more stages, in which the first stage is random, and subsequent stages are purposive. In multi-stage random purposeful sampling the first stage often involves cluster sampling, whereas subsequent stages involve one of the 16 purposive sampling techniques outlined above.

### 3.2.19. *Multi-stage Purposeful Sampling*

Multi-stage purposeful sampling also involves selecting a sample in two or more stages. However, all stages incorporate purposive sampling, unlike the multi-stage purposeful random sampling and random purposeful. Multi-stage purposeful sampling is different than mixed purposeful sampling in that the former is always sequential, whereas the latter typically involves concurrent sampling in which one sample is not a subset of other samples.

Thus, qualitative researchers, as do quantitative researchers, presently have 24 sampling schemes from which to choose. As Kemper et al. (2003) concluded, “the understanding of a wide range of sampling techniques in one’s methodological repertoire greatly increases the likelihood of one’s generating findings that are both rich in content and inclusive in scope” (p. 292).

## 4. **Sample Size Guidelines**

Although there is little consensus about what qualitative research is and how it should be undertaken (Schwandt, 2000; Sandelowski and Barroso, 2003b), there is general agreement that the goal of qualitative research is *not* to generalize beyond a sample to the population. Yet, some qualitative researchers find it difficult to resist the temptation to generalize findings to some population (Onwuegbuzie and Daniel, 2003; Onwuegbuzie and Leech, 2005). Such practices are flawed unless a representative sample has been selected. Whenever a theory is being developed, some type of generalization clearly has taken place. If generalization is not the goal, then he/she should only outline a theory in terms of the particular participant(s), setting, context, location, time, event, incident, activity, experience, and/or processes, as well as with respect to the specific researcher (assuming that the researcher is serving as the instrument). If interpretations and theories remain strictly localized, then the size of the sample is not as crucial. As noted by Onwuegbuzie and Leech (2005), many qualitative researchers

seemingly select the size of their samples in an arbitrarily fashion. As doctoral advisors, over the years, we have observed students who are in the early stages of proposing qualitative research studies being unable to provide a rationale for choice of sample sizes. Unfortunately, students are not the only ones to select their sample sizes arbitrarily. Many beginning and even some experienced qualitative researchers appear to select their sample sizes without paying attention to the role that sample sizes play in the quest to attain data saturation.

The lack of sample size consideration in qualitative research likely stems from the scant discussion in this area. In general, sample sizes in qualitative research should not be too small that it is difficult to achieve saturation. At the same time, the sample should not be too large that it is difficult to undertake a deep, case-oriented analysis (Sandelowski, 1995). More specifically, Creswell (2002) has recommended that qualitative researchers should, (a) study one cultural-sharing group in an ethnography, (b) examine three to five cases in a case study, (c) interview 15–20 people during a grounded theory study, (d) explore the narrative stories of one individual in narrative research (p. 197). In addition, Creswell (1998) recommended interviews with up to 10 people in phenomenological research and interviews with 20–30 people in grounded theory.

Johnson and Christensen (2004) surmise that focus groups usually contain 6–12 persons, whereas Langford et al. (2002) and Morgan (1997) recommend 6–10 individuals. Krueger (2000) recommends 6–9 focus group members and groups with more than 12 participants tend to “limit each person’s opportunity to share insights and observations” (p. 78). Further, Morgan contends that focus groups with less than 6 participants make it difficult to sustain a discussion, whereas groups containing more than 12 members make it difficult for the moderator. According to Krueger, focus groups must be small enough for everyone to have their voices represented but yet large enough to capture a range of voices. Morgan also states that 3–5 focus groups typically are sufficient to reach saturation. Kuzel (1992) recommends that 6–8 data sources or sampling units often will be sufficient when homogeneous samples are selected in qualitative research and that 12–20 data sources generally are necessary. Morse (1994) suggests that qualitative researchers use at least six participants in investigations where the goal is to understand the essence of experience. Morse also recommends 30–50 interviews and/or observations for ethnographies and grounded theory research, and approximately 100–200 units of observation in qualitative ethological studies. Although these guidelines are helpful, the authors did not state how they arrived at these estimates. Thus, metasummaries and metasyntheses are needed to gather evidenced-based data regarding suitable sample sizes. Such data should help guide qualitative researchers

to establish minimum sample sizes based on the number of participants needed.

Furthermore, sampling involves more than the number of participants included in the study; sampling is a process that incorporates the number of participants, the number of contacts with each participant, and the length of each contact. It is important for researchers to consider how much contact will be required in order to reach saturation. Many students and beginning researchers underestimate the amount of time needed with each participant in order to understand their truth space. Thus, it is imperative when considering sampling that these issues are considered.

## 5. Summary and Conclusions

According to The American Heritage College Dictionary (1993: 1072), power is “the ability or capacity to perform or act effectively.” Thus, when combined with the word qualitative, power represents “the ability or capacity to perform or act effectively” in the qualitative research context. Thus, the phrase *qualitative power analysis* represents an analysis of “the ability or capacity to perform or act effectively” with respect to sampling. We contend that power analyses should be used to assess the appropriateness of the units sampled. Here, units could be people, words, text, observations, events, incidents, activities, experiences, social processes, or any other object of study.

Therefore, the purpose of this paper was to advocate always making sample size and sampling considerations in qualitative research. We contend that *all* qualitative research studies should involve some form of qualitative power analysis. Such an analysis would help qualitative researchers to select sample sizes and sampling designs that are most compatible with their research purposes.

Sampling in qualitative research involves making decisions not only about which individuals to study, but also about several study parameters, including settings, contexts, locations, times, events, incidents, activities, experiences, and/or social processes. Studies involving more than one case also require that the researcher make specific choices about which types of cases to include, depending on the research objective, purpose, and question(s). Qualitative inquiry, in general, necessitates continuous re-focusing, re-examination, and re-drawing of study parameters. Nevertheless, some initial sampling decisions still are required prior to the beginning of the study (Miles and Huberman, 1994). Sampling is an issue even in single case studies, which involve the choice of a particular case rather than others, and which necessitates that the researcher makes sampling decisions within the case itself (Miles and Huberman, 1994; Sandelowski, 1995; Maxwell, 1996). Further, as concluded by Stake (2000), in case studies,

researchers are unable to avoid generalizations because at the very least, they are generalizing “to happenings of their cases at times yet to come and in other situations” (p. 439).

We view generalizations as including a process of reflection (Greenwood and Levin, 2000). As surmised by Greenwood and Levin, because all findings are context-bound, (a) any interpretations stemming from these findings should be made after first being appropriately aware of the context under which the findings were constructed and (b) generalizations of any interpretations to another context should be made after first being adequately cognizant of the new context and how this new context differs from the context from which the interpretations were generated, as well as reflecting on the consequences that such a generalization may have. Therefore, generalization represents an active process of reflection (Greenwood and Levin, 2000).

We recommend that before deciding on an appropriate sample size, qualitative researchers should consider identifying a corpus of interpretive studies that used the same design as in the proposed study (e.g., grounded theory, ethnography) and wherein data saturation was reached. The researcher then could examine the sample sizes used in these studies with a view to selecting a sample size that is within the range used in these investigations. In addition, when observations or interviews are the data collection methods of choice, researchers should consider using the extant literature to determine an appropriate number of observations/interviews and an adequate length of time for each observation/interview that are consistent with Lincoln and Guba’s (1985) recommendations of persistent observation and prolonged engagement, respectively, as well as allow the researcher to reach data saturation, theoretical saturation, and/or informational redundancy.

Additionally, it is important that in making their sampling decisions, qualitative researchers should leave an audit trail, in which these decisions are documented at every step of the research process (Halpern, 1983; Lincoln and Guba, 1985). Unfortunately, such methodological information often is omitted from qualitative research reports. Constat (1992) remarked that the procedures used by many qualitative researchers “often remain private and unavailable for public inspection” (p. 254). Such omissions make it difficult for the reader to assess the rigor of the methods used and, subsequently, the trustworthiness of the findings and interpretations. In the words of Constat (1992: 255), unless information about procedures used in qualitative research are delineated, “the research community will be entitled to question the analytical rigor of qualitative research.” We leave the final statements to Miles and Huberman (1984):

just thinking in sampling-frame terms is healthy methodological medicine. If you are talking with one kind of informant, you need to consider why this kind of informant is important, and, from there, which other people should be interviewed.

Remember that you are not only sampling people, but also settings, events, and processes. It is important to line up these parameters with the research questions as well, and to consider whether your choices are doing a representative, time-efficient job of answering them. The settings, events, or processes that come rapidly to mind at the start of the study may not be the most pertinent or data-rich ones. A systematic review can sharpen early and later choices. (p. 41)

## References

- Charmaz, K. (2000). Grounded theory: objectivist and constructivist methods. In N. K. Denzin & Y. S. Lincoln (eds.), *Handbook of qualitative research* (2nd ed.) (pp. 509–5535). Thousand Oaks, CA: Sage.
- Connolly, P. (1998). ‘Dancing to the wrong tune’: ethnography generalization and research on racism in schools. In: P. Connolly & B. Troyna (eds.), *Researching Racism in Education: Politics, Theory, and Practice*. Buckingham, UK: Open University Press, pp. 122–139.
- Constas, M. A. (1992). Qualitative data analysis as a public event: the documentation of category development procedures. *American Educational Research Journal* 29:253–266.
- Crowley, E. P. (1994/1995). Using qualitative methods in special education research. *Exceptionality* 5: 55–70.
- Creswell, J. W. (1998). *Qualitative Inquiry and Research Design: Choosing Among Five Traditions*. Thousand Oaks, CA: Sage.
- Creswell, J. W. (2002). *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. Upper Saddle River, NJ: Pearson Education.
- Denzin, N. K. & Lincoln, Y. S. (2000). The discipline and practice of qualitative research. In: N. K. Denzin & Y. S. Lincoln (eds.), *Handbook of Qualitative Research*. 2nd ed., Thousand Oaks, CA: Sage, pp. 1–28.
- Donmoyer, R. (1990). Generalizability and the single-case study. In: E. W. Eisner & A. Peshkin (eds.), *Qualitative Inquiry in Education: The Continuing Debate*. New York: Teachers College Press, pp. 175–200.
- Flick, U. (1998). *An Introduction to Qualitative Research: Theory, Method and Applications*. London: Sage.
- Fontana, A. & Frey, J. H. (2000). The interview: from structured questions to negotiated text. In: N. K. Denzin & Y. S. Lincoln (eds.), *Handbook of Qualitative Research*. 2nd ed., Thousand Oaks, CA: Sage, pp. 645–672.
- Greenwood, D. J. & Levin, M. (2000). Reconstructing the relationships between universities and society through action research. In: N. K. Denzin & Y. S. Lincoln (eds.), *Handbook of Qualitative Research*. 2nd ed., Thousand Oaks, CA: Sage, pp. 85–106.
- Halpern, E. S. (1983). *Auditing Naturalistic Inquiries: The Development and Application of a Model*. Unpublished doctoral dissertation, Indiana University.
- Johnson, B. & Christensen, L. (2004). *Educational Research: Quantitative, Qualitative, and Mixed Approaches*. 2nd ed., Needham Heights, MA: Allyn & Bacon.
- Jones, S. R. (2002). Writing the word: methodological strategies and issues in qualitative research. *Journal of College Student Development* 43: 461–473.

- Kemper, E. A., Stringfield, S. & Teddlie, C. (2003). Mixed methods sampling strategies in social science research. In: A. Tashakkori & C. Teddlie (eds.), *Handbook of Mixed Methods in Social and Behavioral Research*. Thousand Oaks, CA: Sage, pp. 273–296.
- Krueger, R. A. (2000). *Focus Groups: A Practical Guide for Applied Research*. 3rd ed., Thousand Oaks, CA: Sage.
- Kuzel, A. J. (1992). Sampling in qualitative inquiry. In: B. F. Crabtree & W. L. Miller (eds.), *Doing qualitative research. Research Methods for Primary Care*. Vol. 3. Newbury Park, CA: Sage, pp. 31–44.
- Langford, B. E., Schoenfeld, G. & Izzo, G. (2002). Nominal grouping sessions vs. focus groups. *Qualitative Market Research* 5: 58–70.
- Lincoln, Y. S. & Guba, E. G. (1985). *Naturalistic Inquiry*. Beverly Hills, CA: Sage.
- Maxwell, J. A. (1995, February). *Diversity and Methodology I a Changing World*. Paper presented at the Fourth Puerto Rican Congress of Research in Education, San Juan, Puerto Rico.
- Maxwell, J. A. (1996). *Qualitative Research Design*. Newbury Park, CA: Sage.
- Miles, M. B. & Huberman, A. M. (1984). *Qualitative Data Analysis: A Sourcebook of New Methods*. Beverly Hills, CA: Sage.
- Miles, M. & Huberman, A. M. (1994). *Qualitative Data Analysis: An Expanded Sourcebook*. 2nd ed., Thousand Oaks, CA: Sage.
- Morgan, D. L. (1997). *Focus Groups as Qualitative Research*. 2nd ed., Qualitative Research Methods Series 16. Thousand Oaks, CA: Sage.
- Morse, J. M. (1994). Designing funded qualitative research. In: N. K. Denzin & Y. S. Lincoln (eds.), *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage, pp. 220–235.
- Morse, J. M. (1995). The significance of saturation. *Qualitative Health Research* 5: 147–149.
- Onwuegbuzie, A. J. (2003). Effect sizes in qualitative research: a prolegomenon. *Quality & Quantity: International Journal of Methodology* 37: 393–409.
- Onwuegbuzie, A. J. & Daniel, L. G. (2003, February 12). Typology of analytical and interpretational errors in quantitative and qualitative educational research. *Current Issues in Education* [On-line], 6(2). Available: <http://cie.ed.asu.edu/volume6/number2/>.
- Onwuegbuzie, A. J., Jiao, Q. G. & Bostick, S. L. (2004). *Library Anxiety: Theory, Research, and Applications*. Lanham, MD: Scarecrow Press.
- Onwuegbuzie, A. J. & Leech, N. L. (2005, March). *Generalization Practices in Qualitative Research: Trends in the Literature*. Paper presented at the annual meeting of the Eastern Educational Research Association, Sarasota, FL.
- Onwuegbuzie, A. J. & Leech, N. L. (2005). Taking the “Q” out of research: teaching research methodology courses without the divide between quantitative and qualitative paradigms. *Quality & Quantity: International Journal of Methodology* 39: 267–296.
- Patton, M. Q. (1990). *Qualitative Research and Evaluation Methods*. 2nd ed., Newbury Park, CA: Sage.
- Pelto, P. & Pelto, G. (1975). Intracultural diversity: some theoretical issues. *American Ethnologist* 2: 1–18.
- Poggie, J. J., Jr. (1972). Toward control in key informant data. *Human Organizational* 31: 23–30.
- Sandelowski, M. (1995). Focus on qualitative methods: sample sizes in qualitative research. *Research in Nursing & Health* 18: 179–183.
- Sandelowski, M. & Barroso, J. (2003a). Creating metasummaries of qualitative findings. *Nursing Research* 52: 226–233.
- Sandelowski, M. & Barroso, J. (2003b). Classifying the findings in qualitative studies. *Qualitative Health Research* 13: 905–923.



- Sankoff, G. (1971). Quantitative aspects of sharing and variability in a cognitive model. *Ethnology* 10: 389–408.
- Schwandt, T. A. (2000). Three epistemological stances for qualitative inquiry. In: N. K. Denzin & Y. S. Lincoln (eds.), *Handbook of Qualitative Research*. 2nd ed., Thousand Oaks, CA: Sage, pp. 189–213.
- Stake, R. E. (2000). Case studies. In: N. K. Denzin & Y. S. Lincoln (eds.), *Handbook of Qualitative Research*. 2nd ed., Thousand Oaks, CA: Sage, pp. 435–454.
- Strauss, A. & Corbin, J. (1990). *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Newbury Park, CA: Sage.
- The American Heritage College dictionary*, 3rd ed. (1993). Boston, MA: Houghton Mifflin.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.