Research Methods in Anthropology

THIRD EDITION

Qualitative and Quantitative Methods

H. Russell Bernard

A Division of Rowman and Littlefield Publishers, Inc.
Walnut Creek • Lanham • New York • Oxford

10

Structured Interviewing: Questionnaires

This is the fist of two chapters about **structured interviews**. In a structured interview, each informant or respondent is exposed to the same stimuli. The stimuli are often questions, but they may also be carefully constructed vignettes, lists of words or photos, clips of music or video, a table full of physical artifacts, or a garden full of plants. The idea in structured interviewing is always the same: to control the input that triggers people's responses so that their output can be reliably compared.

I'll cover two broad categories of methods for structured interviewing: **questionnaires** and a range of methods used in **cultural domain analysis**. We begin in this chapter with questionnaires and survey research. I review some of the important lessons concerning the wording of questions, the format

of questionnaires, the management of survey projects, and the maximizing of response rates. (Refer to Chapter 9 again for more discussion of response effects.)

At the end of this chapter, I'll introduce you to some of the interesting and unusual methods that people are using to get at complex and/or touchy topics. In Chapter 11, I'll introduce you to the structured data-gathering methods for cultural domain analysis.

Questionnaires and Survey Research

Survey research goes back over 200 years (take a look at John Howard's monumental 1792 [1973] survey of British prisons), but it really took off in the mid-1930s when quota sampling was first applied to voting behavior studies and to the targeting of consumer messages by advertisers. Over the years, government agencies in all the industrialized countries have developed an insatiable appetite for information about various "target populations" (poor people, public housing residents, users of private health care, etc.). Japan developed an indigenous survey research industry soon after World War II and India, South Korea, Jamaica, Greece, Mexico, and many other countries have since developed their own survey research capabilities.

Today, survey research is big business. In the United States alone, the top 50 market research companies employed over 50,000 people in 1996 (including at least 5,000 professional social scientists) and had revenues of $5.5 billion dollars (Honimichl 1997). And that doesn't count all the university-based survey research facilities or the big international companies like Gallup and Roper that conduct public opinion polls.

Anthropologists are finding more and more that good survey technique can add a lot of value to ethnography. In the 1970s, Sylvia Scribner and Michael Cole studied literacy among the Vai of Liberia. Some Vai are literate in English, others are literate in Arabic, and some adult Vai men use an indigenous script for writing letters. As part of their project, Scriber and Cole ran a survey with 650 respondents. Michael Smith, the cultural anthropologist on their team, was skeptical about using this method with the Vai. He wrote the project leaders about his experience in administering the survey there:

> I was surprised when I first saw how long it [the questionnaire] was. I didn't think that anyone would sit down for long enough to answer it, or, if they did, that they would answer it seriously. . . . Well, I was wrong—and it fascinates me why the Vai should, in the busiest season of the year—during two of the worst farming years one could have picked . . . spent a lot of time answering questions which

had little to do with the essential business at hand. . . . Not only did the majority of people eventually come, but when they got there they answered with great deliberation. How many times does one remember someone saying, "I don't know, but I'll come back and tell you when I've checked with so-and-so." (Scribner and Cole 1981:47)

Face-to-Face, Self-Administered, and Telephone Interviews

There are three methods for collecting survey questionnaire data: (1) personal, **face-to-face** interviews, (2) **self-administered** questionnaires, and (3) **telephone** interviews. Self-administered questionnaires are usually mailed to respondents, but they may also be dropped off and picked up later or they may be given to people in a group all at once. Self-administered questionnaires can also be programmed into a computer or onto an interactive Internet site.

Each of the data-collection methods has its own advantages and disadvantages. There is no conclusive evidence that one method of administering questionnaires is better, overall, than the others. Your choice of a method will depend on your own calculus of things like cost, convenience, and the nature of the questions you are asking.

Personal, Face-to-Face Interviews

Face-to-face administration of questionnaires offers some important advantages, but they have some important disadvantages as well.

Advantages of Face-to Face Interviews

1. They can be used with people who could not otherwise provide information—respondents who are illiterate, blind, bedridden, or very old, for example.
2. If a respondent doesn't understand a question in a personal interview, you can fill in, and, if you sense that the respondent is not answering fully, you can probe for more complete data.

   Carry a notebook that tells you exactly how to respond when people ask you to clarify an unfamiliar term. If you use more than one interviewer, be sure each of them carries a copy of the same notebook. Good interview schedules are pretested to eliminate terms that are unfamiliar to intended respondents. Still, there is always someone who asks "What do you mean by 'income'?" or "How much is 'a lot'?"
3. You can use several different data collection techniques with the same respondent in a face-to-face survey interview. Part of the interview can consist of open-ended questions; another part may require the use of visual aids, such as graphs or cue cards; and in still another, you might hand the respondent a self-administered questionnaire booklet and stand by to help clarify potentially ambiguous items. This is a useful technique for asking really sensitive questions in a face-to-face interview.
4. Personal interviews at home can be much longer than telephone or self-administered questionnaires. An hour-long personal interview is relatively easy, and even two- and three-hour interviews are common. It is next to impossible to get respondents to devote two hours to filling out a questionnaire that shows up in the mail, unless you are prepared to pay well for their time; and it requires exceptional skill to keep a telephone interview going for more than 20 minutes, unless respondents are personally interested in the topic.

   Note, however, that **street-intercept** or **mall-intercept** interviews (where you interview people on the fly), though face to face, usually have to be very quick.
5. Face-to-face respondents get one question at a time and can't flip through the questionnaire to see what's coming. If you design an interview to start with general questions (how people feel about using new technologies at work, for example) and move on to specific questions (how people feel about using a particular new technology), then you really don't want people flipping ahead.
6. With face-to-face interviews, you know who answers the questions.

Disadvantages of Face-to-Face Interviews

1. They are intrusive and **reactive**. It takes a lot of skill to administer a questionnaire without subtly telling the respondent how you hope he or she will answer your questions. Other methods of administration of questionnaires may be impersonal, but that's not necessarily bad. Furthermore, the problem of reactivity increases when more than one interviewer is involved in a project.
2. Personal interviews are costly in both time and money. In addition to the time spent in interviewing people, locating respondents in a representative sample may require going back several times. In urban research especially, count on making up to half a dozen callbacks to get the really hard-to-find respondents.

   It's important to make all those callbacks in order to land the hard-to-get interviews. Survey researchers sometimes use the **sampling by convenient replacement** technique—going next door or down the block and picking up a replacement for an interviewee who happens not to be home at the time. As I mentioned in Chapter 6, this tends to homogenize your sample and make it less and less representative of all the variation in the population you're studying.
3. If you are working alone, without assistants, in an area that lacks good roads, don't plan on doing more than around 200 survey interviews in a year. If you're working in major cities in Europe or North America you can do more, but it gets really, really tough to maintain a consistent, positive attitude long before you get to the 200th interview. With mailed and telephone questionnaires, you can survey thousands of respondents.
4. Personal interview surveys conducted by lone researchers over a long period of

time run the risk of being overtaken by events. A war breaks out, a volcano erupts, or the government decides to cancel elections and imprison the opposition. It sounds dramatic, but these sorts of things are actually quite common across the world. Far less dramatic events can make the responses of the last 100 people you interview radically different from those of the first 100 to the same questions. If you conduct a questionnaire survey over a long period of time in the field, it is a good idea to reinterview your first few respondents and check the stability (reliability) of their reports.

Self-Administered Questionnaires

Self-administered questionnaires also have some clear advantages and disadvantages.

Advantages of Self-Administered Questionnaires

1. Mailed questionnaires put the post office to work for you in finding respondents. If you cannot use the mail (because sampling frames are unavailable, or because you cannot expect people to respond, or because you are in a country where mail service is unreliable), you can use cluster and area sampling (see Chapter 6), combined with the **drop-and-collect** technique. This involves leaving a questionnaire with a respondent and going back later to pick it up. In either case, self-administered questionnaires allow a single researcher to gather data from a large, representative sample of respondents, at relatively low cost per datum.
2. All respondents get the same questions with a self-administered questionnaire. There is no worry about interviewer bias.
3. You can ask more complex questions with a self-administered questionnaire than you can in a personal interview. Questions that involve a long list of response categories, or that require a lot of background data, are hard to follow orally but are often challenging to respondents if worded right.
4. You can ask long **batteries** of otherwise boring questions on self-administered questionnaires that you just couldn't get away with in a personal interview. Look at Figure 10.1. Imagine trying to ask someone to sit still while you recited, say, 30 items and asked for their response.
5. Response effects are absent in self-administered questionnaires that arrive by mail. Questions about sexual behavior (including family planning) and scales that measure attitudes toward women or men or members of particular ethnic/racial groups are particularly susceptible to response effects based on variation in gender, age, race, or ethnicity. The perceived sexual orientation of the interviewer, for example, affects how supportive respondents are of homosexuality (Kemph and Kasser 1996).

   Respondents may be more willing to report socially undesirable behaviors and traits in self-administered questionnaires (and in telephone interviews) than they are in face-to-face interviews. They aren't trying to impress interviewers,

Here is a list of things that people say they'd like to see in their high school. For each item, check how you feel <u>this</u> high school is doing.

	WELL	OK	POORLY	DON'T KNOW
1. High-quality instruction	___	___	___	___
2. Good pay for teachers	___	___	___	___
3. Good mix of sports and academics	___	___	___	___
4. Preparation for college entrance exams	___	___	___	___
5. Safety	___	___	___	___
6. Music program	___	___	___	___
7. Good textbooks	___	___	___	___

Figure 10.1. A battery item in a questionnaire. Batteries can consist of many items.

and anonymity gives people a sense of security, which produces more reports of things like premarital sexual experiences, constipation, arrest records, alcohol dependency, interpersonal violence, and so on (Hochstim 1967; Bradburn 1983).

This does *not* mean that *more* reporting of behavior means *more accurate* reporting. We know better than that now. But, as I've said before, more is usually better than less. If Chicanos report spending 12 hours per week in conversation with their families at home, while Anglos (as white, non-Hispanic Americans are known in the American Southwest) report spending 4 hours, I wouldn't want to bet that Chicanos *really* spend 12 hours, on average, or that Anglos *really* spend 4 hours, on average, talking to their families. But I'd find the fact that Chicanos reported spending three times as much time talking with their families pretty interesting.

6. Self-administered questionnaires can be programmed into a computer, or even given automatically by e-mail. (You may see this called **CAI** or **CAQ** in the social research literature, for **computer-assisted interviewing** or **computer-assisted questionnaire**.) Respondents take quickly to this format—it's used a lot in market research—and often find it to be a lot of fun. Fun is good because it cuts down on fatigue. Fatigue is bad because it sends respondents into robot mode and they stop thinking about their answers (O'Brien and Dugdale 1978; Barnes et al. 1995).

   In 1988, I ran a computer-based interview study as part of an ongoing project to estimate the size of uncountable populations. One member of our team, Christopher McCarty, programmed a laptop to ask respondents in Mexico City and Jacksonville, Florida, about their acquaintanceship networks. (Actually, "laptop"

is not quite the right word for the computer we used. In those days, a portable computer was one that fit under an airline seat—just barely. They were known as "lug-ables.") Those respondents said they enjoyed the experience. "Wow, this is like some kind of computer game," one respondent said.

In some cases, computer-based interviews may produce more honest answers about sensitive topics than do traditional paper-and-pencil tasks. Peterson et al. (1996) randomly assigned two groups of 57 Swedish Army veterans to fill out the Beck's Depression Inventory (see Beck et al. 1961). One group used the pencil-and-paper version, while the other used a computer-based version. Those who used the computer-based version had significantly higher mean scores on really sensitive questions about depression.

Disadvantages of Self-Administered Questionnaires

Despite these advantages, there are some hefty disadvantages to self-administered questionnaires.

1. You have no control over how people interpret questions on a self-administered instrument. There is always the danger that, no matter how much background work you do, no matter how hard you try to produce culturally correct questions, respondents will be forced into making culturally inappropriate choices in closed-ended questionnaires. If the questionnaire is self-administered, you can't answer people's questions about what a particular item means.
2. If you are not working in a highly industrialized nation, or if you are not prepared to use Dillman's Total Design Method (discussed below), you are likely to see response rates of 20–30% from mailed questionnaires. It is entirely reasonable to analyze the data statistically and to offer conclusions about the correlations among variables among those who responded to your survey. But response rates like these are unacceptable for drawing conclusions about larger populations.
3. Even if a questionnaire is returned, you can't be sure that the respondent who received it is the person who filled it out.
4. Mailed questionnaires are prone to serious sampling problems. Sampling frames of addresses are almost always flawed, sometimes very badly. For example, if you use a phone book to select a sample, you miss all those people who don't have phones or who choose not to list their numbers. Face-to-face administration of questionnaires is often based on an area cluster sample, with random selection of households within each cluster. This is a much more powerful sampling design than most mailed questionnaire surveys can muster.
5. In some cases, you may want respondents to answer a question without knowing what's coming next. This is impossible in a self-administered questionnaire.
6. Self-administered questionnaires are simply not useful for studying nonliterate or illiterate populations, or for studying people who can't see. This problem will eventually be solved by voice recognition software, but we're just at the beginning of that particular revolution.

Telephone Interviews

Once upon a time, not too long ago, telephone surveys were considered a poor substitute for face-to-face surveys. Today, telephone interviewing has become the most widely used method of gathering survey data across the industrialized nations of the world where so many households have their own phones. Administering questionnaires by phone has some very important advantages.

Advantages of Telephone Interviews

1. Research has shown that, in the United States at least, answers to many different kinds of questions asked over the phone are as valid as those to questions asked in person or through the mail (Dillman 1978).
2. Phone interviews combine the impersonal quality of self-administered questionnaires with the personal quality of face-to-face interviews. So, telephone surveys are unintimidating (like self-administered questionnaires), but allow interviewers to probe or to answer questions dealing with ambiguity of items (like personal interviews).
3. Telephone interviewing is inexpensive and convenient to do. It's not without effort, though. Professional survey organizations routinely do at least three callbacks to numbers that don't answer, and many survey researchers insist on 10 callbacks to make sure that they get an unbiased sample. As it is, in most telephone surveys, you can expect 30–40% refusals. You can also expect nearly 100% sample completion, because it's relatively easy to replace refusers with people who will cooperate. But remember to keep track of the refusal rate and to make an extra effort to get at least some of the refusers to respond so you can test whether cooperators are a biased sample.
4. Using **random digit dialing** (RDD), you can reach almost everyone who has a phone. In the United States, that means you can reach almost everybody. One recent survey found that 28% of completed interviews using RDD were with people who had unlisted phone numbers (Taylor 1997:424). There are huge regional differences, though, in the availability of telephones (see below).
5. Unless you do all your own interviewing, interviewer bias is an ever-present problem in survey research. It is relatively easy to monitor the quality of telephone interviewers' work by having them come to a central place to conduct their operation. (If you don't monitor the performance of telephone interviewers, though, you invite cheating. This is one of the disadvantages listed below.)
6. There is no reaction to the *appearance* of the interviewer in telephone surveys, although respondents *do* react to accents and speech patterns of interviewers.

Oskenberg et al. (1986) found that telephone interviewers who had the lowest refusal rates had higher-pitched, louder, and clearer voices than most. And, as with all types of interviews, there are gender-of-interviewer and race-of-interviewer effects in telephone interviews, too. Respondents try to figure out the race or ethnicity of the interviewer and then tailor responses accordingly.

In the National Black Election Study, 872 African Americans were polled before and after the 1984 presidential election. Since interviewers were assigned randomly to respondents, some people were interviewed by a white person before the election and an African American after the election, and vice versa: Some people were interviewed by an African American before the election and a white person on the second wave.

Darren Davis (1997) looked at data from this natural experiment. When African American interviewers in the preelection polls were replaced by white interviewers in the postelection surveys, African American respondents were more likely to say that blacks don't have the power to change things, that black people can't make a difference in local or national elections, that black people cannot form their own political party, and that whites are not responsible for keeping blacks down—very powerful evidence of a race-of-interviewer effect.

7. Telephone interviewing is safe. You can talk on the phone to people who live in urban neighborhoods where many professional interviewers (most of whom are women) would prefer not to go. Telephones also get you past doormen and other people who run interference for the rich.

Disadvantages of Telephone Interviewing

The disadvantages of telephone surveys are obvious.

1. If you are doing research in Haiti or Bolivia or elsewhere in the developing world, telephone surveys are out of the question, except in some urban centers, and then only if your research is about relatively well-off people.

   Even in highly industrialized nations, not everyone has a telephone. About 94% of all households in the United States have telephones. This makes *national* surveys a cinch to do and highly reliable. But the distribution of telephones is uneven, which makes some *local* surveys impossible to do by phone.

   In Westchester County, New York, for example, the median per capita income is over $35,000 per year and over 98% of all households have phones. The 2% of households that don't have phones in Westchester County are rental properties. In Coconino County, Arizona, the median per capita income is less than $15,000 per year, and only about 75% of all housing units—owner-occupied and rentals alike—have phones. Be very certain about the demographics of the area you're thinking of studying before deciding on a telephone survey. If having a telephone is associated with having a certain amount of income and/or membership in an ethnic group, as it is in Coconino County, Arizona, then a telephone interview of that population will produce statistically ungeneralizable results.

2. Telephone interviews must be relatively short, or people will hang up. There is some evidence that once people agree to give you their time in a telephone interview, you can keep them on the line for a remarkably long time (up to an hour) by developing special "phone personality" traits. Generally, however, you should not plan a telephone interview that lasts for more than 20 minutes.

3. And finally, this: It has long been known that, in an unknown percentage of occasions, hired interviewers willfully produce inaccurate data. When an interviewer who is paid by the completed interview finds a respondent not at home, the temptation is to fill in the interview and get on to the next respondent. This saves a lot of calling back, and introduces garbage into the data.

   Unless there is continual monitoring, it's particularly easy for interviewers to cheat in telephone surveys—from failing to probe, to interviewing unqualified respondents, to fabricating an item response, and even to fabricating whole interviews. Kiecker and Nelson (1996) hired 33 survey research companies to do eight interviews each, ostensibly as "mop-up" for a larger national market survey. The eight respondents were plants—graduate students of drama, for whom this must have been quite a gig—and were the same eight for each of the surveys. Of the 33 interviewers studied, 10 fabricated an entire interview, 32 fabricated at least one item response, and all 33 failed to record responses verbatim.

The technology of telephone interviewing has become very sophisticated. **Computer-assisted telephone interviewing** (CATI) makes it harder to do things like ask questions out of order, but a determined cheater on your interviewing team can do a lot of damage. The good news is that once you eliminate cheating (with monitoring), the main thing left that can go wrong is inconsistency in the way interviewers ask questions. Unstructured and structured interviews each have their own advantages, but for structured interviews to yield reliable results, they have to be really, really structured. That is, the questions have to be read verbatim so that every respondent is exposed to the same stimulus.

Repeated verbatim readings of questions is boring to do and boring to listen to. When respondents (inevitably) get restless, it's tempting to vary the wording to make the interview process seem less mechanical. This turns out to be a bigger problem in face-to-face interviews (where interviewers are generally working alone, without any monitoring) than in telephone interviews. Presser and Zhao (1992) monitored 40 trained telephone interviewers at the Maryland Survey Research Center. For the 5,619 questions monitored, interviewers read the questions exactly as worded on the survey 91% of the time. Training works.

Still, no matter how much you train interviewers . . . Johnstone et al. (1992) studied 48 telephone interviews done entirely by women and found that female respondents elicited more sympathy, while male respondents elicited more joking. Men, say Johnstone et al., may be less comfortable than women are with being interviewed by women and wind up trying to subvert the interview by turning it into teasing or banter.

Sampling for telephone surveys is also aided by computer. There are companies that sell telephone numbers for surveys. The numbers are chosen to represent businesses or residences and to represent the varying saturation of phone service in different calling areas.

Even the best sample of phone numbers, though, many not be enough to keep you out of trouble. During the 1984 U.S. presidential election, Ronald Reagan's tracking poll used a list of registered voters, Republicans and Democrats alike. The poll showed Reagan comfortably ahead of his rival, Walter Mondale, except on Friday nights. Registered Republicans, it turned out, being wealthier on average than Democrats, were out Friday nights more than Democrats were, and simply weren't available to answer the phone (Begley et al. 1992:38).

When to Use What

There is no perfect data-collection method. However, self-administered questionnaires are preferable to personal interviews when three conditions are met: (1) you are dealing with literate respondents; (2) you are confident of getting a high response rate (at least 70%); and (3) the questions you want to ask do not require a face-to-face interview or the use of visual aids such as cue cards, charts, and the like. Under these circumstances, you get much more information for your time and money than from the other methods of questionnaire administration.

When you really need complete interviews—answers to all or nearly all the questions in a particular survey—then face-to-face interviews are the way to go. Caserta et al. (1985) interviewed recently bereaved respondents about adjustment to widowhood. They interviewed 192 respondents—104 in person, at home, and 88 by mailed questionnaire. Both groups got identical questions.

On average, 82% of those interviewed at home three to four weeks after losing their spouse answered any given question. Just 68% of those who responded to the mailed questionnaire answered any given question. As Caserta et al. explain, the physical presence of the interviewer helped establish the rapport needed for asking sensitive and personal questions about the painful experience of bereavement (ibid.:640).

If you are working in a highly industrialized country, and if a very high proportion (at least 80%) of the population you are studying has telephones, then consider doing a phone survey whenever a self-administered questionnaire would otherwise be appropriate.

If you are working alone or in places where the malls and the phone system are inefficient for data collection, the drop-and-collect technique is a good alternative (see above, p. 244).

Finally, there is no rule against using more than one type of interview. Mauritius, an island nation in the Indian Ocean, is an ethnically complex society. Chinese, Creoles, Franco-Mauritians, Hindus, Muslims, and other groups make up a population of about a million. Ari Nave (1997) was interested in how Mauritians maintain their ethnic group boundaries, particularly through their choices of whom to marry. A government office on Mauritius maintains a list of all people over 18 in the country, so it was relatively easy for Nave to get a random sample of the population.

Contacting the sample was another matter. Nave got back just 347 out 930 mailed questionnaires, but he was able to interview another 296 by telephone and face-to-face, for a total of 643, or 69% of his original sample—a respectable completion rate.

Using Interviewers

There are several advantages to using multiple interviewers in survey research. The most obvious is that you can increase the size of the sample. Multiple interviewers, however, introduce several disadvantages, and whatever problems are associated with interviewer bias are increased with more than one interviewer.

Just as important, multiple interviewers increase the cost of survey research. If you can collect 100 interviews yourself and maintain careful quality control in your interview technique, then hiring one more interviewer would probably not improve your research by enough to warrant both spending the extra money and worrying about quality control. Recall that for estimating population proportions or means, you have to quadruple the sample size to halve the sampling error. If you can't afford to hire three more interviewers (beside yourself), and to train them carefully so that they at least introduce the *same* bias to every interview as you do, you're better off running the survey yourself and saving the money for other things.

This only goes for surveys in which you interview a random sample of respondents in order to estimate a population parameter. If you are studying

the experience of a group of people, or are after cultural data (as in "How are things usually done around here?"), then getting more interviews is better than getting fewer, whether you collect the data yourself or have it collected by others.

Training Interviewers

If you hire interviewers, be sure to train them—and monitor them throughout the research. A colleague used a doctoral student as an interviewer in a project in Atlanta. The senior researcher trained the student, but listened to the interview tapes that came in. At one point, the interviewer asked a respondent, "How many years of education do you have?" "Four," replied the respondent. "Oh," said the student researcher, "you mean you have four years of education?" "No," said the informant, bristling and insulted, "I've had four years of education beyond high school." The informant was affluent, the interview was conducted in his upper-middle-class house, and he had already told the interviewer that he was in a high-tech occupation. So monitor interviewers.

If you hire a *team* of interviewers, you have one extra chore besides monitoring their work. You need to get them to act as a team. Be sure, for example, that they all use the same probes to the various questions on the interview schedule. Especially with open-ended questions, be sure to do random spot checks, *during* the survey, of how interviewers are coding the answers they get. The act of spot checking keeps coders alert. When you find discrepancies in the way interviewers code responses, bring the group together and discuss the problem openly.

Billiet and Loosveldt (1988) found that asking interviewers to tape all their interviews produces a higher response rate, particularly to sensitive questions about things like sexual behavior. Apparently, when interviewers know that their work can be scrutinized (from the tapes) they probe more and get informants to open up more.

Carey et al. (1996) studied the beliefs of 51 newly arrived Vietnamese refugees in upstate New York about tuberculosis. The interviews consisted of 32 open-ended questions on beliefs about symptoms, prevention, treatment, and the social consequences of having TB. The two interviewers in this study were bilingual refugees who participated in a three-day workshop to build their interview skills. They were told about the rationale for open-ended questions and about techniques for getting respondents to open up and provide full answers to the questions. The training included a *written manual* (this is very important) to which the interviewers could refer during the actual study. After the workshop, the trainees did 12 practice interviews with Vietnamese adults who were not in the study.

William Axinn ran the Tamang Family Research Project, a comparative study of villages in Nepal (Axinn 1991). Axinn and his coworkers trained a group of interviewers using the *Interviewer's Manual* from the Survey Research Center at the University of Michigan (University of Michigan 1976). That manual contains the distilled wisdom of hundreds of interviewer training exercises in the United States, and Axinn found the manual useful in training Nepalese interviewers, too.

Axinn recruited 32 potential interviewers. After a week of training (five days at eight hours a day, plus two days of supervised field practice), the 16 best interviewers were selected, 10 men and 6 women. The researchers hired more interviewers than they needed and after three months, four of the interviewers were fired. "The firing of interviewers who clearly failed to follow protocols had a considerable positive effect on the morale of interviewers who had worked hard to follow our rules" (Axinn et al. 1991:200). No one has accused Axinn of overstatement.

Whom to Hire

In general, when hiring interviewers, look for professional interviewers first. Next, look for people who are mature enough to accept the need for rigorous training and who can work as part of a team. If need be, look for interviewers who can handle the possibility of going into some rough neighborhoods and who can answer the many questions that respondents will come up with in the course of the survey.

If you are running a survey based on personal interviews in a developing country, consider hiring college students, and even college graduates, in the social sciences. "Social sciences," by the way, does not mean the humanities. In Peru, Donald Warwick and Charles Lininger found that "some students from the humanities . . . were reluctant to accept the 'rigidities' of survey interviewing." Those students felt that "as educated individuals, they should be allowed to administer the questionnaire as they saw fit in each situation" (Warwick and Lininger 1975:222).

I would not use anyone as an interviewer who had that kind of attitude. But many undergraduate students in the developing world get lots of real research experience, since graduate training in the social sciences is not widely available outside the major industrialized nations. Students who are experienced interviewers have a lot to contribute to the design and content of questionnaires. Remember, you are dealing with colleagues who will be justly resentful if you treat them as merely employees of your study. By the same token, college

students in developing nations—particularly those in public universities—are likely to be members of the elite, who may find it tough to establish rapport with peasant farmers or the urban poor (Hursh-César and Roy 1976:308).

Make It Easy for Interviewers to Do Their Job

If you use interviewers, be sure to make the questionnaire booklet easy to use. Leave enough space for interviewers to write in the answers to open-ended questions—but not too much space. Big spaces are an invitation to some interviewers to develop needlessly long answers (Warwick and Lininger 1975:152).

Also, use two different type faces for questions and answers; put instructions to interviewers in capital letters and questions for respondents in normal type. Figure 10.2 is an example:

5. INTERVIEWER: CHECK ONE OF THE FOLLOWING
☐ R HAS LIVED IN CHICAGO MORE THAN FIVE YEARS. SKIP TO QUESTION 7.
☐ R HAS LIVED IN CHICAGO LESS THAN FIVE YEARS. ASK QUESTION 6 AND CONTINUE WITH QUESTION 7.
6. Could you tell me where you were living five years ago?
7. Where were you born?

SOURCE: Adapted from D. P. Warwick and C. A. Lininger, *The Sample Survey: Theory and Practice*, p. 153. © 1975, McGraw-Hill. Reprinted by permission.

Figure 10.2. Using two different type faces in a survey instrument.

Closed vs. Open-Ended: The Problem of Threatening Questions

The most often-asked question about survey research is whether **forced-choice** (also called **closed-ended**) or **open-ended** items are better. Schuman and Presser (1979) tested this. They asked one sample of people this question: "Please look at this card and tell me which thing you would most prefer in a job." The card had five items listed: (1) high income; (2) no danger of being fired; (3) working hours are short—lots of free time; (4) chances for advancement; and (5) the work is important and gives a feeling of accomplishment. Then they asked another sample the open-ended question: "What would you most prefer in a job?"

About 17% of the respondents to the closed-ended question chose "chances for advancement" and over 59% chose "important work." By contrast, fewer than 2% of the respondents who were asked the open-ended question mentioned "chances for advancement, and just 21% said anything about "important" or "challenging" or "fulfilling" work.

When the questions get really threatening, the problem gets worse. Masturbation, alcohol consumption, and drug use are reported with 50–100% greater frequency in response to open-ended questions (Bradburn 1983:299). For reporting these kinds of behavior, people are apparently least threatened when they can offer their own answers to open-ended questions on a self-administered questionnaire, rather than being forced to choose among a set of fixed alternatives (e.g., once a month, once a week, once a day, several times a day) and are most threatened by a face-to-face interviewer (Blair et al. 1977).

Ivis et al. (1997) found that at least one pretty embarrassing question was better asked in a closed-ended format—over the phone. People in their survey who were asked: "How often in the last 12 months have you had five or more drinks on one occasion?" Then, later in the interview, they were asked the same question, but were given nine fixed choices: (1) every day; (2) about once every other day; . . . (9) never in the last year.

The closed-ended format produced significantly more positive responses. Many people in the industrialized countries have become accustomed to telephone surveys and are comfortable with the anonymity of those surveys. Notice that the anonymity of telephone surveys lets the *interviewer*, as well as the respondent, off the hook. You can ask people things you might be squeamish about if the interview were face to face, and respondents feel that they can divulge very personal matters to disembodied voices on the phone.

Experiments are continuing on the efficacy of survey format (self-administered, face-to-face, telephone) and question format (open vs. closed ended), and the jury is still out (see Wentland and Smith 1993). Overall, since closed-ended items are so efficient, most survey researchers prefer them to open-ended questions and use them whenever possible.

But there is no rule that prevents you from mixing question types. Many survey researchers use the open-ended format for really intimidating questions and the fixed-choice format for everything else, even on the phone. Even if there are no intimidating questions in a survey, it's a good idea to stick in a few open-ended items. The open-ended questions break the monotony for the respondent, as do tasks that require referring to visual aids (like a graph).

The responses to fixed-choice questions are unambiguous for purposes of analysis. Be sure to take full advantage of this and precode fixed-choice items on a questionnaire. Put the codes right on the instrument so that computer input of the data is made as easy (and as error free) as possible.

Question Wording and Format

There are some well-understood rules that all survey researchers follow in constructing questionnaire items. Here are 15 of them.

1. *Be unambiguous.* If respondents can interpret a question differently from the meaning you have in mind, they will. In my view, this is the source of most response error in closed-ended questionnaires.

The problem is not easy to solve. A simple question like "How often do you visit a doctor?" can be very ambiguous. Are acupuncturists, chiropractors, chiropodists, and public clinics all doctors? If you think they are, you'd better tell people that, or else you leave it up to them to decide. In some parts of the southwestern United States, people may be visiting native curers and herbalists. Are those practitioners doctors? In Mexico, many community clinics are staffed by nurses. Does "going to the doctor" include a visit to one of those clinics?

If you ask, "How long have you lived in Mexico City?" does "Mexico City" include the 20 million people who live in the urban sprawl, or just the 8 million who live in the Federal District? And how "near" is "near Nairobi"?

Words like "lunch," "community," "people," and hundreds of other innocent lexical items have lurking ambiguities associated with them, and phrases like "family planning" will cause all kinds of mischief. Half the respondents in the 1985 General Social Survey were asked if they agreed that there was too little spending for "assistance to the poor," while half were asked if there was too little spending for "welfare." A whopping 65% agreed with the first wording; just 19% agreed with the second (T. W. Smith 1987:77).

Even the word "you," as Payne (1951) pointed out, can be ambiguous. Ask a nurse at the clinic "How many patients did you see last week?" and you might get a response like "Who do you mean, me or the clinic?" Of course, if the nurse is filling out a self-administered questionnaire, she'll have to decide for herself what you had in mind. Maybe she'll get it right; maybe she won't.

2. *Use a vocabulary that your respondents understand, but don't be condescending.* This is a difficult balance to achieve. If you're studying a narrow population (sugar cane cutters, midwives, leather workers), then proper ethnography and pretesting with a few knowledgeable informants will help ensure appropriate wording of questions.

But if you are studying a more general population, even in a small town of just 3,000 people, then things are very different. Some respondents will require a low-level vocabulary; others will find that vocabulary insulting. This is one of the reasons often cited for doing personal interviews: You want the opportunity to phrase your questions differently for different segments of the population. Realize, however, that this poses risks in terms of reliability of response data.

3. *Remember that respondents must know enough to respond to your questions.* You'd be surprised at how often questionnaires are distributed to people who are totally unequipped to answer them. I get questionnaires in the mail all the time asking for information I simply don't have.

Most people can't recall with any acceptable accuracy how long they spent in the hospital last year, how many miles they drive each week, or how much they've cut back on their use of air conditioning. They *can* recall whether they own a television, have ever been to Cairo, or voted in the recent elections. And they can tell you whether they *think* they got a fair price for the land they vacated when the dam was built or *believe* the local member of parliament is doing a better job than her predecessor at giving equal time to rich people and poor people who come to her with complaints.

4. Make sure there's a *clear purpose for every question* you ask in a survey. When I say "clear purpose," I mean clear to respondents, not just to you. And once you're on a topic, stay on it and finish it. Respondents can get frustrated, confused, and annoyed at the tactic of switching topics and then coming back to a topic that they've already dealt with on a questionnaire. Some researchers do exactly this just to ask the same question in more than one way and to check respondent reliability. This underestimates the intelligence of respondents and is asking for trouble—I have known respondents to sabotage questionnaires that they found insulting to their intelligence.

You can (and should) ask questions that are related to one another at different places in a questionnaire, so long as each question makes sense in terms of its placement in the overall instrument. For example, if you are interviewing labor migrants, you'll probably want to get a labor history—by asking where the respondent has worked during the past few years. Later, in a section on family economics, you might ask whether a respondent has ever sent remittances and from where.

As you move from one topic to another, put in a transition paragraph that makes each shift logical to the respondent. For example, you might say: "Now that we have learned something about the kinds of food you like, we'd like to know about . . ." The exact wording of these transition paragraphs should be varied throughout a questionnaire.

5. Pay careful attention to **contingencies** and **filter questions**. Many question topics contain several contingencies. Suppose you ask someone if they are married. If they answer "no," then you probably want to ask whether they've ever been married. You may want to know whether they have children, irrespective of whether they are married or have ever been married. You may want to know what they think is the ideal family size, irrespective of whether they've been married, plan to be married, have children, or plan to have children.

You can see that the contingencies can get very complex. The best way to ensure that all contingencies are accounted for is to build a **contingency flow chart** like that shown in Figure 10.3 (Sirken 1972; Sudman and Bradburn 1982).

6. *Use clear scales.* There are some commonly used scales in survey research—things like: Excellent-Good-Fair-Poor; Approve-Disapprove; Oppose-Favor; For-Against; Good-Bad; Agree-Disagree; Better-Worse-About the Same; etc. Just because these are well known, however, does not mean that they are clear and unambiguous to respondents.

To cut down on the ambiguities associated with these kinds of scales, explain the meaning of each potentially ambiguous scale when you introduce it. With self-administered questionnaires, use five scale points rather than three, if you can. For example, use Strongly Approve, Approve, Neutral, Disapprove, Strongly Disapprove, rather than Approve, Neutral, Disapprove. This will give people the opportunity to make finer-grained choices. If your sample is large enough, you can distinguish during analysis among respondents who answer, say, "strongly approve" vs. "approve" on some item.

For smaller samples, you'll have to aggregate the data into three categories for analysis. Self-administered questionnaires allow the use of seven-point scales, like the semantic differential scale shown in Figure 10.4. Telephone interviews usually require three-point scales. (More on how to construct scales in Chapter 12.)

7. Try to *package questions in self-administered questionnaires*, as shown earlier in Figure 10.1. This is a way to get a lot of data quickly and easily, and, if done properly, it will prevent respondents from getting bored with a survey. For example, you might say "Please indicate how close you feel to each of the persons on this chart" and provide the respondent with a list of relatives (mother, father, sister, brother, etc.) and a scale (very close, close, neutral, distant, very distant, etc.).

Be sure to make scales unambiguous (if you are asking how often people think they do something, don't say "regularly" when you mean "more than

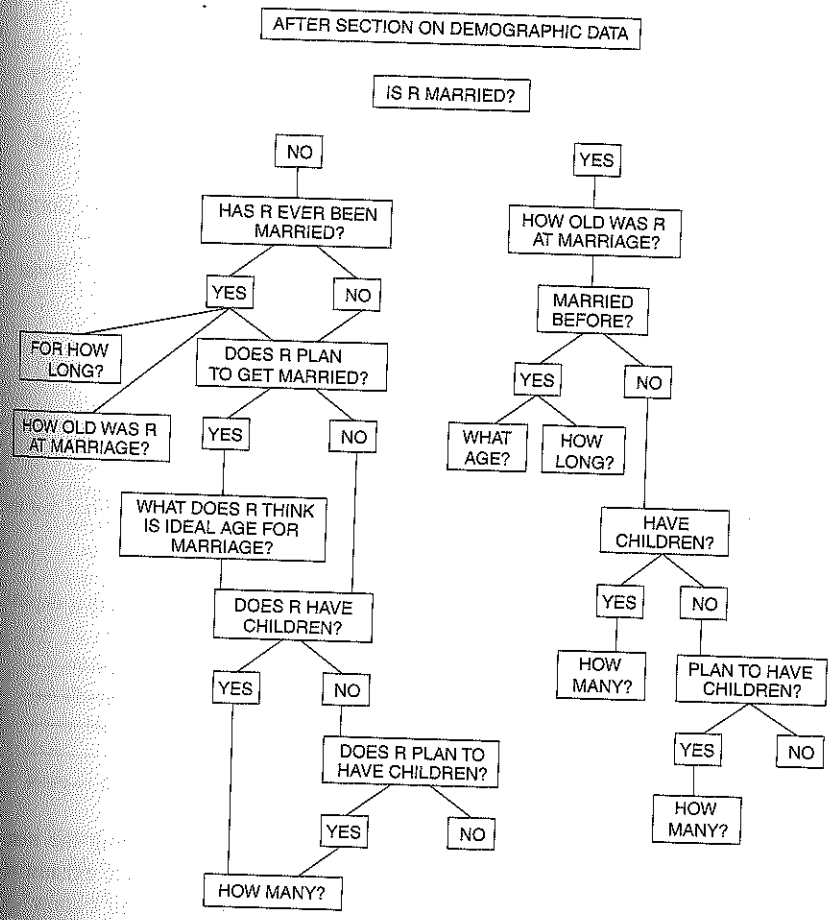

Figure 10.3. Flow chart of filter questions for part of a questionnaire.

once a month") and limit the list of activities to no more than seven. Then introduce a question with a totally different format, to break up the monotony and to keep the respondent interested.

Packaging is best done in self-administered questionnaires. If you use these kinds of lists in a personal interview, you'll have to repeat the scale for at least the first three items or activities you name, or until the respondent gets the pattern down. This can get very tiring for both interviewers and respondents.

8. If you want respondents to check just one response, then be sure to *make the possible responses to a question exhaustive and mutually exclusive.*

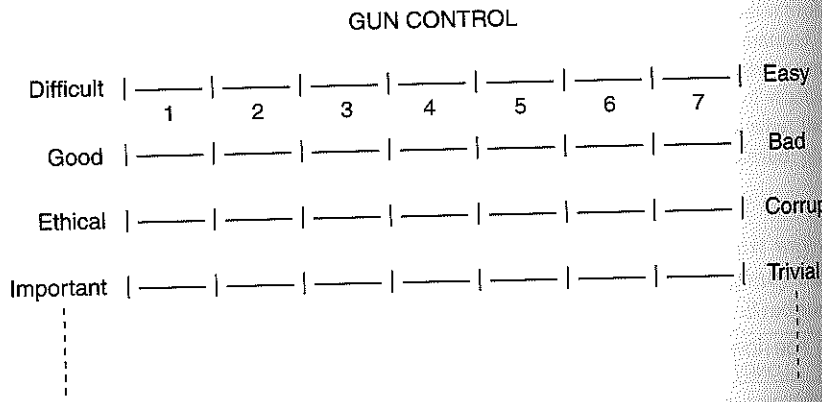

Figure 10.4. A seven-point semantic differential scale.

Here is an example (taken from a questionnaire I received) of what *not* to do:

> How do you perceive communication between your department and other departments in the university? [check one]
> There is much communication _____
> There is sufficient communication _____
> There is little communication _____
> There is no communication _____
> No basis for perception _____

The "no basis for perception" response took care of making the item exhaustive. You can always make questionnaire items like this one exhaustive by giving respondents the option of saying some variant of "don't know"—like "no basis for perception." Some researchers feel that this just gives respondents a lazy way out—that people need to be made to work a bit. But if there is a good chance that some of your respondents really won't have the information you ask for, then I think the "don't know" option is too important to leave out. In consumer preference surveys, though, where you actually give someone a taste of a cracker and ask them to tell you if they like it, the "don't know" option is a bad idea.

The problem for me on the questionnaire item above was that I wanted to check both "little communication" and "sufficient communication." For me, at least, these two categories were not mutually exclusive—I didn't think there was a lot of communication, and I wasn't at all bothered by that—but the author of the survey asked me to "check one."

9. *Keep unthreatening questions short.* Questions that are likely to intimidate respondents should have long preambles to lessen the intimidation effect. The questions themselves, however, should contain as few words as possible.
10. Always *provide alternatives*, if appropriate. Suppose people are being asked to move off their land to make way for a new highway. The government offers to compensate people for the land, but people are suspicious that the government won't evaluate fairly how much compensation landowners are entitled to. If you take a survey and ask "Should the government offer people compensation for their land?" respondents can answer yes or no for very different reasons. Instead, let people check whether they agree or disagree with a set of alternatives, like: "The government should offer people compensation for their land" and "An independent board should determine how much people get for their land."
11. *Avoid loaded questions.* Any question that begins "Don't you agree that . . ." is a **loaded question.** Sheatsley (1983) points out, however, that asking loaded questions is a technique you can use to your advantage, on occasion, just as leading or baiting informants can be used in unstructured interviewing. A famous example comes from Alfred Kinsey's landmark study of sexual behavior of American men (Kinsey et al. 1948). Kinsey asked men, "How old were you the first time you masturbated?" This made respondents feel that the interviewer already *knew* about the fact of masturbation, and was only in search of additional information.
12. *Don't use* **double-barreled** questions. Here is one I found on a questionnaire: "When did you leave home and go to work on your own for the first time?" There is no reason to assume, of course, that someone had to leave home in order to go to work, or that they necessarily went to work if they left home.

Here is another bad question:

> Please indicate if you agree or disagree with the following statement:
> Marijuana is no more harmful than tobacco or alcohol, so the personal use of marijuana should be legal.

Suppose a respondent agrees (or disagrees) with the first part of the statement—the assertion that marijuana is no more harmful than tobacco or alcohol. He or she may agree or disagree with the second part of the statement. If respondents answer "yes" or "no," how do you know if they are indicating agreement with both parts of the statement or just one part? Which part? How can you tell? You can't. That's why it's a bad question.

13. *Don't put false premises into questions.* I once formulated the following question for a survey in Greece: "Is it better for a woman to have a house and cash as a dowry, or for her to have an education and a job that she can bring to the marriage?" This question was based on a lot of ethnographic work in a community, during which I learned that many families were sinking their resources into

getting women educated and into jobs and offering this to eligible bachelors as a substitute for traditional material dowries.

My question, however, was based on the false premise that all families respected the custom of dowry. The question did not allow respondents to state a third alternative—namely, that they didn't think dowry was a custom that ought to be maintained in any form, traditional or modern. In fact, many families were deciding to reject the dowry custom altogether—something that I missed for some time because I failed to pretest the item (see Pretesting, p. 270).

14. *Don't take emotional stands in the wording of questions.* Here's an example of the sort of question you see on surveys all the time—and that you should never ask: "Should the legislature raise the drinking age to 21 in order to reduce the carnage among teens on our highways?" Another example of a bad question is: "Don't you agree with the president when he says . . . ?"
15. When asking for opinions *on controversial issues, specify the referent situation* as much as possible. Instead of asking "Do you approve of abortion?" ask "Under what conditions do you approve of abortion?" Then give the respondent as exhaustive a list of circumstances as possible to check. If the circumstances are not exclusive (rape and incest are not necessarily exclusive, for example), then let respondents check as many circumstances as they think appropriate.

Translation and Back Translation

All the tips given here about writing good survey questions continue to apply when you are working in another culture. They are just a lot more difficult to implement because you have to deal with phrasing questions properly in another language as well. The best way to deal with this is through back translation (Brislin 1970).

Back translation is the standard method for adapting social and psychological measurement scales—like the Locus of Control scale, Beck's Depression Inventory, and so on. First, write any questionnaire in *your* native language, paying attention to all the lessons of this chapter. Then have the questionnaire translated by a bilingual person who is a native speaker of the language you a working in. Work closely with the translator, so that she or he can fully understand the subtleties you want to convey in your questionnaire items.

Next, ask another bilingual person, who is a native speaker of *your* language, to translate the questionnaire back into the original language. This back translation should be almost identical to the original questionnaire you wrote. If it isn't, then something was lost in one of the two translations. You'd better find out which one it was and correct the problem. This is what Axinn et al.

(1991) did in their study of fertility in Nepal. In fact, they cross-checked the meaning of each question in the back translation against the original and went through the whole exercise several times until they were satisfied that questions in Nepalese were sensible.

You can also use back translation to check the content of open-ended interview, but be warned: This is tough work. Daniel Reboussin interviewed Diola women who had come from southwestern Senegal to Dakar in search of work (1995). All the women used French at work, but they preferred Diola for interviews. Reboussin, who speaks French, spoke very little Diola, so he worked with an interpreter—a man named Antoine Badji—to develop an interview schedule in French, which Badji translated into Diola.

During the interviews, Badji asked questions in Diola and Reboussin audiotaped the responses. After each interview, Badji translated each tape (orally) into French. Reboussin transcribed the French translations, translating into English as he went. Then he read his English transcriptions *back* to Badji, translating (orally) into French as he read. That way, Badji could confirm or disconfirm Reboussin's understanding of Badji's French rendering of the tapes.

As I said, this was tough work. It took Reboussin and Badji 17 weeks to conduct 30 interviews and get them all down into English.

The Response Rate Problem

Mailed questionnaires can be very, very effective, but there is one problem with them that all survey researchers watch for: getting enough of them back. In 1936, the *Literary Digest* sent out 10 million straw poll ballots in an attempt to predict the winner of the presidential election. It got back 2.3 million ballots and predicted Alf Landon over Franklin Delano Roosevelt in a landslide. Roosevelt got 61% of the vote.

Now, you'd think that 2.3 million ballots would be enough for anyone, but two things caused the *Digest* debacle. First, the surveyors selected their sample from automobile registries and telephone books. In 1936, this favored richer people, who tend to be Republican. Second, the 2.3 million ballots were only 23% of the 10 million sent out. The low response rate biased the results in favor of the Republican challenger since those who didn't respond tended to be poorer and less inclined to participate in surveys (Squire 1988).

How to Adjust for Nonresponse

Skip to 1991. The American Anthropological Association sent questionnaires to a sample of its 1,229 members. The sample was stratified into several

cohorts who had received their Ph.D. degrees beginning in 1971–1972 and 1989–1990. The 1989–1990 cohort comprised 306 then-recent Ph.D.s. The idea was to find out what kinds of jobs those anthropologists had.

The AAA got back 840 completed questionnaires, or 68% of the 1,229. The response rate is not high for this kind of survey, where respondents are being asked for information from their own professional organization. (The U.S. Office of Management and Budget demands a minimum 75% response rate from survey contract researchers [Fowler 1984:48] and, as we saw earlier, this is not an excessive demand.)

The results of the survey were reported in the *Anthropology Newsletter* in May 1991. Of those responding from the 1989–1990 cohort, 41% said they had academic jobs. The *Anthropology Newsletter* didn't report the response rate by cohort, but suppose that 68% of the 1989–1990 cohort—the same percentage as applies to the overall survey—sent back their questionnaires. That's 208 out of 306 responses. The 41% who said they had academic jobs would be 85 of the 208 respondents; the other 123 had nonacademic jobs.

Suppose that everyone who didn't respond (32%, or 98 out of 306) had nonacademic jobs. (Maybe that's why they didn't bother to respond.) In that case, 98 + 123 = 221 out of the 306 people in the cohort, or 72%, had nonacademic jobs that year—not the 59% (100% − 41%) as reported in the survey.

It's unlikely, of course, that *all* the nonresponders were in nonacademic jobs. To handle the problem of nonresponse, the AAA might have run down a random grab of 10 of the nonresponders and interviewed them by telephone. Suppose that seven said they had nonacademic jobs. You'll recall from Chapter 7 on sampling theory that the formula for determining the 95% confidence limits of a point estimator is:

$$P(\text{the true proportion}) = 1.96\sqrt{PQ/n} \qquad \textbf{Formula 10.1}$$

For our hypothetical example, this means that

$$1.96\sqrt{(.7)(.3)/10} = .28$$

The probable answer for the 10 holdouts is .70 ± .28. Somewhere between 42% and 98% of the 98 nonresponders from the 1989–1990 cohort probably had nonacademic jobs. We guess that 41–96 of those 98 nonresponders had nonacademic jobs. We can now make a reasonable guess: 123 of the responders plus 41–96 of the nonresponders, that is, 164–216 of the 306 people in the cohort, or 54–71%, probably had nonacademic jobs.

Low response rate can be a disaster. People who are quick to fill out and return mailed questionnaires tend to have higher incomes and consequently tend to be more educated than the later respondents. Any dependent variables that co-vary with income and education, then, will be seriously distorted if you get back only 50% of your questionnaires. And what's worse, there is no accurate way to measure nonresponse bias. With a lot of nonresponse, all you know is that you've got bias, but you don't know how to take it into account.

Improving Response Rates: Dillman's Total Design Method

Fortunately, a lot of research has been done on increasing response rates to mailed questionnaires. Yammarino et al. (1991) reviewed 184 controlled experiments, done between 1940 and 1988, on maximizing the return of mailed questionnaires. Don Dillman, of the Survey Research Laboratory at Washington State University, has synthesized the research on maximizing return rates and has developed what he calls the Total Design Method of mail and telephone surveying (Dillman 1978, 1983; Salant and Dillman 1994).

Professional surveys done in the United States following Dillman's method achieve an average return rate of around 73%, with many surveys reaching an 85%–90% response rate. In Canada and Europe, around 79% of personal interviews are completed, and the response rate for mailed questionnaires is around 75% (Dillman 1978, 1983). Of course, those numbers are for the usual kinds of surveys about consumer behaviors, political attitudes, and so on. What happens when you ask people really threatening questions? In the Netherlands, Nederhof (1985) conducted a mail survey on attitudes toward suicide and achieved a 65% response rate. Pretty impressive.

The average response rate for face-to-face interviews in the United States was between 80% and 85% during the 1960s, but fell to less than 70% in the early 1970s (American Statistical Association 1974). It has apparently recovered somewhat, as more is learned about how to maximize cooperation by potential respondents.

Willimack et al. (1995), for example, report a refusal rate of 28% for face-to-face interviews in the annual Detroit Area Study conducted by the University of Michigan. They also report that giving people a gift-type ballpoint pen (a small, nonmonetary incentive) *before* starting a face-to-face interview lowered refusal rates and increased the completeness of responses to questions. The bottom line is that, with everything we've learned over the years, the gap between the response rate to personal interviews and mailed questionnaires is now insignificant, but the rate of refusals is still high in all forms of surveys.

Many scholars are concerned with the response rate problem (Roth and BeVier 1998; Turley 1999; Synodinos and Yamada 2000). We do not, however, have good data about the extent to which refusals make a difference in

the *outcome* of survey research. Some scholars, in fact, challenge the conventional wisdom that a low response rate vitiates the usefulness of a survey (for a review, see Krosnick 1999).

This does not in any way reduce the value of personal interviews, especially for anyone working in developing nations. It does mean, however, that if you are conducting *mailed* survey research in the United States, Canada, Western Europe, Australia, New Zealand, or Japan, you should use Dillman's method.

Steps in Dillman's Method

1. **Professionalism**: Mailed questionnaires must look thoroughly professional. Jaded, hard-bitten, oversurveyed people simply don't respond to amateurish work. Fortunately, with today's word processing, making attractive questionnaire booklets is easy. Use standard-size paper: 8.5 × 11 in the United States and slightly longer A4 paper in the rest of the world. Several researchers have found that light green paper produces a higher response rate than white paper (Fox et al. 1988). I wouldn't use any other colors until controlled tests are made.

You must be thinking: "Controlled tests of *paper color*?" Absolutely. It's because social scientists have done their homework on these little things that a response rate of over 70% is now routine—provided you're willing to spend the time and money it takes to look after all the little things. Read on and you'll see how small-but-important those "little things" are.

2. **Front and back covers**: Don't put any questions on either the front or back covers of the booklet. The front cover should contain a title that provokes the respondent's interest and some kind of eye-catching graphic design. By "provoking interest," I don't mean "threatening." A title like "The Greenville Air Quality Survey" is fine. "Polluted Air Is Killing Us" isn't.

Graphic designs are better than photographs on survey covers. Photos contain an enormous amount of information, and you never know how respondents will interpret the information. If a respondent thinks a photo contains an editorial message (in favor of or against some pet political position), then the survey booklet goes straight into the trash.

The front cover should also have the name and return address of the organization that's conducting the survey.

The back cover should contain a *brief* note thanking the respondent and inviting open-ended comments about the questionnaire. Nothing else.

3. **Question order**: Pay careful attention to question order. Be sure that the first question is directly related to the topic of the study (as determined from the title on the front of the booklet), that it is interesting and easy to answer, and that it is nonthreatening. Once someone starts a questionnaire or an interview, they are very likely to finish it. Introduce threatening questions well into the instrument, but don't cluster them all together.

Put general socioeconomic and demographic questions at the end of a questionnaire. These seemingly innocuous questions are threatening to many respondents who fear being identified (Sudman and Bradburn 1982). Once someone has filled out a questionnaire, they are unlikely to balk at stating their age, income, religion, occupation, etc.

4. **Formatting**: Construct the pages of the questionnaire according to standard conventions. Use uppercase letters for instructions to respondents and mixed upper and lower case for the questions themselves. Never allow a question to break at the end of a page and continue on another page. Mailed surveys have to look good and be easily readable or they get tossed out.

Use plenty of paper; don't make the instrument appear cramped. Line answers up vertically rather than horizontally, if possible. This, for example, is not so good:

Strongly Approve Approve Neutral Disapprove Strongly Approve

This is better:

Strongly Approve
Approve
Neutral
Disapprove
Strongly Disapprove

It pays to spend time on the physical format of a questionnaire. The general appearance, the number of pages, the type of introduction, and the amount of white (or green) space—all can affect how people respond, or whether they respond at all (Akkerboom and Schmeets 1998). Once you've gone to the expense of printing up hundreds of survey instruments, you're pretty much stuck with what you've got.

Use lots of open space in building schedules for personal interviews, too. Artificially short, crowded instruments only result in interviewers missing items and possibly in annoying respondents (imagine yourself sitting for 15 minutes in an interview before the interviewer flips the first page of an interview schedule.)

5. **Length**: Keep mailed questionnaires down to 10 pages, with no more than 125 questions. Beyond that, response rates drop (Dillman 1978). Herzog and Bachman (1981) recommend splitting questionnaires in half and alternating the order of presentation of the halves to different respondents to test for response effects of questionnaire length.

It is tempting to save printing and mailing costs and to try to get more questions into a few pages by reducing the amount of white space in a self-administered questionnaire. Don't do it. Respondents are never fooled into thinking that a thin-but-crowded questionnaire is anything other than what it seems to be: a long questionnaire that has been forced into fewer pages and is going to be hard to work through.

6. **The cover letter**: A one-page cover letter should explain, in the briefest possible terms, the nature of the study, how the respondent was selected, who should fill out the questionnaire (the respondent or the members of the household), who is funding the survey, and why it is important for the respondent to send back the questionnaire. ("Your response to this questionnaire is very important. We need your response because . . .")

The one thing that increases response rate more than any other is university sponsorship (Fox et al. 1988). University sponsorship, though, is not enough. If you want a response rate that is not subject to bias, be sure to address the cover letter directly and personally to the respondent—no "Dear Respondent" allowed—and sign it using a blue ballpoint pen. Ballpoints make an indentation that respondents can see—yes, some people do hold those letters up to the light to check. This marks the letter as having been individually signed.

The cover letter must guarantee confidentiality and must explain the presence of an identification number (if there is one) on the questionnaire. Some survey topics are so sensitive that respondents will balk at seeing an identification number on the questionnaire, even if you guarantee anonymity. In this case, Fowler (1984) recommends eliminating the identification number (thus making the questionnaire truly anonymous) and telling the respondents that they simply cannot be identified.

Enclose a printed postcard with the respondent's name on it and ask the respondent to mail back the postcard *separately* from the questionnaire. Explain that this will notify you that the respondent has sent in the questionnaire so that you won't have to send the respondent any reminders later on. Fowler found that people hardly ever send back the postcard without also sending back the questionnaire.

7. **Packaging**: Package the questionnaire, cover letter, and reply envelope and postcard in another envelope for mailing to the respondent. Type the respondent's name and address on the mailing envelope. Avoid mailing labels. Use first-class postage on the mailing envelope and on the reply envelope. Some people respond better to real stamps than to metered—even first-class metered—postage. Hansley (1974) found that using bright commemorative stamps increased response rate.

8. **Inducements**: What about sending people money as an inducement to complete a survey? Mizes et al. (1984) found that offering respondents $1 to complete and return a questionnaire resulted in significantly increased returns; but offering respondents $5 did not produce a sufficiently greater return to warrant using this tactic. In 1984, $5 was close to the value of many respondents' time for filling out a questionnaire. This makes responding to a survey more like a strictly economic exchange and, as Dillman pointed out, makes it easier for people to turn it down (1978:16).

Inflation will surely have taken its toll by this time (sending people $1 in the mail to answer a survey can't possibly buy as much response today as it did in the 1980s), but the point is clear: There is a Goldilocks solution to the problem of how much money to send people as an incentive to fill out and return a survey. If you send people too much money or too little, they throw the survey away. If you send them just the right amount, they are likely to fill out the survey and return it.

Warriner et al. (1996) confirmed this finding. They offered people in Ontario, Canada, $2, $5, or $10 to send back a mailed survey. Factoring in all the costs of follow-up letters, the $5 incentive produced the most returns for the money.

First-class postage and monetary incentives may seem expensive, but they are cost-effective because they increase the response rate. Whenever you think about cutting corners in a survey, remember that all your work in designing a representative sample goes for nothing if your response rate is low. Random samples cease to be representative unless the people in them respond. Also remember that small monetary incentives may be insulting to some people. This is a cultural and socioeconomic class variable that only you can evaluate in your specific research situation. (See Church 1993 for a meta-analysis of the effect of inducements on response rates to mailed surveys.)

9. **Contact and follow-up**: Pay careful attention to contact procedures. Send a letter to each respondent explaining the survey and informing the respondent that a questionnaire will be coming along soon. Send a postcard reminder out to all potential respondents a week after sending out the questionnaire. Don't wait until the response rate drops before sending out reminders. Some people hold onto a questionnaire for a while before deciding to fill it out or throw it away. A reminder after one week stimulates response among this segment of respondents.

Send a second cover letter and questionnaire to everyone who has not responded two weeks later. Finally, four weeks later, send another cover letter *and questionnaire*, along with an additional note explaining that you have not yet received the respondent's questionnaire, and stating how important it is that the respondent participate in the study. Heberlein and Baumgartner (1978, 1981) found that sending a second copy of the questionnaire increases response rate 1%–9%. Since there does not appear to be any way to predict whether the increase will be 1% or 9%, the best bet is to send the extra questionnaire.

When you send out the second copy of the questionnaire, send the packet by certified mail. House et al. (1977) showed that certified mail made a big difference in return rate for the second follow-up.

Does All This Really Make a Difference?

Thurman et al. (1993) were interested in the attitudes and self-reported behaviors of people who admit to drunk driving. Using Dillman's Total Design method, they sent out questionnaires to a national sample of 1,310 and got back 765, or 58%. Not bad for a first pass.

Unfortunately, for lack of time and money, Thurman et al. couldn't follow through with all the extra mailings. Of the 765 respondents, 237 said they were nondrinkers. This left 525 eligible questionnaires for analysis.

Of the 525 respondents who said they were consumers of alcohol, 133 admitted driving while drunk in the past year. Those 133 respondents provided data of intrinsic interest, but the 765 people who responded from the nationally representative sample of 1,310 may have been a biased sample on which to base any generalizations. I say "may be" a biased sample because there is no way to tell. And that's the problem.

The bottom line: The last interview you get in any survey—whether you're sending out questionnaires, doing a phone survey, or contacting respondents for face-to-face interviews—is always the most costly and it's almost always worth it. If you really care about representative data, you won't think of all the chasing around you have to do for the last interviews in a set as a nuisance but as a necessary expense of data collection. And you'll prepare for it in advance by establishing a realistic budget of both time and money.

Pretesting and Learning from Mistakes

There is no way to emphasize sufficiently the importance of **pretesting** any survey instrument you prepare. No matter how much you do to prepare a culturally appropriate questionnaire, it is absolutely guaranteed that you will have forgotten something important or that you will have poorly worded one or more vital element. These glitches can be identified only by pretesting.

If you are building a self-administered questionnaire, bring in at least 6–10 pretest respondents and sit with them as they fill out the entire instrument. Encourage them to ask questions about each item. Your pretest respondents will make you painfully aware of just how much you took for granted, no matter how much ethnographic research you did or how many focus groups you ran before making up a questionnaire.

For face-to-face interviews, do your pretesting under the conditions you will experience when the survey is under way for real. If respondents are going to come to your office, then pretest the instrument in your office. If you are going to respondents' homes, then go to their homes for the pretest.

Never use any of the respondents in a pretest for the main survey. If you are working in a small community, where each respondent is precious (and you don't want to use up any of them on a pretest), take the survey instrument to another community and pretest it there. This will also prevent the pretest respondents in a small community from gossiping about the survey before it actually gets under way. A "small community," by the way, can be "the 27 students from Taiwan at your university" or all the residents of an Indonesian peasant village.

Use all your interviewers in any pretest of a face-to-face interview schedule and be sure to do some of the pretesting yourself. After the interviewers have done the pretests, bring them together for a discussion on how to improve the survey instrument. As you conduct the actual survey, ask respondents to tell you what they think of the study and of the interview they've just been through.

At the end of the study, bring all the interviewers back together for an evaluation of the project. If it is wise to learn from your mistakes, then the first thing you've got to do is find out what the mistakes are. If you give them a chance, your respondents and interviewers will tell you.

Cross-Sectional and Longitudinal Surveys

Most surveys are **cross-sectional**. The idea is to measure some variables at a single time. Of course, people's attitudes and reported behaviors change over time, and you never know if a single sample is truly representative of the population. Many surveys are conducted again and again to monitor changes and to ensure against picking a bad sample. Multiple cross-sectional polls use **longitudinal design**. The daily—even hourly—tracking polls in U.S.

presidential elections are an extreme example, but in many industrialized countries some questions have been asked of representative samples for many years.

The Gallup Organization, for example, has been asking Americans to list "the most important problem facing this country today" for about 60 years. The data track the concerns of Americans about unemployment, the quality of education, drugs, street crime, the federal deficit, taxes, health care costs, poverty, racism, AIDS, abortion. . . . In March 1992, 67% of Gallup's sample said that unemployment was the most important problem in America. In June 2000, just 3% gave that answer. Of course, in 1992, the United States was in a recession, while in 2000 the country was experiencing a booming economy.

No surprises there. Still, these data are important because they were collected with the same instrument. People were asked the same question again and again over the years. After several generations of effort, longitudinal survey data have become a treasured resource in the highly industrialized nations.

Panel Studies

Multiple cross-sectional surveys have their own problems. If the results from two successive samples are very different, you don't know if it's because people's attitudes or reported behaviors have changed, or the two samples are very different, or both. To deal with this problem, survey researchers may use the powerful **panel design**. In a panel study, you interview the exact same people over again. Panel studies are like true experiments: Randomly selected participants are tracked for their exposure or lack of exposure to a series of interventions in the real world.

The Panel Study on Income Dynamics, for example, has tracked thousands of American families since 1969. Some families have been tracked for 30 years, and new families are added as others drop out of the panel (families are interviewed every other year). Among many other things, the data track the effect of living in different kinds of neighborhoods on the social and economic mobility of men and women, of whites and African Americans, of old people and young.

This panel study is run by the Institute for Social Research at the University of Michigan, which also makes the data available to the public at http://www.isr.umich.edu/src/psid/. Hundreds of papers have been written from the data of this study (for a good recent example, see South and Crowder 2000).

Perhaps the most famous panel study of all time is the Framingham Heart Study. In 1948, medical researchers began tracking 5,209 men and women between 30 and 62 years old from one small town—Framingham, Massachusetts. In 1971, as the original panel began to die off, another 5,124 panelists were added—this time, the original panelists' adult children and their spouses. Every two years, all the panelists go in for a complete medical checkup. This study has identified and nailed down the major risk factors for heart disease, which include, of course, behaviors (exercise, smoking) and inner states (attitudes, stress)—things that anthropologists, as well as epidemiologists, are interested in. Basic information about the Framingham study is available from the National Heart, Lung, and Blood Institute at http://www.nhlbi.nih.gov/about/framingham/index.html.

Respondent Mortality

Panel studies may suffer from what's called the **respondent mortality** problem. This is where people drop out between successive waves of the panel survey. If this happens, and the results of successive waves are very different, you can't tell if that's because of (1) the special character of the dropout population, (2) real changes in the variables you're studying, or (3) both. For example, if dropouts tend to be male or poor, your results in successive waves will overrepresent the experiences of those who are female or affluent.

Respondent mortality is not always a problem. Roger Trent and I studied riders on the Morgantown, West Virginia, "People Mover," an automated transport system that was meant to be a kind of horizontal elevator. You get on a little railway car (they carry only 8 seated and 12 standing passengers), push a button, and the car takes you to your stop—a block away or eight miles across town. The system was brought on line a piece at a time between 1975 and 1980. Trent and I were tracking public support as the system went more places and became more useful (Trent and Bernard 1985). We established a panel of 216 potential users of the system when the system opened in 1975 and reinterviewed the members of that panel in 1976 and 1980 as more and more pieces of the system were added.

All 216 original members of the panel were available during the second wave and 189 were available for the third wave of the survey. Note, though, that people who were unavailable had moved out of Morgantown and were no longer potential users of the system. What counted in this case was maintaining a panel large enough to represent the attitudes of people in Morgantown about the People Mover system. The respondents who stayed in the panel still represented the people whose experiences we hoped to learn about.

Some Specialized Survey Techniques

Factorial Surveys

In a **factorial survey** (see Rossi and Nock 1982), respondents are presented with **vignettes** that describe hypothetical social situations and are asked for

their judgments about those situations. The General Social Survey is an annual survey of about 1,600 adults in the United States (for more on the GSS, go to http://www.icpsr.umich.edu/GSS99/index.html). Here's a vignette that was in the 1992 GSS:

> This family has four children, the youngest is 6 months old, living with their mother. The mother is divorced. The mother has a college degree and is unemployed and not looking for work because she has no ready means of transportation. The father has remarried and is permanently disabled. The family is likely to face financial difficulties for a couple of years. Her parents cannot afford to help out financially. The family has $1,000 in savings. All in all, the family's total income from other sources is $100 per week.
>
> What should this family's weekly income be? Include both the money already available from sources other than the government, and any public assistance support you think this family should get.

Amount already received by this family	*Average U.S. family income*
X	X

0 50 100 150 200 250 300 350 400 450 500 550 600

SOURCE: "The Dimensions of Poverty: Public Perceptions of the Deserving Poor" by J. A. Will, 1993, *Social Science Research*, p. 322. Reprinted with permission of Academic Press.

There are 10 variables in this vignette (number of children, marital status of the mother, how much savings the family has, the total income of the family, etc.) with 1,036,800 possible combinations. That seems just about right to me. The calculus for any individual's opinion about how much money to award the deserving poor on welfare is really that complicated. Now, each of the 1,600 respondents in the GSS saw seven vignettes each, so the survey captured

$$(1{,}600 \text{ people})\ (7 \text{ vignettes})\ (10 \text{ variables}) = 112{,}000$$

combinations, or a sample of about 11% of all the factors that probably go into people's opinions on this issue.

The results of that survey were very interesting. Respondents awarded people who were looking for work a lot more than they awarded people who weren't looking for work. But mothers only got an extra $6 per week for seeking work, while fathers got more than $12. And if mothers were unemployed because they wouldn't take minimum-wage jobs, they had their allotments

reduced by $20 per week, on average, compared to what people were willing to give mothers who were working full time (Will 1993). (For comparison, it took about $150 in 2000 to buy what $100 bought in 1986.)

The factorial survey combines the validity of randomized experiments with the reliability of survey research and lets you measure subtle differences in opinion.

Time Budgets and Diaries

Time budget surveys have been done all over the world to track how ordinary human beings spend most of their days (Szalai 1972). The idea is to learn about the sequence, duration, and frequency of behaviors and about the contexts in which behaviors take place. Some researchers ask respondents to keep diaries; others conduct "yesterday interviews," in which respondents are asked to go over the last 24 hours and talk about everything they did. Some researchers combine these methods, collecting diaries from respondents and then following up with a personal interview.

Perhaps the earliest time budget in anthropology was done during two and a half years of fieldwork between 1930 and 1934 by Audrey Richards among the Bemba of Zimbabwe (Northern Rhodesia back then). Richards went across the country, spending from three to six weeks each in a series of villages. She pitched her tent in the middle of each village so she could watch people's activities, and in two of those villages, Kasaka and Kapamba, she kept daily calendars of all the adults, both men and women (Richards 1939:10–11). Richards asked several informants to estimate how long they took to accomplish various tasks—planting a garden, chasing locusts, cutting trees. Then she averaged what people told her and got a rough estimate for each task. Later she applied those figures to her observations of people in the two villages where she kept records of the daily activities (ibid.:395 and Appendix E).

Susan Shaw's study of family activities in Canada (1992) is typical of the use of time budgets in modern societies. Shaw studied 46 middle- and working-class couples who had children living at home. All the fathers were employed full time; among the mothers, 12 were employed full time, 9 were employed part time, and 25 were full-time homemakers. Both parents kept time diaries for one day during the week and for one day on a weekend. Then, Shaw interviewed the parents separately, for one to two hours in their homes. For each activity that they had mentioned, parents were asked if they considered the activity to be work or leisure, and why.

Shaw calculated the amount of time that each parent reported spending with their children—playing with them, reading to them, and so on. The rather dramatic results are in Table 10.1. For these Canadian families, at least, the more

TABLE 10.1
Average Amount of Time Fathers and Mothers Report Spending with Children, by the Mother's Employment Status

		Time with children per day (in minutes)	
Mother's employment status	*N*	*Mothers*	*Fathers*
Employed full time	12	97	71
Employed part time	9	144	52
Full-time homemaker	25	241	23
Total	46		

SOURCE: "Dereifying Family Leisure: An Examination of Women's and Men's Everyday Experiences and Perceptions of Family Time" by S. Shaw, 1992, *Leisure Studies*, p. 279. Reproduced by permission of Taylor and Francis.

the women worked outside the home, the more time the fathers spent with their children.

Diaries and time-budget interviews, particularly with the aid of checklists, appear to be more accurate than 24-hour recall of activities. But no matter what you call them, time budgets and diaries are still methods for collecting self-reports of behavior. They may be *less* inaccurate than simply asking people to tell you what they did over the past day or week, but they are not perfect. A lot of work remains to be done on testing the accuracy of activity diaries against data from direct observation. In Chapter 15, we'll look at methods for direct observation and measurement of behavior.

Randomized Response

Randomized response is a technique for estimating the amount of some socially negative behavior in a population—things like shoplifting, extramarital sex, child abuse, being hospitalized for emotional problems, and so on. The technique was introduced by Stanley Warner in 1965 and is particularly well described in Williams 1978:73. It is a simple, fun, and interesting tool. Here's how it works.

First, you formulate two questions, *A* and *B*, that can be answered "yes" or "no." One question, *A*, is the question of interest (perhaps, "Have you ever shoplifted?"). The possible answers to this question (either "yes" or "no") do not have known probabilities of occurring. That is what you want to find out.

The other question, *B*, must be innocuous and the possible answers (again "yes" or "no") must have known probabilities of occurring. For example, if you ask a someone to toss a fair coin and ask, "Did you toss a heads?" then

the probability that they answer "yes" is 50%. If the chances of being born in any given month were equal, then you could ask respondents, "Were you born in April, May, or June?" and the probability of getting a "yes" would be 25%. Unfortunately, births are seasonal, so the coin-toss question is preferable.

Let's assume you use the coin toss for question *B*. You ask someone to toss the coin and to note the result *without letting you see it*. Next, you have them pick a card, from a deck of 10 cards in which each card is marked with a single integer from 1 to 10. The respondent *does not tell you what number he or she picked*, either. The secrecy associated with this procedure makes people feel secure about answering question *A* (the sensitive question) truthfully.

Next, hand the respondent a card with the two questions, marked *A* and *B*, written out. Tell them that if they picked a number between one and four from the deck of 10 cards, they should answer question *A*. If they picked a number between 5 and 10, they should answer question *B*.

That's all there is to it. You now have the following: (1) each respondent knows they answered "yes" or "no" and which question they answered; (2) you know *only* that a respondent said "yes" or "no" but not which question, *A* or *B*, was being answered.

If you run through this process with a sufficiently large, representative sample of a population, and if people cooperate and answer all questions truthfully, then you can calculate the percentage of the population that answered yes to question *A*. Here's the formula:

$$P_{A \text{ or } B} = [(P_{A^+} \times P_A) + (P_{B^+} \times P_B)] \quad \textbf{Formula 10.2}$$

The percentage of people who answer "yes" to *either A* or *B* equals (the percentage of people who answer "yes" to question *A*) times (the percentage of times that question *A* is asked) plus (the percentage of people who answered "yes" to question *B*) times (the percentage of times question *B* is asked).

The only unknown in this equation is the percentage of people who answered "yes" to question *A*, the sensitive question. We know, from our data, the percentage of "yes" answers to *either* question. Suppose that 33% of all respondents said "yes" to *something*. Since respondents answered question *A* only if they chose a number from 1 to 4, then *A* was answered 40% of the time and *B* was answered 60% of the time. Whenever *B* was answered, there was a 50% chance of it being answered "yes" because that's the chance of getting a heads on the toss of a fair coin. The problem now reads:

$$.33 = A\ (.40) + .50\ (.60)$$

or

$$.33 = .40\,A + .30$$

which means that $A = .075$. That is, given the parameters specified in this experiment, if 33% of the sample says "yes" to either question, then about 8% of the sample answered "yes" to question *A*.

There are two problems associated with this technique. First, no matter what you say or do, some people will not believe that you can't identify them and will therefore not tell the truth. Bradburn et al. (1979) report that 35% of known offenders would not admit to having been convicted of drunken driving in a randomized response survey. Second, like all survey techniques, randomized response depends on large, representative samples. Since the technique is time-consuming to administer, this makes getting large, representative samples difficult.

Still, the evidence is mounting that for some sensitive questions—Did you smoke dope in the last week? Have you ever bought a term paper? Have you stolen anything from your employer?—when you want the truth, the randomized response method is worth the effort (see Scheers and Dayton 1987, Nordlund et al. 1994, and Clark and Desharnais 1998 for some examples).

Every time I read in the newspaper that self-reported drug use among adolescents has dropped by such-and-such an amount since whenever-the-last-self-report-survey-was-done, I think about how easy it is for those data to be utter nonsense. And I wonder why the randomized response technique isn't more widely used.

Dietary Recall

Studies of diet and human nutrition mostly rely on informants to recall what they've eaten over the last 24 hours or what they usually eat for various meals. These **24-hour recall** tests, or **dietary recall** interviews, are a specialized kind of structured interview. The problem is, they produced dreadfully inaccurate results.

C. J. Smith et al. (1996) compared the responses of 575 Pima and Papago Indians in Arizona to a 24-hour recall instrument about food intake with responses to a very detailed survey called the Quantitative Food Frequency Questionnaire. In the QFF, interviewers probe for a list of regularly consumed foods in a community. Smith et al. also assessed the energy expenditure of 21 people in the research group using the Doubly Labeled Water technique. The DLW technique involves giving people special water to drink—water with isotopes that can be tracked in blood and urine samples—and then testing, over time, their actual intake of nutrients.

The correlation, across the 21 participants, between the energy intake measured by the DLW technique and the energy intake estimated by the informants' responses to the QFF, was 0.48. This correlation is statistically significant, but it means that just 23% (0.48^2) of the variation in actual energy intake across the 21 people was accounted for by their responses to a very detailed interview about their food consumption. The correlation of actual energy intake with estimates from the 24-hour recall data were much worse. (For an overview of methods in nutritional anthropology, see Pelto et al. 1989. For some recent research on measuring food intake, see Chapman et al. 1997, Schoenberg 1997, and Melnik et al. 1998.)

Johnson et al. (1996) also found no useful relation between individual 24-hour recall measurements of energy intake among children in Vermont and measurements of those same children by the DLW technique. But, in the all-is-not-lost department, Johnson et al. found that averaging the data for energy intake across *three* 24-hour recalls in 14 days (on day 1, day 8, and day 14) produced results that were very similar to those produced by the DLW technique.