

# Can Qualitative Research Produce Reliable Quantitative Findings?

EDWARD C. GREEN  
Washington, D.C.

*This article asks whether qualitative anthropological research using rapid methods can produce results as reliable and valid as those derived from expensive sample surveys, at least under certain circumstances. The discussion focuses on nonformal sampling or informant selection techniques on which anthropologists rely. To help answer the question, the author describes two relatively low-cost, rapid, qualitative studies in applied medical anthropology in Swaziland and in the Palestinian Territories. Although random sample surveys have become the gold standard for most policy makers dealing with KAP (knowledge, attitudes, and practices) data, qualitative research can form the basis of policy or program decisions under certain conditions.*

Can qualitative, anthropological research using rapid methods produce results as reliable and valid as those derived from expensive sample surveys, at least under certain circumstances? Can interviewees selected nonrandomly for qualitative research be considered—or be proven to be—nationally representative to an acceptable degree? Can quantitative estimates based on nonrandomly derived, purposive samples yield valid quantitative data? I provide some answers to these questions by reporting experiences from two relatively low-cost, rapid, qualitative studies in applied medical anthropology, one conducted some years ago in Swaziland (Green 1982, 1985, 1987a, 1987b) and the other conducted recently in the Palestinian Territories (Green, Abu-Khalid, and Omari 2000). Both focused on health-related beliefs and behavior. Both were national in scope and were funded by the U.S. Agency for International Development (USAID). Experience in both countries suggests affirmative answers to the questions posed, and preliminary evidence supporting these conclusions is provided from random sample surveys conducted at roughly the same times.

---

*I thank the U.S. Agency for International Development (USAID)/West Bank and Gaza Strip, CARE International, and Macro International for financial and other support during the Palestine research. Most of the Palestinian research was funded under USAID subcontract no. 31306-0S-271, and all research was conducted by the Palestinian Central Bureau of Statistics. The Swaziland research was funded under USAID contract no. 465-0087.*



*Field Methods*, Vol. 13, No. 1, February 2001 3–19  
© 2001 Sage Publications, Inc.

The discussion focuses on informal sampling or informant selection on which anthropologists typically rely. Arcury and Quandt (1999:128–29) observed that anthropologists rarely specify how interviewees or informants are selected in qualitative research. They argued that typical selection techniques such as “convenience” and “snowball” samples raise serious questions about representativeness. In fact, most anthropologists do not seem to concern themselves greatly with formal sampling techniques for a variety of reasons, including the following:

- They typically conduct qualitative research rather than sample surveys.
- They often conduct exploratory research, seeking to identify patterns rather than measure patterns that are already quite well established.
- They are usually interested in collective aspects of culture rather than intercultural variation, which is the property of individuals.
- They focus more on meaning, rules, and emics rather than on observable behavior and etics. As Handwerker and Wozniac (1997) noted, “The socially constructed nature of cultural phenomena makes the classical sampling criterion of independent case selection not only impossible to attain but also undesirable” (p. 874).
- They often conduct research with populations for whom sampling frames are lacking.
- They typically review all relevant ethnographic literature on the group to be studied, as well as neighboring groups. This serves as a reality check and helps overcome some limitations of informal sampling (Bernard 1994:361; Bauer and Wright 1996:184). If preliminary findings conflict with prior ethnography, the anthropologist can always modify or completely rethink and revise instruments, informant selection, and overall methods, an option not feasible in survey research.<sup>1</sup>

The iterative, explorative method characteristic of qualitative research is highly productive in ways that sample surveys cannot be, and this method is not based on formal sampling. Yet, sampling considerations have a place in qualitative research because questions about the representativeness of qualitative findings inevitably arise (Denzin and Lincoln 1994).

## BACKGROUND TO THE STUDY IN SWAZILAND

In March 1981, I was hired by the Academy for Educational Development to assist in the implementation of the USAID-funded Rural Water-Borne Disease Control Project in Swaziland.<sup>2</sup> The project called for a KAP (knowl-

edge, attitudes, and practices) study relating to water and sanitation in Swaziland. The primary purpose of the study was to provide baseline data for the design of a national health education strategy aimed at reducing the incidence of water-borne diseases. The research findings were to guide not only the main project component, health education, but two other components as well, namely, public health engineering and environmental sanitation.

The project budgeted for a sample survey design and methodology for the KAP study. Yet, I had doubts about the appropriateness of such an approach. In addition to the general problem of obtaining valid data through survey methods in rural Africa (see Cohen 1973), the information I was after related to sensitive areas such as personal hygiene, excretory behavior, and ethnomedical beliefs. The impersonal, precoded questionnaire typical of survey research is notoriously deficient in eliciting information of this kind, even if it has value in measuring patterns that are reasonably well established.

After considerable discussion, review of previous surveys in Swaziland, and preliminary ethnographic fieldwork, I presented a modified research plan to the USAID mission in Swaziland. I proposed that an informal study of health beliefs and behavior, relying on local-level health workers, traditional healers, and their patients as key informants, be carried out before attempting the sample survey. I hoped that the study would provide a fund of qualitative information that would be valuable in the design and interpretation of surveys and might well provide information that surveys could never discover. My plan was to rely on traditional anthropological methods of key informant interviewing and participant observation, at least in the early stages of research. It was during the study relying on local-level health workers that I first attempted to come up with a quantitative national estimate from qualitative research. The health workers were called rural health motivators (RHMs).

RHMs are women who are chosen by their communities to receive about eight weeks of training in preventive health care at a regional clinic. After training, RHMs work among their neighbors, promoting homestead sanitation, the purification of drinking water, proper infant nutrition, and other practices related to disease prevention. I felt that RHMs would make good key informants—even culture brokers—since they were insiders in their communities who understood and promoted public health goals. I expected that they would be likelier than their neighbors to give candid and truthful replies to sensitive and even embarrassing questions about what their neighbors thought and did. Furthermore, I could use a flexible, open-ended interview schedule, which I felt was the only appropriate instrument.

### Value of Community-Level Health Workers as Key Informants

Why the RHM study? I have discovered from research in many parts of the world that there is very little, if anything, that village health workers (as RHMs are more generally known) and other community-level health workers do not know about the health-related beliefs, attitudes, and behavior of the people they work with, especially if they come from the same community. This is true in any society. Therefore, an effective, quick, inexpensive and easy way to discover health-related beliefs and behavior is to systematically interview local-level health workers as key informants. There is even the advantage that health workers are more likely than the general public to speak frankly about beliefs and behaviors that are not approved or are in conflict with formal health education, or are considered superstitious or backward. Thus, interviews with health workers can serve as validity checks on what mothers and others in the general public say about their own behavior—a benefit not available in sample survey research. Moreover, it is often useful to know what health promoters promote and which types of health services are provided, and by whom, when interpreting answers about health beliefs and practices.

Finally, health workers have at least some idea about what is important epidemiologically and etiologically (by using such terms, I am referring, of course, to Western biomedical ideas), so they can help interpret meaning and significance for the researcher. They might even volunteer important health information not specifically asked for by interviewers.

### Rapid RHM Study

Returning to research in Swaziland, I submitted a revised research proposal to USAID and the Ministry of Health. This was accepted with minimal delay. Over the next eight weeks or so, and using an interpreter, I began my qualitative research by interviewing a convenience sample of forty-two RHMs in eight regional clinics. The sample was nonrandom, yet I attempted to interview RHMs in roughly equal proportions from each of the four major topographic zones in an attempt to achieve rough geographic balance. I located RHMs at regional clinics where they came periodically to collect their pay. The nurses holding the paychecks asked the RHMs to cooperate with me, which helped. The whole process was quick and easy. No RHM declined to be interviewed. If I hoped to interview the RHM from a particular community, but she failed to appear for her paycheck on payday, I usually did not return to a particular regional clinic to locate that RHM again. No doubt it would have added rigor to have randomly selected RHMs for interviewing and to have decided on a minimal number of RHMs in the sample. I did not

know how many RHMs I would interview when the process began. Time was not so much a factor here as my lack of faith in my informal methods. They seemed improvised and far less scientific than the planned sample survey.

My interpreter helped me develop effective probing techniques that were used when RHMs gave stereotyped and self-serving answers such as “The people around here boil their drinking water because we have been teaching them how important this is.” I knew I was on the right track when I received comments such as “I have tried and tried, but many people around here just refuse to boil their drinking water!”

Since each RHM visited approximately 40 homesteads, the interviews provided information based on nearly 1,680 homesteads, representing some 3% of the estimated 50,000 homesteads in Swaziland. It seemed possible to regard our RHM-visited homesteads as reasonably representative of Swaziland as a whole, since they were situated throughout the country’s four major topographic zones. I attempted to obtain some quantitative data on the number of pit latrines built or under construction in homesteads, since we needed baseline data on latrine and construction, and I would be asking about it in the sample survey. RHMs had no trouble providing the information, since they were required to report monthly on latrine construction progress.

I estimated the percentage of Swaziland homesteads that had a pit latrine in 1981 in a simple manner. I divided the total number of pit latrines reported to me by all RHMs by the total number of homesteads covered by all forty-two RHMs interviewed. RHMs reported 412 latrines. Dividing that by the total number of homesteads (1,680) suggested that some 24.7% of homesteads in the areas covered by RHMs had a latrine. I saw no reason why this should not be a good approximate *national* figure for the proportion of homesteads with a latrine.

When we conducted the relatively expensive and time-consuming national random sample survey some months later (Green 1982), we found that 21.9% of homesteads sampled had a pit latrine. But another 3.9% had one under construction, and 0.2% had two or more latrines (distinctions not made in the RHM interviews; those under construction were counted), bringing the total proportion of homesteads in the sample with some kind of latrine to 26% (Green 1982:23). In fact, 24.7% was within the margin of statistical error of our sample survey, meaning the two figures were essentially the same. My quick study using RHM key informants cost about \$300, while the later sample survey cost about \$30,000 (comparing direct costs only).

I attempted to quantify only the latrine estimate in 1981–82, but I always wondered if I had stumbled on something important concerning an unrealized (or undiscussed?) potential of qualitative research—quite apart from the widely acknowledged value of qualitative findings for many other reasons

and purposes. A few years after the Swaziland research, I was asked to conduct a “national qualitative study” of diarrheal disease KAP in Bangladesh. Using a purposive sample of 240 Bangladeshis (I am almost embarrassed to report the national population at the time—about 100 million), our quantifiable findings proved to be within 4%–16% of key findings from the National Oral Rehydration Program random sample survey (Green 1986:358). Because these parallel findings are already published, I will do no more than mention them here.

### BACKGROUND TO THE STUDY IN THE PALESTINIAN TERRITORIES

In 1999, USAID (again) requested a national, in-depth, qualitative study, this time of the health of women and infants in the Palestinian Territories, as part of its Maternal and Child Health Pilot Initiative. The purpose of the study was to provide preliminary information about knowledge, attitudes, and practices related to perinatal health (e.g., pregnancy, breastfeeding, and infant nutrition, fertility decisions, child spacing and contraception, and use or nonuse of perinatal health services) as part of a USAID-funded project aimed at improving perinatal health.

In response to the USAID request, the Palestinian Central Bureau of Statistics and I developed a study combining several qualitative research methods—namely, in-depth and key informant interviewing, focus group discussion, and limited rapid ethnographic methods—between February and December 1999. We interviewed a total of 328 Palestinians, and an additional 30 participated in informal focus group discussions. We conducted interviews in all significant geographic areas of the West Bank and the Gaza Strip (with a smaller number in the East Jerusalem area).

We used a flexible, open-ended, semistructured interview schedule (or topic guide in focus group discussions) for in-depth interviewing. The instrument was modifiable, which is to say we were able to add questions—indeed whole lines of inquiry—as we accumulated information and patterns began to emerge. We were also able to drop questions after we felt that a pattern had become sufficiently well established. The qualitative research process is inductive and creative in that it tests, discards, and confirms hypotheses, in a very preliminary way, as the interviewing proceeds. Survey research, by contrast, tends to confirm and measure the details of what may already be fairly well known, allowing for anticipation and categorization of responses.

The interview schedule for the Palestine study consisted of some fifty questions, many of which contained more specific subquestions or topics to

explore. It was developed in accordance with a scope of work prepared by USAID/WBG (West Bank and Gaza Strip), which contained many (in fact, too many) maternal and child health topics of interest to USAID's Maternal and Child Health Pilot Initiative. During the first research phase, we modified the interview schedule as patterns began to be established. It was not modified during phase 2 due to the exigencies of fieldwork.

### Fieldwork

The first phase was conducted under CARE International. During phase 1, which took place primarily during July and August 1999, 31 Palestinian key informants (mostly community-level health workers) were interviewed. About the same number participated in informal focus group discussions. Phase 2 fieldwork was conducted during December 1999 under the sponsorship of Macro International. In this phase, 297 Palestinians were interviewed, bringing the total number of Palestinians either interviewed or in focus groups to 358. Prior to phase 1, I trained two Palestinian researchers (already experienced in survey research) in qualitative and anthropological research methods. They, in turn, trained eleven interviewers to conduct phase 2 interviews.

### Selection of Those Interviewed

Interviews were conducted with a range of informants and in a variety of settings in order to capture variation that might be expected to occur along lines of geography, gender, socioeconomic status (SES), age, West Bank or Gaza residence, refugee camp or non-refugee camp residence, degree of religious conservatism, and rural/urban residence. While we tried to interview Palestinians who represent the variety of opinion and behavior found nationally, we did not have precise percentages or numbers of interviewees within the categories just described.

Most (about 85%) interviewees were chosen because they represent the general public. The term here refers to those who represented the target audience in the anticipated USAID-supported maternal and child health project, namely, women of reproductive age, men (husbands), and older women (such as mothers-in-law, known to influence reproductive and contraceptive decisions). Other Palestinians representing relevant or divergent viewpoints were also interviewed. Such key informants were regarded as special target audiences for the future project; they included religious and political leaders, traditional birth attendants (*dayas*), traditional healers, and health workers of all types (but particularly midwives, nurse midwives, and village health workers).

TABLE 1  
Number and Location of Phase 2 Interviews

<i>Locality Type</i>	<i>West Bank</i>	<i>Gaza</i>	<i>Total</i>
City	62	63	125
Village	59	0	59
Camp	55	58	113
Total	176	121	297

Geographically, interviews were spread between the West Bank and the Gaza Strip in proportions that reflected relative population size. Jerusalem is part of the middle West Bank in Table 1, which shows the number and location of phase 2 interviews. It should be noted that in phase 2 fieldwork, during which most interviews were conducted, there was a degree of random sampling. A sampling expert from the Palestinian Central Bureau of Statistics (PCBS) chose localities for interviews on the basis of random sampling of census tract enumeration areas.

Once a locality was selected, interviewers used purposive selection to ensure that at least some informants of special interest (e.g., traditional birth attendants, traditional healers, local health workers) were among those included. We did not use strict quota sampling to ensure certain numbers of informants were included; rather, we used what has been called judgment sampling (Bernard 1994:95), simply ensuring that *some* informants of a certain type were included. We also took care to include what Babbie (1983:252) and others call deviant cases: informants likely to represent views most likely to deviate from the norm. For example, we included a ninety-year-old *imam* and an eighty-year-old traditional birth attendant, anticipating that they would hold archconservative views. We felt these views would provide a useful counterbalance to the views of younger, more educated, and urban Palestinians.

It should be stressed that we placed more emphasis on capturing the *range* of viewpoints than on achieving a sample projectable to the Palestinian Territories as a whole. USAID planned a major Demographic and Health Survey (DHS) soon after our qualitative study—indeed, our study was partly to inform the questions, wording, and interpretation of the DHS. We left sampling and measurement to an anticipated DHS.<sup>3</sup>

Phase 1 interviews with thirty-one key informants are not included in the table because haphazard sampling was used during the exploratory first stage; we modified the interview schedule as we proceeded, and no interviews were conducted in the southern West Bank.

### Quantifying Health Patterns Derived from Qualitative Research

The objective of the Palestinian qualitative study was not to quantify patterns, since this is the purpose and strength of sample surveys and a DHS was planned. Yet some questions, or lines of questioning, began logically with a query about numbers of things. For example, we would first ask about the age of female marriage in the local area before going on to related questions, such as whether this age has fallen or risen in recent years. Thus, we ended up with quantifiable answers for a few of our fifty-plus questions.

I noticed while reviewing answers that the trends in numbers in our study seemed very similar to what was known from recent sample surveys in the Palestinian Territories. I decided to test the degree of quantitative concordance between our study and the 1996 PCBS national health survey (and for one question, a 1999 study done in the Gaza Strip) by comparing findings on five questions that provided quantifiable answers: age of marriage, number of desired children, percentage of women who attend prenatal care, percentage of women who attend postnatal care, and preference or behavior with regard to cousin marriage. The areas of comparison were limited by available data from recent sample surveys in the region, as well as by the number of answers in our study that produced quantifiable answers.

We made the comparisons by first calculating the mean averages of the answers given in the qualitative study. Note that we asked questions about general patterns rather than what a particular respondent said about his or her behavior. That is, we asked a question such as “At what age do Palestinian women usually get married in this community?” rather than “At what age were *you* (first) married?” This difference alone ought to produce differences in quantifiable responses. We then compared the means from our five questions with data on similar questions from sample surveys conducted by the PCBS (1997). In the case of the question on cousin marriage, the comparison was with a sample survey conducted by a nongovernmental organization called the Women’s Affairs Center (WAC) (1999).

As can be seen, findings between the qualitative study and the two sample surveys are highly compatible, and any differences are expected and explainable (see Table 2).

For age of marriage and number of children desired, qualitative findings are virtually identical with, and within the statistical margin of error of, the PCBS random sample survey. Furthermore, our figure for cousin marriage preference in Gaza is fully compatible with the WAC study in Gaza, which provides figures for self-reported *behavior* rather than *preference*. The 43% in the latter study was derived from examining court records showing marriage between cousins; the 57% was derived from self-reported behavior.

TABLE 2  
Comparison of Findings between the Qualitative Study and Sample Surveys

<i>Findings Compared</i>	<i>Qualitative Study (1999)</i> (N = 297)	<i>Palestinian Central Bureau of Statistics (PCBS)</i>
		<i>Health Survey (1996)</i> (N = 3,349)
Age of marriage	17.1	18
Number of children desired	5.08	5.5
Prenatal care attendance	98%	92.9%
Postnatal care attendance	58.5%	20%
Prefer "cousin marriage" <sup>a</sup>	48% <sup>b</sup>	43%–57% <sup>c</sup>

a. Gaza only.

b. Refers to self-reported preference, not actual behavior.

c. Refers to marriage records and to self-reported behavior. Data are from the Women's Affairs Center (1999), not the PCBS.

Our figure is only slightly below the mean between the two WAC study figures and, therefore, is fully compatible with the findings from the WAC survey. If anything, our figure slightly lower than the mean between the two WAC figures is to be expected, since there has been recent health education directed against cousin marriage, and our study was conducted somewhat later than the WAC study.

In two other comparison areas, there are some differences, yet they are in fact differences that might be expected due to the impact of recent health education and improved health services. First, reported prenatal care levels are higher by five percentage points in our study. One year after the PCBS health survey, Hundt et al. (1997) found that 95% of randomly sampled women in Gaza reported attending prenatal care. Thus, the trend seems to have been toward higher levels of prenatal care due to health education focused on increasing levels of prenatal care. Our study results suggest that the trend continues and might have reached 98% by 1999, a modest increase during the four-year period 1996–99.

Postnatal care attendance appears significantly higher after three years, rising from 20% to 58.5%, if our sample can be considered representative. In the past three years, there has been emphasis on persuading women to attend postnatal care, given the disappointing findings in the 1996 PCBS health survey that only 20% of women used this service. Moreover, in our study, we counted *home visits* by village health workers who provided postnatal care, while the PCBS did not, and that alone could account for most of the difference between the two findings. I would not be surprised if the next national

Palestinian health survey finds a postnatal care attendance level close to 60% (depending, of course, on how the question is asked and whether home visits are counted).

## DISCUSSION

How do we explain the parallels between qualitative and quantitative research in Swaziland and Palestine? It may be noted that both populations are relatively homogeneous ethnolinguistically, that both are relatively concentrated geographically (of course we are excluding expatriate Swazis and Palestinians here), and that research focused on health-related KAP in both studies (although the second study included questions about marriage- and family-related attitudes and behavior). The Swaziland study might be considered a special case in which a convenience sample of key informants provided numerical information about something they knew well (the number of latrines in the areas they visited) and about which they had no reason to mislead, except that they might have exaggerated the numbers to appear to have been doing their jobs well.<sup>4</sup>

In any case, some countable things are intrinsically easier to estimate with accuracy than others; using village health worker counts of all latrines in their areas is not the same as estimating population frequencies of normative events. Let us therefore consider the Palestinian case in more detail, since sampling and representativeness issues are more pertinent here, and we have more quantitative findings to compare.

Considering that our qualitative study in the Palestinian Territories relied on a purposive, convenience sample (yet covering all major geographic areas) of interviewees rather than a random sample, and considering the difference in the size of the two samples (nearly 300 compared with 3,349), it is of considerable methodological interest that the quantitative answers are so similar between the qualitative study and recent surveys. To begin, our sample size was decidedly less than what a random sample should have been for a national Palestinian population of three million. Bernard (1994:77–79, citing Krejcie and Morgan 1970) suggested a sample size of 384 for a population of one million at the 5% confidence level for dichotomous variables. Yet, the similarity between our findings and those of sample surveys using much larger random samples suggests that our sample size was adequate.

Could the parallels between the studies amount to some sort of coincidence? Perhaps on one or two findings but probably not on four or five (difference found in postnatal care, the fourth question, is explainable and predictable). What, therefore, might account for the similar findings? Perhaps

our purposive sampling with quota features, while not as rigorous as the procedures recommended, was just good enough to approximate a national random sample and therefore account for the similarities in findings.

Or might there be explanations apart from sampling? Another possible factor is the domain investigated. Behavior and attitudes related to health, marriage, and family tend to be highly patterned, and the variations one finds (in ethnolinguistically homogeneous populations) tend to be on the basis of predictable factors, factors known from countless health and demographic surveys and quantitative studies (e.g., SES, level of education, age, gender, rural/urban residence). If these factors are taken into account in informant selection, strict random sampling may not be crucial in order to discover basic health-related KAP patterns, or those related to marriage and child preferences, and even to quantify them.<sup>5</sup> I suspect that this may work better when dealing with health, marriage, and/or procreation than in some other domains, those in which elements of the system investigated may be more heterogeneous and variable.

Handwerker and Wozniak (1997) showed that random sampling and sample size become less important when researching cultural data, as distinct from what they called life experience data. The latter are elicited by questions such as “How old are you?” or “What is your race?” Answers from any one person have no necessary relationship with the answer of any other person. Cultural data are elicited by questions such as “What distinguishes family members from other people?” Answers to this type of question depend on social interactions and cultural constructions (Handwerker and Wozniak 1997:869–70). Handwerker and Wozniak showed that small, convenience samples of informants can produce consistent (reliable and valid) results pertaining to cultural data if informants are knowledgeable about a cultural domain and there is a high level of agreement among them. The need to expand sample size depends, in this case, on informants’ “experiential variability” within a particular domain and on levels of agreement and identified cultural boundaries (Handwerker and Wozniak 1997:875).

This distinction might help explain parallels in Palestinian findings between answers to our question about *general patterns* (“At what age do Palestinian women usually get married in this community?”) and answers to the more rigorous *respondent-centered* question in the PCBS survey (“At what age were *you* (first) married?”). Although the second question sounds like a “life experience” question, perhaps it is really a cultural data question that therefore requires less rigorous sampling in order to capture variation. After all, the answer depends on social interaction and cultural construction (marital patterns including values and norms). Still, I admit to being sur-

prised that these divergent interviewing approaches did not make more difference in quantifiable response patterns.

In fact, qualitative and quantitative methods are not so categorically different from one another as generally assumed. Bernard (1996) noted, "When it comes right down to it, qualitative data (text) and quantitative data (numbers) can be analyzed by quantitative and qualitative methods" (p. 10). He illustrated this with the following simple table:

<i>Analysis</i>	<i>Data</i>	
	<i>Qualitative</i>	<i>Quantitative</i>
Qualitative	a	b
Quantitative	c	d

Cell a is the qualitative analysis of qualitative data, and cell d is the quantitative analysis of quantitative data. This, in the popular imagination, is how the different types of data are supposed to be analyzed. Less well understood and acknowledged is the qualitative analysis of quantitative data (cell b) and the quantitative analysis of qualitative data (cell c). Cell b involves "the search for, and the presentation of, meaning in the results of quantitative data processing" (Bernard 1996:10). It is not uncommon for sample surveys to minimize or overlook this type of analysis, perhaps in the belief that it is "speculative" and no longer grounded in data. For example, the 1996 PCBS health survey, with which I compare our Palestinian findings, found that in three key health indicators, health status in rural villages was superior to health status in cities. But this went unmentioned. One has to reach this conclusion by noticing these findings reported on different pages and observing that they contradict earlier findings that showed that health status was consistently inferior in rural areas. Clearly, much to do with meaning and significance can be lost by not attempting qualitative analysis of quantitative data.

And finally we come to cell c, quantitative analysis of qualitative data, which "involves turning the data from words or images into numbers" (Bernard 1996:10). This can be done by coding ethnographic or other data and looking for emerging patterns. I suppose it might also characterize making numeric estimates based on information supplied by knowledgeable informants, or on findings from convenience or other nonrandom samples. The point is that contrary to the dichotomous model popularly held in at least international applied research circles, there is no clear-cut distinction between quantitative analysis and qualitative analysis when it comes to data and analysis.

Another question arises: For practical purposes of policy making or program design or evaluation, how much variation can there be in quantifiable findings, whether from sample surveys or estimates from essentially qualitative research? As health planners and program directors know, it does not matter whether 45% or 50% prefer or practice cousin marriage, whether 93% or 98% attend prenatal care, or whether 22% or 24% of homesteads have a pit latrine. The first becomes “about half,” the second becomes “almost all,” and the third becomes “not nearly enough” for policy or program purposes.<sup>6</sup> Decision makers are usually familiar enough with survey research to know that differences of 2%–8% or slightly more arise in all sampling endeavors and may reflect differences in sample size, sampling error, and population variance. Still, small differences that may suggest *trends* (e.g., prenatal care rising from 92% to 98% levels) can be important to policy makers even if differences are due largely to the vagaries of sampling.

Until we know more about how and why and to what degree and under what circumstances certain types of qualitative research (defined broadly as nonrandom sample survey KAP-type research) can usefully or reliably be quantified, it is unlikely that program planners or policy makers will base decisions on studies generally regarded as “qualitative.” Random sampling has become the gold standard for most policy makers dealing with KAP data, which means it is the standard for applied researchers.<sup>7</sup> Standard qualitative findings of anthropological research in the form of descriptive behavioral patterns are unlikely to form the basis of policy or program decisions unless researchers can provide some sort of measure showing how widespread or common a pattern appears to be. Another option might be to make the distinction between life experience and cultural data better understood and demonstrate—as Handwerker and Wozniak (1997) did—that small, convenience samples can be adequate for cultural data.

In any case, it becomes important for those of us using essentially qualitative methods to (1) make the distinction between life experience and cultural data; (2) make informant selection in qualitative research more explicit (Why was a particular informant chosen? How is she or he knowledgeable about a particular domain? What perspectives and biases might be expected from such an informant?); (3) attempt some quantitative analysis of qualitative data (e.g., quantification of data from smaller and nonrandomly derived samples when the samples seem, and can be justified as, reasonably representative); (4) attempt to verify numerical estimates derived by comparing these estimates with findings from sample surveys, if available (providing comparison with “gold standard” research); and (5) characterize the domain of investigation (e.g., health, agriculture) along with the probable variables influencing knowledge, attitudes, and practices (e.g., SES, age, level of edu-

cation) and how these variables have been taken into account in sampling or informant selection.

What the data presented in this article show is that, sometimes, a nonrandomly selected sample can avoid systematic biases and some degree of nonrandom error to produce results that will match general population trends. But at which times? From the examples described here, favorable conditions may include ethnolinguistic homogeneity and perhaps geographic concentration of the population studied, investigation of "cultural data," and investigation focused on domains in which elements of the system are relatively homogeneous and vary in association with predictable factors. When these conditions prevail, it may be possible to use small, convenience samples, purposive sampling with quota features, or a similar method that is far quicker and less expensive than a random sample survey to arrive at valid estimates of defined populations.

Why not just use strict random sampling and put to rest questions about representativeness? There are a number of reasons, including lack of an available sampling frame and lack of adequate time and funds. Strict random sampling usually raises costs of research exponentially and takes more time.<sup>8</sup> By strict, I mean something like a multistage, stratified cluster sample in which a preselected household or person, if not available at the time of first interviewer visit, is revisited at least two times. In practice, interviewers in sample surveys usually settle for the nearest coherent and sober neighbor, which results in sampling a little closer to the qualitative-quantitative (or purposive sample-random sample) middle ground just described in our Palestinian study. Recognizing this, it may be time to change prevailing thinking (among nonanthropologists) that quantitative methods are scientific and precise while qualitative methods are merely preliminary and suggestive, there being no middle ground.

These examples from my applied field studies are intended to raise questions about the relative usefulness of different types of research strategies in certain contexts, not to convince anyone that qualitative studies can reliably replace quantitative surveys. It should be obvious that both research strategies have their strengths and weaknesses, and that either or a combination of both might be better depending on circumstances. We still have more to learn about what those circumstances are, but they seem to arise when cultural data are sought in culturally homogeneous groups and when the investigation focuses on domains in which elements of the system are relatively homogeneous and vary in association with predictable factors. The problem of knowing when these conditions exist and nonrandom samples are appropriate remains. I hope other researchers will draw on their experience and confirm or refute this hypothesis. If confirmed, we ought to see more studies that

combine the best features of quantitative and qualitative research in ways that are increasingly time- and cost-efficient.

## NOTES

1. In more than twenty years of applied research in developing countries, I have rarely seen anyone but an anthropologist review existing ethnography prior to attempting research.

2. Background material to the Swaziland research is adapted from Green (1987a).

3. Yet, in the end, U.S. Agency for International Development decided not to conduct the Palestinian Demographic and Health Survey, at least as of this writing.

4. This raises the question: How does one know, in short-term fieldwork, that one is not being misled? The only insurance against this is to do what anthropologists usually do: Check all information with different informants (including those likely to have very different biases and perspectives), check findings with existing literature, and try to find an explanation for any discrepancies. Bernard (1994:361) emphasized the value of the “constant validity check.” Yet, such commonsense cross-checking is often considered “outside the scope” of sample surveys.

5. A recent health-related study of U.S. farmers used informal purposive sampling based on many of the same factors that we used (geography, gender, age) but also included farmers’ herd size, acreage farmed, and income—for the same reason of achieving representativeness of the total particular population under investigation (Perry and Bloom 1998:344).

6. Descriptive rather than more sophisticated inferential statistics seem to inform most policy and program users of survey research, at least in applied international health circles.

7. At least one major interdisciplinary journal, *Studies in Family Planning*, has adopted the policy of not publishing studies based on any sampling procedure other than random sampling (with the possible exception of focus group research). Ten years ago, it was possible to publish central location intercept studies based on exit interview strategies of finding and interviewing people at service delivery sites.

8. Bernard’s (1994) *Research Methods in Anthropology* discusses alternative, potentially less expensive strategies for sampling other than strict random sampling.

## REFERENCES

- Arcury, T. A., and S. A. Quandt. 1999. Participant recruitment for qualitative research: A site-based approach to community research in complex societies. *Human Organization* 58:128–41.
- Babbie, E. 1983. *The practice of social research*. 3rd ed. Belmont, CA: Wadsworth.
- Bauer, M. C., and A. L. Wright. 1996. Integrating qualitative and quantitative methods of model infant feeding behavior among Navajo mothers. *Human Organization* 55:183–205.
- Bernard, H. R. 1994. *Research methods in anthropology: Qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.
- . 1996. Qualitative data, quantitative analysis. *Cultural Anthropology Methods Journal* 8 (1): 9–11.

- Cohen, R. 1973. Warning epistemologies: Quality and quantity in African research. In *Survey research in Africa: Its implications and limitations*, edited by W. O'Barr et al., 36–47. Evanston, IL: Northwestern University Press.
- Denzin, N. K., and Y. S. Lincoln, eds. 1994. *Handbook of qualitative research*. Thousand Oaks, CA: Sage.
- Green, E. C. 1982. *A knowledge, attitudes, and practices survey of water and sanitation in Swaziland*. Mbabane: Swaziland Ministry of Health; Washington, DC: Academy for Educational Development.
- . 1985. Factors relating to the presence and use of sanitary facilities in rural Swaziland. *Tropical and Geographical Medicine* 37 (1): 81–85.
- . 1986. Diarrhea and the social marketing of oral rehydration salts in Bangladesh. *Social Science and Medicine* 21:357–66.
- . 1987a. Anthropology and the planning of health education strategies in Swaziland. In *Anthropological praxis: Translating knowledge into action*, edited by B. Wulff and S. Fiske, 15–25. Boulder, CO: Westview Press.
- . 1987b. Beliefs and behavior related to water in rural Swaziland. *International Journal of Water Resources Development* 2 (4): 29–42.
- Green, E. C., K. Abu-Khalid, and M. Omari. 2000. *Palestinian maternal and child health: A qualitative national study*. Ramallah: Palestinian Central Bureau of Statistics; Calverton, MD: Macro International. Available through the U.S. Agency for International Development/West Bank and Gaza Strip.
- Handwerker, W. P., and D. F. Wozniak. 1997. Sampling strategies for the collection of cultural data: An extension of Boas's answer to Galton's problem. *Current Anthropology* 38:869–75.
- Hundt, G., Y. Abed, A. El'Alem, and R. Shawaa. 1997. *Evaluation and improvement of maternal and child preventive health resources and services of Palestinians in the Gaza Strip—Summary of main findings*. Gaza: Avicenne Initiative.
- Palestinian Central Bureau of Statistics. 1997. *The health survey in the West Bank and Gaza Strip: Main findings*. Ramallah: Palestinian Central Bureau of Statistics.
- Perry, M. J., and F. R. Bloom. 1998. Perceptions of pesticide associated cancer risks among farmers: A qualitative assessment. *Human Organization* 58:342–49.
- Women's Affairs Center. 1999. *Choice and power: A discussion of early marriage in the Gaza Strip*. Gaza: Women's Affairs Center.

*EDWARD C. GREEN is a Washington, D.C.-based applied medical anthropologist who consults for a number of international development organizations. He has worked in Africa, Asia, the Middle East, Eastern Europe, Latin America, and the Caribbean. He has worked as an adviser to the ministries of health in Swaziland (1981–83) and Mozambique (1994–95). He is the author or editor of five books and numerous journal articles and book chapters. His most recent book is Indigenous Theories of Contagious Disease (AltaMira Press, 1999).*