

Descriptions of Sampling Practices Within Five Approaches to Qualitative Research in Education and the Health Sciences

Timothy C. Guetterman

Key words:

qualitative sample size; qualitative sampling; qualitative inquiry; saturation; choosing cases; case study; ethnography; grounded theory methodology; narrative inquiry; phenomenology

Abstract: Although recommendations exist for determining qualitative sample sizes, the literature appears to contain few instances of research on the topic. Practical guidance is needed for determining sample sizes to conduct rigorous qualitative research, to develop proposals, and to budget resources. The purpose of this article is to describe qualitative sample size and sampling practices within published studies in education and the health sciences by research design: case study, ethnography, grounded theory methodology, narrative inquiry, and phenomenology. I analyzed the 51 most highly cited studies using predetermined content categories and noteworthy sampling characteristics that emerged. In brief, the findings revealed a mean sample size of 87. Less than half of the studies identified a sampling strategy. I include a description of findings by approach and recommendations for sampling to assist methodologists, reviewers, program officers, graduate students, and other qualitative researchers in understanding qualitative sampling practices in recent studies.

Table of Contents

- [1. Introduction](#)
 - [1.1 Why examine qualitative sample sizes?](#)
- [2. Review of Related Literature](#)
- [3. Method](#)
 - [3.1 Sampling](#)
 - [3.2 Sample size](#)
 - [3.3 Data collection](#)
 - [3.4 Data analysis](#)
- [4. Findings](#)
 - [4.1 Case study](#)
 - [4.2 Ethnography](#)
 - [4.3 Grounded theory methodology](#)
 - [4.4 Narrative inquiry](#)
 - [4.5 Phenomenology](#)
- [5. Discussion](#)
 - [5.1 Sample size patterns](#)
 - [5.2 Lacking procedural detail](#)
 - [5.3 Saturation](#)
 - [5.4 Sampling practices by qualitative approach](#)
 - [5.5 Sampling practices by discipline](#)
 - [5.6 Limitations](#)
- [6. Recommendations](#)
- [7. Conclusion](#)

[Acknowledgments](#)

[References](#)

[Author](#)

[Citation](#)

1. Introduction

How large should a qualitative sample size be? This question seems to plague new qualitative researchers who often turn to sampling suggestions (BERNARD, 2000; CHARMAZ, 2006; CRESWELL, 2013; GUEST, BUNCE & JOHNSON, 2006; MORSE, 1994; SANDELOWSKI, 1995) that are, perhaps too often, taken as canonical (MORSE, 2000). However, before discussing sample sizes further, it is necessary to discuss sampling generally, and although we should resist the urge to think about qualitative sampling from a quantitative viewpoint (EMMEL, 2013), it is helpful to compare the two approaches. [1]

Sampling in quantitative research typically follows random sampling procedures (CRESWELL, 2015). Researchers calculate the required sample size before beginning the study and that size remains a constant target throughout the study. They can turn to the literature for sample size guidelines for particular analyses that will have appropriate power to detect effects. Qualitative sampling, however, is less direct. As EMMEL (2013) explained, qualitative sampling is not a single planning decision, but it is an iterative series of decisions throughout the process of research. A reflexive researcher then makes adjustments and considers the implications of sampling on interpretation. Although qualitative sampling is substantially more complicated than sample sizes and sites (ibid.), sample sizes and sampling practices are the focus of this article. Section 1.1 discusses reasons to conduct research on qualitative sample sizes and summarizes the aims of the study, and Section 2 reviews relevant literature on qualitative sampling. Section 3 describes the method, including the sample of studies, data collection, and data analysis methods. Next, the findings appear in Section 4 with a description of sample size and practices within published studies in education and health sciences by research design. Section 5 summarizes the major procedural themes emerging across the corpus of studies and the key limitations of this article. Section 6 presents recommendations for sampling procedures, followed by conclusions in Section 7. [2]

1.1 Why examine qualitative sample sizes?

Although recommendations exist for determining qualitative sample sizes (e.g., CRESWELL, 2013; MORSE, 1994), the literature appears to contain few instances of research on qualitative sample sizes. Practical guidance is needed to determine sample sizes for rigorous qualitative research. The lack of guidance poses a problem because researchers planning qualitative studies need to estimate sample sizes in order to 1. allocate resources and budget, 2. develop proposals for funding, 3. develop proposals for institutional review boards, and 4. conduct rigorous and systematic qualitative research. A systematic examination

of the sampling practices in published qualitative studies adds to the sampling conversation and may help to refine sampling recommendations by providing an empirical basis. The findings of this study may be useful to methodologists, reviewers, program officers, graduate students, and other qualitative researchers in understanding qualitative sampling practices in recent studies. [3]

One way to approach the topic of qualitative sampling is to focus on methods by describing sampling practices in recent studies. A description provides insights into how qualitative sampling works in actual research. Thus, the purpose of this article is to describe sample size and sampling practices within highly cited published studies in education and the health sciences. To achieve this purpose, three research questions guided the study. The research questions were: 1. What overall sampling procedures do researchers describe and cite when conducting qualitative studies within five approaches to qualitative inquiry in the two disciplines? The five approaches are defined by John CRESWELL's (2013) typology: case study, ethnography, grounded theory methodology, narrative inquiry, and phenomenology. 2. What specific sample sizes do researchers report within the five approaches to qualitative inquiry? 3. How do researchers describe their determination of sample size? This article is not intended to suggest rules for appropriate sample sizes within approaches to qualitative research. Rather, it is meant to describe sample size and extensiveness across published studies in both health sciences and education that explicitly use one of five approaches to qualitative inquiry. The sample sizes in published studies might provide a baseline for the researcher to then tailor estimates to particular circumstances when planning a study. [4]

2. Review of Related Literature

This literature review discusses three areas related to qualitative sample size: summary of sampling recommendations, literature related to sample size, and existing studies of qualitative sample sizes. As EMMEL (2013) noted, with theoretical or purposive sampling, the researcher is reflexive and makes decisions in response to empirical findings and theoretical developments that occur in the study. The process follows the iterative nature of qualitative research. A key qualitative feature is that research questions are typically limited, studying a central phenomenon in a particular context. The researcher's intent is not to generalize from the sample to a population, but to explain, describe, and interpret (MAXWELL, 2013) this phenomenon. Consequently, sampling is not a matter of representative opinions, but a matter of information richness. Appropriateness and adequacy are paramount in qualitative sampling (MORSE & FIELD, 1995). Scholars have provided sampling procedures (e.g., CRESWELL, 2013; MARSHALL & ROSSMAN, 2011; MAXWELL, 2013; MERRIAM, 2009; MORSE, 1994; PATTON, 2015). For example, CRESWELL's (2013) comprehensive qualitative text devoted four pages to sampling. He presented three considerations of the purposeful sampling strategy: deciding the participants or sites, selecting the sampling strategy, and determining the sample size. [5]

Some researchers have addressed the challenges of determining sample size. PATTON (2015) explained that purposeful sampling involves selecting information rich cases. In addition to the purpose of the inquiry, PATTON acknowledged the role of resource limitations in determining a qualitative sample size. MERRIAM (2009) also discussed the process for selecting a sample and determining sample size. She noted that it depends on the research questions, the data collected, the data analysis, and the availability of resources. To the specific question of how many, MERRIAM wrote, "there is no answer" (p.80). She recommended including approximate numbers, subject to change, when developing proposal. SANDELOWSKI (1995) discussed the aspects of the study that the researcher should take into account when determining a sample size or evaluating its adequacy. Determining sample size involves "judgment and experience in evaluating the quality of the information against uses to which it will be put" (p.183). Clearly, sample size is contingent on many considerations. [6]

In the midst of these cautions, several scholars have addressed sample size directly. CRESWELL (2013) suggested collecting extensive details about a few sites or individuals. He provided observations and some recommendations of sample size ranges for the five approaches: case study, no more than four to five cases; ethnography, a single culture sharing group; grounded theory methodology, 20 to 30 cases; narrative inquiry, one to two cases observed unless developing a collective story; and phenomenology, three to ten cases, with observed sample sizes from one to 325. MORSE (1994) also provided suggestions for various qualitative approaches: to understand the essence of an experience, six participants; ethnography, 30 to 50; and grounded theory methodology, 30 to 50. These two texts, MORSE's (1994) and CRESWELL's (2013), provide researchers with concrete numbers. Others, such as EMMEL (2013), have cautioned against reliance on these suggested sizes and urged researchers to consider additional factors. Furthermore, MORSE (2000) has since clarified the assumptions behind the sample sizes and recommends that researchers account for the scope, the topic, the quality of data, the design, and the use of shadowed data (i.e., participants' reports about others). [7]

A related topic is how to assess the adequacy of a sample in terms of the size of the sample and the appropriateness of the particular individuals, other data sources, or sites (ONWUEGBUZIE & LEECH, 2007). Methodologists have discussed the use of the grounded theory concept of *theoretical saturation* (GUEST et al., 2006) as the marker of a sufficient sample size. Theoretical saturation is the point at which the qualitative analyst does not see new information in the data related to the codes, themes, or theory (ibid.). Saturation, however, might not be the best marker of an adequate sample. O'REILLY and PARKER (2012) questioned the relevance of theoretical or thematic saturation beyond grounded theory methodology and argued for more transparency in how researchers achieved saturation. [8]

Several researchers have analyzed studies published to examine sample sizes. SOBAL (2001) conducted a content analysis of studies published in a nutrition education journal to examine sample extensiveness. For studies using individual

interviews, the mean number of participants was 45; for group interviews, the mean was 15 groups and 141 participants. SAFMAN and SOBAL (2004) conducted a similar study of sample extensiveness in health education research. Studies using individual interviews had an average sample size of 103 (SD 134) and ranged from 2 to 720. Finally, MASON (2010) conducted a thorough examination of qualitative sample sizes in PhD dissertations. He found a mean sample size of 31 and reported the most typical sample sizes were 10, 20, 30, and 40. These studies provided valuable guidance for the present study. [9]

3. Method

I conducted a literature search for qualitative studies published within these two disciplines, health sciences and education, over the last five years and categorized the studies into one of the qualitative research designs, using CRESWELL's (2013) five design typology. The purpose of the data collection was to adequately describe sampling strategies associated with the five approaches. [10]

3.1 Sampling

This study employed a combination of two purposive sampling strategies: critical case and stratified sampling. Critical case sampling involves selecting a small number of important cases to "yield the most information and have the greatest impact on the development of knowledge" (PATTON, 2015, p.276). The sample was limited to published, peer-reviewed journal articles for two reasons: they generally attain higher quality standards, and they are a principal source of scholarly evidence (CRESWELL, 2015). Next, the sample involved critical case sampling through the selection of the most highly cited articles, based on the Web of Science "times cited" count. The reason for using Web of Science is because it is a fully indexed, "curated database of published, peer-reviewed content that is selected according to publicly available standards" (<http://wokinfo.com/>). The Web of Science covers over 12,000 journals and serves as the basis for impact factor calculation. It also has a citation tracking feature, which provides a common measure of the influence of research output (VAN AALST, 2010). The times cited count provides a measure of the use of the article, and perhaps, its influence in the research community. However, it should be noted that many have criticized the Web of Science for its national bias and the lack of transparency in determining which articles are citable (PLOS ONE MEDICINE EDITORS, 2006). In addition, researchers have suggested alternatives, such as Google Scholar, as a more comprehensive measure of research output (VAN AALST, 2010). [11]

Ostensibly, the articles I examined based on the Web of Science times cited employed sampling that generated findings and interpretations that were useful to other researchers. In addition, I employed a stratified purposive sampling strategy to allow for comparison (PATTON, 2015). I selected the five most highly cited articles within education and the health sciences for each of CRESWELL's (2013) five approaches to qualitative research. The rationale for this approach was to capture variations between two fields in which qualitative inquiry is widespread. A

sample of studies that span the five approaches will further assist in understanding sampling based on design, given the desired outcome to describe sample sizes, based on published research, for each approach and to develop recommendations for researchers. [12]

3.2 Sample size

My target sample of 50 articles contained the five most highly cited papers in education and the health sciences within each approach from 2008 through 2012. The reason for ending sampling at 2012 was to have a year lag to account for newer publications that have not been read or cited less often. The study considers each of the five approaches as cases of sampling. Thus, it consists of five cases, which is within CRESWELL's (2013) recommendation for the number of cases to include in a case study. I selected ten published articles within each approach because I felt this number would be adequate to yield sufficient depth. [13]

3.3 Data collection

To search for relevant articles, I entered the approach (e.g., "grounded theory") into the [Web of Science database](#) and limited the timespan to 2008-2012. To limit articles to education and the health sciences, I placed criteria on the search results for journal articles in education¹ and health sciences² disciplines. [14]

Not all categories (i.e., disciplines) were represented in all searches. In instances where two articles met inclusion criteria and tied for the times cited, I included both articles, resulting in 51 studies. [15]

3.4 Data analysis

I employed a qualitative text analysis, as described by Udo KUCKARTZ (2014). Each article was the unit of analysis. Following KUCKARTZ (2014) and MAYRING (2000), the overall steps in the analytic process were to: 1. determine my goal for developing categories based on research question, 2. determine the degree of differentiation between categories, 3. set the abstraction level, 4. begin with first text passage and create categories, 5. read full text of article line by line and construct additional categories, 6. assign categories and create additional categories, 7. rearrange categories when necessary and move to next article, 8. revise the category system. [16]

1 The exact limits for education were:

web of science categories=(education educational research)

timespan=2008-2012, indexes=Science Citation Index (sci) expanded, Social Science Citation Index (ssci), Arts & Humanities Citation Index (a&hci), Book Citation Index-Science (bkci-s), Book Citation Index-Social Sciences and Humanities (bkci-ssh)

2 The exact limits for health sciences were:

web of science categories=(medicine general internal or nursing or health policy services or clinical neurology or nursing sci or nursing ssci or rehabilitation or health care sciences services or neurosciences or psychiatry or oncology)

timespan=2008-2012. indexes=sci-expanded, ssci, a&hci, bkci-s, bkci-ssh.

Step 1 was to determine my goal for developing categories, based on the research questions (Section 1.1). The goal was to categorize sampling procedures, sample sizes, and descriptions of the determination of sample size. In Step 2, I determined that the differentiation of categories needed to be specific enough to describe differences in procedures among the five approaches and disciplines. In Step 3, I determined the level of abstraction to be the precise procedures described by the authors of articles. Because this study focused on sampling procedures, many of which are described in existing qualitative methods literature (e.g., CRESWELL, 2013; PATTON, 2015), the categories were theoretically derived (KUCKARTZ, 2014) from the research questions and set a *priori*. I created a codebook, which consisted of the following categories: the qualitative approach, sample size, the sites, identified sampling strategy, sampling procedure discussion and authors cited, the data sources, observation sessions, and funding. However, other categories were allowed to emerge throughout the analysis. In Step 4, I began with the first article and focused on the method section. Although much of the data was written into the method section of each article, it was necessary to examine the full text, as additional details appeared in the introduction, discussion, tables, and even the abstract. In Step 5, I proceeded to analyze each article for other noteworthy sampling procedures and coded those procedures as they emerged. Examples of new categories were the length of interviews and number of sites. In Step 6, I coded within each article and refined the codebook to be more specific. In Step 7, I rearranged the codebook as more fine-grained categories emerged, such as discussions of theoretical saturation and funding sources. Finally, in Step 8, the category system became fixed after reviewing approximately one third of the articles. I proceeded with coding the remaining articles. [17]

After coding all articles, I compiled these categories in three ways: for the entire corpus of articles, for health sciences compared to education, and for each of the five qualitative approaches. I used MAXQDA to count frequencies of codes and develop a series of contingency tables to compare descriptive statistics of sampling characteristics among the disciplines and among the five approaches. [18]

4. Findings

The mean sample size across all studies was 87 participants with a minimum of one and a maximum of 700, across two sites on average. For studies that used interviews, researchers typically conducted one interview per person. The maximum number of interviews was four per person. The mean number of observations was 115. Table 1 presents a summary of the descriptive statistics of the sampling characteristics for all studies and by discipline. The following section presents a summary of the findings, including a description of the sampling characteristics for the studies reviewed, organized by discipline within each of the five approaches to qualitative inquiry (see Table 2).

	<i>N</i>	Interviews per participant	Observations	Sites
Health sciences				
<i>N</i>	26	26	6	16
Mean	93	1	178	2
<i>SD</i>	141	1	186	1
Min	1	1	20	1
Max	586	4	532	5
Education				
<i>N</i>	25	23	5	17
Mean	80	1	40	3
<i>SD</i>	169	1	45	2
Min	1	1	6	1
Max	700	4	111	8
Total				
<i>N</i>	51	49	11	33
Mean	87	1	115	2
<i>SD</i>	153	1	153	2
Min	1	1	6	1
Max	700	4	532	8

Table 1: Summary of sampling characteristics of all studies³

Table 2: Summary of sampling characteristics by qualitative approach. Please click [here](#) to download the PDF file. [19]

4.1 Case study

Ten published case studies had a mean sample size of 188 at three sites on average. Half of the studies also detailed their use of observations with a mean of 176 observations. Of note, sample sizes in the reviewed case study articles appeared to relate to how the case was bounded. The sample size was typically inclusive of all participants within the case. Researchers' choices of cases to study, however, were more elusive to determine. Half of the studies identified a sampling strategy or rationale. Two of five studies included a literature citation with the sampling discussion. Finally, none of the case studies included a discussion of saturation or the adequacy of the sample. [20]

³ *N*=number of studies, *SD*=standard deviation, Min=Minimum, Max=Maximum

4.1.1 Education

Sample sizes within education focused studies ranged from 12 (PARK ROGERS & ABELL, 2008) to 700 and seemed to reflect the size of the case. The largest sample was in the BOALER and STAPLES (2008) 5-year longitudinal study, which involved three high schools, interviews with approximately 60 students in each year of attendance, and 600 hours of observation. The reason for selecting the particular schools was to observe and study three different approaches to mathematics teaching, also noting schools were similar in size, had a similar philosophy to hire committed, knowledgeable math teachers, but differed in location and demographics. BOALER and STAPLES labeled their sampling as purposive and cited YIN (1994). Two other studies labeled their sampling strategy: CONOLE, De LAAT, DILLON and DARBY (2008) described their sampling as purposive to select "information-rich case studies" (p.513) and cited MAYES (2006). Their sampling purpose was to find students with "a lot of experience in using technology to support their learning" (CONOLE et al., 2008, p.513). In addition, KARAMUSTAFAOGLU (2009) randomly selected 40 physics teachers from secondary schools in a single city to study teachers' perceptions of student-centered physics learning and conducted 30 minute interviews with a subsample of six. Finally, although they did not label their sampling strategy, GROSSMAN et al. (2009) included a variety of data sources across eight sites and described their reason for the particular classes they chose to observe, as related to the purpose to study clinical practice education. [21]

4.1.2 Health sciences

Sample sizes were smaller in the health science's case studies, ranging from 2 to 420. The smallest sample (AGYEPONG & ADJEI, 2008) involved the two authors as participants, fully documenting their experiences with national health insurance in Ghana. The study with the largest sample, GREENHALGH et al. (2008) involved multiple data sources: 250 staff interviews, 1,500 hours of ethnographic observation, interviews and focus groups with 170 patients and their caretakers, 2,500 pages of correspondence and documentary evidence, and other relevant surveys and statistical data collected elsewhere. GREENHALGH et al. reported selecting four particular sites because they were the early adopters of the summary care record program in the U.K. Two studies identified a specific sampling strategy: BRIXLEY et al. (2008) conducted a case study of activities and interruptions that physicians and nurses experienced in a Level 1 trauma center in the U.S., using observations and candidly describing their strategy as a convenience sampling. Additionally, in their study of the bedside handovers that occur among nurses, CHABOYER, McMURRAY and WALLIS (2010, p.29) reported a "purposive sample of nursing staff" for interviews. [22]

4.2 Ethnography

The ten ethnographic studies had a mean sample size of 128. The reported sample size for education studies was drastically lower than health sciences. The overall sample size seemed to be determined by the size of the culture sharing group. In instances when the group size was a realistic number to interview or include in focus groups, the researchers generally seemed to include the entire group within the sample. Four studies identified the sampling strategy, one referencing CRESWELL's (1998) qualitative text. None of the articles included discussion of saturation or the adequacy of the sample. [23]

4.2.1 Education

The ethnographic education studies had sample sizes ranging from 6 to 33 with an average of 23. One study did not list the exact sample size. Although none of the studies explicitly labeled the sampling strategy, several discussed the rationale and approach. For instance, CARLONE, HAUN-FRANK and WEBB's (2011) comparative ethnography of two classrooms that promoted reform-based science practices described a comprehensive search of teachers within 70 schools to select the teachers for their commitment to reform-based science. Although they did not label it, they seemed to use a critical case sampling strategy (PATTON, 2015). Another ethnography (STEVENS, O'CONNOR, GARRISON, JOCUNS & AMOS, 2008) sample arose out of a larger study. The sample consisted of the engineering students as the culture-sharing group at four universities in the U.S. [24]

4.2.2 Health sciences

The ethnography articles in health sciences had larger samples and contained more sampling procedure details. The smallest sample (n=19) was in FUDGE, WOLFE and McKEVITT (2008) ethnographic study of the user involvement policy in health services organizations in the U.K. The authors included an exemplary table that detailed all data sources (e.g., participant observation sessions, interviews with each group, and particular documents). They noted purposive selection of service users to account for variations in gender, age, and medical severity. The particular numbers appeared to reflect all of their encounters during their participant observation experience, which fully represented the culture-sharing group. In another example, DAVIS and JOHNSON's (2008) ethnography focusing on patterns of prescription opioid use, misuse, and diversion contained the largest sample (n=586). They explained likely instances of oversampling and undersampling from their recruitment strategy, and the sampling strategy seemed more quantitatively oriented. [25]

The other studies included the sampling strategy and an explanation. Exploring the meaning of having epilepsy in sociocultural contexts in communities in China and Vietnam, JACOBY et al. (2008) described their sampling as purposive and employing maximum variation. In contrast, the STORENG et al. (2008) ethnographic study of 82 women to understand their experiences with emergency

obstetric care in Burkina Faso was nested within a larger prospective cohort study using quantitative stratified sampling (CRESWELL, 2015). They noted oversampling a particular group to better understand those experiences. Finally, WARE et al. (2009) conducted an ethnography of adherence to antiretroviral therapy for individuals infected with HIV in sub-Saharan Africa. They described their use of purposeful sampling, citing CRESWELL's (1998) text, to represent the perspectives of patients, treatment partners, and health care professionals. In addition, WARE et al. employed random sampling to draw the potential participants in two sites and invited participants from an ongoing prospective study at the third site. Thus, their sampling reflected a blend of quantitative random sampling and purposeful sampling. In general, the description of sampling strategies in health science ethnographies incorporated quantitative sampling traditions, which may account for the larger sample sizes. [26]

4.3 Grounded theory methodology

Among the ten studies that used a grounded theory approach, the mean sample was 59. Two noteworthy features are unique to grounded theory methodology: theoretical sampling and the origin of the concept of saturation. Of the ten studies, one mentioned use of theoretical sampling. Seven identified the specific sampling strategy, the most of any of the five approaches. Six mentioned saturation and procedures to achieve saturation. The grounded theory studies contained the most references to saturation. [27]

4.3.1 Education

Sample sizes ranged from 6 to 134. One study, LIN, LIN and HUANG (2008), epitomized theoretical sampling and saturation. Overall, they seemed to account for about 100 to 200 participants in the study. They described their use of theoretical sampling, citing ORLIKOWSKI (1993) and STRAUSS and CORBIN (1990). Additionally, they discussed their plan to achieve theoretical saturation through a three-phase study. They used each phase of the study to add additional data, assess for saturation, and ultimately achieve saturation. At each stage, they reported their assessment of whether additional data were needed. Other studies also accounted for saturation. SARGEANT et al. (2010), the largest sample, was a grounded theory study, using the approach of CORBIN and STRAUSS (2008). They used purposive sampling to select eight programs based on degrees of structure/rigor, a continuum of learners, and nationality. Participants seemed to fully entail each program. They conducted an additional group to "increase data saturation for general practitioners" (SARGEANT et al., 2010, p.1213) for a total of 17 focus groups. In addition, STRAUSS, CHATUR and TAYLOR (2009) identified their sampling strategy as stratified purposive sampling to ensure representation for two institutions and across gender and reported sampling until saturation occurred. Finally, other studies (BARTON & TAN, 2009; PUTWAIN, 2009) detailed their sampling and case selection procedures without naming a particular strategy. Studying funds of knowledge and discourse in a science class, BARTON and TAN (2009) determined their sample by collaborating with a teacher to select four girls that reflected a range of science

interest, success, and participation. They also included a boy who asked to participate and the teacher. [28]

4.3.2 Health sciences

The health science's samples were larger with a minimum of 20 and a maximum of 147. Four of the studies labeled their sampling strategy, and all included a description of sampling procedures. Furthermore, three studies discussed the procedures to ensure saturation. For example, FRIED, BULLOCK, IANNONE and O'LEARY (2009) conducted a grounded theory study (STRAUSS & CORBIN, 1998) of the process of health behavior change among a group of older adults living in a community. The sampling strategy was purposeful to ensure diverse ethnic/racial and socioeconomic statuses among participants. They reported enrolling participants until the point of theoretical saturation. Most of the studies described purposeful sampling, yet one indicated convenience sampling. MORSE, EDWARDSSEN and GORDON (2008) researched empathic patient-physician communication by analyzing consultations between patients diagnosed with lung cancer and their thoracic surgeons or oncologists. They reported selecting 20 of these medical encounters at one hospital based on convenience sampling. Next, they analyzed the transcripts from each encounter and developed codes until achieving saturation. Finally, HORN, TZANETOS, THORPE and STRAUS (2008) provided a strong example of a saturation discussion. HORN et al. conducted a grounded theory study as part of a larger study. The authors noted that they achieved saturation by the fourth focus group and decided not to hold additional focus groups, citing DENZIN and LINCOLN (2000) and LINGARD and KENNEDY (2007). [29]

4.4 Narrative inquiry

Of the ten narrative studies, two presented a constellation of teacher or school stories and the particular numbers were unavailable. The mean sample size was 18 at an average of two sites. Two of the narrative studies included a discussion of sampling strategy. Both cited literature to support their assertions. One narrative study included mention of saturation; it was in the in health sciences. [30]

4.4.1 Education

Three studies reported specific sample sizes, which ranged from 1 to 24. The larger sample sizes represented collective narrative inquiry. For example, CRAIG (2009, p.604) described her work with several teachers and administrators and presented the work as a "constellation of stories." Concerning depth of inquiry, KIRKPATRICK and BYRNE (2009) included an interesting discussion of their process of returning to participants to build upon previous interviews and share narrative summaries with each participant. In general, the studies in this group did not label a sampling strategy, or cite literature regarding sampling decisions. [31]

4.4.2 Health sciences

Sample sizes ranged from 1 to 52 at an average of two sites. Reported sample sizes were larger for the health sciences as compared to education studies. In addition, the articles tended to identify the sampling strategy. An example is the HAINES, POLAND and JOHNSON (2009, p.69) narrative study to understand smoking among young women, in which the sampling strategy was "purposive and theory-driven, seeking a range of adolescent smoking experiences and participants," citing both PATTON (1990) and CRESWELL (1998). They further supplemented their sample through snowball sampling. [32]

Two additional studies provided a noteworthy sampling description. First, HOPFER and CLIPPARD (2011) interviewed 36 college women and two clinicians regarding human papillomavirus (HPV) vaccination decisions. The authors labeled their sampling strategy purposive and clearly described the rationale, explaining that they recruited all eligible participants from those enrolled in a college course that required participation in a research study. Furthermore, HOPFER and CLIPPARD (2011, p.265) detailed their reason for stratifying the sample in a particular manner: "Because this study was designed to guide a future HPV vaccine campaign aimed at reaching the unvaccinated, a greater number of women were interviewed who were not yet vaccinated." Interviews were scheduled for 40 minutes with the women enrolled in college and one hour for the clinicians. Second, PINNOCK et al. (2011) conducted a narrative study, using REISSMAN's (1993) approach of individuals living with chronic obstructive pulmonary disease (COPD). The authors described purposeful sampling. While they did not cite general qualitative sampling literature, they did cite their previous studies, which indicated a set of 16 to 20 interview sets (i.e., patient and their caregivers) was sufficient to achieve saturation. Patients (n=21) and their informal (n=13) and professional (n=18) caregivers participated in up to four interviews (40 to 150 minute duration) over a six to nine month period. [33]

4.5 Phenomenology

Analyzing 11 phenomenological studies, the mean sample size was 21 participants at a single site. Six articles referenced the sampling strategy. Interestingly, four phenomenological studies mentioned saturation; three were in health sciences. [34]

4.5.1 Education

The mean sample size was 15 and ranged from 8 to 31. The study with the largest sample (EDELBRING, DASTMALCHI, HULT, LUNDBERG & DAHLGREN, 2011) reported students' experiences with virtual patients in clinical education in Sweden. The authors conducted interviews in groups for a total of 13 interviews with 31 students at a single site where they were completing a rheumatology rotation. Furthermore, two of the five studies labeled the sampling strategy. OTTENBREIT-LEFTWICH, GLAZEWSKI, NEWBY and ERTMER (2010) studied the value beliefs that underlie teachers' uses of technology using a

hermeneutical phenomenology approach. The authors described their use of "convenient purposeful sampling procedures" (p.1324) and cited CRESWELL's (1998) "Qualitative Inquiry and Research Design: Choosing Among Five Traditions" as reference. They did not provide a rationale for the sample size but noted a larger sample was not feasible for the researcher team. Data sources included an interview, observation, and electronic portfolio. In addition, DE WET (2010) conducted an interesting phenomenological study of principal-on-teacher bullying in South Africa. She employed snowball sampling, citing PATTON (2002), to access hard to reach individuals. DE WET (2010, p.1452) also discussed saturation in several places, writing: "Interviews were conducted until definite categories and themes became evident and the information became saturated." Other studies did not mention saturation. [35]

4.5.2 Health sciences

The mean sample size was 25 with a minimum of 8 and maximum of 52. Three studies specified the sampling strategy (purposive), and two mentioned saturation. For instance, MARTINS (2008) conducted a phenomenological study of the experiences with the health care system among persons who were homeless in the U.S. She described recruitment procedures and mentioned a purposive sampling strategy. MARTINS conducted 30 to 60 minute interviews with 15 adults who were homeless and receiving care at a free clinic. She made a clear reference to the concept of saturation, explaining that she interviewed new participants until achieving saturation. Saturation she defined as sufficient quality, completeness, and amount of information in addition to no evidence of new themes in the interviews. MARTINS reported reaching saturation after 12 interviews but completing three additional interviews to further ensure no new themes emerged. In addition, BECK and WATSON's (2008, p.231) study of breast-feeding experiences among 52 women who experienced birth trauma noted the number exceeded "what was necessary to achieve saturation of data." [36]

5. Discussion

The following discussion summarizes the major procedural themes emerging across the corpus of studies included in this review of published qualitative studies. Specifically, these data illuminated patterns related to sample size, sampling procedural details, saturation, qualitative approach, and discipline. The patterns are a unique contribution to the qualitative literature because they provide further insight into the sampling procedures and publishing practices used by qualitative researchers. It can inform future research as well as methodology concerning qualitative sampling. [37]

5.1 Sample size patterns

The mean sample size exceeded the observed and recommended sample sizes suggested by CRESWELL (2013) and MORSE (1994). A potential explanation is that sample sizes have increased over time, as these sample size recommendations originated 15 and 20 years ago, respectively. Perhaps, an increase occurred along with the growth of qualitative research or as researchers attempted to align their studies with accepted quantitative standards. The sample sizes found also exceed those reported by MASON (2010) in his study of PhD theses, which is interesting considering the different corpus of studies he reviewed. The consistent findings provide evidence that sample sizes may tend to exceed what may be needed. Superfluous sampling brings several concerns. First, as data tend to become repetitive, the qualitative analysis will lose depth. Second, the study will consume more resources than needed. Finally, I question the ethical implications of burdening more research participants than we actually need as researchers. [38]

Unfortunately, many of the studies reviewed do not give insight into the rationale for the particular sample size. Of the 51 studies, 24 included a reference to a sampling strategy. The majority of those references, however, were "purposeful" or "purposive" sampling without specifying a particular type (e.g., maximum variation). In a few articles, the authors described their sampling approach without labeling it specifically. Nevertheless, the sample size discussion is partly related to the particular journal and its policy. [39]

5.2 Lacking procedural detail

The procedural detail surrounding sampling, case selection, and the extent of data collection was vague in general. Most articles did not report the duration of interviews or observations. Those that included this information described it differently. About half reported the duration of interviews as a range only (e.g., interviews were 30 to 60 minutes in length). Another common technique was to report an approximate duration. The most thorough reporting included both. For instance, EK and TERNESTEDT (2008, p.472) wrote that interviews "lasted from 20 to 90 minutes, with an average of 55 minutes." Of all studies with reported interview duration, the shortest was 15 minutes and the longest was 180 minutes. The study with the shortest average duration had one of the highest sample sizes (over 500). Aside from that study, no pattern between sample size and duration of interviews was evident among the studies. Observation descriptions also differed among studies. Some reported the duration of each observation session. Others reported the sum of all observation sessions as an indicator of the depth of data collection. For example, BRIXLEY et al. (2008, p.234) reported: "Five attending ED [Emergency Department] physicians were observed for a total of 29 h, 31 min. Eight RNs [Registered Nurses] were shadowed for a total of 40h, 9 min." In general, however, few studies described observations in detail. No patterns regarding observations emerged. [40]

5.3 Saturation

Studies that included an assessment of saturation had lower sample sizes. Several studies contained excellent examples of discussions of the adequacy of the sample and saturation. Of the 51 studies, 11 contained a discussion of saturation and achieving saturation. Of those 11 studies, six used grounded theory methodology, four phenomenological approaches, and one was a narrative inquiry. The mean sample size in the 11 studies discussing saturation was 53 (minimum 10, maximum 147, median of 28), which is smaller than the mean for all studies. GUEST et al. (2006) conducted an experiment with a data set to determine when saturation is achieved and found that 12 interviews were optimal. Interestingly, except for one study, the studies reviewed exceeded this size. The finding is also consistent with MASON's (2010) conclusion that PhD students employed a large sample size relative what was needed to achieve saturation, perhaps because they do not understand saturation or think that a larger sample will appear more rigorous to supervisory committees and peer reviewers. [41]

5.4 Sampling practices by qualitative approach

Several patterns were evident when examining sampling by approach. The sample size was highest for case studies at 189 on average and seemed dependent on how the case was bound. For grounded theory methodology, the mean sample size was 59, which far exceeds the sample sizes recommended by qualitative authors (e.g., CRESWELL, 2013). In addition, I searched for journal policies on sample size for comparison and found one recently imposed by the journal *Archives of Sexual Behavior* (DWORKIN, 2012). The mean sample size also exceeded that policy of a minimum of 25-30 participants for qualitative studies involving interviews. In contrast, the sample sizes were lowest for phenomenology with a mean of 20 and narrative inquiry with a mean of 21 participants, which likely reflects the collective stories developed. Ethnography varied in terms of sample size (mean of 128 participants). The sample generally reflected a single culture-sharing group. Perhaps, the size of the group determines whether full participation is realistic or whether a smaller sample is needed. When the group size was large, the researchers seemed to draw a sample. Overall, the patterns across qualitative approach were consistent with MASON's (2010) study of sample size and saturation in PhD theses. [42]

5.5 Sampling practices by discipline

Additional patterns within the studies are notable. Studies in the health science had a slightly higher mean number of participants. Health sciences sample sizes exceeded education in all qualitative approaches but case studies. Moreover, with the exception of grounded theory methodology, studies with no disclosed external funding had the smallest sample sizes. Finally, some samples seemed to be derived from quantitative traditions, such as drawing smaller qualitative studies from larger clinical trials. In these cases, the sample for the qualitative study was a subset of a quantitative random sample. [43]

5.6 Limitations

Before offering recommendations, it is necessary to discuss some limitations of this article. First, it was challenging to find sampling details in published studies. This systematic review is based only on what is written in these studies. Next, the scope of this study only includes the five approaches to qualitative inquiry, as defined by CRESWELL (2013). Obviously, many other qualitative traditions and variations exist and warrant their own sample size description. Moreover, the scope of articles is limited to the most highly cited studies identified through the Web of Science and the journals included in that database. The Web of Science does tend to over-represent journals from the United States and United Kingdom (VAN AALST, 2010) and may not adequately represent journals in other nations. Another problem with the Web of Science, according to the PLoS ONE MEDICINE EDITORS (2006), is that it does not transparently disclose its procedures to determine what articles are included in citation counts, which makes the articles included somewhat arbitrary. Among other weaknesses, the Web of Science has been found to contain citation errors that skew the actual number of citations (VANCLAY, 2012). Because of these issues, other search engines will likely generate different articles and different times cited. Additionally, I must acknowledge that sampling five studies in each discipline per approach was largely based on CRESWELL's (2013) case study recommendation and resource availability. However, clear patterns and trends in sample size and sampling procedure discussions appeared to emerge well before analyzing all 51 studies. Based on this observation, the sample of studies appeared adequate for the intended purpose to describe sampling practices. [44]

6. Recommendations

Sample size considerations appeared to involve two concerns: the size of the sample (i.e., extensiveness) and the appropriateness (i.e., relevance) of the sample, discussions of which were missing from most studies. Addressing these concerns requires procedures prior to the study (while planning), during the study, and after completing analysis and interpretation. As a planning step, the researcher should identify a specific sampling strategy (e.g., selecting extreme cases), determine how many individuals are necessary, and document a rationale. The researcher should remain reflexive throughout the research process, continually assessing and exploring sampling issues including theoretical saturation. It seems particularly critical to assess the adequacy of the sample. MARTINS (2008) provided an excellent example of discussing when she reached saturation and her procedure to collect three additional interviews to ensure no new themes emerged. After completing the analysis and interpretation, the researcher should address the adequacy of the sample. ONWUEGBUZIE and LEECH (2007, p.117) recommended that researchers conduct qualitative power analysis of "'the ability or capacity to perform or act effectively' with respect to sampling." Qualitative power analysis is a technique to synthesize other studies of similar phenomenon and provide a basis for the sampling decisions (ibid.). Regardless of the stage of the research process, sampling in qualitative research should be intentional, and the text should indicate the rationale for procedural

decisions. The researcher can accomplish that through a discussion of sampling strategy and an assessment of whether the sample was appropriate. For instance, studies should include a discussion of why particular individuals were relevant in addition to the reason for the particular sample size. [45]

If the researcher is unable to describe the reason for selecting a sample size or if the reason is resource limitations, it is important to critically reflect on the sample in the limitations section. Many studies did include a discussion of limitations, but the limitations tended to address the limits of generalizability due to the small sample or site. The discussions were brief and seemed cursory. The intent of qualitative research is to explain, describe, and interpret in depth (MAXWELL, 2013). Thus, the discussion of limitations should focus on whether the researcher achieved the intended depth rather than generalizability. [46]

Based on this systematic review of published qualitative studies and existing sampling literature, I recommend the following:

- When planning a study, specify a sampling strategy and a sample size based on similar published literature. Specify that the sample size is tentative and likely to change through the process of research (MERRIAM, 2009).
- When writing:
 - report sampling details thoroughly (ONWUEGBUZIE & LEECH, 2007);
 - describe the sampling strategy and reasons for selecting a sample size. This description will allow the community of researchers to learn about qualitative sampling methodology and evaluate the quality of findings;
 - assess whether the sample was large enough beyond mention of generalizability. Ask whether the sample was adequate to achieve a sufficient level of depth, and report the procedure to assess adequacy or saturation;
 - assess and report whether the sample was appropriate (O'REILLY & PARKER, 2012). Ask how well the actual sample represented the cases or individuals necessary to address the research questions.
- Journal reviewers and editors should critique the sampling discussion thoroughly and require additional details when relevant. [47]

7. Conclusion

The unique contribution of this study is the insight it provides into the actual sampling practices in the highly cited studies in published literature. By comparing two disciplines and five approaches to qualitative research, it gives us an indication of where sampling practices are tailored and where they are relatively static. An overarching finding was the lack of procedural details about sampling practices in the studies. This information seems necessary both for the reader to become a co-researcher while reading the study and to permit advancements of qualitative inquiry. Simply, when considering sampling, researchers need to move beyond "how many?" to address the questions of "how?" and "why?" [48]

Acknowledgments

Wayne BABCHUK provided valuable feedback to the manuscript and throughout the study.

References

- Agyepong, Irene A. & Adjei, Sam (2008). Public social policy development and implementation: A case study of the Ghana National Health Insurance scheme. *Health Policy and Planning, 23*(2), 150-160. doi: 10.1093/heapol/czn002 [*]⁴
- Barton, Angela C. & Tan, Edna (2009). Funds of knowledge and discourses and hybrid space. *Journal of Research in Science Teaching, 46*(1), 50-73. doi: 10.1002/tea.20269 [*]
- Beck, Cheryl T. & Watson, Sue (2008). Impact of birth trauma on breast-feeding—A tale of two pathways. *Nursing Research, 57*(4), 228-236. doi: 10.1097/01.nnr.0000313494.87282.90 [*]
- Bernard, H. Russell (2000). *Social research methods: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.
- Boaler, Jo & Staples, Megan (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside school. *Teachers College Record, 110*(3), 608-645. [*]
- Brixey, Juliana J.; Tang, Zhihua H.; Robinson, David J.; Johnson, Craig W.; Johnson, Todd R.; Turley, James P.; Patel, Vimla L. & Zhang, Jiajie J. (2008). Interruptions in a level one trauma center: A case study. *International Journal of Medical Informatics, 77*(4), 235-241. doi: 10.1016/j.ijmedinf.2007.04.006 [*]
- Carlone, Heidi B.; Haun-Frank, Julie & Webb, Angela (2011). Assessing equity beyond knowledge- and skills-based outcomes: A comparative ethnography of two fourth-grade reform-based science classrooms. *Journal of Research in Science Teaching, 48*(5), 459-485. doi: 10.1002/tea.20413 [*]
- Chaboyer, Wendy; McMurray, Anne & Wallis, Marianne (2010). Bedside nursing handover: A case study. *International Journal of Nursing Practice, 16*(1), 27-34. doi: 10.1111/j.1440-172X.2009.01809.x [*]
- Charmaz, Kathy (2006). *Constructing grounded theory: A practical guide through qualitative analysis*. Thousand Oaks, CA: Sage.
- Comber, Barbara & Nixon, Helen (2009). Teachers' work and pedagogy in an era of accountability. *Discourse-Studies in the Cultural Politics of Education, 30*(3), 333-345. doi: 10.1080/01596300903037069 [*]⁵
- Conole, Grainne; de Laat, Maarten; Dillon, Teresa & Darby, Jonathan (2008). "Disruptive technologies", "pedagogical innovation": What's new? Findings from an in-depth study of students' use and perception of technology. *Computers & Education, 50*(2), 511-524. doi: 10.1016/j.compedu.2007.09.009 [*]
- [Corbin, Juliet M.](#) & Strauss, Anselm L. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Thousand Oaks, CA: Sage.
- Craig, Cheryl J. (2009). Research in the midst of organized school reform: Versions of teacher community in tension. *American Educational Research Journal, 46*(2), 598-619. doi: 10.3102/0002831208330213 [*]
- Creswell, John W. (1998). *Qualitative inquiry and research design: Choosing among five traditions*. Thousand Oaks, CA: Sage.
- Creswell, John W. (2013). *Qualitative inquiry and research design: Choosing among five approaches* (3rd ed.). Thousand Oaks, CA: Sage.
- Creswell, John W. (2015). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (5th ed.). Boston, MA: Pearson.
- Curtis, J. Randall; Engelberg, Ruth; Young, Jessica P.; Vig, Lisa K.; Reinke, Lynn F.; Wenrich, Marjorie D.; McGrath, Barbara; McCown, Ellen & Back, Anthony L. (2008). An approach to understanding the interaction of hope and desire for explicit prognostic information among

4 "[*]" denotes the articles included in the review sample for this article.

5 "[**]" denotes the articles included in the review sample for this article but not cited in the text.

- individuals with severe chronic obstructive pulmonary disease or advanced cancer. *Journal of Palliative Medicine*, 11(4), 610-620. doi: 10.1089/jpm.2007.0209 [**]
- Davis, W. Ross & Johnson, Bruce D. (2008). Prescription opioid use, misuse, and diversion among street drug users in New York City. *Drug and Alcohol Dependence*, 92(1-3), 267-276. doi: 10.1016/j.drugalcdep.2007.08.008 [*]
- de Wet, Corene (2010). The reasons for and the impact of principal-on-teacher bullying on the victims' private and professional lives. *Teaching and Teacher Education*, 26(7), 1450-1459. doi: 10.1016/j.tate.2010.05.005 [*]
- Denzin, Norman K. & Lincoln, Yvonna S. (Eds.) (2000). *Handbook of qualitative research* (2nd ed.). Thousand Oaks, CA: Sage.
- Dworkin, Shari L. (2012). Sample size policy for qualitative studies using in-depth interviews. *Archives of Sexual Behavior*, 41(6), 1319-1320. doi: 10.1007/s10508-012-0016-6
- Edelbring, Samuel; Dastmalchi, Maryann; Hult, Håkan; Lundberg, Ingrid E. & Dahlgren, Lars Owe (2011). Experiencing virtual patients in clinical learning: A phenomenological study. *Advances in Health Sciences Education*, 16(3), 331-345. doi: 10.1007/s10459-010-9265-0 [*]
- Ek, Kristina & Ternstedt, Britt-Marie (2008). Living with chronic obstructive pulmonary disease at the end of life: a phenomenological study. *Journal of Advanced Nursing*, 62(4), 470-478. doi: 10.1111/j.1365-2648.2008.04611.x [*]
- Emmel, Nick (2013). *Sampling and choosing cases in qualitative research: A realist approach*. London: Sage.
- Fields, Deborah A. & Kafai, Yasmin B. (2009). A connective ethnography of peer knowledge sharing and diffusion in a tween virtual world. *International Journal of Computer-Supported Collaborative Learning*, 4(1), 47-68. doi: 10.1007/s11412-008-9057-1 [**]
- Fried, Terri R.; Bullock, Karen; Iannone, Lynne & O'Leary, John R. (2009). Understanding advance care planning as a process of health behavior change. *Journal of the American Geriatrics Society*, 57(9), 1547-1555. doi: 10.1111/j.1532-5415.2009.02396.x [*]
- Fudge, Nina; Wolfe, Charles D.A.; & McKeivitt, Christopher (2008). Assessing the promise of user involvement in health service development: Ethnographic study. *British Medical Journal*, 336(7639), 313-321. doi: 10.1136/bmj.39456.552257.BE, <http://www.bmj.com/content/336/7639/313> [Accessed: May 9, 2015]. [*]
- Greenhalgh, Trisha; Stramer, Katja; Bratan, Tanja; Byrne, Emma; Mohammad, Yara & Russell, Jill (2008). Introduction of shared electronic records: Multi-site case study using diffusion of innovation theory. *British Medical Journal*, 337, a1786-a1796. doi: 10.1136/bmj.a1786, <http://www.bmj.com/content/337/bmj.a1786> [Accessed: May 9, 2015]. [*]
- Greenhalgh, Trisha; Stramer, Katja; Bratan, Tanja; Byrne, Emma; Russell, Jill & Potts, Henry W.W. (2010). Adoption and non-adoption of a shared electronic summary record in England: A mixed-method case study. *British Medical Journal*, 340, c3111-c3122. doi: 10.1136/bmj.c3111, <http://www.bmj.com/content/340/bmj.c3111> [Accessed: May 9, 2015]. [**]
- Grossman, Pam; Compton, Christa; Igra, Danielle; Ronfeldt, Matthew; Shahan, Emily & Williamson, Peter W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, 111(9), 2055-2100. [*]
- Guest, Greg; Bunce, Arwen & Johnson, Laura (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59-82.
- Haines, Rebecca J.; Poland, Blayne D. & Johnson, Joy L. (2009). Becoming a "real" smoker: Cultural capital in young women's accounts of smoking and other substance use. *Sociology of Health & Illness*, 31(1), 66-80. doi: 10.1111/j.1467-9566.2008.01119.x [*]
- Hopfer, Suellen & Clippard, Jessie R. (2011). College women's HPV vaccine decision narratives. *Qualitative Health Research*, 21(2), 262-277. doi: 10.1177/1049732310383868 [*]
- Horn, Leora; Tzanetos, Katina; Thorpe, Kevin & Straus, Sharon E. (2008). Factors associated with the subspecialty choices of internal medicine residents in Canada. *BMC Medical Education*, 8, Art. 37. doi: 10.1186/1472-6920-8-37, <http://www.biomedcentral.com/1472-6920/8/37> [Accessed: May 9, 2015]. [*]
- Jacoby, A.; Wang, W.; Vu, T. D.; Wu, J.; Snape, D.; Aydemir, N.; Parr, J.; Reis, R.; Begley & Baker, G. (2008). Meanings of epilepsy in its sociocultural context and implications for stigma: Findings from ethnographic studies in local communities in China and Vietnam. *Epilepsy & Behavior*, 12(2), 286-297. doi: 10.1016/j.yebeh.2007.10.006 [*]

- Karamustafaoglu, Orhan (2009). Active learning strategies in physics teaching. *Energy Education Science and Technology Part B-Social and Educational Studies*, 1(1-2), 27-50. [*]
- Karlsson, Agneta; Arman, Maria & Wikblad, Karin (2008). Teenagers with type 1 diabetes—A phenomenological study of the transition towards autonomy in self-management. *International Journal of Nursing Studies*, 45(4), 562-570. doi: 10.1016/j.ijnurstu.2006.08.022 [**]
- Kelchtermans, Geert (2009). Who I am in how I teach is the message: Self-understanding, vulnerability and reflection. *Teachers and Teaching*, 15(2), 257-272. doi: 10.1080/13540600902875332 [**]
- Kendall, Marilyn; Murray, Scott A.; Carduff, Emma; Worth, Allison; Harris, Fiona; Lloyd, Anna; Carvers, Debbie; Grant, Liz; Boyd, Kirsty & Sheikh, Aziz (2009). Use of multiperspective qualitative interviews to understand patients' and carers' beliefs, experiences, and needs. *British Medical Journal*, 339, 196-199. doi: 10.1136/bmj.b4122, <http://www.bmj.com/content/339/bmj.b4122> [Accessed: May 10, 2015].
- Kirkpatrick, H. & Byrne, C. (2009). A narrative inquiry: Moving on from homelessness for individuals with a major mental illness. *Journal of Psychiatric & Mental Health Nursing*, 16(1), 68-75. doi: 10.1111/j.1365-2850.2008.01331.x [*]
- Kuckartz, Udo (2014). *Qualitative text analysis: A guide to methods, practice and using software*. London: Sage.
- Lin, Fu-ren; Lin, Shang-cheng & Huang, Tzu-ping (2008). Knowledge sharing and creation in a teachers' professional virtual community. *Computers & Education*, 50(3), 742-756. doi: 10.1016/i.compedu.2006.07.009 [*]
- Lingard, Lorelei & Kennedy, Tara J. (2007). *Qualitative research in medical education*. Edinburgh: Association for the Study of Medical Education.
- Marshall, Catherine & Rossman, Gretchen B. (2011). *Designing qualitative research* (5th ed.). Thousand Oaks, CA: Sage.
- Martins, Diane Coccozza (2008). Experiences of homeless people in the health care delivery system: A descriptive phenomenological study. *Public Health Nursing*, 25(5), 420-430. doi: 10.1111/j.1525-1446.2008.00726.x [*]
- Mason, Mark (2010). Sample size and saturation in PhD studies using qualitative interviews. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 11(3), Art. 8, <http://nbn-resolving.de/urn:nbn:de:0114-fqs100387> [Accessed: May 1, 2015].
- Maxwell, Joseph A. (2013). *Qualitative research design: An interactive approach* (3rd ed.). Thousand Oaks, CA: Sage.
- Mayes, Terry (2006). *LEX: The learner experience of e-learning methodology report*. Glasgow Caledonian University, http://www.webarchive.org.uk/wayback/archive/20140615232533/http://www.jisc.ac.uk/media/documents/programmes/elearningpedagogy/lex_method_final.pdf [Accessed: May 9, 2015].
- Mayes, Rachel; Llewellyn, Gwynnyth & McConnell, David (2008). Active negotiation: Mothers with intellectual disabilities creating their social support networks. *Journal of Applied Research in Intellectual Disabilities*, 21(4), 341-350. doi: 10.1111/j.1468-3148.2008.00448.x [**]
- Mayring, Philipp (2000). Qualitative content analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 1(2), Art. 20, <http://nbn-resolving.de/urn:nbn:de:0114-fqs0002204> [Accessed: May 9, 2015].
- Merriam, Sharan B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: John Wiley.
- Montgomery, C.M.; Lees, S.; Stadler, J.; Morar, N. S.; Ssali, A.; Mwanza, B.; Mntambo, M.; Phillip, J.; Watts, C. & Pool, R. (2008). The role of partnership dynamics in determining the acceptability of condoms and microbicides. *AIDS Care. Psychological and Socio-Medical Aspects of AIDS/HIV*, 20(6), 733-740. doi: 10.1080/09540120701693974 [**]
- Morse, Diane S.; Edwardsen, Elizabeth A. & Gordon, Howard S. (2008). Missed opportunities for interval empathy in lung cancer communication. *Archives of Internal Medicine*, 168(17), 1853-1858. doi: 10.1001/archinte.168.17.1853 [*]
- Morse, Janice M. (1994). Designing funded qualitative research. In Norman K. Denzin & Yvonna S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 220-235). Thousand Oaks, CA: Sage.
- Morse, Janice M. (2000). Determining sample size. *Qualitative Health Research*, 10(1), 3-5. doi: 10.1177/104973200129118183

- Morse, Janice M. & Field, Peggy Anne (1995). *Qualitative research methods for health professionals* (2nd ed.). Thousand Oaks, CA: Sage.
- O'Reilly, Michelle & Parker, Nicola (2012). "Unsatisfactory saturation": A critical exploration of the notion of saturated sample sizes in qualitative research. *Qualitative Research*, 13(2), 190-197. doi: 10.1177/1468794112446106
- Onwuegbuzie, Anthony J. & Leech, Nancy L. (2007). A call for qualitative power analyses. *Quality & Quantity*, 41(1), 105-121. doi: 10.1007/s11135-005-1098-1
- Orlikowski, Wanda J. (1993). CASE tools as organizational change: Investigating incremental and radical changes in systems development. *MIS Quarterly*, 17, 309-340.
- Ottenbreit-Leftwich, Anne T.; Glazewski, Krista D.; Newby, Timothy J. & Ertmer, Peggy A. (2010). Teacher value beliefs associated with using technology: Addressing professional and student needs. *Computers & Education*, 55(3), 1321-1335. doi: 10.1016/j.compedu.2010.06.002 [*]
- Owens, Christabel; Lambert, Helen; Lloyd, Keith & Donovan, Jenny (2008). Tales of biographical disintegration: How parents make sense of their sons' suicides. *Sociology of Health & Illness*, 30(2), 237-254. doi: 10.1111/j.1467-9566.2007.01034.x [**]
- Park Rogers, Meredith A. & Abell, Sandra K. (2008). The design, enactment, and experience of inquiry-based instruction in undergraduate science education: A case study. *Science Education*, 92(4), 591-607. [*]
- Patton, Michael Quinn (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage.
- Patton, Michael Quinn (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks CA: Sage.
- Patton, Michael Quinn (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Thousand Oaks, CA: Sage.
- Pinnock, Hilary; Kendall, Marilyn; Murray, Scott A.; Worth, Allison; Levack, Pamela; Porter, Mike; MacNee, William & Sheikh, Aziz (2011). Living and dying with severe chronic obstructive pulmonary disease: multi-perspective longitudinal qualitative study. *British Medical Journal*, 342, d142-d152. doi: 10.1136/bmj.d142, <http://www.bmj.com/content/342/bmj.d142> [Accessed: May 9, 2015]. [*]
- PLoS Medicine Editors (2006). The impact factor game. *PLoS Medicine*, 3(6), 0707-0708. doi: 10.1371/journal.pmed.0030291, <http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.0030291> [Accessed: May 9, 2015].
- Putwain, David William (2009). Assessment and examination stress in key stage 4. *British Educational Research Journal*, 35(3), 391-411. doi: 10.1080/01411920802044404 [*]
- Reid, Joanne; McKenna, Hugh; Fitzsimons, Donna & McCance, Tanya (2009). The experience of cancer cachexia: A qualitative study of advanced cancer patients and their family members. *International Journal of Nursing Studies*, 46(5), 606-616. doi: 10.1016/j.ijnurstu.2008.10.012 [**]
- Reissman, Catherine K. (1993). *Narrative analysis*. Newbury Park, CA: Sage.
- Rose, Ellen (2011). The phenomenology of on-screen reading: University students' lived experience of digitised text. *British Journal of Educational Technology*, 42(3), 515-526. doi: 10.1111/j.1467-8535.2009.01043.x [**]
- Rosedale, Mary (2009). Survivor loneliness of women following breast cancer. *Oncology Nursing Forum*, 36(2), 175-183. doi: 10.1188/09.onf.175-183 [**]
- Safman, Rachel M. & Sobal, Jeffery (2004). Qualitative sample extensiveness in health education research. *Health Education & Behavior*, 31(1), 9-21. doi: 10.1177/1090198103259185
- Sandelowski, Margarete (1995). Sample size in qualitative research. *Research in Nursing and Health*, 18, 179-183.
- Sargeant, Joan; Armson, Heather; Chesluk, Ben; Dornan, Timothy; Eva, Kevin; Holmboe, Eric; Lockyer, Jocelyn; Loney, Elaine; Mann, Karen & van der Vleuten, Cees (2010). The processes and dimensions of informed self-assessment: A conceptual model. *Academic Medicine*, 85(7), 1212-1220. doi: 10.1097/ACM.0b013e3181d85a4e [*]
- Smith, Brett & Sparkes, Andrew C. (2008). Changing bodies, changing narratives and the consequences of tellability: A case study of becoming disabled through sport. *Sociology of Health & Illness*, 30(2), 217-236. doi: 10.1111/j.1467-9566.2007.01033.x [**]
- Sobal, Jeffery (2001). Sample extensiveness in qualitative nutrition education research. *Journal of Nutrition Education*, 33(4), 184-192. doi: 10.1016/s1499-4046(06)60030-4

- Stevens, Reed; O'Connor, Kevin; Garrison, Lari; Jocuns, Andrew & Amos, Daniel M. (2008). Becoming an engineer: Toward a three dimensional view of engineering learning. *Journal of Engineering Education*, 97(3), 355-368. [*]
- Storeng, Katerini Tagmatarchi; Baggaley, Rebecca F.; Ganaba, Rasmañe; Ouattara, Fatoumata; Akoum, Melanie S. & Filippi, Véronique (2008). Paying the price: The cost and consequences of emergency obstetric care in Burkina Faso. *Social Science & Medicine*, 66(3), 545-557. doi: 10.1016/j.socscimed.2007.10.001 [*]
- Straus, Sharon E.; Chatur, Fatima & Taylor, Mark (2009). Issues in the mentor-mentee relationship in academic medicine: A qualitative study. *Academic Medicine*, 84(1), 135-139. [*]
- Strauss, Anselm L. & Corbin, Juliet M. (1990). *Basics of qualitative research: Grounded theory, procedures, and techniques*. Newbury Park, CA: Sage.
- Strauss, Anselm L. & Corbin, Juliet M. (1998). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (2nd ed.). Thousand Oaks, CA: Sage.
- Vähäsantanen, Katja; Hökkä, Päivi; Eteläpelto, Anneli; Rasku-Puttonen, Helena & Littleton, Karen (2008). Teachers' professional identity negotiations in two different work organisations. *Vocations and Learning*, 1(2), 131-148. doi: 10.1007/s12186-008-9008-z [**]
- van Aalst, Jan (2010). Using Google Scholar to estimate the impact of journal articles in education. *Educational Researcher*, 39(5), 387-400. doi: 10.3102/0013189X10371120
- Vanclay, Jerome K. (2012). Impact factor: Outdated artefact or stepping-stone to journal certification? *Scientometrics*, 92(2), 211-238. doi: 10.1007/s11192-011-0561-0
- Vavrus, Frances (2009). The cultural politics of constructivist pedagogies: Teacher education reform in the United Republic of Tanzania. *International Journal of Educational Development*, 29(3), 303-311. doi: 10.1016/j.ijedudev.2008.05.002 [**]
- Ware, Norma C.; Idoko, John; Kaaya, Sylvia; Biraro, Irene Andia; Wyatt, Monique A.; Agbaji, Oche; Chalamilla, Guerino & Bangsberg, David R. (2009). Explaining adherence success in sub-Saharan Africa: An ethnographic study. *PLoS Medicine*, 6(1), 39-47. doi: 10.1371/journal.pmed.1000011, <http://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1000011> [Accessed: May 9, 2015]. [*]
- Xu, Yueting & Liu, Yongcan (2009). Teacher assessment knowledge and practice: A narrative inquiry of a Chinese college EFL teacher's experience. *TESOL Quarterly*, 43(3), 493-513. [**]
- Yin, Robert K. (1994). *Case study research: Design and methods* (2nd ed.). Beverly Hills, CA: Sage.

Author

Tim GUETTERMAN is an applied research methodologist in the University of Nebraska-Lincoln's International Mixed Methods Research and Training Academy. His research interests, scholarship, and teaching are in research methodology, namely mixed methods research and qualitative inquiry. He has extensive professional experience in the field of evaluation with a focus on healthcare programs.

Contact:

Timothy C. Guetterman

Department of Educational Psychology
University of Nebraska-Lincoln
114 Teachers College Hall
Lincoln, NE 68588, USA

E-mail: tcguetterman@gmail.com

Citation

Guetterman, Timothy C. (2015). Descriptions of Sampling Practices Within Five Approaches to Qualitative Research in Education and the Health Sciences [48 paragraphs]. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, 16(2), Art. 25, <http://nbn-resolving.de/urn:nbn:de:0114-fqs1502256>.