

Research on Aging

<http://roa.sagepub.com>

Sampling in Qualitative Research: Rationale, Issues, and Methods

Mark R. Luborsky and Robert L. Rubinstein

Research on Aging 1995; 17; 89

DOI: 10.1177/0164027595171005

The online version of this article can be found at:
<http://roa.sagepub.com/cgi/content/abstract/17/1/89>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

Additional services and information for *Research on Aging* can be found at:

Email Alerts: <http://roa.sagepub.com/cgi/alerts>

Subscriptions: <http://roa.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations (this article cites 9 articles hosted on the
SAGE Journals Online and HighWire Press platforms):
<http://roa.sagepub.com/cgi/content/refs/17/1/89>

Sampling in Qualitative Research

Rationale, Issues, and Methods

MARK R. LUBORSKY
ROBERT L. RUBINSTEIN
Philadelphia Geriatric Center

In gerontology the most recognized and elaborate discourse about sampling is generally thought to be in quantitative research associated with survey research and medical research. But sampling has long been a central concern in the social and humanistic inquiry, albeit in a different guise suited to the different goals. There is a need for more explicit discussion of qualitative sampling issues. This article will outline the guiding principles and rationales, features, and practices of sampling in qualitative research. It then describes common questions about sampling in qualitative research. In conclusion it proposes the concept of qualitative clarity as a set of principles (analogous to statistical power) to guide assessments of qualitative sampling in a particular study or proposal.

Questions of what is an appropriate research sample are common across the many disciplines of gerontology, albeit in different guises. The basic questions concern what to observe and how many observations or cases are needed to assure that the findings will contribute useful information. Throughout the history of gerontology, the most recognized and elaborate discourse about sampling has been associated with quantitative research, including survey and medical research. But concerns about sampling have long been central to social and humanistic inquiry (e.g., Mead 1953). The authors argue such concerns remained less recognized by quantitative researchers be-

AUTHORS' NOTE: Support for the first author by the National Institute of Child Health and Human Development (#RO1 HD31526) and the National Institute on Aging (#RO1 AG09065) is gratefully acknowledged. Ongoing support for the second author from the National Institute of Aging is also gratefully acknowledged.

RESEARCH ON AGING, Vol. 17 No. 1, March 1995 89-113
© 1995 Sage Publications, Inc.

cause of differing focus, concepts, and language. Recently, an explicit discussion about concepts and procedures for qualitative sampling issues has emerged. Despite the growing numbers of textbooks on qualitative research, most offer only a brief discussion of sampling issues, and far less is presented in a critical fashion (Gubrium and Sankar 1994; Werner and Schoepfle 1987; Spradley 1979, 1980; Strauss and Corbin 1990; Trotter 1991; but cf. Denzin and Lincoln 1994; DePoy and Gitlin 1993; Miles and Huberman 1994; Peltó and Peltó 1978).

The goal of this article is to extend and further refine the explicit discussion of sampling issues and techniques for qualitative research in gerontology. Throughout the article, the discussion draws on a variety of examples in aging, disability, ethnicity as well as more general anthropology.

The significance of the need to understand qualitative sampling and its uses is increasing for several reasons. First, emerging from the normal march of scientific developments that builds on prior research, there is a growing consensus about the necessity of complementing standardized data with insights about the contexts and insiders' perspectives on aging and the elderly. These data are best provided by qualitative approaches. In gerontology, the historical focus on aging pathology obscured our view of the role of culture and personal meanings in shaping how individuals at every level of cognitive and physical functioning personally experience and shape their lives. The individual embodying a "case" or "symptoms" continues to make sense of, manage, and represent experiences to him- or herself and to others. A second significance to enhancing our appreciation of qualitative approaches to sampling is related to the societal contexts of the scientific enterprise. Shifts in public culture now endorse the inclusion of the experiences and beliefs of diverse and minority segments of the population. A reflection of these societal changes is the new institutional climate for federally funded research, which mandates the inclusion and analysis of data on minorities. Qualitative approaches are valuable because they are suited to assessing the validity of standardized measures and analytic techniques for use with racial and ethnic subpopulations. They also permit us to explore diversities in cultural and personal beliefs, values, ideals, and experiences.

This article will outline the guiding principles and rationales, features, and practices of sampling in qualitative research. It describes the scientific implications of the cultural embeddedness of sampling issues as a pervasive feature in wider society. It then describes common questions about sampling in qualitative research. It concludes by proposing an analog to statistical power, *qualitative clarity*, as a set of principles to guide assessments of the sampling techniques in a study report or research proposal. The term *clarity* was chosen to express the goal of making explicit the details of how the sample was assembled, the theoretical assumptions, and the practical constraints that influenced the sampling process. Qualitative clarity should include at least two components, theoretical grounding and sensitivity to context. The concept focuses on evaluating the strength and flexibility of the analytic tools used to develop knowledge during discovery procedures and interpretation. These can be evaluated even if the factors to be measured cannot be specified.

A wide range of opinions about sampling exists in the qualitative research community. The authors take issue with qualitative researchers who dismiss these as irrelevant or even as heretical concerns. The authors also disagree with those quantitative practitioners who dismiss concerns about qualitative sampling as irrelevant in general on the grounds that qualitative research provides no useful knowledge. It is suggested that such a position is untenable and uninformed.

This article focuses only on qualitative research; issues related to combined qualitative and quantitative methods are not discussed. The focus is on criteria for designing samples; qualitative issues related to suitability of any given person for research are not addressed. The criteria for designing samples constitute what Johnson (1990) labels as "Criteria One issues," the construction and evaluation of theory and data-driven research designs. Criteria Two issues relate to the individual subjects in terms of cooperativeness, rapport, and suitability for qualitative study methods.

Although this article may appear to overly dichotomize qualitative and quantitative approaches, this was done strictly for the purposes of highlighting key issues in a brief space. The authors write here from the perspective of researchers who work extensively with both orientations, singly and in combination, in the conduct of major in-depth and longitudinal research grants that employ both methods. It is the

authors' firm belief that good research requires an openness to multiple approaches to conceptualizing and measurement phenomena.

Contributions, Logic and Issues in Qualitative Sampling

Major contributions. Attention to sampling issues has usually been at the heart of anthropology and of qualitative research since their inception. Much work was devoted to evaluating the appropriateness of theory, design strategies, and procedures for sampling. Important contributions have been made by research devoted to identifying and describing the nature of sample universes and the relevant analytic units for sampling. For example, the "universe of kinship" (Goodenough 1956) has been a mainstay of cross-cultural anthropological study. Kinship studies aim to determine the fundamental culturally defined building blocks of social relationships of affiliation and descent (e.g., Bott 1971; Fortes 1969). Ethnographic investigations document the diversity of kinship structures, categories of kith and kin, and terminologies that give each culture across the globe its distinctive worldview, social structure, family organization, and patterns to individual experiences of the world.

Concerns with sampling in qualitative research focus on discovering the scope and the nature of the universe to be sampled. Qualitative researchers ask, "What are the components of the system or universe that must be included to provide a valid representation of it?" In contrast, quantitative designs focus on determining how many of what types of cases or observations are needed to reliably represent the whole system and to minimize both falsely identifying or missing existing relationships between factors. Thus the important contributions of qualitative work derived from concerns with validity and process may be seen as addressing core concerns of sampling, albeit in terms of issues less typically discussed by quantitative studies. Two examples may clarify this; one concerns time allocation studies of Peruvian farmers and the other addresses a census on Truk Island in the South Pacific.

The Andes mountains of Peru are home to communities of peasants who farm and tend small herds to garner a subsistence living. To help guide socioeconomic modernization and to improve living conditions,

refined time allocation studies (see Gross 1984) were conducted in the 1970s to assess the rational efficiency of traditional patterns of labor, production, and reproduction. Seemingly irrational results were obtained. A systematic survey of how villagers allocated their time to various activities identified a few healthy adults who sat in the fields much of the day. Given the marginal food supplies, such "inactivity" seemed irrational and suggested a possible avenue for the desired interventions to improve village economic production. Only after interviewing the farmers to learn why the men sat in the fields and then calculating the kilocalories of foods gained by putting these men to productive work elsewhere was an explanation uncovered. It was discovered that crop yields and available calories would *decline*, not increase, due to foraging birds and animals. Because the farmers sat there, the events of animal foraging never occurred in the data universe. Here, judgments about the rationality of behaviors were guided by too narrow a definition of the behavioral universe, shaped by reliance on analytic factors external to the system (e.g., biases in industrial economies that equate "busyness" with production). An important message here is that discovery and definition of the sample universe and of relevant units of activity must precede sampling and analyses.

On Truk Island in the South Pacific, two anthropologists each conducted an independent census using the same methods. They surveyed every person in the community. Statistical analyses of these total universe samples were conducted to determine the incidence of types of residence arrangements for newlywed couples. The researchers reached opposite conclusions. Goodenough (1956) argued that his colleague's conclusion that there are no norms for where new couples locate their residence clearly erred by classifying households as patrilocal (near the father), matriloc, or neoloc (not near either parent) at one time as if isolated from other social factors. Goodenough used the same residence typology as did his colleague in his analysis, but identified a strong matrilineal pattern (wife's extended family). Evidence for this pattern becomes clear when the behaviors are viewed in relation to the extended family and over time. The newlyweds settle on whatever space is available but plan to move later to the more socially preferred (e.g., matrilineal) sites. This later aspect was determined by combining survey-based observations of behavior with

interviews to learn “what the devil they think they are doing” (Geertz 1973). Thus different analytic definitions of domestic units led to opposite conclusions, despite the use of a sample of the *total universe* of people! Social constructions of the lived universe, subjectively important temporal factors have to be understood to identify valid units for analyses and interpretation of the data.

The Peruvian and the Truk Island examples illustrate some of the focal contributions of qualitative approaches to sampling. Altering the quantitatively oriented sampling interval, frequency, or duration would not have produced the necessary insights. The examples also suggest some of the dilemmas challenging sampling in qualitative research. These will be addressed in a later section. Both cases reveal the influence of deeply ingrained implicit cultural biases in the scientific construction of the sampling universe and the units for sampling.

The Cultural Embeddedness of the Concept of Sampling

Sampling issues are not exclusive to science. Widespread familiarity with sampling and related issues is indicated by the pervasive popular appetite for opinion and election polls, surveys of consumer product prices and quality, and brief reports of newsworthy scientific research in the mass media. Sampling issues are at the heart of jury selection, which aims to represent a cross section of the community; frequent debates erupt over how to define the universe of larger American society (e.g., by race and gender) to use for juror selection in a specific community. We can shop for sampler boxes of chocolates to get a tasty representation of the universe of all the candies from a company. Debates about the representativeness, size, and biases in survey results because of the people selected for study or the small size of samples are a part of everyday conversation. Newspapers frequently report on medical or social science research, with accounts of experts’ challenging the composition or size of the sample or the wording of the survey questions. Critical skills in sampling are instilled during schooling and on-the-job training.

Such widespread familiarity with basic sampling issues suggests a deep cultural basis for the fascination and thus the need for a more critical understanding. The concept and practices of sampling resonate

with fundamental cultural ideals and taboos. It is perhaps the case that sampling is linked, in American culture, to democratic ideals and notions of inclusion and representation.

What does that mean for qualitative researchers designing sampling strategies? We need to be aware that the language of science is laden with cultural and moral categories. Thus gerontological research may potentially be shaped by both cultural themes masked as scientific principles. Basic terms for research standards can simultaneously apply to ideals for social life (Luborsky 1994). We construct and are admonished by peers to carefully protect independent and dependent variables; we design studies to provide the greatest statistical power and speak of controlling variables. At the same time, psychosocial interventions are designed to enhance these same factors of individual independence and senses of power and control. We examine constructs and data to see if they are valid or invalid; the latter word also is defined in dictionaries as referring to someone who is not upright but physically deformed or sickly. Qualitative research, likewise, needs to recognize that we share with informants in the search for themes and coherence in life, and normatively judge the performance of others in these terms (Luborsky 1994, 1993b).

The ideals of representativeness and proportionality are not, in practice, unambiguous or simple to achieve as is evidenced in the complex jury selection process. Indeed, there is often more than one way to achieve representativeness. Implicit cultural values may direct scientists to define some techniques as more desirable than others. Two current examples illustrate how sampling issues are the source of vitriolic debate outside the scientific community: voting procedures, and the construction or apportionment of voting districts to represent minority, ethnic, or racial groups. Representing "the voice of the people" in government is a core tenet of American democracy, embodied in the slogan "one person one vote." Before women's suffrage, the universe was defined as "one *man* one vote." A presidential nomination for U.S. Attorney General Dr. Lani Guinier, was withdrawn, in part, because she suggested the possibility of an alternative voting system (giving citizens more than one vote to cast) to achieve proportional representation for minorities. We see in these examples that to implement generalized democratic ideals of equal rights and representation can be problematic in the context of the democratic

ideal of majority rule. Another example is the continuing debate in the U.S. Supreme Court over how to reapportion voting districts so as to include sufficient numbers of minority persons to give them a voice in local elections. These examples indicate the popular knowledge of sampling issues, the intensity of feelings about representativeness, and the deep dilemmas about proportional representation and biases arising within a democratic society. The democratic ideals produce multiple conflicts at the ideological level.

It is speculated that the association of sampling issues with such core American cultural dilemmas exacerbates the rancor between qualitative and quantitative gerontology; whereas in disciplines that do not deal with social systems, there is a tradition of interdependence instead of rancor. For example, the field of chemistry includes both qualitative and quantitative methods but is not beset by the tension found in gerontology. Qualitative chemistry is the set of methods specialized in identifying the types and entire range of elements and compounds present in materials or chemical reactions. A variety of discovery-oriented methods are used, including learning which elements are reacting with one another. Quantities of elements present may be described in general ranges as being from a trace to a substantial amount. Quantitative chemistry includes measurement-oriented methods attuned to determining the exact quantity of each constituent element present. Chemists use both methods as necessary to answer research problems. The differences in social contextual factors may contribute to the lower level of tension between quantitative and qualitative traditions within the European social sciences situated as they are within alternative systems for achieving democratic representation in government (e.g., direct plebiscites or multiparty governments rather than the American electoral college approach to a two-party system).

Ideals and Techniques of Qualitative Sampling

The preceding discussion highlighted the need to first identify the ideal or goal for sampling and second to examine the techniques and dilemmas for achieving the ideal. The following section describes several ideals, sampling techniques, and inherent dilemmas. Core

ideals include the determination of the scope of the universe for study and the identification of appropriate analytic units when sampling for meaning

Defining the universe. This is simultaneously one of qualitative research's greatest contributions and greatest stumbling blocks to wider acceptance in the scientific community. As the examples of the Peruvian peasants and Trukese postmarital residence norms illustrated, qualitative approaches that can identify relevant units (e.g., of farming activity or cultural ideals for matrilineal residence) are needed to complement behavioral or quantitative methods if we are to provide an internally valid definition of the scope of the universe to be sampled. Probability-based approaches do not capture these dimensions adequately.

The problem is that the very nature of such discovery-oriented techniques runs counter to customary quantitative design procedures. This needs to be clearly recognized. Because the nature of the units and their character cannot be specified ahead of time, but are to be discovered, the exact number and appropriate techniques for sampling cannot be stated at the design stage but must emerge during the process of conducting the research. One consequence is that research proposals and reports may appear incomplete or inadequate when in fact they are appropriately defined for qualitative purposes. One technique in writing research proposals has been to specify the likely or probable number of subjects to be interviewed.

Evidence that a researcher devoted sufficient attention to these issues can be observed in at least two dimensions. First, one finds a wealth of theoretical development of the concepts and topics. In qualitative research, these serve as the analytic tools for discovery and aid in anticipating new issues that emerge during the analyses of the materials. Second, because standardized measurement or diagnostic tests have not yet been developed for qualitative materials, a strong emphasis is placed on analytic or interpretive perspectives to the data collection and data analyses.

Expository styles, traditional in qualitative studies, present another dilemma for qualitative discussions of sampling. An impediment to wider recognition of what constitutes an adequate design is customary, implicit notions about the "proper" or traditional formats for writing research proposals and journal articles. The traditional format for grant

applications places discussions of theory in the section devoted to the general significance of the research application separate from the methods and measures. However, theoretical issues and conceptual distinctions are the research tools and methods for qualitative researchers, equivalent to the quantitative researchers' standardized scales and measures. As the authors have observed it written reviews of grant applications over many years, reviewers want such "clutter" in qualitative documents placed where it belongs elsewhere in the proposal, not in the design section (Rubinstein 1994). Qualitative researchers look for the analytic refinement, rigor, and breadth in conceptualization linked to the research procedures section as signs of a strong proposal or publication. Thus basic differences in scientific emphases, complicated by expectations for standardized scientific discourse, need to be more fully acknowledged.

Appropriate analytic units: Sampling for meaning. The logic or premises for qualitative sampling for meaning is incompletely understood in gerontology. Although it appears that, in the last decade, there has been an improved interdisciplinary acceptance and communication within gerontology, gerontology is largely driven by a sense of medicalization of social aging and a bias toward survey sampling and quantitative analysis based on "adequate numbers" for model testing and other procedures. At the same time, and partly in reaction to the dominance of the quantitative ethos, qualitative researchers have demurred from legitimating or addressing these issues in their own work.

Understanding the logic behind sampling for meaning in gerontological research requires an appreciation of how it differs from other approaches. By sampling for meaning, the authors indicate the selection of subjects in research that has as its goal the understanding of individuals' naturalistic perceptions of self, society, and the environment. Stated in another way, this is research that takes the insider's perspective. Meaning is defined as the process of reference and connotation, undertaken by individuals, to evoke key symbols, values, and ideas that shape, make coherent, and inform experience (D'Andrade 1984; Good & Good 1982; Luborsky and Rubinstein 1987; Mishler 1986; Rubinstein 1990; Williams 1984). Clearly, the qualitative approach to meaning stands in marked contrast to other approaches to assessing meaning by virtue of its focus on naturalistic data and the discovery of the informant's own evaluations and cate-

gories. For example, one approach assesses meaning by using standardized lists of predefined adjectives or phrases (e.g., semantic differential scale methods, Osgood, Succi, and Tannenbaum 1957); another approach uses diagnostic markers to assign individuals to predefined general types (e.g., depressed, anxious) as a way to categorize people rather than describe personal meaning (e.g., the psychiatric diagnostic manual, DSMIII-R, APA 1987).

The difference between the me of that night and the me of tonight is the difference between the cadaver and the surgeon doing the cutting. (Flaubert, quoted in Crapanzano 1982, p. 181)

It is important to understand that meanings and contexts (including an individual's sense of identity), the basic building blocks of qualitative research, are not fixed, constant objects with immutable traits. Rather, meanings and identities are fluid and changeable according to the situation and the persons involved. Gustave Flaubert precisely captures the sense of active personal meaning-making and remaking across time. Cohler (1991) describes such meaning-making and remaking as the personal life history self, a self that interprets, experiences, and marshals meanings as a means to manage adversity. A classic illustration of the fluidity of meanings is the case presented by Evans-Pritchard (1940) who explains the difficulty he had determining the names of his informants at the start of his fieldwork in Africa. He was repeatedly given entirely different names by the same people. In the kinship-based society, the name or identity one provides to another person depends on factors relative to each person's respective clan membership, age, and community. Now known as the principle of segmentary opposition, the situated and contextual nature of identities was illustrated once the fieldworker discovered the informants were indexing their names to provide an identity at an equal level of social organization. For example, to explain who we are when we travel outside the United States, we identify ourselves as Americans, not as someone from 1214 Oakdale Road. When we introduce ourselves to a new neighbor at a neighborhood block party, we identify ourselves by our apartment building or house on the block, not by reference to our identity as residents at the state or national level.

Themes and personal meanings are markers of processes not fixed structures. Life stories, whose narration is organized around a strongly

held personal theme(s) as opposed to a chronology of events from birth to present day, have been linked with distress and clinical depression (Luborsky 1993b). Williams (1984) suggests that the experience of being ill from a chronic medical disease arises when the disease disrupts the expected trajectory of one's biography. Some researchers argue that a break in the sense of continuity in personal meaning (Becker 1993), rather than any particular meaning (theme), precedes illness and depression (Atchley 1988; Antonovsky 1987).

Another example of fluid meaning is ethnicity. Ethnic identity is a set of meanings that can be fluid and vary according to the social situation, historical time period, and its personal salience over the lifetime (Luborsky and Rubinstein 1987, 1990). Ethnic identity serves as a source of fixed, basic family values during child socialization; more fluidly, as an ascribed family identity to redefine or even reject as part of psychological processes of individuation in early adulthood; sometimes a source of social stigma in communities or in times of war with foreign countries (e.g., "being Italian" during World War II); and a source of continuity of meaning and pride in later life that may serve to help adapt to bereavement and losses.

From the qualitative perspective, there are a number of contrasts that emerge between sampling for meaning and more traditional, survey-style sampling, which has different goals. Those who are not familiar with the sampling-for-meaning approach often voice concerns over such aspects as size (Lieberman 1992), adequacy and, most tellingly, purpose of the sampling. Why, for example, are sample sizes often relatively small? What is elicited and why? What is the relationship between meanings and other traditional categories of analyses, such as age, sex, class, social statuses, or particular diseases?

What is perhaps the most important contrast between the sampling-for-meaning approach and more standard survey sampling is found in the model of the person that underlies elicitation strategies. The model of the person in standard research suggests that important domains of life can be tapped by a relatively small number of standardized "one size fits all" questions, organized and presented in a scientific manner, and that most responses are relatively objective, capable of being treated as a decontextualized trait, and are quantifiable (Mishler 1986; Trotter 1991). From this perspective, individuals are viewed as sets of fixed traits and not as carriers and makers of meaning.

Sampling for meaning, in contrast, is based on four very distinct notions. The first is that responses have contexts and carry referential meaning. Thus questions about events, activities, or other categories of experience cannot be understood without some consideration of how these events implicate other similar or contrasting events in a person's life (Scheer and Luborsky 1991). This is particularly important for older people.

Second, individuals often actively interpret experience. That is to say, many people—but not all—actively work to consider their experience, put it in context, and understand it. Experience is not a fixed response. Further, the concern with meanings or of remaking meaning can be more emergent during some life stages and events or attention to certain kinds of meanings than others. Examples of this include bereavement, retirement, ethnic identity, and personal life themes in later life.

Third, certain categories of data do not have a separable existence apart from their occurrences embodied within routines and habits of the day and the body. Although certain categories of elicited data may have a relatively objective status and be relatively "at hand" for a person's stock of knowledge, other topics may never have been considered in a way that enables a person to have ready access to them (Alexander, Rubinstein, Goodman, and Luborsky 1992). Consequently, qualitative research provides a context and facilitates a process of collaboration between researcher and informant.

Fourth, interpretation, either as natural for the informant or facilitated in the research interview, is basically an action of interpretation of experience that makes reference to both sociocultural standards, be they general cultural standards or local community ones, as well as the ongoing template or matrix of individual experience. Thus, for example, a person knows cultural ideals about a marriage, has some knowledge of other people's marriages, and has intimate knowledge of one's own. In the process of interpretation, all these levels come into play.

These issues occur over a variety of sampling frames and processing frameworks. There are three such sampling contexts. First, sampling for meaning occurs in relation to individuals as representatives of experiential types. Here, the goal is the elucidation of particular types of meaning or experience (personal, setting-based, sociocultu-

ral), through inquiry about, discussion of, and conversation concerning experiences and the interpretation of events and social occurrences. The goal of sampling, in this case, is to produce collections of individuals from whom the nature of experience can be elicited through verbal descriptions and narrations.

Second, sampling for meaning can occur in the context of an individual in a defined social process. An example here could include understanding the entry of a person into a medical practice as a patient, for the treatment of a disorder. Qualitatively, we might wish to follow this person as she moves through medical channels, following referrals, tests, and the like. Even beginning this research at a single primary physician, or with a sample of individuals who have a certain disorder, the structure of passage through a processing system may vary widely and complexly. However, given a fixed point of entry (a medical practice or a single disease), sampling for meaning is nested in ongoing social processes. Researchers wish to understand not only the patient's experience of this setting as she moves through it (e.g., Esteroff 1982) but also the perspectives of the various social actors involved.

Finally, researchers may wish to consider sampling for meaning in a fixed social setting. In a certain way, sampling for meaning in a fixed social setting is what is meant, in anthropology and other social sciences, by "participant observation." The social setting is more or less fixed, as is the population of research informants. An example might be a nursing home unit, with a more or less fixed number of residents, some stability but some change, and regular staff of several types representing distinctive organizational strata and interests (administration, medicine, nursing, social work, aides, volunteers, family, or environmental services).

It is important to note that even though qualitative research focuses on the individual, subjectivity or individuality is not the only goal of study. Qualitative research can focus on the macrolevel. One basic goal of qualitative research in aging is to describe the contents of people's experiences of life, health, and disability. It is true that much of the research to date treats the individual as the basic unit of analysis. Yet, the development of insights into the cultural construction of life experiences is an equal priority because cultural beliefs and values instill and shape power-

ful experiences, ideals, and motivations and shape how individuals make sense of and respond to events.

Studying how macrolevel cultural and community ideologies pattern the microlevel of individual life is part of a tradition stretching from Margaret Mead, Max Weber, Robert Merton, Talcott Parsons, to studies of physical and mental disabilities by Edgerton (1967), Esteroff (1982), and Murphy (1987). For example, Stouffer's (1949) pioneering of survey methods revealed that American soldiers in World War II responded to the shared adversity of combat differently according to personal expectations based on sociocultural value patterns and lived experiences. These findings further illustrate Merton's theories of relative deprivation and reference groups, which point to the basis of individual well-being in basic processes of social comparison.

The notion of stigma illustrates the micro- and the macrolevels of analyses. For example, stigma theory's long reign in the social and political sciences and in clinical practice illustrates the micro- and macroqualitative perspectives. Stigma theory posits that individuals are socially marked or stigmatized by negative cultural evaluations because of visible differences or deformities, as defined by the community. Patterns of avoidance and denial of the disabled mark the socially conditioned feelings of revulsion, fear, or contagion. Personal experiences of low self-esteem result when negative messages are internalized by, for example, persons with visible impairments, or the elderly in an ageist setting. Management of social stigma by individuals and family is as much a focus as is management of impairments. Stigma is related significantly to compliance with prescribed adaptive devices (Zola 1982; Luborsky 1993a). A graphic case of this phenomenon are polio survivors who were homebound due to dependence on massive bedside artificial ventilators. With the recent advent of portable ventilators, polio survivors gained the opportunity to become mobile and travel outside the home, but they did not adopt the new equipment, because the new independence was far outweighed by the public stigma they experienced (Kaufert and Locker 1990).

A final point is that sampling for meaning can also be examined in terms of sampling within the data collected. For example, the entire corpus of materials and observations with informants needs to be examined in the discovery and interpretive processes aimed at describing relevant units for analyses and dimensions of meaning. This is in

contrast to reading the texts to describe and confirm a finding without then systematically rereading the texts for sections that may provide alternative or contradictory interpretations.

Techniques for selecting a sample. As discussed earlier, probability sampling techniques cannot be used for qualitative research by definition, because the members of the universe to be sampled are not known a priori, so it is not possible to draw elements for study in proportion to an as yet unknown distribution in the universe sampled. A review of the few qualitative research publications that treat sampling issues at greater length (e.g., Depoy and Gitlin 1993; Miles and Huberman 1994; Morse 1994; Ragin and Becker 1992) identify five major types of nonprobability sampling techniques for qualitative research. A consensus among these authors is found in the paramount importance they assign to theory to guide the design and selection of samples (Platt 1992). These are briefly reviewed as follows.

First, convenience (or opportunistic) sampling is a technique that uses an open period of recruitment that continues until a set number of subjects, events, or institutions are enrolled. Here, selection is based on a first-come, first-served basis. This approach is used in studies drawing on predefined populations such as participants in support groups or medical clinics. Second, purposive sampling is a practice where subjects are intentionally selected to represent some explicit predefined traits or conditions. This is analogous to stratified samples in probability-based approaches. The goal here is to provide for relatively equal numbers of different elements or people to enable exploration and description of the conditions and meanings occurring within each of the study conditions. The objective, however, is not to determine prevalence, incidence, or causes. Third, snowballing or word-of-mouth techniques make use of participants as referral sources. Participants recommend others they know who may be eligible. Fourth, quota sampling is a method for selecting numbers of subjects to represent the conditions to be studied rather than to represent the proportion of people in the universe. The goal of quota sampling is to assure inclusion of people who may be underrepresented by convenience or purposeful sampling techniques. Fifth, case study (Ragin and Becker 1992; Patton 1990) samples select a single individual, institution, or event as the total universe. A variant is the key-informant approach (Spradley 1979), or intensity sampling

(Patton 1990) where a subject who is expert in the topic of study serves to provide expert information on the specialized topic. When qualitative perspectives are sought as part of clinical or survey studies, the purposive, quota, or case study sampling techniques are generally the most useful.

How many subjects is the perennial question. There is seldom a simple answer to the question of sample or cell size in qualitative research. There is no single formula or criterion to use. A "gold standard" that will calculate the number of people to interview is lacking (cf. Morse 1994). The question of sample size cannot be determined by prior knowledge of effect sizes, numbers of variables, or numbers of analyses—these will be reported as findings. Sample sizes in qualitative studies can only be set by reference to the specific aims and the methods of study, not in the abstract. The answer only emerges within a framework of clearly stated aims, methods, and goals and is conditioned by the availability of staff and economic resources.

Rough "rules of thumb" exist, but these derive from three sources: traditions within social science research studies of all kinds, commonsense ideas about how many will be enough, and practical concerns about how many people can be interviewed and analyzed in light of financial and personnel resources. In practice, from 12 to 26 people in each study cell seems just about right to most authors. In general, it should be noted that Americans have a propensity to define bigger as better and smaller as inferior. Quantitative researchers, in common with the general population, question such small sample sizes because they are habituated to opinion polls or epidemiology surveys based on hundreds or thousands of subjects. However, sample sizes of less than 10 are common in many quantitative clinical and medical studies where statistical power analyses are provided based on the existence of very large effect sizes for the experimental versus control conditions.

Other considerations in evaluating sample sizes are the resources, times, and reporting requirements. In anthropological field research, a customary formula is that of the one to seven: for every 1 year of fieldwork by one researcher, 7 years are required to conduct the analysis. Thus, in studies that use more than one interviewer, the ability to collect data also increases the burden for analyses.

An outstanding volume exploring the logic, contributions, and dilemmas of case study research (Ragin and Becker 1992) reports that survey researchers resort to case examples to explain ambiguities in their data, whereas qualitative researchers reach for descriptive statistics when they do not have a clear explanation for their observations. Again, the choice of sample size and group design is guided by the qualitative goal of describing the nature and contents of cultural, social, and personal values and experiences within specific conditions or circumstances, rather than of determining incidence and prevalence.

Who and who not? In the tradition of informant-based and of participatory research, it is assumed that all members of a community can provide useful information about the values, beliefs, or practices in question. Experts provide detailed, specialized information, whereas nonexperts do so about daily life. In some cases, the choice is obvious, dictated by the topic of study, for example, childless elderly, retirees, people with chronic diseases or new disabilities. In other cases, it is less obvious, as in studies of disease, for example, that require insights from sufferers but also from people not suffering to gain an understanding for comparison with the experiences and personal meanings of similar people without the condition. Comparisons can be either on a group basis or matched more closely on a one-to-one basis for many traits (e.g., age, sex, disease, severity), sometimes referred to as yoked pairs. However, given the labor-intensive nature of qualitative work, sometimes the rationale for including control groups of people who do not have the experiences is not justifiable.

Homogeneity or diversity. Currently, when constructing samples for single study groups, qualitative research appears to be about equally split in terms of seeking homogeneity or diversity. There is little debate or attention to these contrasting approaches. For example, some argue that it is more important to represent a wide range of different types of people and experiences in order to represent the similarities and diversity in human experience, beliefs, and conditions (e.g., Kaufman 1987, 1989) than it is to include sufficient numbers of people sharing an experience or condition to permit evaluation of within-group similarities. In contrast, others select informants to be relatively homogeneous on several characteristics to strengthen com-

parability within the sample as an aid to identifying similarities and diversity.

Summary and Reformulation for Practice

To review, the authors suggest that explicit objective criteria to use for evaluating qualitative research designs do exist, but many of these focus on different issues and aspects of the research process, in comparison to issues for quantitative studies. This article has discussed the guiding principles, features, and practices of sampling in qualitative research. The guiding rationale is that of the discovery of the insider's view of cultural and personal meanings and experience. Major features of sampling in qualitative research concern the issues of identifying the scope of the universe for sampling and the discovery of valid units for analyses. The practices of sampling, in comparison to quantitative research, are rooted in the application of multiple conceptual perspectives and interpretive stances to data collection and analyses that allow the development and evaluation of a multitude of meanings and experiences.

This article noted that sampling concerns are widespread in American culture rather than in the esoteric specialized concern of scientific endeavors (Luborsky and Sankar 1993). Core scientific research principles are also basic cultural ideals (Luborsky 1994). For example, "control" (statistical, personal, machinery), dependence and independence (variables and individual), a reliable person with a valid driver's license matches reliability and validity concerns about assessment scales. Knowledge about the rudimentary principles of research sampling is widespread outside of the research laboratory, particularly with the relatively new popularity of economic, political, and community polls as a staple of news reporting and political process in democratic governance. Core questions about the size, sources, and features of participants are applied to construct research populations, courtroom juries, and districts to serve as electoral universes for politicians.

The cultural contexts and popular notions about sampling and sample size have an impact on scientific judgments. It is important to acknowledge the presence and influence of generalized social sensi-

bilities or awareness about sampling issues. Such notions may have less direct impact on research in fields with long-established and formalized criteria and procedures for determining sample size and composition. The generalized social notions may come to exert a greater influence as one moves across the spectrum of knowledge-building strategies to more qualitative and humanistic approaches. Even though such studies also have a long history of clearly articulated traditions of formal critiques (e.g., in philosophy and literary criticism), they have not been amenable to operationalization and quantification.

The authors suggested that some of the rancor between qualitative and quantitative approaches is rooted in deeper cultural tensions. Prototypic questions posed to qualitative research in interdisciplinary settings derive from both the application of frameworks derived from other disciplines' approaches to sampling as well as those of the reviewers as persons socialized into the community where the study is conceived and conducted. Such concerns may be irrelevant or even counterproductive.

Qualitative Clarity as an Analog to Statistical Power

The guiding logic of qualitative research, by design, generally prevents it from being able to fulfill the assumptions underlying statistical power analyses of research designs. The discovery-oriented goals, use of meanings as units of analyses, and interpretive methods of qualitative research dictate that the exact factors, dimensions, and distribution of phenomena identified as important for analyses may not always be specified prior to data analyses activities. These emerge from the data analyses and are one of the major contributions of qualitative study. No standardized scales or tests exist yet to identify and describe new arenas of cultural, social, or personal meanings. Meaning does not conform to normative distributions by known factors. No probability models exist that would enable prediction of distributions of meanings needed to perform statistical power analyses.

Qualitative studies however can, and should, be judged in terms of how well they meet the explicit goals and purposes relevant to such research.

The authors have suggested that the concept of *qualitative clarity* be developed to guide evaluations of sampling as an analog to the concept of statistical power. Qualitative clarity refers to principles that are relevant to the concerns of this type of research. That is, the adequacy of the strength and flexibility of the analytic tools used to develop knowledge during discovery procedures and interpretation can be evaluated even if the factors to be measured cannot be specified. The term *clarity* conveys the aim of making explicit, for open discussion, the details of how the sample was assembled, the theoretical assumptions and the pragmatic constraints that influenced the sampling process. Qualitative clarity should include at least two components, theoretical grounding and sensitivity to context. These are briefly described next.

Rich and diverse theoretical grounding. In the absence of standardized measures for assessing meaning, the analogous qualitative research tools are theory and discovery processes. Strong and well-developed theoretical preparation is necessary to provide multiple and alternative interpretations of the data. Traditionally, in qualitative study, it is the richness and sophistication of the analytic perspectives or "lenses" focused on the data that lends richness, credibility, and validity to the analyses. The relative degree of theoretical development in a research proposal or manuscript is readily apparent in the text, for example, in terms of extended descriptions of different schools of thought and possible multiple contrasting of interpretive explanations for phenomena at hand. In brief, the authors argue that given the stated goal of sampling for meaning, qualitative research can be evaluated to assess if it has adequate *numbers of conceptual perspectives* that will enable the study to identify a variety of meanings and to critique multiple rich interpretations of the meanings.

Sampling within the data is another important design feature. The discovery of meaning should also include sampling within the data collected. The entire set of qualitative materials should be examined rather than selectively read after identifying certain parts of the text to describe and confirm a finding without reading for sections that may provide alternative or contradictory interpretations.

Sensitivity to contexts. As a second component of qualitative clarity, sensitivity to context refers to the contextual dimensions shaping the meanings studied. It also refers to the historical settings of the scien-

tific concepts used to frame the research questions and the methods. Researchers need to be continually attentive to examining the meanings and categories discovered for elements from the researchers' own cultural and personal backgrounds. The first of these contexts is familiar to gerontologists: patterns constructed by the individual's life history; generation; cohort; psychological, developmental, and social structure; and health. Another more implicit contextual aspect to examine as part of the qualitative clarity analysis is evidence of a critical view of the methods and theories introduced by the investigators. Because discovery of the insiders' perspective on cultural and personal meanings is a goal of qualitative study, it is important to keep an eye to biases derived from the intrusion of the researcher's own scientific categories. Qualitative research requires a critical stance as to both the kinds of information and the meanings discovered, and to the analytic categories guiding the interpretations. One example is recent work that illustrates how traditional gerontological constructs for data collection and analyses do not correspond to the ways individuals themselves interpret their own activities, conditions, or label their identities (e.g., "caregiver," Abel 1991; "disabled," Murphy 1987; "old and alone," Rubinstein, 1986; "Alzheimer's disease," Gubrium 1992; "life themes," Luborsky 1993b). A second example is the growing awareness of the extent to which past research tended to define problems of disability or depression narrowly in terms of the individual's ability, or failure, to adjust, without giving adequate attention to the societal level sources of the individual's distress (Cohen and Sokolovsky 1989). Thus researchers need to demonstrate an awareness of how the particular questions guiding qualitative research, the methods and styles of analyses, are influenced by cultural and historical settings of the research (Luborsky and Sankar 1993) in order to keep clear whose meanings are being reported.

To conclude, our outline for the concept of qualitative clarity, which is intended to serve as the qualitatively appropriate analog to statistical power, is offered to gerontologists as a summary of the main points that need to be considered when evaluating samples for qualitative research. The descriptions of qualitative sampling in this article are meant to extend the discussion and to encourage the continued development of more explicit methods for qualitative research.

REFERENCES

- Abel, Emily. 1991. *Who Cares for the Elderly*. Philadelphia: Temple University Press.
- Alexander, Baine, Robert Rubinstein, Marcene Goodman, and Mark Luborsky. 1992. "A Path Not Taken: A Cultural Analysis of Regrets and Childlessness in the Lives of Older Women." *The Gerontologist* 32(5):618-26.
- American Psychiatric Association (APA). 1987. *Diagnostic and Statistical Manual of Mental Disorders DSM-III-R revised*. Washington, DC: APA.
- Antonovsky, Aaron. 1987. *Unraveling the Mystery of Health*. San Francisco: Jossey-Bass.
- Atchley, Robert. 1988. "A Continuity Theory of Aging." *The Gerontologist* 29(2):183-90.
- Becker, Gaylene. 1993. "Continuity After a Stroke: Implications of Life-Course Disruptions in Old Age." *The Gerontologist* 33(2):148-58.
- Bott, Elizabeth. 1971. *Family and Social Networks*. London: Tavistock.
- Cohen, Carl and Jay Sokolovsky. 1989. *Old Men of the Bowery*. New York: Guilford.
- Cohler, Bertram. 1991. "The Life Story and the Study of Resilience and Response to Adversity." *Journal of Life History and Narrative* 1(2&3):169-200.
- Crapanzano, Vincent. 1982. "The Self, the Third, and Desire." In *Psychosocial Theories of the Self*, edited by B. Lee. New York: Plenum.
- D'Andrade, Roy. 1984. "Cultural Meaning Systems." In *Culture Theory: Essays on Mind, Self, and Emotion*, edited by R. Shweder and R. LeVine. New York: Cambridge Press.
- Denzin, Norman and Yolanda Lincoln. 1994. *Handbook of Qualitative Research*. Thousand Oaks, CA: Sage.
- DePoy, Elizabeth and Laura Gitlin. 1993. *Introduction to Research: Multiple Strategies for Health and Human Services*. St. Louis, MO: Mosby.
- Edgerton, Robert. 1967. *The Cloak of Competence*. Berkeley: University of California Press.
- Esteroff, Susan. 1982. *Making It Crazy: An Ethnography of Psychiatric Patients in an American Community*. Berkeley: University of California Press.
- Evans-Pritchard, Edmund E. 1940. *The Nuer: A Description of the Livelihood and Political Institutions of a Nilotic People*. Cambridge, England: Cambridge University Press.
- Fortes, Meyer. 1969. *Kinship and the Social Order*. Chicago: Aldine.
- Geertz, Clifford. 1973. *The Interpretation of Culture*. New York: Basic Books.
- Good, Byron and Mary-Jo Delvechio Good. 1982. "Toward a Meaning-Centered Analysis of Popular Illness Categories." In *Cultural Conceptions of Mental Health and Therapy*, edited by A. Marsella and G. White. Dordrecht, Holland: Reidel.
- Goodenough, Ward. 1956. "Residence Rules." *Southwestern Journal of Anthropology* 12(1):22-37.
- Gross, Daniel. 1984. "Time Allocation: A Tool for the Study of Cultural Behavior." *Annual Review of Anthropology* 13:519-58.
- Gubrium, Jay. 1992. *The Mosaic of Care*. New York: Springer.
- Gubrium, Jaber and Andrea Sankar. 1994. *Qualitative Methods in Aging Research*. Thousand Oaks, CA: Sage.
- Johnson, John. 1990. *Selecting Ethnographic Informants*. Thousand Oaks, CA: Sage.
- Kaufert, Joseph and David Locker. 1990. "Rehabilitation Ideology and Respiratory Support Technology." *Social Science and Medicine* 30(8):867-77.
- Kaufman, Sharon. 1987. *The Ageless Self: Sources of Meaning in Late Life*. Madison: University of Wisconsin Press.
- . 1989. "Long-Term Impact of Injury on Individuals, Families, and Society: Personal Narratives and Policy Implications." In *Cost of Injury in the United States: A Report to Congress*, edited by D. Rich, Ellen MacKenzie and Associates. San Francisco, CA: Institute

- for Health and Aging, University of California Press, and Injury Prevention Center, Johns Hopkins University Press.
- Lieberson, Stanley. 1992. "Small N's and Big Conclusions." In *What is a Case?* edited by C. Ragin and H. Becker. Cambridge, England: Cambridge University Press.
- Luborsky, Mark. 1993a. "Sociocultural Factors Shaping Technology Usage: Fulfilling the Promise." *Technology and Disability* 2(1):71-8.
- . 1993b. "The Romance With Personal Meaning in Gerontology: Cultural Aspects of Life Themes." *The Gerontologist* 33(4):445-52.
- . 1994. "The Identification and Analysis of Themes and Patterns." In *Qualitative Methods in Aging Research*, edited by J. Gubrium and A. Sankar. Thousand Oaks, CA: Sage.
- Luborsky, Mark and Robert Rubinstein. 1987. "Ethnicity and Lifetimes: Self Concepts and Situational Contexts of Ethnic Identity in Late Life." In *Ethnic Dimensions of Aging*, edited by D. Gelfand and C. Barresi. New York: Springer.
- . 1990. "Ethnic Identity and Bereavement in Later Life: The Case of Older Widowers." In *The Cultural Context of Aging: Worldwide Perspectives*, edited by J. Sokolovsky. New York: Bergin and Garvey.
- Luborsky, Mark and Andrea Sankar. 1993. "Extending the Critical Gerontology Perspective: Cultural Dimensions." *The Gerontologist* 33(4):440-4.
- Mead, Margaret. 1953. "National Character." In *Anthropology Today*, edited by A. Kroeber. Chicago: University of Chicago Press.
- Miles, M. and A. Huberman. 1994. *Qualitative Data Analysis*. Thousand Oaks, CA: Sage.
- Mishler, Elliott. 1986. *Research Interviewing*. Cambridge, MA: Harvard University Press.
- Morse, Janet. 1994. "Designing Funded Qualitative Research." In *Handbook of Qualitative Research*, edited by N. Denzin and Y. Lincoln. Thousand Oaks, CA: Sage.
- Murphy, Robert. 1987. *The Body Silent*. New York: Columbia University Press.
- Osgood, Charles, G. Succi, and P. Tannenbaum. 1957. *The Measurement of Meaning*. Urbana: University of Illinois Press.
- Patton, Michael. 1990. *Qualitative Evaluation and Research Methods*. Thousand Oaks, CA: Sage.
- Pelto, Peter and Gertrude Pelto. 1978. *Anthropological Research: The Structure of Inquiry*, 2nd ed. Cambridge, England: Cambridge University Press.
- Platt, Joseph. 1992. "Cases of Cases." In *What is a Case?* edited by C. Ragin and H. Becker. Cambridge, England: Cambridge University Press.
- Ragin, Charles and Howard Becker. 1992. *What is a Case?: Exploring the Foundations of Social Inquiry*. Cambridge, England: Cambridge University Press.
- Rubinstein, Robert. 1986. *Singular Paths: Old Men Living Alone*. New York: Columbia University Press.
- . 1990. "The Environmental Representation of Personal Themes by Older People." *Journal of Aging Studies* 4(2):131-8.
- . 1994. "Proposal Writing." In *Qualitative Research Methods in Aging Research*, edited by J. Gubrium and A. Sankar. Thousand Oaks, CA: Sage.
- Scheer, Jessica and Mark Luborsky. 1991. "The Cultural Context of Polio Biographies." *Orthopedics* 14(11):1173-81.
- Spradley, James. 1979. *The Ethnographic Interview*. New York: Holt, Rinehart & Winston.
- . 1980. *Participant Observation*. New York: Holt, Rinehart & Winston.
- Stouffer, S. A. 1949 [1965]. *The American Soldier*. Vols. 1 & 2. New York: Wiley.
- Strauss, Anselm and Juliet Corbin. 1990. *Basics of Qualitative Research: Grounded Theory Procedures and Techniques*. Thousand Oaks, CA: Sage.

- Trotter, Robert. 1991. "Ethnographic Research Methods for Applied Medical Anthropology." In *Training Manual in Applied Medical Anthropology*, edited by C. Hill. Washington, DC: American Anthropological Association.
- Werner, Oswald and George Schoepfle. 1987. *Systematic Fieldwork*. Vols. 1 & 2. Thousand Oaks, CA: Sage.
- Williams, Gareth. 1984. "The Genesis of Chronic Illness: Narrative Reconstruction." *Sociology of Health and Illness* 6(2):175-200.
- Zola, Irving. 1982. *Missing Pieces: Chronicle of Living With a Disability*. Philadelphia: Temple University Press.

Mark R. Luborsky, Ph.D., is a senior research anthropologist and assistant director of research at the Philadelphia Geriatric Center. Federal and foundation grants support his studies of sociocultural values and personal meanings in early and late adulthood, and how these relate to mental and physical health, and to disability and rehabilitation processes. He also consults and teaches on these topics.

Robert L. Rubinstein, Ph.D., is a senior research anthropologist and director of research at the Philadelphia Geriatric Center. He has conducted research in the United States and Vanuatu, South Pacific Islands. His gerontological research interests include social relations of the elderly, childlessness in later life, and the home environments of old people.