

Qualitative research: Thoughts on how to do it; how to judge it; when to use it

Qualitative methods have been gaining acceptance in biomedical research over the past decades.^{1,2} Nevertheless, skepticism remains about the validity, reliability, generalizability and general value of this type of work. The most enthusiastic acceptance of the qualitative research has been as a way to garner conceptual information about understudied domains in order to generate hypotheses for later quantitative studies. But many qualitative researchers reject this limitation on the perceived value of their approach. They assert that qualitative research asks different sorts of questions, ones that quantitative methods simply cannot answer.

The impetus for writing this editorial was to accompany and help situate an excellent example of the appropriate use of qualitative research methods. In this issue of *Genetics in Medicine*, Geller et al.³ compare how scientists and science writers view and describe the same set of genetics discoveries. The authors are not interested in measuring the amount of convergence on “facts” that exists between those who discover and those who describe discoveries to the public, but rather on obtaining a deeper understanding of the relationship and interaction between them. The study suggests, among other things, that a too quick assumption of misunderstanding and cross-purposes between scientists and science writers is unwarranted.

The purpose of this editorial is briefly to examine what qualitative research does, how it does it, and how a reader unfamiliar with this method may be able to judge the merit and believability of a particular study using qualitative methods.

WHAT'S “TRUTH” GOT TO DO WITH IT?

There is an assumption, not least among qualitative researchers, that the argument between qualitative and quantitative methods is reducible to an argument over positivism. That is, it is assumed that quantitative social and behavioral science presuppose a regularity in social phenomena similar to that found in the natural world and aim to construct theories on a level of generality and regularity similar to that assumed to exist in the natural sciences. Conversely, it is assumed that qualitative methods are based on the view that all events, actors, and experiences are particular, completely situated in a unique experience, socially constructed and therefore immune to prediction or replicable explanation. I would argue that neither quantitative nor qualitative social scientists really hold these views. Psychologists using survey methods, for example, are aware of research participants’ desires for social acceptability and how that may skew responses even on validated instruments. And very few qualitative researchers study the experi-

ence of an individual without the desire to illuminate the experiences of other individuals in a similar situation.

Closer to the heart of the matter is the fact that quantitative methods already entail a position on the nature of social truth. It is contained within the discipline of statistics and bounded by the wisdom of the null hypothesis – a study can never *prove* a point but merely disprove the counter point. Qualitative researchers, however, drawn to the richness and ambiguity of words, must constantly confront their views on the nature of (social) truth. And, in fact, varying among themselves they discuss this issue a great deal.

At one end of the spectrum is the position marked by doubts about the ability of a researcher to very accurately observe any phenomenon outside him/herself. At its most radical, this work is thus focused on the “self” of the researcher, who becomes a mirror through which the phenomenon is glimpsed. It is likely the association of this postmodern stance with qualitative methods that has led to the most skepticism from quantitative researchers. But qualitative researchers who work in the biomedical field are generally doing “applied” social science and rarely take this position. In fact, many believe that this demurral about the ability to ever know anything leads merely to paralysis.

At the other end of the spectrum are qualitative researchers who use semi-structured interview guides with large samples and clearly delineated methods for inter-coder reliability. They tend to keep their analyses very close to the overt level of their participants’ statements, turning participants’ words into codes, which can then be tabulated and counted. But this position has been subject to criticism as well. For example, well-known medical sociologist, Kathy Charmaz, told a group of qualitative researchers that she found it ironic that many of them had moved so far away from the initial roots of qualitative methods – the realization of the complexity of processes of human interaction. Instead, they have come to “argue that people will tell you what most concerns them in the setting. I contend that they often cannot. The most important processes are tacit.”⁴

I would contend that it is between these two poles that the best qualitative research is done and that despite varying methods, qualitative research is unified by its struggle to make useful and accurate statements about the social world while constantly acknowledging that:

- Each person’s “truth” in any situation will be relative, partial, and dependent on his/her current context; thus there will always be multiple truths in any situation.
- Each person has a strong drive to create a coherent narrative of his/her life. This produces rich material which gives

indications of what is most important to an individual in a particular social setting, but the coherence of the story is illusory and misleading if seen as merely an accurate recitation of events.

- Individuals' reasons for their actions and choices are often not directly accessible to them; in addition, people can deeply hold multiple and even conflicting opinions on the same topic.

METHODS OF DATA COLLECTION

It is perhaps a giddy delight at the rightness of these observations that more than anything else bonds qualitative researchers together. However, after the realization comes the conundrum: If there are so many limitations on identifying and eliciting the "truth" about any situation and yet one wants to make a contribution to knowledge, what methods can one possibly use to do one's work? The answer is a wide array of methods ranging from the very time-consuming practice of ethnography to semi-structured interview methods.

Ethnography is the quintessential qualitative method. The researcher is a "participant-observer" who spends long periods of time in the setting of interest, conducting interviews, talking casually to individuals, observing interactions and conversations and being there at various moments of decision as well as reflection. Renee Anspach, for example, spent 16 months observing neonatal intensive care units (NICUs).⁵ Her work richly describes the texture of the conflicts in which nurses and physicians often found themselves around the prognoses of critically ill infants, and it shows how these conflicts were based in part on the fact that what each intuitively counted as "data" differed. She also observed NICU staff move parents to the periphery of decisions and '...employ a set of practices to elicit parents' assent to decisions that had already been made' (p. 166).⁵ Her ethnography is a recreation of an entire world and an illumination of on-the-ground, on-the-spot clinical ethics.

However, ethnography is ultimately a very luxurious method of working and few researchers past the stage of a dissertation have time to do such work. Qualitative researchers thus try to approach the advantages of field research using other methods. One of the most commonly cited approaches is what sociologist, Erving Goffman, called "triangulation." Goffman believed that what people say about a situation often differs from what they actually do in that situation and he thus preferred to observe people's actions or, if doing interviews, to get at least three individuals' "take" on the same situation – thus, "triangulation." The term has expanded to also mean "triangulating" on a problem by bringing different *types* of data to bear.

For example, Carole Browner and I studied the use of maternal serum alpha fetoprotein testing (MSAFP) by an ethnically diverse group of women shortly after the California state mandate to offer this test.⁶ We designed a study with multiple sources of data including observations of prenatal intake appointments and group educational sessions; interviews with women at two points in their pregnancies; interviews with the

nursing staff that did the prenatal intakes; compilation of all the educational materials women were given about MSAFP; and, ultimately, a chart review to calculate the percentage of test uptake and its variability based on sociodemographics. This approach proved crucial to our eventual analysis, especially as we found that well over 80% of the women had accepted MSAFP testing, and their reported reason was most commonly "why not?" and "I want to do everything I can to help my baby." It was only through a process of data triangulation that we arrived at our analysis of how the medicolegal pressures on physicians and nurses to offer testing interacted with the beliefs of both health care providers and pregnant women about such difficult to discuss topics as disability and pregnancy termination. The use of almost precisely the same language to describe the purposes and advantages of routine prenatal care with the very different purposes of prenatal genetic testing was an important clue in our analysis of the "shared silences" of all participants.

Anthropologist, Clifford Geertz, famously stated that the role of ethnography is to provide "thick description" rather than prediction. Ethnography and complex triangulation methods can provide such thick descriptions because they are able to invest time in what might be called "thick observation." The questions such studies pose tend to be large and of the form *what is the experience of X like?* when "X" may be as broad as the construction of a new sense of self following a disease diagnosis, or *How does Y occur* when "Y" is as diffuse as how decisions are made to terminate life supports in a NICU.

Most frequently, however, qualitative research in the biomedical realm depends on interviews alone. In an interview only study, qualitative researchers are cut off from the approaches that best fit with our view of the complex, multiplicity of social truth. Various approaches are taken to try to deal with this challenge.

One technique used by many investigators is to interview the same participant multiple times. This allows the researcher to observe processes unfold over time as themes lose or gain in importance; to "test" predictions made in one interview (e.g. *my husband will not support my decision to not have prenatal testing*) against what has actually occurred by the time of the next interview. It provides another version of "triangulation," as participants provide their own, implicit comparison. Multiple contacts also allow for greater intimacy to develop between researcher and participant, thus leading, it is hoped, to more honest conversation and less "stage-managed" interview behavior.

But multiple interviews also constitute a large commitment of resources and logistics and so many qualitative projects involve only one interview with each participant. At this point the question is what sort of interview to do. The options fall under the broad categories of open-ended and semi-structured interviews. Open-ended interviews have very loose interview guides which generally pose 10-20 general questions. The answers are not meant to be strictly comparable from one participant to the next but rather to stimulate each participant to expand on the same general areas. For example, open-ended

interviews done by Dobscha et al.⁷ with physicians in Oregon who had received requests for physician-assisted suicide, asked each physician to tell the story of the request that stayed most in his/her mind. The interviewer then followed up to get the participant to describe the experience as fully as possible. The researcher who chooses to do open-ended interviews will, as a general rule, have interviews that are longer (two hour interviews are not uncommon) and the number of participants will be fewer, among other reasons in order to keep the data set to a manageable size.

Semi-structured interviews, on the other hand, are generally of shorter length and the sample includes a larger number of participants. Interview guides have a set structure which is generally pilot tested and revised before actual data collection begins and each question will assiduously be asked of each participant. Although the interviewer will be encouraged to follow up intriguing responses, standardized follow-up probes are likely to be included in the interview guide to further encourage comparability of responses.

The purposes to which such interviews can be put are different from more expansive qualitative research. The responses to questions will be comparatively shorter, the language to analyze less rich, and the interview situation more artificial. Samples are thus usually constructed to assure capturing the contrasts which have been determined ahead of time to be of interest in the study. For example, in work Wylie Burke and I did on women's understandings of breast cancer, risk and genetic susceptibility testing, we were particularly interested in the possible differences one might find depending both on women's race/ethnic background and on their personal and familial history of breast cancer. We therefore purposively selected participants to fill a predetermined sampling grid with three race/ethnic groups and four levels of breast cancer history and risk, with 20 women in each category for a planned sample of 240 women.

Once one has reached this distance from ethnography, it is fair to ask, what makes these interview studies qualify as *qualitative* research? One obvious answer is that participants answer questions in their own words and at whatever length they wish. But perhaps a better answer is that qualitative research is as much defined by the way data are analyzed as by the way they are collected.

METHODS OF DATA ANALYSIS

Most qualitative researchers who do interview studies use data analysis methods that are beholden to some extent to the "grounded theory" methods of Glaser and Strauss. When these sociologists wrote *The Discovery of Grounded Theory* in 1967⁸, they were concerned that sociology was suffering from a kind of methodologic *rigor mortis* that focused on verification of existing theory rather than figuring out if the theory itself had much conceptual merit. Glaser and Strauss wanted to reverse this trend by developing a systematic approach to *inductive* methods which could unearth what was in the data themselves. The method they proposed begins by merging the timelines of

data collection and data analysis, beginning analysis concomitant with even the earliest stages of data collection. The intent is to locate themes of interest directly from the data and to be able to assess if the interviews already completed have paid sufficient attention to those themes. If not, further interviews are conducted. This is called "theoretical sampling" and means that the size of the sample is not set ahead of time, a fairly controversial aspect of the method.

However, it is in its meticulous attention to data coding that grounded theory has had its deepest influence on qualitative research. The method provides a step-by-step approach to taking a mass of words (and/or observations) and successively reducing them to manageable units of analysis. Coding begins with an initial set of codes which name the overt content of a sentence (e.g. "breast cancer" or "reaction of husband to illness") and then moves on to more detailed coding, often using the participant's own language. Thus, a participant in talking about changes in her life following the onset of a disability might say, "*there's never enough time to get everything done,*" which could be captured in a short-hand code as "time pressures of disability." This might be an area about which the researchers did not specifically ask. However, by the creation of this code, they could now return to the data from other interviews to look to see if this was, in fact, a common experience – or not. And, if so, with what other characteristics of participant, of situation, of disability might it co-occur and how might it affect other parts of the participant's experience of illness. Returning to already read interviews to, as it were, interrogate the data on the basis of new codes, is referred to as the "constant comparative method." As coding progresses, "theoretical memos" are written, moving away from coding and toward finished analysis. These memos are also meant to leave a "trail" showing how the analysis was built up. All this ends when the researcher believes s/he has reached a point of "saturation" in understanding and a literal picture has emerged of the entire domain under study and how parts of it operate, affect each other, and are connected.

Rigorous methods of coding, however, do not completely solve the issue of epistemology which has never ceased to haunt qualitative research. How, after all, does one know that one's view of the data is correct, especially if one's data consists of only one interview with each participant and no other methods of triangulation? One commonly accepted method is to make data analysis a team activity. A good description of this approach appears in the article by Geller et al.: "*Following review of a subset of the initial interview transcripts by the investigators to identify themes, a codebook was developed through an iterative process of transcript review and refinement of subcodes and definitions. Two coders. . . double-coded sets of transcripts and refined code definitions until intercoder reliability of over 80% was achieved, the remaining interviews were [then] coded by the research assistant, then reviewed by the project coordinator for completeness and accuracy of coding.*"(199)³

Another approach to corroboration of one's analysis is called "respondent validation" in which the researcher presents his/her initial analysis to research participants. The pur-

pose is not specifically to have the participants agree that this is the account they would have given, but rather to recognize themselves, their words and ideas, and to understand and accept the researchers' analysis. This has been found by some researchers to be both a respectful and a valuable approach, leading to improvements in the ultimate data analysis. Nevertheless, limitations of the approach have been noted as well⁹, including participant lack of interest in such exercises, lack of agreement *among* participants about the validity or recognizability of the analysis, temptation of the researcher to alter his/her analysis based on negative, emotional reactions of the participants to the researcher's view, and, perhaps most important, the implication that the participants' analysis is more valid than the researcher's.

The final arbiter of the validity of the research, however, is ultimately the reader. The question thus reappears: Without access to known statistical methods, how can a reader judge the quality of the analysis presented? There are, in fact, an increasing number of ideas being promulgated about how the validity and reliability of qualitative research can be judged. One of the most interesting appeared in a recent article by Walter et al.¹⁰ in the form of a table which pulls together in considerable detail, all the questions one might, as a reader, want to "ask" of any qualitative study. This includes – as covered in this editorial – how the sample was selected, data collected, and analysis performed. There is one crucial element in this table that is worth particular highlighting. Called "justification of data interpretation," it concerns whether sufficient data have been presented to support the descriptive findings. And by "data," what is meant is the words of participants themselves. This point cannot be stressed enough. It is the words of the participants that constitute the data from qualitative interviews. Thus, the researcher must convince the reader of two things. The first is that the data presented in support of the analysis are truly representative. This is a fairly technical matter, although one often addressed inadequately. Thus, quotes should be numbered and identified, and the reader should be told how and why these were selected from the original sample. In support of the representativeness of the quotes presented, negative, unusual or contradictory cases may also be discussed. The second issue is less technical but of equal importance. Does the researcher provide a full discussion of how s/he discovered and built the analysis? This is an absolutely key part of convincing the reader of the validity of one's qualitative research and yet, paradoxically, the conventions of quantitative research reporting – which separate Results sections from Discussion sections – actually make it more difficult to create the trail of analysis and fully expound upon it.

Qualitative research uses skills and methods which are unfamiliar to many science readers. Therefore, checklists of the sort mentioned above have value in helping to judge whether the researcher has done a solid and creditable job. However,

none of them can tell the reader if the analysis is useful, insightful, or even inspired; to know that, one must read the work itself. Qualitative methods have the power to investigate questions that quantitative methods can neither pose nor illuminate in the same way. The work of Geller et al. in this issue is an example of this. Other work which I would strongly recommend in order to comprehend the value of qualitative research includes Mathews, et al.'s¹¹ study of rural, African-American breast cancer patients and how they come to terms with the diagnosis of breast cancer. Through a detailed analysis of these women's narratives, this work goes further toward explaining the phenomenon of late diagnosis of breast cancer among this population than multiple studies with impressive p values. Similarly, Hunt et al.¹² provide a window into the issue of patient 'noncompliance' by creating an analytical framework for contrasting patient and provider goals, strategies, and evaluation criteria in chronic illness management, using examples from research on type 2 diabetes care. Work of this caliber does not merely provide hypotheses for later study. Rather, at its very best, qualitative research can do something that, for me, quantitative studies almost never do, and that is yield an *aha!* moment that lets you know you have just learned something truly new.

Nancy Press, PhD

Schools of Nursing and Medicine
Oregon Health & Science University
Portland, Oregon

References

1. Barbour RS. The newfound credibility of qualitative research: tales of technical essentialism and co-option. *Qualitative Health Research* 2003; 13(9):1019-1027.
2. Higginbottom GMA. Sampling issues in qualitative research. *Nurse Researcher* 2004; 12(1):7-19
3. Geller G, Bernhardt BA, Gardner M et al. Scientists' and science writers' experiences reporting genetic discoveries: Toward an ethic of trust in science journalism. *Genet Med* 2005; 3:198-205
4. Charmaz K. Premises, principles, and practices in qualitative research: revisiting the foundations. *Qualitative Health Research* 2004; 14:976-993.
5. Anspach, RR. *Deciding Who Lives: Fateful Choices in the Intensive-Care Nursery*. Berkeley: University of California Press, 1993.
6. Press N, Browner CH. Why women say yes to prenatal diagnosis. *Social Science and Medicine* 1997; 45:979-989.
7. Dobscha SK, Heintz RT, Press N, Ganzini L. Oregon physicians' responses to requests for assisted suicide: a qualitative study. *Journal of Palliative Medicine* 2004; 7:451-461.
8. Glaser, BG and AL Strauss. *The Discovery of Grounded theory: The Strategies for Qualitative Research*. New York: Aldine de Gruyter, 1967.
9. Bloor MJ. Notes on Member Validation. In: Emerson, RM, editor. *Contemporary Field Research: A Collection of Readings*. Prospect Heights: Waveland Press, 1983: 156-172.
10. Walter FM, Emery J, Braithwaite D, Marteau TM. Lay Understanding of Familial Risk of Common Chronic Diseases: A Systematic Review and Synthesis of Qualitative Research. *Annals of Family Medicine* 2004; 2:583-594.
11. Mathews HF, Lannin DR, Mitchell JP. Coming to terms with advanced breast cancer: black women's narratives from eastern North Carolina. *Social Science and Medicine*. 1994;38:789-800.
12. Hunt LM, Arar NH. An analytical framework for contrasting patient and provider views of the process of chronic disease management. *Medical Anthropology Quarterly*. 2001;15:347-67.