

## Chapter 18

---

---

# Research Procedures

In this chapter we'll discuss the major methods for measuring variables and collecting data to test hypotheses. We will present a brief introduction to the mechanics of these major procedures and we will also discuss them in terms of the advantages and disadvantages they offer to the researcher. There are any number of books devoted to presenting the intricacies of interviewing, questionnaire construction, unobtrusive measurement, etc. The serious communication researcher should refer to these sources to get more complete information about techniques, advantages, and pitfalls of the particular research procedures that are required by her research. Furthermore, as the measurement of communication data in many cases involves acquiring the cooperation of respondents in varying degrees, we will discuss the ethical implications of selecting methods of gathering data.

## Measurement Techniques

### *Behavioral Observation*

Much of the measurement of communication variables falls in the category of behavioral observation. With this technique, a trained person observes a specific set of behaviors exhibited by a research subject and uses trained judgment to assign the correct values to the appropriate variables.

For instance, in the Bandura et al. experiment described in Chapter 14, if the observer sees a child from the experimental group carry out an act of physical or verbal aggression identical to that originally exhibited by the adult model, the child's act is counted in the category "Imitative Aggression", and not in the category called "Partially Imitative Aggression". In the Leavitt study described

in the same chapter, the observer starts a clock at the beginning of a trial, and stops the clock once all group members have thrown their switches. The number of seconds needed for completion is the value assigned to the variable "Time Elapsed" for that group's trial.

In both studies the observer determines the level of the variable. Because a human observer is involved, there can be large differences in the reliability of measurement. Reliability often varies according to the complexity or

difficulty of the judgment required by the measurement task. The instructions issued to the observer in the Leavitt study might have read something like this:

*"Start the clock and give the signal to start the trial. Stop the clock when all five lights have been illuminated. Read the clock's elapsed time and enter the number on line 5 of the coding sheet. Reset the clock."*

This kind of behavioral observation is likely to yield highly reliable measurement. First of all, the same observer should obtain very similar results if he repeats the same task repeatedly under the same conditions. This is test-retest reliability. Second, given the simplicity and straightforwardness of the instructions, several different observers of the same trial should show very high levels of agreement on the amount of elapsed time. Such agreement is referred to as intercoder, interexperimenter or interjudge reliability.

The greater ambiguity associated with the decision making process in the Bandura study will probably result in lower levels of test-retest and intercoder reliability. For instance, the observer must decide to what extent a child must deviate from the model's behavior before the child's behavior is assigned to the category "Partially Imitative Aggression" rather than to the "Imitative Aggression" category. If different observers evaluate the same behavior, there is likely to be some disagreement. If two different measuring instruments (the observers) give two different "readings" about the phenomenon being measured, there is measurement unreliability.

Reliability in behavioral observation can be improved in two ways: (1) make the observation task simple and concrete; and (2) give clear, extensive instructions and training to the observers. If you ask an observer to "Count the number of nice remarks made by each person in the conversation," you are going to get some widely different answers from different observers who hear the same taped conversation, and even from the same observer who codes the same conversation at a two-week interval. The instructions are too vague to produce reliable measurement.

On the other hand, if you ask the observer to "Count the number of times each person commented favorably on the clothing of the other," and "Count the number of times each person said 'thank you' to the other," you will probably get reliable results. By combining a number of such concrete observational tasks, you can measure "niceness" with much more detail and reliability than the first vague question could ever give.

## Obtrusive and Unobtrusive Measurement

Behavioral observation can be either obtrusive or unobtrusive measurement. This distinction refers to the extent to which the respondent or subject is aware that he or she is being evaluated. As we saw in Chapters 13 and 17, this awareness can affect both the internal validity and external validity of a study. Awareness can produce sensitization to the experimental manipulation, enhanced memory effects, reactivity to the research setting, and a host of other artificial effects which will obscure true relationships.

It is almost always the goal of a communication researcher to make observation as unobtrusive as possible. This can be done with careful design of the research setting or by choosing a measurement method that is inherently unobtrusive.

**Reducing Obtrusiveness.** Research settings can often be constructed so that the observer is inconspicuous or completely hidden. For example, in the Bandura study the children were observed through a one-way mirror which prevents the observed person from seeing the observer. The children may not have been aware of the purpose of a one-way mirror, but for older research participants the presence of a one-way mirror will be a dead give-away that they are being observed. This realization may affect behavior in unknown ways. But even if they realize that they are being observed from behind a mirror, there is a tendency to ignore the observer after a time, because there are no movements or noises from the observer to remind the subject that she is being observed.

If the subject suspects that he is being surreptitiously observed, he may actually react more

strongly than if he is told that someone is behind the one-way mirror. The presence of a passive mirror or a small video camera in a discreet box are easily ignored after the novelty wears off, so it is often better to inform a subject that they are being observed that it is to allow them to have unconfirmed suspicions. Even if it is impossible to completely hide the observer, the obtrusive effect can be reduced by placing the observer in an out-of-the-way corner of the room and instructing him to remain as motionless and quiet as possible, to avoid rustling the coding sheets, etc.

There is a privacy issue involved with unobtrusive measurement. It boils down to this: should the research subject be informed that he is being observed, even if that information may affect his behavior during the research? We'll have more to say about this in the final section of this chapter when we discuss various ethical dilemmas in communication research.

**Naturally Unobtrusive Measurement.** Some types of observational measurement are inherently unobtrusive. This data is collected with little or no awareness by the sources of the data that communication research is being conducted.

For example, state and federal governments routinely collect social data. The U.S. Government collects census data which describes the population in regions as large as the whole country and as small as a few square miles. This data is often very useful for comparison with sample demographic data, and for weighting or correcting sample data to make it more representative of the population values. Census data can also be used directly when the researcher uses a social system unit of analysis like the neighborhood or region. For example, the mean income or number of telephones in urban census tracts could be useful variables for a telecommunications researcher. The U.S. Commerce department also collects detailed data about business organizations that can be used for similar aggregate analysis purposes. Governmental data is available at many public libraries and at most university libraries.

There are also archives of individual-level data that are very useful as sources of unobtrusive communication data. Public opinion poll and marketing research data are available from archives like the Institute for Social Research at the University of Michigan and the Roper Center for Public Opinion Research at the University of Connecticut. These archives contain the data and summary results from tens of thousands of questions asked of millions of people. They are particularly useful for overtime research. Similar or identical poll questions taken from surveys conducted over a span of years provides the communication researcher with a very valuable and inexpensive source of over-time data.

For the mass communication researcher, these archives are particularly useful when their information is combined with data from media archives which collect and preserve newspaper and magazine stories, television newscasts, television and radio commercials, and other media messages. Most large libraries carry the New York Times Index which can be used to summarize the frequency that newspaper stories about selected issues or topics appear. The Vanderbilt Television Archives publish an index of network television story coverage and can provide videotapes of stories about selected topics. The researcher can use a media archive to provide the material for a content analysis (described in more detail later in this chapter). The data from the content analysis, combined with data from a public opinion archive, can be used to track the relationship between media messages and aggregate audience response.

Archives of original documents like letters and manuscripts can also be a source of unobtrusive data to the researcher interested in analyzing messages. For example, the organizational researcher might gain access to electronic mail messages in a sample of corporations, and use this information to study communication patterns within different types of organizations. She might also collect all the interoffice mail envelopes and code the sender's and recipient's departments to unobtrusively measure interdepartmental communication. This kind of measurement produces no demand characteristics and no sensitization of research subjects.

Reusing the data collected by other researchers (secondary analysis) is often a very efficient way to collect information. This measurement may or may not be considered obtrusive. For example, an interpersonal communication researcher might be able to gain access to interviews and transcripts at a research center for family communication and therapy. Since the research subjects were probably aware that their responses were being recorded, the data will be subject to some sensitization and social demand contamination. But if the subject of the interviews was, for example, establishment of rules for adolescents, and the communication researcher is interested in the dominance of the conversation by the mother or father, he can consider his dominance measurement as being unobtrusive.

There are many, many other sources of data for secondary analysis. Commercial research organizations often maintain databases that can be made available to academic researchers after their business value has disappeared. Electronic media ratings organizations like Arbitron and A.C. Nielsen are an example.

## Self-Report Measurement

In self-report measurement, an individual respondent directly provides the value of the measured variable. There are no observers involved. This is a very different situation from behavioral observation. Self-report measurement is almost always obtrusive and subject to social demand characteristics.

In the Leavitt study in Chapter 14, data on satisfaction with group membership and individual perceptions of leadership in their group was collected directly from the research subjects. Group members were asked to respond to the question “How did you like your job in the group” by placing a check on a rating scale that was marked “disliked it” on one end and “liked it” on the other. In this case, the individual respondent ultimately controlled the value that was assigned to his liking of his role. This type of measurement frequently occurs in communication research when subjects are asked to fill out scales (in a paper-and-pencil test) to measure their perceptions about topics or toward others who are involved in the communication process.

Self-report measurement, like behavioral observation, is subject to unreliability. Since the research subjects cannot be extensively trained, the researcher must insure that the instructions for filling out the scales or questionnaire items are clear and unambiguous (we’ll discuss this in more detail later, in the section on questionnaire design). The measurement task must be as simple as possible, too, and the measurement items should be as specific as possible. You’ll get better measurement if you ask a subject a series of questions about specific behaviors, like “About how many times have coworkers asked your advice about job-related problems in the past week” and “How many times in the past year have you been appointed to head work group committees by your employer” than if you just ask the global question “How good a leader are you”.

Often self-report measurements require that the research subject remember some past information or behaviors. This was the case in the example in the previous paragraph. This kind of retrospective measurement is difficult to achieve because it relies on fallible human memory. As a quick example of this, consider your ability to accurately report an activity that most of you have probably spent a lot of time doing recently: Try to write down the names and the plots of all the television programs you’ve seen in the past week. Not too easy, is it?

One way to improve self-report in retrospective measurement is to restrict the time range that the respondent must consider. It may be better to ask “What TV programs did you watch last night” than to ask “What TV programs have you watched this Fall?” There will be more random variation in the first question, since some respondents who are normally heavy viewers will have watched no TV because they were visiting their chatty great-aunt, while other light viewers might have spent the evening as couch potatoes because they had a hard day at the office. But these uncharacteristic responses will cancel out over the entire sample, and the more specific question should lead to better recall of the actual programs viewed.

It is also easier to remember specific things than general things. Asking a respondent to estimate the number of times in the past six months that she has had a disagreement with her spouse about money that resulted in both of them shouting is better than just asking how many arguments she has had with her spouse. Once again, many specific items (like arguments about money, child care, sex) can be added together to give measurement of a general class of items (like arguments).

## Surveys and Interviews

Interviews and surveys (which may or may not use an interviewer) are very commonly applied tools in communication research. These are obtrusive measurement techniques that can produce either self-report or behavioral observation data.

The term survey is usually applied to observational research which uses a questionnaire as the primary measurement instrument. The unit of analysis is almost always the individual. Surveys are frequently done at a single time point, but panel or equivalent samples designs are also used to conduct over-time research.

Surveys can be generally classified as in-person, mail, telephone, mall intercept, or computer-administered. We'll discuss the characteristics of each briefly. Some of the differences among the types of surveys are summarized in Table 18-1.

Table 18-1 Characteristics of Different Types of Surveys

Survey Type	Obtaining Sample	Cooperation Rate	Cost per Respondent	Strength	Weakness
<i>In-Person</i>	Difficult	Medium	High	Visual & manipulative measurement, interviewer rapport	Respondent fear, expensive
<i>Telephone</i>	Easy	High	Medium	Fast, plus in-person advantages	Limited time for interview
<i>Mall Intercept</i>	Easy	Medium	Medium	Fast, plus in-person advantages	Unrepresentative sample
<i>Mail</i>	Easy	Low	Low	Large sample is inexpensive	Unrepresentative sample, respondent errors
<i>Computer-Administered</i>	Difficult	High	Medium	Automatic data entry	Respondent must have computer

## In-Person Survey

In this kind of survey, the questionnaire questions are presented to the respondents by an interviewer who is speaking to them face-to-face. The interviewer records the answers.

Having an interviewer speaking to the respondent has some real advantages. A good interviewer will make a personal connection with the respondent that will increase the motivation of the respondent to answer the questions fully and truthfully. It is more difficult to refuse to cooperate with a real, live person standing in front of you than to say "no" to a voice on the telephone or a letter in the mail. It is also harder to quit cooperating in the middle of the questionnaire when the interviewer is sitting a few feet away, expectantly awaiting an answer.

In-person surveys can be used with measurement procedures that use complicated visual or measurement aids. Interviewers can show respondents photographs, videotapes, or objects like product packages or sample magazines. They can use packs of cards for ranking or sorting tasks like multidimensional scaling procedures, small poster-type displays of scale points to aid the respondent in replying to verbal questions, and other visual aids that help the respondent to provide more accurate answers.

In-person surveys have some significant disadvantages, however. The first problem is in obtaining a representative sample. Area or households are frequently the primary sampling units, but these have real problems because of differences in housing density in different areas, different numbers of persons per household, etc. (see Chapter 6). There are also some neighborhoods or regions into which interviewers are reluctant to go. For example, inner city tenements and sparsely populated rural areas that require lots of driving between interviews are particularly difficult. It is hard to meet the equal likelihood requirement of a representative sample unless some complicated and expensive steps in selecting the sample (and motivating the interviewers) are used.

Selecting the units is not the only problem in obtaining a representative sample. There is a further problem in getting participation from the sampled persons. People are increasingly reluctant to allow strangers into their homes, even those who identify themselves as noncommercial communication researchers.

At one time in-person surveys had a very high cooperation rate. Cooperation rate, or response rate, is the proportion of sampled units who agree to complete the survey. Response rate for in-person surveys has been falling dramatically in recent years. In the 1960's, response rates near 90% were not uncommon, but today the figure is closer to 60%.

Cooperation can be improved by using any means which establishes the identity and authenticity of the interviewer. The lowest response rate will be seen when interviewers knock unannounced on a door and ask to come in for an interview. More cooperation will be achieved when the respondent is contacted in advance by mail and/or telephone, and an appointment for conducting the interview is set up. Contacting respondents by mail also provides a chance for the interviewer to present authentic credentials, like a university letterhead, which will help reassure the respondent that she is not opening the door to an ax murderer.

But of course contacting the respondent several times costs much more money than just knocking on the door, so there is a trade-off between cooperation and cost. In fact, the cost per completed interview for in-person surveys is the largest drawback to this research procedure. The cost of obtaining a completed interview includes a complex sampling procedure, travel time and expenses for the interviewer, incentives to go into scary neighborhoods, and other incidental expenses. The cost of an in-person interview is frequently 2-5 times as much as the cost of a telephone interview.

## Telephone Survey

In a telephone survey, the interview is conducted over the telephone. This eliminates interviewer travel time and expenses. Obtaining a representative sample with techniques like random digits dialing is relatively simple, as we noted in Chapter 6. This reduces the cost of a telephone survey.

But there are drawbacks to telephone interviewing. The loss of personal face-to-face contact between the interviewer and respondent can have negative consequences for cooperation and for the accuracy of the answers. No visual aids can be used, so this kind of survey cannot be used for communication research which relies on visual judgments or responses.

Also, a telephone interview cannot be as long as an in-person interview. While an in-person interview might easily last 30 minutes or longer, the upper limit for a telephone interview is about half that time, and many experts recommend 10 minutes as the target time for a telephone interview. This is not very long, and it places a stringent limitation on the amount of information that the researcher can obtain with this kind of survey.

Cooperation rates for telephone interviews are relatively high, although they have also been falling over the years. There is not much of a fear factor in cooperating with a telephone interviewer for the respondent. But requests for telephone interviews have become more frequent, and phony telephone surveys have been used as a ploy by telemarketers and crank callers, so that the suspicions of a typical respondent may be aroused when he is contacted for a legitimate survey. Both these factors have lowered the willingness of respondents to cooperate with telephone surveys. In the 1960's, response rates were typically over 80%, while many telephone surveys now fail to reach even a 70% cooperation rate.

## Mall Intercept Survey

This is an in-person survey that is conducted by selecting the respondents from the persons who are present in some public place, such as a shopping mall. Respondents are randomly selected from passers-by. This eliminates the expense of complicated sampling procedures and interviewer travel, as well as the problem of dealing with bad neighborhoods and rural areas.

These advantages must be balanced against the significant drawback of obtaining an unrepresentative sample. As we saw in Chapter 6, the persons present in a shopping mall or a public square are not going to be typical of the entire population, so the results of this kind of survey have poor external validity when they are generalized to the general population.

For some kinds of research, however, a mall intercept provides a good approximation of a representative sample. For example, a marketing communication researcher conducting research on advertising appeals for clothing may be willing to define her population as being those people who frequent shopping malls, where most clothing is sold. In this case, if a good random sampling procedure is used at the mall intercept—people are selected using some fixed rule, like every 50th

person who passes, interviewing is done on all days of the week, and at all times the mall is open, etc.—then the resulting sample may be quite representative of the population of consumers. But using a mall intercept for political communication research is probably unwise.

Cooperation rates for mall intercepts are somewhat better than in-person interviews, as the mall is a neutral and safe meeting ground. But the length that the interview can take is shorter. People in malls and public places are frequently pressed for time, and will literally not sit still for a very long interview. Cooperation can be improved by offering small incentives for respondents like discounts at a mall store, free ice cream cones from a nearby shop, etc. These add a little to the cost of the interview, but some businesses may provide such incentives to researcher as a low cost way of increasing their foot traffic.

There are practical problems with mall interviewing, too. Many malls do not permit such activity, or they may have a single commercial research firm which is licensed to conduct all survey research on the premises. Gaining permission to interview may be difficult or expensive.

## Mail Survey

The mail survey is unlike the other kinds of survey discussed above in one very important way: it does not involve an interviewer. The survey is self-administered.

This introduces some significant problems. There is no interviewer who can exert subtle social pressure to convince the respondent to cooperate. If the respondent has trouble understanding the questionnaire instructions or questions, there is no one to ask for help. Since most questionnaires will be written in English, the respondent must be functionally literate in that language. This will eliminate almost 20% of the general population immediately.

Because of these problems, the response rate for mail surveys is usually very low. Rates of 20% or less for a general mailing are typical. The response rate will increase if the respondent perceives the survey as important, or holds strong opinion about the issues explored by the questionnaire, or identifies strongly with the sponsoring organization. A survey of attitudes toward guns conducted by the National Rifle Association will have a very high response from gun fanciers, while a survey of preferences for muffins conducted by Acme Foods will have a very low response rate from the general public.

Questionnaire design and the length of the questionnaire are very important. Long questionnaires and those with confusing instructions or complex scales will be thrown in the wastebasket. So will those with poorly written or confusing questions. The single best thing the researcher can do to increase response rate is to pay lots of attention to creating a very good, clean questionnaire.

Another way to improve response rate is to provide some incentive for the respondent. Even small incentives, like a \$1 gift certificate, will help. Of course, even small incentives are expensive when they are mailed to hundreds or thousands of respondents. The researcher can sometimes provide a no-cost “psychological incentive” by stressing the value of the information for creating public good in the introduction. This works particularly well for targeted audiences. For example, a researcher studying use of public libraries by active cardholders can stress the value of the results in improving the library services being used by the respondent.

Response rates can be improved by contacting respondents who do not reply to the first mailing. A typical way to do this is to mail out a reminder post card or a second copy of the questionnaire to those members of the sample who do not respond to the original mailing. This might be followed by a telephone call if the second mailing does not produce results. A rough rule of thumb is that each additional contact will add another 5% or so to the response rate.

Multiple contacts are very expensive. Say the original mailing to a sample of 1000 cost \$1000 and produced a 25% response rate. Each completed questionnaire in this mailing costs the researcher \$4. The second mailing of the complete questionnaire to the 750 people who did not return the first questionnaire will cost \$750 and produce perhaps 50 more questionnaires, at a cost of \$15 per questionnaire. But the extra questionnaires may be worth the cost, as the sample will be much more representative with their addition.

Multiple contacts require that the respondents be identified, so the researcher can tell who has returned the questionnaire and who has not. This can be a problem if the questionnaire deals with sensitive topics where respondent anonymity is desirable. Even innocent questionnaires with identification are avoided by some respondents who are afraid that the information might be used to target them for junk mail or sales calls.

One method to deal with this problem is to identify each questionnaire with a code number, and to include some phrase in the questionnaire introduction like "Each response is identified with a confidential code number. This number will be used for tabulation purposes only, and your responses will be kept completely private." Some researchers attach this code number to a return envelope and mention nothing about it in the questionnaire. This practice falls in an ethical gray area, and we'll discuss it further in the last section of this chapter.

Mail surveys are relatively cheap because there are no interviewer expenses. But with low return rates and multiple mailings, they may actually cost more per completed questionnaire than telephone surveys. The low return rates usually mean that the researcher cannot make a strong case for a representative sample, and so must qualify his results with cautions about the generalizability of the findings.

## Computer Administered Survey

This kind of survey relies on a computer to conduct the interview. This provides some advantages and disadvantages.

On the positive side, a very extensive set of questions can be prepared, but only a relevant subset will be asked of a given respondent. If the researcher is studying relationship communication, an early question might be "Are you married?" If the respondent answers yes, she might be given one set of questions; if no, she would get another set of questions aimed at unmarried persons. Because the computer can immediately analyze replies and use this information to select the correct questions, the respondent does not have to follow complex instructions on a printed questionnaire. All he sees is a set of simple questions presented on a computer screen. This is called adaptive interviewing, and some very sophisticated computer software is available which does extensive statistical data analysis during the interview and uses the results to structure the interview.

Since the reply to the questionnaire questions is made directly on a computer keyboard, the data can be immediately analyzed. Complete data analysis can be obtained a few minutes after the end of interviewing. The cost of keying in the data is eliminated.

The largest drawback to computer-administered surveys is the lack of respondent access to computer systems. This kind of survey is not very useful for reaching a general population. It can be used in doing business related research, since most businesses, even very small ones, have personal computer systems.

## Disk by Mail

These surveys are a combination of mail and computer administered surveys. The respondent is mailed a computer disk with a programmed set of questions. The respondent replies to the questions using her own computer system, then mails the disk back to the researcher. Since computer disks can be reproduced in quantities for \$1 to \$2 each, the cost is not a lot higher than printed questionnaires, and the difference is offset by the elimination of data input costs.

Computer administered surveys can also be used in in-person or mall intercept settings. The interviewer can present the respondent with a laptop computer, show him how to respond to the questions, and then stand by to help out if computer operation problems occur.

Some researchers feel that computer administered interviews are superior to interviews conducted by humans when the questions involve very sensitive personal issues. People are surprisingly willing to reveal personal details to the very impersonal computer, apparently because it is perceived as being anonymous and not judgmental. Socially desirable responses are less frequent. Who cares what a computer thinks? The respondent can complete the survey in private (the researcher can leave the room), and embarrassing responses disappear from the screen into the depths of computer memory, rather than residing rather permanently on paper.

## Calculating Response Rates

We have mentioned cooperation or response rates repeatedly. They are a critical part of the decision about which survey method to use, but they are a little more complicated than we've admitted.

Let's take a telephone survey as an example. Suppose a political communication researcher wishes to study voting behavior in a state. He then defines his population as all persons over the age

of 18 in the state. Using a random digits procedure, he selects a sample of 1000 telephone numbers. The results of calling these numbers are shown in Table 18-2. Only 400 of the original 1000 numbers result in a completed interview. What is the response rate?

Table 18-2 Response Rates in a Hypothetical Telephone Survey

<b>Original Sample</b>	<b>1000</b>
Disconnected Number	100
Business or Fax Number	200
No Answer after 5 Call Backs	100
Answering Machine	50
No One in Household over 18	25
Did Not Complete Full Interview	50
Refused to Cooperate	75
<b>Total Completed Interviews</b>	<b>400</b>

The simplest answer is 40%, since that is the percentage of the original numbers which gave good data. But it's probably incorrect to consider the disconnected numbers and business numbers as part of the original population. If we eliminate them, we have a response rate of 400 / 700, or 57%.

But what about the 100 numbers that never answered? Some telephone numbers are used for computer data transmission, so they never answer. Others may be the back rooms of businesses, or be used for automatic alarm systems, or other non-personal uses. But they may also be the households of persons who are just not at home very often. It is very hard to guess how many persons in the universe are missed in this category. If the interviewer tried calling at different times of day on different days of the week, the researcher may be justified in eliminating these no answers from the population. The more call backs that are tried, the better the justification for eliminating the no answer numbers from the population. If he eliminates them, the response rate is now 400 / 600, or 67%.

If an answering machine picked up the telephone every time the interviewer called back, there is probably some evidence that this number should be part of the universe. But the answering machine might be in a work-at-home office or a teenager's bedroom. Some researchers would eliminate these from the population, and increase the response rate to 400 / 550, or 73%.

The next set of persons agree to participate, but do not complete the interview, because they tell the interviewer that there is no one over 18 in the household. This eliminates them from the population, and increases the response rate to 400 / 525, or 76%.

The last two sets are clearly part of the universe. They qualify, but do not complete the interview because they either refuse at the beginning to participate, or quit part of the way through the interview. However, some of those who refuse may be under 18, so including them in the universe may underestimate the actual response rate. But there is no way to know this.

As you can see, we can get response rates ranging from 40% to 76% from the same study. Some of the differences in the response rates reported by different researchers are due to differences in the way the rate is calculated. To give the best description of the response, the researcher must consider each group which did not cooperate fully, or which did not clearly refuse to cooperate. She must then decide whether they are more likely to be inside or outside of the population, and adjust the calculation accordingly. It is very important that the researcher report the details of who was included and excluded in the write-up of the research, so that the reader can tell exactly how the response rate was computed.

## Choosing the Type of Survey

The characteristics of each type of survey should be matched to the requirements of the research problem. For instance, if obtaining a representative sample is very important, mail surveys are ruled out. But if obtaining a large number of responses at reasonable cost is a requirement, then mail surveys may be preferable to the other types.

Generally, the combination of relatively high cooperation rate and low cost makes telephone surveys attractive for many research applications. But telephone surveys cannot be used whenever any visual aids are required, so either mail, mall intercept, or in-person surveys may be required.

The researcher must weigh the various pro's and con's of each survey type carefully. Making up a more extensive matrix like that shown in Table 18-1 may help with the decision.

## Interviewer Requirements

The primary rule for interviewers is this: Intrude in the research process as little as possible. This means that the interviewer should never direct the respondent toward an answer, should never be judgmental, and should not interpret the respondent's answers according to her own beliefs or values. The interviewer must be consistent in her communication style and language, so that each respondent is exposed to the same kind of measurement environment.

This is easier said than done. Some types of interviewing require judgments and decisions from the interviewer, while others do not. A telephone interview with a scripted introduction and closed-ended questions will give the interviewer little chance for interpretation or choice of language or instructions that might guide the respondent toward a particular answer. But in semi-structured interviews like focus groups (see below), the interviewer must exercise immediate judgment about the kind of question to ask next, and may ask questions differently of different respondents. This gives the interviewer the great potential to bias the results.

Even in a very structured interview, a poor interviewer can lead a respondent to particular answers by vocal inflection or by off-hand comments ("Are you SURE you want to say that TV censorship is OK?"). The only answer is training, and interviewer training is crucial in all forms of communication research that require measurement which involves interaction between the researcher and the research subject. The highest-quality questionnaire or most sophisticated research design can be ruined if the researcher neglects sufficient training for the people who collect the data.

The first step in training interviewers is to give them a detailed set of instructions on how to approach the research subject and what is expected of them during the interview. The interviewer must share the researcher's appreciation for non-direction and pleasantness, and have some idea about how to start the interview and set its tone.

The next step in interviewer training is the dry-run interview. The interviewer runs through the full interview process with someone playing the part of a typical respondent, while the researcher observes the interview. The researcher then discusses the interviewer's performance, notes the areas that need improvement, and repeats the process until the researcher is convinced that the interviewer is sufficiently low-key and consistent that she will not bias the results.

## In-Person Interviews

There are some particular requirements for interviewers who must interact face-to-face with respondents. First, the interviewer must establish his legitimacy. Showing some identification, such as a photo identification card with the interviewing organization prominently shown, is often the first step. The interviewer must then gain the respondent's cooperation by being polite, but persistent. The interviewer can stress the importance of the research, its ease, etc. This is essentially an exercise in salesmanship. The interviewer is asking the respondent to contribute something valuable (her time) and he needs to give the respondent some persuasive reason to make this contribution.

After cooperation is obtained, the interviewer must establish a personal connection with the respondent. There is no good prescription that will spell out how to do this. Basically, the interviewer must be friendly, polite, and speak respectfully to the respondent. Some interviewers do this automatically, others must learn it, and some never achieve it. Those that do not must be eliminated as possible interviewers.

The formation of a relationship between the interviewer and the respondent carries some dan-

gers, too. The interviewer must guard against introducing any influence that might create social demand in respondents' answers. The interviewer must also strive to treat all respondents in a consistent fashion, so that any bias accidentally introduced is at least consistent across all respondents. To be clearer about this, suppose a female interviewer is able to easily establish a warm relationship with female respondents, but is rather stiff with male respondents. This might bias the answers given the interviewer along sex lines, and introduce false conclusions about gender differences on the measured variables.

## Telephone Interviews

The telephone interviewer does not have the advantage of the personal immediacy provided by face-to-face interaction. He must identify himself very quickly, without the aid of a physical artifact like an ID card. He must also quickly assure the respondent that he is not selling something or making an obscene telephone call.

The telephone interviewer must also quickly gain cooperation from the respondent and establish some personal connection. Generally, this connection will be "cooler" than a face-to-face relationship. This actually may aid the interviewer in asking sensitive questions, but it may hinder her in gaining replies to difficult or challenging questions.

The telephone interview will often generate more requests for clarification of questions from respondents, since they will have no visual aids. For example, an in-person marketing communication interviewer may show a respondent a picture of a new product, then ask questions about it, while a telephone interviewer will have to describe the product verbally. The telephone respondent is then likely to ask for clarifications, like "how big is it". The telephone interviewers must be instructed on how to reply to such questions. It is particularly critical that all interviewers apply the same rules to their replies. If one interviewer replies "I'm not allowed to say," while another replies "It weighs 4.75 pounds, comes in white and red, and it fits on your desktop", response bias will be introduced by the different interviewers.

## Focus Group Interviews

One particular type of interviewing that has been used more and more extensively in recent years is the focus (or "focused") group interview. This term has been applied to a wide range of interviewing techniques and settings, so it actually refers to a class of procedures, rather than one type of interview. But we can describe a typical focus group interview process in general terms.

First, the researcher determines the topic or topics he wishes to investigate. This may be very general ("What are the perceptions of my political candidate by the electorate") or it may be a specific list of topics ("...perceptions of the candidate's stand on abortion; reaction to campaign commercials stressing family values; evaluation of candidate's response to charges by staffers of sexual harassment..."). The focus group interviewer (often called the moderator) uses this list to guide the interviewing process. There are usually no pre-written questions, so this list serves the same purpose as a questionnaire in a survey. It structures the interviewing process, but to a much lesser degree than a questionnaire. The focus group interview is more like a journalistic interview than like a survey.

Second, a group of respondents is chosen. The group size is typically 4-10 people. The researcher often determines the characteristics of the group ("...white males over 35..."; "...suburban Midwest residents..."). Convenience samples are the norm in current research practice. This limits the generalizability of the results, but convenience sample are often used for pragmatic (i.e., cost) reasons. Multiple groups may be involved in the research project, but the total number of participants is usually fairly small.

Third, the group is assembled at some interviewing location where the interviewer and the group can be placed in comfortable surroundings that will encourage them to talk freely. The interview room usually is equipped with some means by which the group can be observed. This is frequently a one-way mirror, behind which observers watch or summarize the conversation of the group members. The focus group may also be audio or videotape recorded. The group interview process often requires an hour or more. At the beginning, the group must have time to relax and warm up to the discussion, and then the discussion must be allowed to keep flowing to get at more subtle views or opinions.

In some cases the focus group interviewer also summarizes the interactions among group members, but this is generally not a good practice. The interviewer has to listen carefully to responses and allow them to lead her to the next question, while still considering how to steer the discussion to cover the issues outlined by the researcher. If she also has to summarize the conversation, the flow of interaction in the group may be interrupted while she struggles to write down notes. This choppy discussion may give much less information.

Finally, the group's discussion must be analyzed and conclusions drawn from it. This is often done by making a transcript of the discussion from the audio or videotape recordings. This transcript can be analyzed impressionistically by one or more judges, or it can be subjected to a formal content analysis. For exploratory research, the impressionistic procedure is often used. For formal hypothesis testing, full content analysis is often used.

Focus group interviewing is expensive. It requires highly trained interviewers, expensive or scarce facilities, and lots of time from the respondents, who often have to be paid for their participation. For this reason, focus group research often uses very small numbers of respondents, as we noted above. But it produces a very rich set of data in the transcripts. Because the researcher does not completely determine the structure of the interview beforehand, focus group interviews can produce unexpected and important results. An interview that starts discussing a political candidate's problems with his personal life may end up showing that the group is much more concerned about his stand on gun control.

Focus groups are frequently used as a first step in constructing a more formal questionnaire. This is an excellent, if expensive, strategy. The strength of focus group interviewing is in the wide range of qualitative information that is produced by its lack of structure. But the small *N*'s, the convenience samples, and the lack of consistency in the way questions are asked when multiple groups are interviewed makes the external validity of focus group research generally poor. (The makers of a popular soft drink who changed its formula as a result of focus group results found this out to their very expensive dismay.)

On the other hand, structured questionnaires, such as those described in the next section, can be applied to large representative samples under controlled interviewing conditions. But if the researcher asks the wrong questions on the questionnaire, the results will be poor. By using focus groups to identify the critical variables and processes that characterize a phenomenon, then using more formal quantitative techniques to test hypotheses about the phenomenon, the communication researcher can produce superior results.

Training a focus group moderator is no small task. A telephone survey interviewer might be trained in a few hours, but a focus group interviewer will require days of practice and many dry runs before his basic group interviewing skills are sufficient. Once trained in the basic processes, the interviewer can take on subsequent focus groups more quickly. But the interviewer must have good knowledge of both the research question and of the group characteristics in order to ask relevant questions and pick up on the subtle hints that group members exhibit. Asking a young male business executive who eats most meals in restaurants to moderate a group of middle-aged female housewives discussing baking products is not going to work well. The researcher should always try to find an interviewer whose background and life-style matches the group's as much as possible. The interviewer must either be trained extensively about the discussion topic, or have some prior knowledge of it.

Because of the difficulty of training and knowledgeable moderators and of assembling groups in appropriate interviewing facilities, there are a number of commercial research firms which specialize in providing focus group interviewing services. These firms frequently specialize in certain topics or ethnic groups. There are focus group suppliers that specialize in political communication research, in consumer and marketing communication research, in Spanish-language research, etc. Using these firms for focus group research is not cheap, but the odds of obtaining good data are usually improved.

It is not impossible for the individual researcher to use focus groups, however. But the researcher should recognize that they are quite demanding and will require extensive preparation and analysis to produce useful results. Once again, it's very easy to do a poor job, and very hard to do a good one.

## Questionnaire Design

Many communication research projects rely on data collection from questionnaires. Experimental, field and observational research designs frequently use questionnaires. Public opinion surveys, political surveys, surveys of product purchases, and readership and viewership surveys are all general examples of research which is conducted almost exclusively with questionnaires. But an interpersonal communication researcher might use questionnaires to get descriptions of the communication behavior of individuals, their demographic characteristics, and so forth. Good questionnaire design and construction is critical to a wide range of communication research.

Writing a questionnaire is a little like playing the guitar. It's easy to just jump in and do it badly, but it's very hard work to do it well. At first glance, the process looks like simplicity itself. Just ask the respondents some questions, and record their responses. But 50 years of study of questionnaire construction has shown that the way one asks questions often determines the quality of the information and may even determine the results that one gets.

### *The Questionnaire Introduction*

The first thing a questionnaire must do is convince the respondent to cooperate in answering the questions. This task is particularly demanding if self-administered questionnaires are being used. This is always the case in mail surveys, and might occur in experimental research if the dependent variables are measured with scaled variables, or if the researcher wishes to measure demographic variables that are easily operationalized with questionnaire items.

The introduction to the questionnaire must be very clear and simple. It should state the purpose of the research and it must disarm any fears or suspicions about the motives of the researcher. It must establish a rapport with the respondent and convince him or her to cooperate. Questionnaires have been misused as sales devices, for political solicitations, and for other non-research purposes, and the respondents are justifiably suspicious about them.

The introduction to a mail questionnaire might be in a cover letter, or it might be the first paragraph of the questionnaire form. For telephone or in-person interviewing, the introduction might be read or spoken to the respondent. It is best to defuse the suspicions of the respondent by immediately identifying the person or organization doing the research. For example, a telephone interview might begin with an introduction like this:

*"Hello, I'm John Jones. I'm conducting a survey of audience reactions to television advertising for the Opinion Research Institute at the University of Calisota. This is a research project, and the information will not be used for sales purposes. Your identity will not be recorded. Can you spare a few minutes to answer some questions?"*

There may be times when stating the full purpose of the research or even the organization doing the research at the beginning of the interview might bias responses to later questions. Withholding information in the questionnaire introduction becomes an ethical issue, and we'll discuss this later in this chapter.

### *Question Wording*

Writing individual question items on the questionnaire demands that the researcher pay attention to a number of important things.

### *Vocabulary and Style*

Questions must be written in the simplest language possible. Use of polysyllabic words like "polysyllabic" should be avoided, as should jargon and slang. Using long sentences with many dependent clauses is also poor practice. The questionnaire writer should strive for "journalistic" prose: short declarative sentences using active tenses. If a 12-year-old can't read and understand the question, it probably should be rewritten.

## Leading Questions

Items can lead respondents to a certain response by the way that they are phrased. This biases the responses, and it will probably annoy the observant respondent. Asking a question like “Do you support the efforts of the Independent Party to improve the working conditions of all Americans?” will produce no useful information. Rephrasing the question as “How would you rate the performance of the Independent Party on workplace regulation?” allows the possibility of a valid negative response.

Another way that questions can lead the respondent to an answer involves associating some response with an influential group. “Most Church leaders deplore televised violence. How do you feel about the amount of violence on television?” Obviously, the questionnaire writer must avoid this kind of question.

Leading questions are sometimes unethically used to produce the results desired by the organization which sponsors the research. A critical reader of research results will be sensitive to the way in which questions are asked, and will reject the results of any research which uses such questions. The questionnaire writer should be aware of the leading nature of some questions, and be very careful to phrase all questions in a neutral fashion that makes all responses equally easy for the respondent.

## Social Desirability

The response to some questions may be biased because of prevailing ideas about “correct” opinions or behavior. Such questions are said to have demand characteristics, because social norms demand that the respondent who does not see himself as socially deviant (i.e., everyone) give a particular “correct” answer. The respondent then gives an answer because he thinks it is expected or socially appropriate, rather than because it accurately represents his position. “Are you in favor of equal employment opportunities for minorities?” will be overwhelmingly answered in the affirmative by nearly all respondents, but it’s unlikely that such a large number of respondents really hold that opinion. Phrasing questions that avoid demand characteristics is often difficult when dealing with emotionally charged issues.

To compensate, the questionnaire writer can sometimes reframe an issue that has demand characteristics so that a range of responses is possible. The basic strategy is to reassure the respondent that she is not alone in expressing an opinion at either end of the measurement scale. For example, the equal employment question in the previous paragraph might be recast like this: “Some people feel that equal employment programs have produced problems for the average worker, while others feel that they have opened up opportunities for all workers. On a 1 to 7 scale, where 1 is Strongly Unfavorable and 7 is Strongly Favorable, how would you rate your own feelings about equal employment programs?”

## Vague Questions

If a questionnaire item is not phrased carefully, different respondents may interpret the question differently, and the resulting measurement will be invalid. “Would you call your relationship with your spouse good or bad?” is a very general question. Some respondents may interpret the word relationship to mean sexual relationship, while others may consider it to mean communication of private feelings, and still others will think about all the items that make up their personal definition of a relationship, and mentally average them. The researcher will actually be getting a measurement of different concepts from each of these respondents (that is, invalid measurement). To avoid this, questions must be as concrete and explicit as possible. Their meaning should be unambiguous.

Since questionnaire items are part of the operationalization of a theoretical concept, the researcher can refer to the theoretical definition (see Chapter 2) as a guide to the meaning elements that should be included in a question. In the above example, if the theoretical definition of “relationship” included both sexual compatibility and selfdisclosure of personal feelings, the single vague question could be expanded to two more explicit questions about each of these elements.

## Double Items

Sometimes a poorly constructed questionnaire item may really ask two or more questions. The respondent must then base his response on one of the two questions and ignore the other, or he may simply throw up his hands and refuse to answer. Suppose a manager is asked this question: "Are you in favor of installing computer mediated conferencing equipment in your organization, or do you think existing communication facilities are adequate?" It is quite possible that she thinks existing communication facilities are poor, but does not feel that computer conferencing will solve the problem (she supports installing additional voice mail equipment). She cannot respond accurately to this item.

Double items must be split into simple items that measure only a single concept: "Are you in favor of installing computer mediated conferencing equipment?" and "Do you feel existing communication facilities in your organization are adequate?"

## Structuring Questions

Aside from the issues of question wording discussed above, the questionnaire author must also deal with some structural issues in each question.

## Complex Instructions

Some questions or measurements require explanation before the respondent can reply. These should be as short as possible, and as simple as possible. This is particularly the case with telephone or in-person interviewing, as the respondent cannot back up and reread the instructions. Even in mail and self-administered questionnaires, long and complex instructions will reduce the number of respondents who complete the item or the whole questionnaire.

Any set of instructions that has more than three or four sentences should be critically reviewed. Of course, there is a trade-off here. Respondents must clearly understand what is expected of them. But if explaining the procedure requires a long and complex set of instructions, the researcher probably should review and revise his measurement procedure. It may be too complicated to be successful.

## Long Lists of Alternatives

Respondents deal best with a fairly small set of comparative judgments. Humans have trouble making more than 5 to 10 simultaneous comparisons. Questionnaire items that ask respondents to rank order 30 alternatives are going to give poor results. So are questions that force the respondent to choose from a large set of items that differ on more than a few basic dimensions.

If you present a respondent with two different sample layouts for a magazine article, he will have little trouble choosing the one he prefers. But if you present him with 25 samples, you may overwhelm his ability to make fine distinctions and to consider all the differences among the choices. Research has shown that people presented with this situation tend to focus on just a few dimensions in order to simplify their task. Unfortunately, the dimensions they focus on may differ among respondents, so the resulting measurement may have poor validity. For example, some respondents may focus on attractive typefaces in judging magazine layouts, while others may use pictures or graphics. This will give preferences which are not really comparable.

Comparisons of complicated situations or objects should be broken down to comparisons on basic dimensions whenever possible, in order to increase the validity of the judgments, and the number of comparisons reduced to the minimum possible. There are advanced methods, such as conjoint analysis which address the problem of making complicated distinctions among alternative choices on a number of dimensions without overloading the respondent.

## "Strange" Response Requirements

Some measurement procedures rely on responses from the research subjects that are not familiar. For example, some scaling procedures ask respondents to rate ideas or objects on unfamiliar dimensions. A semantic differential item might ask a respondent to rate another person on a scale with "hot" at one end and "cold" at the other. Or a multidimensional scaling item might request the

“distance” between an apple and an orange. This is not the way most people normally make distinctions, and the researcher is going to have to take some pains in the instructions for such items to insure that the respondent feels comfortable with such responses.

Instructions that disarm the respondent’s potentially negative reaction to the measurement task are useful in these cases. Examples are usually necessary, too. The instructions might be phrased like this: “Although you may find it a little strange to fill out these scales, just check the box that seems best to you. There are no right or wrong answers. For example....”

## Open- vs. Closed-Ended Questions

Closed-ended questions require the respondent to choose from a fixed set of alternatives or to give a single numerical value. Open-ended questions allow the respondent to choose their own replies. Some examples are shown in Figure 18-3.

Choosing one type of question or the other is not strictly a question of style or preference. Some measurement techniques will require scales or numerical responses, while others will require free responses.

Closed-ended questions are much easier (i.e., cheaper) to process. Responses can be precoded (the data processing codes are placed right on the questionnaire) so that data entry is simplified. There is also a uniformity to the response that improves the reliability of the responses. But closed-ended questions can be restrictive. If the questionnaire author has not included all relevant alternatives, the respondent may be frustrated in trying to answer.

Open-ended questions are much harder to handle, because each must be examined by a trained coder, and the responses classified for analysis. The researcher must conduct a separate content analysis for each question. (Content analysis procedures are described later in this chapter.) Highly verbal respondents will give long replies; respondents with low levels of literacy may skip the questions entirely. But open-ended answers can provide a very rich source of information, particularly about topics that have a wider range of possible answers than the researcher can anticipate.

The general prescription is to use open-ended questions sparingly, and reserve them for measurements that have many alternatives, or for exploratory purposes. Closed-ended questions are efficient and reliable. But it’s a mistake to force all questions into a closed-ended format, if open-ended will give better information.

## Questionnaire Architecture

The entire questionnaire should be seen as a communication between the researcher and the respondent. Simply creating a good set of questionnaire items is not enough. The items must be placed in the questionnaire so that the whole presentation is effective.

## Length and Fatigue

Filling out a questionnaire is work. Since participation in communication research is almost always voluntary, a long, demanding questionnaire can significantly cut down on the number of people willing to finish it. A questionnaire should be as short as possible, and the items as simple as possible. The fatigue that comes from filling out a long questionnaire may interfere with the responses. While a careful researcher can reverse questionnaire items on half the questionnaires (but not always—see the next section) to remove systematic effects of fatigue, the fatigue will still cause errors in measurement that will make it harder to detect real relationships among variables.

Although the amount of interest and commitment by the respondents will determine the length of questionnaire that they are willing to complete, any self-administered questionnaire that requires more than 15 minutes to fill out in a pilot test should be critically reviewed to see if it can’t be reduced in size. A telephone questionnaire should be even shorter, about 10 minutes.

A primary rule is this: never ask anything that your research design does not require. Novice questionnaire authors often decide to include questions “just in case” they might be useful later. Avoid this temptation! Each questionnaire item should be part of the operational definition of a theoretical concept laid out during your analysis of the research questions. If there’s no theoretical variable, there doesn’t need to be an operational questionnaire item. The damage done by a too-long questionnaire will far outweigh the insurance of asking extra questions that may not relate to any theoretical construct of interest.

Table 18-3 Examples of Closed-Ended and Open-Ended Questions with Precoding

Closed-Ended										
Generally speaking, do you feel that your communication with your spouse is										
1) VERY SATISFACTORY										
2) SOMEWHAT SATISFACTORY										
3) NEITHER SATISFACTORY NOR UNSATISFACTORY	(4:22)									
4) SOMEWHAT UNSATISFACTORY										
5) VERY UNSATISFACTORY										
Approximately how many hours a day do you spend watching television?	_____ Hours (4:23-26)									
The people in the advertisement appear to be										
Happy	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table> Sad (4:27)									
Good	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table> Bad (4:28)									
Excited	<table border="1"><tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table> Calm (4:29)									
	<table border="1"><tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td><td>8</td><td>9</td></tr></table>	1	2	3	4	5	6	7	8	9
1	2	3	4	5	6	7	8	9		
Of the program types listed below, mark the ONE program category that you like to watch the most										
<input type="checkbox"/> News programs (1)	<input type="checkbox"/> Sports (5)									
<input type="checkbox"/> Movies (2)	<input type="checkbox"/> Public affairs (6)									
<input type="checkbox"/> sitcoms (3)	<input type="checkbox"/> Dramatic series (7) (4:30)									
<input type="checkbox"/> Musical/variety (4)	<input type="checkbox"/> Other (8)									
Open-ended										
Think back to yesterday. What topics did you discuss with your children?	(5:6-7)									
Please list, as fully as you can, everything you remember about the diet soda commercial.	(5:8-11)									
What is the most important issue facing the American public today?	(5:12-13)									

## Question Order

Filling out a questionnaire takes time, so different concepts are measured at different time points over a short period. Measurement of one concept at one point in the questionnaire may affect later measurement of other concepts. Sensitization effects occur within the time span of the measurement, and they may affect the results.

Suppose a researcher is interested in the public's concern about the amount of sexually-oriented material on cable television. She may want to start with an open-ended question like "What do you think is the most important issue or problem facing the country today?". This can be followed by a more focused question, "What is the most important issue or problem related to the mass media?" and finally by specific questions about depictions of sexual activity in the media in general and then on cable television. By ordering the questions from general to specific, she can get a measurement of the respondent's concern in the context of other issues that the respondent thinks are important. If she had reversed the order of the questions, the questions about cable television would probably have prompted respondents to consider this issue more strongly when they responded to the questions about mass media and about general problems faced by the country.

Many attitude or opinion questions suffer from sensitizing or context effects that are introduced by other questions or by the measurement setting itself. For example, valid measurement is sometimes impossible if the intent of the researcher is known by the respondents. Although concealing intent introduces some ethical concerns (see the final section in this chapter), it can be done by "masking" sensitive questions with questions about other issues. To do this, the questionnaire author arranges the critical questions randomly in a set of other unrelated questions. If a family communication researcher wishes to find out respondents' attitudes toward several aspects of divorce (child custody, alimony, religious prohibitions, etc.), he can sprinkle questions about these issues among dozens of other questions about other political and personal matters. Placing all the questions about divorce together in a block might sensitize the respondents to the issue, and make their responses more extreme.

Another important consideration on item placement is where to put sensitive questions like those about age, income or personal habits. In general, these are best placed near the end of the questionnaire. Respondents are more likely to cooperate in answering these questions after some rapport has been established with the interviewer during in-person and telephone interviews, or some time commitment has already been made by filling out the first part of mail or self-administered questionnaires.

Personal questions are always potentially dangerous, so they should be asked only if absolutely necessary. These questions may cause the respondent to become suspicious of the researcher's or interviewer's motives or they may be considered offensive or an invasion of privacy. If a respondent feels this way, he or she may refuse to answer any further questions, so information from questions that appear after the sensitive ones will be lost.

## Question Skips

Many questionnaires require "skips", where the interviewer or respondent is instructed to skip over a set of questions, if he had a particular answer to some prior question. For example, if a marketing communication researcher is studying automotive advertisements, he might want to ask two different sets of questions of persons who have purchased an automobile in the past year, and those who have not:

1. Have you purchased a new or used automobile in the past 12 months?

Yes  No

→ IF YOU ANSWERED YES, SKIP TO QUESTION 4

2. Do you expect to purchase a new automobile in the next 12 months?

Yes  No

3. Please rate the importance to you of each of the following features of a new car

<i>Very important</i>				<i>Very unimportant</i>				
1	2	3	4	5	6	7	<i>Color</i>	
1	2	3	4	5	6	7	<i>Auto. trans.</i>	
1	2	3	4	5	6	7	<i>Four-wheel drive</i>	

4. Are you a subscriber to any of the following magazines?

Yes	No	
___	___	<i>Road and Track</i>
___	___	<i>Car and Driver</i>
___	___	<i>Motorcar</i>

Skips are a necessary evil. They are a particular problem in self-administered questionnaires, because the researcher cannot insure that the respondent has answered the right set of questions. Even when an interviewer is reading the questions, skips complicate the training and slow down the interview process. Skips should be as short as possible, and the instructions should be very clear. You should avoid complicated skip patterns (“If you answered ‘yes’ to Question 14 and are a male who checked the scale in Question 5b 6 or higher, skip to Question 33...”). Many skips with complex instructions will insure many errors by respondents.

## Esthetics

The appearance of a questionnaire can be very important for self-administered questionnaires. Respondents are more likely to think that the research being conducted with a professional-looking questionnaire is important, and to spend the necessary time to complete it. Even in questionnaires filled out by interviewers, a simple layout can cut down on the number of interviewer errors.

It is important to make the questionnaire as visually simple and clean as possible. Use lots of white space in laying out the questionnaire, and space or indent the items and sections so that related sets of questions, like skip blocks, are visually offset. Adding white space may add to the number of pages in the questionnaire, so there is some tradeoff, especially in the cost of postage if the questionnaire is to be mailed. But a questionnaire with tiny type and cramped response lines and boxes is much more likely to be thrown in the trash.

Desktop publishing and advanced word processing packages make the task of producing a good-looking questionnaire much less difficult than it once was. These software packages can produce very professional looking questionnaires on simple desktop computers equipped with inexpensive laser printers. The questionnaire author should strive for a consistent visual style throughout the questionnaire, and avoid “typeface diddling” with many fonts. Select a basic typeface, with one or two sparingly used alternates (like italics) for instructions, important questions, etc.

## Precoding and Data Processing

The data analysis for almost all surveys will involve entering the data into a computer file so that it can be processed by a tabulation or statistical package. This is an expensive step, and the questionnaire author should consider ways to cut costs.

Responses from questionnaires are usually coded with numbers. If a scale is used, the numbers are the scale points. Numbers are arbitrarily assigned to nominal categories for questionnaire items that do not use magnitude-type measurement. These numbers are typically placed into fixed columns on one or more data lines. These data lines are still often called “cards,” because data used to be punched into IBM cards). These data lines contain all the responses for a single questionnaire, which are collectively called an observation.

One way to cut data entry costs is through precoding the responses. Table 18-3 shows a sample of precoding. The code and location of the data are placed in parentheses right next to the item. For example, the scaled responses (1 to 5) to the question about communication with spouse is to be entered on line or card 4, in column 22 of the observation. The number of hours of television viewing is to be placed in line 4, columns 23 to 26.

The open-ended responses must be manually coded, using some content analysis scheme. The content codes are normally written by the content coder in the margin beside the written response. The questionnaire author has specified that the coded response to the question about discussions with children should be entered into line 5, columns 6 and 7, of the observation.

It is very important to reserve enough columns to enter the full range of codes, including signs and decimal points. The question about hours of television viewing reserved four columns for the response, which allows it to range from 0.00 to 24.0, i.e., the full range. If the author had left only two columns, the response would have had to be rounded to the nearest hour (00 to 24 hours), and this would have caused lost precision when a respondent use any fractional hours in her reply, such as "2 1/2 hours."

Likewise, room must be reserved for the largest number of open-ended codes that is likely to show up in the responses. Since the questionnaire author is not likely to know this number in advance, she should leave plenty of room. The first open-ended question in Table 18-3 is limited to 100 categories of discussion, because the author has reserved only two columns to hold codes 00 to 99. The second openended question can have 10,000 categories of response because the author has reserved four columns.

In large-scale self-administered surveys, precoding can be taken a step further by having the survey questions and responses printed on an optical scanner sheet. The respondents mark their responses in "bubbles" that the scanner can read and translate directly into a data file. This cuts the data entry costs greatly.

## *Pilot and Pretest Studies*

All questionnaires should be pilot tested and revised before they are used in a research project. The researcher should first have a convenience sample serve as pilot subjects and fill out the questionnaire. Then the researcher should go over each instruction and question on the questionnaire and ask each pilot subject "Did you understand this?" and "What did this mean? How did you interpret it?". By probing for understanding and having the pilot subjects explain why they filled out the items as they did, the researcher can find instructions or questions that need to be rewritten, areas of misunderstanding, architectural problems that caused the pilot subject confusion, etc.

If at all possible, a full pretest study with a small sample from the target population should also be conducted. This test should address the same problems as the pilot test, but with a sample of persons who are similar to the persons who will participate in the full research study. It is usually harder to probe for language or question order problems in a pretest, especially if the questionnaire is used in a mail or telephone survey, but architectural problems like confusing skips will show up, as will fatigue problems.

The data from the pretest can be used to create and test data processing procedures. The best technique is to carry out all the procedures necessary to test the research hypotheses, using the pretest data. Often this will point out fatal problems like missing questionnaire items ("How could I forget to ask the respondent's age!"), or errors in the precoding scheme ("The scale goes from 1 to 10, but the coding says the response is supposed to be placed in Column 29 of the data line. It's hard to type a 10 in one column!").

Pilot and pretesting, followed by a thorough revision of the questionnaire, will produce a much better measurement instrument that will cause far fewer problems for the researcher, and will give more accurate results.

## **Content Analysis**

Content analysis is a measurement procedure used in a wide variety of communication research. Essentially, content analysis is the measurement of constructs which can be observed within the messages produced in a communication process. It is used in almost all subdisciplines of communication research. Content analysis is used to analyze media content, the transcripts of interper-

sonal conversations or group discussions, persuasive messages, organizational memos, and even nonverbal interchanges.

The first task in conducting a content analysis is to define the units of analysis. In order of decreasing size, these can be the medium, the message or story carried by that medium, or the individual assertions that are found in a given story. These units can be aggregated at lower levels to produce higher-level units. That is, the content of assertions can be added together to define message content, and the content of messages can be added together to define medium content.

Only mass media researchers would be likely to choose the medium as unit of analysis. The Stempel study described in Chapter 16 is an example of this unit. The content produced by individual newspaper messages is summarized in a single observation of medium content. This unit of analysis would be used to contrast the content of different television networks, or to compare print versus television on some common aspect of content.

The message or story unit of analysis is very common. Here each observation is a complete story or intact message. In interpersonal or group communication research, this unit might be a complete conversation or meeting. In organizational communication research it might be a memo, electronic mail message, or transcript of a meeting.

The assertion is the smallest unit of analysis. It is a part of a complete message, and it may be defined as a paragraph, sentence, phrase or clause, or even individual word. Content analysis of assertions is more difficult to conduct as the assertion unit becomes smaller, but smaller units can be used to make finer distinctions in content and they sometimes improve the reliability of measurement.

The next step in a content analysis is to define the content categories. These are the content variables which take on some value when applied to a sample of assertions or messages or media. For example, a small group communication researcher might analyze the assertions made by each person during a group discussion about alcohol use according to some set of categories like "asked question about medical effects," "provided information about alcoholism rates," "agreed with previous assertion", and so forth. By counting the number of times an assertion was made in each category, the researcher can characterize the content of the entire discussion.

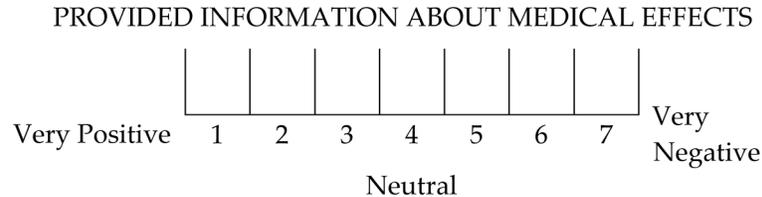
An important point should be made here. If the set of categories is applied independently to each participant in the communication, the resulting set of values can be considered to be properties of the individual. Thus content analysis, which really operates at the message unit of analysis, can be shifted to the individual unit of analysis, and used in research which focuses on individual behaviors or differences.

Content categories must be mutually exclusive so that any content unit (assertion or message, for example) will properly fall into only one. They must be exhaustive so that every message unit can be placed in a category. Often the researcher will meet this requirement by providing an "other" category into which any odd message unit can be placed. The alert reader will recognize these requirements as the same ones that nominal variables must meet (see Chapter 7).

The level of measurement that these content variables take on depends upon the definition of the coding procedure. The simplest is a nominal level, in which each content category is simply marked as "present" or "absent." Coding the content of the group discussion of alcohol use in this way will give a picture of the topics that were discussed, but it will not provide any information about how much they were discussed. Nominal coding is quick, but it provides only minimal information about the content, as it does not distinguish between topics which dominate the discussion and those which are only mentioned in passing.

To get some measurement of the amount of content in each category, the frequency of appearance of an assertion or message in each category can be counted. This will give a ratio level of measurement for each content category. The value for each category (actually, it's now a content variable) now contains information about the amount of a particular content which is present. In the group discussion, for example, the researcher might find that the categories relating to the social effects of alcohol use are mentioned far more frequently than those relating to the medical effects, although both are mentioned at least once. This information is important in inferring the relative importance of the two areas to the group members.

The intensity (amount) and valence (evaluative direction) of message units can be coded to give even more information. Instead of just counting each message unit that falls in a category, the content coder assigns it some numerical weight that reflects the unit's intensity and possibly some positive or negative sign that reflects its valence. In the alcohol use discussion, the assertion that



falls in the category “provided information about medical effects” might be coded on a 7-point scale like this:

An assertion taken from the transcript that read “...alcohol can cause death by damaging the liver...” might be coded a ‘1’, while an assertion that “...people who have two drinks a day are less likely to die from heart disease...” might be coded a ‘6’.

While this kind of coding provides more information, it relies heavily on the judgment of the content coders. This will probably reduce the reliability of the measurement. To improve measurement reliability, the categories must be as concrete as possible. Very general or vague categories will make accurate classification difficult. Categories must be clearly defined so that the content coders can be sure about the kinds of message unit that should be placed in each category. If intensity or valence scales are used, the scale end points and units should be defined as clearly as possible. Using examples in coder training may help in this definition.

### *Intercoder Reliability*

Human judges are almost always used in content analysis (although computer content analysis is being increasingly used for some straightforward content summaries), so the researcher must determine the amount of unreliability which is introduced by differences in judgment. A basic reliability check comes from computing an intercoder reliability coefficient. To compute this coefficient, a set of identical message units is coded, using the same set of categories, by two or more coders. If the measurement is perfectly reliable, every coder will produce the same results. To the extent that the results from different coders are not the same, the measurement is unreliable.

There are a number of ways to compute intercoder reliability, and they depend on the level of measurement of the content categories. If the categories are nominal, a simple percentage agreement can be computed. For example, suppose the researcher defines 50 content categories. She has two coders code an identical set of 10 messages using these categories. There are 50 judgments which must be made for each message (for each of the 50 categories, the coder is answering the question “does the message contain this content?”). By dividing the number of categories about which the coders agree that the content is either present or absent, the researcher can get a reliability figure for each message, and for the entire set of messages. Table 18-4 illustrates this process.

Intercoder correlations can be computed to get the intercoder reliability for intensity and valence scales, since in this case, the categories are coded at the interval level. Table 18-5 shows how the data might be set up if the 50 categories were measured on interval scales. There are 500 judgments made by each coder (50 categories, with one scale per category  $\times$  10 messages). The correlation between the responses for Coder A and Coder B on this set of  $N=500$  observations is an indication of the reliability of the coding. If the coders give identical ratings, the correlation will be 1.00. If the coding measurement is totally unreliable, the correlation will be 0.00. (See the next chapter for details on calculating a correlation coefficient).

Generally, the percent agreement for a good set of content categories should be above 80%. Likewise, a rule of thumb for the size of an intercoder correlation that indicates acceptable reliability is .80.

Poor reliability may indicate that the content categories are poorly defined or are too general. But it may also indicate that the content coders are not well trained. Coder training is critical to good measurement.

Table 18-4 Calculation of Percent Agreement Reliability for Nominal Categories

Message	No. of Agreements	No. of Disagreements	% Agreement
1	46	4	92
2	40	10	80
3	49	1	98
4	45	5	90
5	43	7	86
6	41	9	82
7	48	2	96
8	39	11	78
9	45	5	90
10	44	6	88
OVERALL % AGREEMENT = 88%			

### Coder Training

The first step in coder training is to provide the coders with detailed written definitions and descriptions of the content categories, and instructions on how to carry out the coding task. After the coders have studied this information, they should each code an identical sample of messages. From this data, the researcher should compute an intercoder reliability for each content category. This information will pinpoint poorly defined or vague categories. The researcher should then re-define the categories, and repeat the process until the intercoder reliabilities for the categories (and thus the overall reliability) is satisfactory.

Poor intercoder reliability might be the result of a poor coder, rather than a flaw in the coding categories or definitions. To isolate poor coders, the researcher will need more than two coders. For example, suppose three coders (A, B, and C) code the 10 messages in Table 18-5, and the intercoder correlations are:

$$r_{AB} = .80$$

$$r_{AC} = .50$$

$$r_{BC} = .40$$

Coder C's ratings are inconsistent with those of Coder A and Coder B, but both Coder A and B produce similar ratings. Coder C may not understand the coding task, or he may be unmotivated, or possibly he's just sloppy. Poor coders have to be retrained or eliminated from the research study so that their data does not contaminate the results.

Table 18–5 Calculation of Intercoder Correlation for Categories Coded at Interval Level

Message	Category	Coder A	Coder B
1	1	+5	+4
1	2	0	–1
1	3	+2	+1
...	...		
1	50	–4	–5
2	1	–2	–1
2	2	–6	–4
2	3	+4	+1
...	...		
2	50	+2	+5
3	1	–1	+1
3	2	0	0
3	3	+7	+7
...	...		
...	...		
9	50	–3	–5
10	1	–1	+1
10	2	0	0
10	3	+7	+7
...	...		
10	50	–2	0
Correlation between coder values: $r_{AB} = .82$			

### Computer Content Analysis

As we mentioned above, some content analysis can be conducted by computer program. The computer program scans the messages, which are contained in computer text files, and produces a summary. The text might be newspaper or magazine articles or conversation transcripts which have been typed into computer files, or it might be information from databases or computer archives.

Computer content analysis is restricted to very simple content judgments. For instance, a researcher might have a computer program retrieve all the Associated Press national coverage for the past year, and count the number of times the word “Poland” appears. This would give a one measurement of the amount of coverage of Poland that has appeared. Word scans can be combined to give rough definitions of more detailed content. The researcher might instruct the computer program to count the number of stories in which both “Poland” and one of the following words appears: “Communism”, “Prime Minister”, “Congress”, “Treaty”. This would give an indication of the amount of coverage of political events in Poland, and should eliminate most stories on travel or arts, or other non-political events in Poland.

If you look critically at the preceding example, you can see the difficulty of computer content analysis. It is very hard to define a set of words that will clearly differentiate Polish political stories from all others, and will still include all relevant stories. Stories about U.S. relations with Poland will probably appear in the analysis, as will unrelated stories about the Polish Communist government's repression of the arts in the 1960's, etc. Some relevant stories that do not mention Communism, prime ministers, Congress, or treaties will be missed, too.

Human coders would have no problem in dealing with these distinctions, as they have a wealth of "common knowledge" to make judgments. Computers are very thorough and very reliable, but they lack this body of common knowledge. The researcher is essentially trading expensive, sophisticated judgment (which may be somewhat unreliable) for very reliable, cheap, but simple computer analysis.

## Ethical Issues

Communication research practices often raise ethical issues. Some of these are easy to handle, while others require some difficult decisions.

The primary ethical imperative for the researcher is this: "Tell the Truth". This immediately instructs the researcher to refrain from fabricating data, or from deliberately structuring a research study so that demand characteristics or social pressure will produce the results that the researcher wants for some personal or political reason. It also demands that the researcher interpret the data objectively, and report all results accurately, whether they be good or bad (from the perspective of the researcher).

But there are other ethical issues raised by particular communication research procedures that are less clearly resolved. One of these involves concealing the researcher's intent from the research subject. "Tell the Truth" cuts both ways on this issue. In order to avoid sensitizing effects that will distort the results, the researcher must conceal something from the subject. In other words, in order to tell the truth (about the phenomenon), the researcher must avoid telling the truth (about the research study).

It is usually ethical to withhold some information from a research subject if two important conditions are met: (1) the subject is not deliberately misled about anything that could cause him stress or damage; (2) the subject is fully informed about the research project at the end of the measurement procedure, when there is no further danger of sensitization. This is sometimes called debriefing the research subject.

There is an ethical continuum at work. Very few people would have any problem with a researcher masking the intent of a set of questionnaire items by mixing them with a set of dummy items. There is no direct attempt to mislead the research subject.

Withholding the identity of the organization conducting the research is somewhat less acceptable. However, if knowing that identity is likely to bias the response, it is probably acceptable to avoid mentioning the organization's name, and, if the respondent asks, to say that "I can't tell you that, because it might affect your answers".

Another ethical gray area is identification of respondents in surveys. Again, the policy is to tell the truth, but not necessarily the whole truth. For example, if the respondent's identity must be recorded (so that a second interview can be conducted at later time, or follow-up mailings can be made), you must not explicitly promise that the survey responses are anonymous. But it is probably ethical to code the response envelopes and not mention the subject of anonymity in the cover letter. But there is a fine line between being somewhat vague in explaining the research procedures and being misleading, and the researcher must consider the question very carefully.

Concealed observation of research subjects is also an area of ethical concern. An individual's privacy should not be invaded without his consent. But in some areas of research, such as nonverbal communication, a subject's knowledge that she is being observed will completely invalidate the measurements. What should the ethical researcher do in order to make unobtrusive behavioral observations?

First, recognize the difference between observations made in public places and those which invade privacy. If a person walks down the street, he is expecting to be observed, and he has tacitly given up his right to privacy. In this case, the researcher can make unobtrusive observations without ethical concern. But a person does not expect to be observed in her home, so using a telescope to

peer through someone's windows is clearly unethical. Even quasi-public places provide some expectations of privacy that a researcher must respect. In one infamous case, a social psychologist made hidden observations of persons using a public rest room. This was generally condemned as an unethical practice. Even though the rest room was open to the public, there was a general expectation that one's behavior there was private.

In laboratory settings, there is also an implied consent by the research subject that his behavior can be observed and recorded. But it is always better to make this explicit. For example, the subject may be asked to sign a consent form which contains some language like "... Your actions and behaviors may be observed as part of the research..." By signing such a form, the subject consents to be unobtrusively observed, so there is no problem with placing a research assistant behind a one-way mirror to record behavior, for example.

Permanent recording of behavior on audio or videotape adds another wrinkle. The implied or direct consent to be observed by the subject probably covers taping, but most researchers would feel more comfortable if the subjects consented directly to the recording. Keeping a permanent record seems to be one step beyond mere observation. If the research design will not be damaged by revealing that the recording is taking place, the prescription is clear: ask for the subject's permission to make recordings in the consent form.

But in many cases, the knowledge that cameras are rolling will make subjects self-conscious or overly formal in their behavior, and this may affect the measurements. A possible compromise used by some researchers is to secretly tape the research sessions, then after the measurement is completed, ask for permission to keep the recordings. If the subject refuses, the tapes are immediately erased.

Further along the ethical continuum is deliberate deception of research subjects. It is usually not ethical to deliberately mislead or lie to the subject. We can't be categorical about this, however. Some kinds of research require that the subject be lead to believe that something is true when it is not. For example, a "conversational partner" may actually be a research assistant who is deliberately behaving in some fashion to introduce an experimental manipulation. The assistant must be presented to the subject in a false light. That's a polite way of saying that the researcher must lie to the subject. And this fact must be made clear to the respondent during debriefing.

The extent of the deception and the magnitude of its impact on the subject probably determines its ethical status. A "white lie" about a research assistant is not too bad, if the subject is told the truth at the end of the measurement. But an interviewer who identifies himself as being an FBI agent before asking a series of personal questions is probably acting unethically, even if he reveals his true identity at the end of the interview.

Once again, the two conditions mentioned above must hold before deception can ever be considered to be within the bounds of acceptability: the subject cannot be damaged by the deception, and must be informed about the deception at the end of the measurement. Even then, deliberately misleading research subjects should be a last resort. The positive benefits of the research must be weighed carefully against the negative ethical cost of misleading another person.

At the far end of the continuum is research which carries the possibility of physical or mental damage to the subject. Fortunately, this occurs infrequently in communication research. But it is not unheard of. Researchers studying effects of pornography or graphic violence, for example, must consider the possibility that the material may have a damaging impact on subjects. In these cases it is absolutely mandatory to fully inform the subject about the kinds of material or procedures to which they will be exposed. Subjects must have all pertinent information so that they can make an informed decision about whether or not to participate in the research.

Most academic institutions have a Human Subjects Committee which reviews research procedures that involve individuals. These committees normally require that subjects be informed about the research procedures, and that they sign a consent form to indicate that they have been fully informed and have decided to participate. These committees usually take a dim view of deception, and require detailed justification of its necessity before they will approve such a procedure. They usually require that a subject be informed that they can decline to participate, or can cease participation at any time. While these instructions may dismay the researcher who is trying to obtain a representative sample of subjects, they are necessary to protect individuals.

Professional organizations such as the American Association for Public Opinion Research, the American Marketing Association, the American Psychological Association, and others have guidelines for the ethical treatment of research subjects. The communication researcher should consult

these when considering ethical issues.

## Summary

Measurement techniques can be classified in several ways. One major type of measurement is behavioral observation. This kind of measurement relies on a human observer to assign values to variables. Behavioral observation can be either obtrusive or unobtrusive. In obtrusive measurement, the research subject is aware that he is being observed, while the opposite is true in unobtrusive measurement.

Obtrusiveness increases the possibility of interference between the measurement or research setting and the natural operation of the phenomenon being studied. It may also introduce sensitization that distorts research subjects' behaviors or their responses to measurement. The researcher must always strive to reduce measurement obtrusiveness whenever possible.

Behavioral observation introduces questions about measurement reliability, as the basic measurement instrument relies on human observation and judgment. Measurement procedures must be designed to make the observer's judgments as easy as possible. Clear instructions, behavioral variables that are concrete and clearly defined, and observer training all help to improve measurement reliability.

In contrast to behavioral observation, self-report measurements do not rely on the judgment of trained observers. The research subject directly provides the information which assigns values to variables. Self-report measurements can be either obtrusive or unobtrusive, but are normally the latter, since the subject is directly involved in the measurement procedure. An important exception involves naturally unobtrusive measurements, such as those obtained from archives, libraries, or databases.

Surveys and interviews are commonly used communication research measurement techniques. Surveys can be generally classified as in-person, mail, telephone, mall intercept, or computer-administered. Each has different characteristics such as response rate and cost, and the selection of the appropriate type of survey must be matched to the research problem being investigated.

Interviewers for in-person or telephone surveys must be trained carefully. Focus group interviewers require even more training, as they do not have a questionnaire to structure the interview. Focus group interviews can range freely about many topics, so they are particularly useful for exploratory research. The data from a focus group is unlike the coded data from a survey. It is usually a transcript which is summarized by judges or content coders.

Surveys use questionnaires as their primary research tool. Writing a good questionnaire is a demanding task. The author must construct an introduction which convinces the respondent to cooperate, and she must be very sensitive to issues like question wording, vocabulary, leading questions, questions which produce socially desirable responses, vague questions, and double item questions.

The presentation and type of questions must be considered, too. Complex instructions and long lists of alternative answers must be avoided. "Strange" measurement procedures must be explained clearly. The author must choose between open-ended questions whose responses must be classified and coded after respondents have replied to them, and closed-ended questions which have a fixed set of alternative responses. Open-ended questions allow more freedom of response, but they are harder to analyze, while closed-ended questions restrict responses but are easy to analyze.

The questionnaire architecture as a whole must also be considered. The author must keep the questionnaire short enough that respondent fatigue is not a factor. The sensitizing effects of question order must be considered, and question skips must be constructed so that respondents are not confused. The physical appearance of the questionnaire must be such that respondents are not confronted with a confusing or ugly document. Questionnaires must be tested before being used on the final research sample.

Content analysis is another major communication research procedure. Content can be analyzed according to the units of analysis of medium, message or story, or assertion within a message. Mutually exclusive and exhaustive content categories are first defined, and the content of a sample of units is measured by coding the appearance, frequency, or intensity and valence of the units that fall in each category.

Since content analysis, like behavioral observation, relies on human judgment, it is important to assess the reliability of measurement. This is done by computing intercoder reliability coefficients. These summarize the agreement among content coders about the presence or amount of content in each category. During the development of the content coding categorical scheme and the training of coders, these coefficients are particularly useful to detect poor categories or problems with individual coders.

In many communication research procedures, the researcher must confront ethical decisions. The primary ethical concerns of telling the truth, protecting the privacy of individuals, and protecting them from harm may conflict with the desire to conduct communication research that is free from the spurious effects introduced by the research setting or measurement procedures. The latter desire may require withholding information, or even outright deception of research subjects.

While subjects should always be informed as much as possible about research procedures, it may be ethical to temporarily withhold information or mislead subjects, if the subject is not damaged, and if she is fully informed of the truth at the end of the measurement procedure. Informed consent for taping observations is desirable. For participation in research studies which have some potential for long-term effects, this consent is mandatory.

## References and Additional Readings

- Babbie, E.R. (1992). *The practice of social research* (6th ed.). Belmont, CA: Wadsworth. (Chapter 10, "Survey Research").
- Beauchamp, T.L., Fadden, R.R., Wallace, R.J., & Walters, L. (1982). *Ethical issues in social science research*. Baltimore: Johns Hopkins University Press. (Chapter 2, "Ethical Issues in Different Social Science Methods").
- Converse, J.M. & Presser, S. (1986). Survey questions: Handcrafting the standardized questionnaire. *Sage University Paper series on Quantitative Applications in the Social Sciences*, 07-063. Beverly Hills, CA: Sage.
- Greenberg, B.S. (1981). Ethical issues in communication research. In G.H. Stempel & B.H. Westley (Eds.), *Research methods in mass communication* (pp. 255-277). Englewood Cliffs, NJ: Prentice Hall.
- Hsia, H.J. (1988) *Mass communications research methods: A step-by-step approach*. Hillsdale, NJ: Lawrence Erlbaum Associates. (Chapter 7, "Questionnaire Construction"; Chapter 8 "Mail and Telephone Surveys"; Part IV, "Historical, Qualitative, and Secondary Research").
- Lavrakas, P.J. (1987). *Telephone survey methods: sampling, selection, and supervision*. Beverly Hills, CA: Sage.
- Moser, C.A. & Kalton, G. (1972). *Survey methods in social investigation* (2nd Ed.). New York: Basic Books.
- Oppenheim, A.N. (1966). *Questionnaire design and attitude measurement*. New York: Basic Books. (Chapter 2, "Problems of Questionnaire Construction").
- Reynolds, P.D. (1982). *Ethics and social science research*. Englewood Cliffs, NJ: Prentice Hall.
- Stempel, G.H., III (1981). Statistical designs for content analysis. In G.H. Stempel & B.H. Westley (Eds.), *Research methods in mass communication* (pp. 132-143). Englewood Cliffs, NJ: Prentice Hall.
- Webb, E.J, Campbell, D.T., Schwartz, R.D., & Sechrest, L. (1972). *Unobtrusive measures: Nonreactive research in the social sciences*. Chicago: Rand McNally.