12

Collecting, Analyzing, and Reporting Quantitative Data

Chapter Contents

The best survey is easy to answer, easy to analyze, easy to report, and meaningful. Although most of the work goes into the writing of questions themselves, research managers also need to think ahead to analysis and reporting strategies. Designing surveys and training interviewers with data entry in mind can improve the speed and reliability of results. Planning what results need to be reported can aid analysis and ensure that data are collected in a way that makes it possible to create necessary tables and figures. Once the results become available, the manager can think about how to share good, bad, or surprising news with the client.

DESIGNING SURVEYS FOR EASY DATA ENTRY

Several formats can be used to create easily analyzed questionnaires. Many survey research centers, for example, use CATI (computer-assisted

telephone interviewing) systems, in which telephone or personal interviewers enter data directly into the computer as they talk with respondents. Managers also can use do-it-yourself systems, some of which now are available online from sites such as Zoomerang (info.zoomerang.com) and Survey Monkey (www.surveymonkey.com), which provide question templates for survey designers that range in sophistication (and cost) and can be customized. Statistical packages such as Statistical Package for the Social Sciences provide more sophistication, supplying software for questionnaire design, data entry, and powerful analyses. The cost for these packages ranges from several hundred dollars to tens of thousands of dollars.

CATI or computer-assisted data collection (CADAC) can eliminate the need for cumbersome data-entry procedures. Interviewers using a CATI system typically wear a headset and ask questions prompted by the computer. As they type in the answers from each participant, the computer automatically updates the database to incorporate the new information. This type of system also is growing more popular for personal interviews, particularly for sensitive issues. Survey computer-aided data entry can reduce data entry errors by 77%. In addition, it can reduce confusion on complicated questionnaires that include conditional, or branching, questions. *Conditional questions* separate participants into groups depending on their response. For example, consumers who answer "yes" to a question about having previously purchased a company's product can go on to answer questions about the product's quality and the company's responsiveness. Consumers who answer "no" answer a different set of questions that may focus on the potential interest in the company's or the competitors' products. CADAC or CATI systems now are common among professional research firms. Even when researchers do not have access to fully automated data collection systems, they often can make use of scanable response sheets on which respondents mark answers by filling in the appropriate circles.

If survey questionnaires must be entered manually, several design issues can make data entry more efficient. Even if researchers have access to automated data entry, these principles can make it easier to analyze and interpret the results.

Directionality

As mentioned in chapter 11, attention to directionality can make results more intuitive to report. *Directionality* refers to the match between the numbers that represent answers to questions in the computer and the idea they symbolize. For example, on a Likert scale, a 5 is used to represent "strongly agree" and a 1 is used to represent "strongly disagree" so that a high score means more agreement instead of less agreement. Always

use a high number to indicate a positive response, even if the question is negatively worded. For example, a higher number should represent "very often" on a scale measuring "how often have you had difficulty receiving your order on time?"

Coding and Code Books

Some questionnaire designers include the numbers for data entry purposes on the questionnaires themselves, whereas others rely on a separate code book. The *code book* is a copy of the questionnaire with annotations that direct data entry personnel and analysts. The annotations include short variable names—abbreviations for the questions—along with information on where in a computer data set the answers appear and what numbers represent each answer. We address each of these issues separately.

The advantage to including numbers on the questionnaire is that data entry personnel see a constant reminder of the codes they must enter, making memory lapses less likely. This can help considerably on complicated questionnaires. Numbers also can help respondents follow along when answers include a lot of response options. In the example below, note how much easier it is to find the center of the scale when it includes numbers instead of blanks:

Which of the following best describes our customer service representatives?

Incompetent	1	2	3	4	5	6	7	8	9	Competent
Impolite	1	2	3	4	5	6	7	8	9	Polite
Unhelpful	1	2	3	4	5	6	7	8	9	Helpful

Which of the following best describes our customer service representatives?

Incompetent	_	_	_	_	_	_	_	_	_	Competent
Impolite	_	_	_	_	_	_	_	_	_	Polite
Unhelpful	_	_	_	_	_	_	_	_	_	Helpful

Questionnaire designers must be careful, however, to ensure the questionnaire does not become so cluttered with numbers that it becomes difficult to read or seemingly impersonal. Usually, some codes appear only in the code book. Telephone questionnaires can include more codes, as long as the instrument remains easy for interviewers to read. It helps both interviewers and data entry personnel to have sets of answers clustered in small groups, such as in the following example:

OK, thinking again about some election issues, please think now about the public schools where you live, and tell me how good a job your schools are

doing educating students in each of the following areas. Use a scale from 1 to 5, with 1 meaning *very poor* (VP) and 5 meaning *very good* (VG):

	VP				VG	
26. The first area is basic subjects such as reading, writing, and math	1	2	3	4	5	DK/RF
27. Training students for a specific job?	1	2	3	4	5	DK/RF
28. Preparing students for college?	1	2	3	4	5	DK/RF
29. Teaching moral values?	1	2	3	4	5	DK/RF

Note that the example includes abbreviations for the interviewers (such as DK/RF for "don't know or refused to answer") that respondents will not need to hear or see.

Edge Coding

Usually, variable names and locator information for the data set in the computer appear only in the code book and in the margins. The use of margins for coding information is known as *edge coding*. These codes include the following:

1. *Tracking information.* Often, survey supervisors want to keep a record of which interviewer completed each questionnaire, on which day the survey took place if data collection occurs over several days or weeks, and any other information of background importance. Each set of answers is given an identification number, which makes it easier to check for and correct mistakes in data entry. This information usually appears in the top margin of a questionnaire, as shown below:

Caller: _____ Date:_____ Phone No.: _____ ID #:_____

College: _____ Year:_____ Last Gift to:_____

Phase III Questions for University Foundation Fall 2005

First, I am going to read a short list of ways you might hear about Superior University. For each one, please tell me if you have heard about SU through this source *often, sometimes, rarely*, or *never* during the past year.

2. *Question information.* Most information regarding where answers to each question will appear in the computer go in the left margin next to the questions, as shown below:

We're already halfway through our survey, and we really appreciate your help. Now, if you have declined to give money to SU at any time in the past, please tell me if each of the following issues have been *very important,*

somewhat important, not very important, or *not at all important* reasons for your decision.

				__(8)__		(N/A, Have never refused to give)	
		4	3	2	1	8	9
(32) GETTO	27. You just didn't get around to it	VI	SI	NVI	NAI	N/A	RF/DK
(33) AFFORD	28. You felt you couldn't afford it	VI	SI	NVI	NAI	N/A	RF/DK
(34) DISSAT	29. You are dissatisfied with the quality of education at SU	VI	SI	NVI	NAI	N/A	RF/DK
(35) NODIFF	30. You feel your gift would not make a difference to students at SU	VI	SI	NVI	NAI	N/A	RF/DK

Note that the abbreviations for each question appear in all capital letters and require no more than eight letters. Many computer programs cannot accommodate variable names longer than eight letters. The information in parentheses refers to the location of the question responses in the data set. Most often this number points to the cell in a spreadsheet.

3. *Answer codes.* The numbers at the top of each column refer to the codes used by the computer to indicate the answers to the questions regarding barriers to giving. A 2 indicates "not very important," a 4 indicates "very important," and so on.

It is important to be able to find a particular questionnaire quickly in case of data entry mistakes. The identification number makes it possible to isolate the questionnaire that needs to be reentered. Without an identification number on each questionnaire (usually in the top right-hand corner, where the example says "ID #"), corrections can become an enormous headache. Keep in mind that some corrections always will be needed, thus the process of fixing mistakes in the data set has a name: *data cleaning*.

Data Entry Conventions

You can assign any number to any answer, but many people are accustomed to seeing things a certain way. For example, it is common to assign a 1 to a "yes" answer and a 0 to a "no" answer. In a series of items for which respondents "check all that apply," each checked response would get a 1, and each blank response would get a 0.

Most questions do not need more than seven or eight response categories. As a result, another conventional number code is 8, which indicates "don't know." For refusals, some people use a 9, and others use a blank.

Remember that the code book must account for each answer. Some questions will have multiple answers ("check all that apply"), each of which will need to be coded as a separate item, as shown below:

50. What is your race or ethnicity? (MARK ALL THAT APPLY)

_____1. African American
_____2. Asian
_____3. Hispanic
4. Native American_____
5. White_____
6. Other_____

[CODE AS 1/0] 9. RF

Open-Ended Questions

Most computer programs do not lend themselves to *qualitative analysis*, meaning the analysis of words or pictures instead of numbers, so that data entry of open-ended data usually requires a special program. Open-ended questions usually are skipped during data entry of closed-ended questions. A separate word processing file can be used to keep track of identification numbers and open-ended comments for each question. A file or notebook also must be kept for interviewers and data-entry personnel to log problems or anomalies that will require decisions from the research manager, such as what to do with a respondent who has checked the space between two responses instead of checking one response clearly.

TRAINING INTERVIEWERS

Interviewers must act polite even if respondents do not, read slowly and clearly, avoid lengthy conversations, and read each question in a consistent way. Interviewer training is an essential part of reliable data collection. Interviewer instructions vary slightly depending on the organization, the project, and the facilities, but some general guidelines apply to most situations. In their classic survey research text, Warwick and Lininger (1975) suggested several keys to standardized personal interviewing, which apply equally well to telephone interviews.

1. *Use the questionnaire carefully, but informally.* The questionnaire is a tool for data collection. Interviewers must be familiar with the purposes of the study and the questionnaire, including question order, question wording, skip patterns, and the like. Interviewers who are well prepared can take a relaxed, informal approach to their work. This helps maximize interviewers' ability to collect high-quality data and avoid carelessness in the interview process.
2. *Know the specific purpose of each question.* Interviewers need to understand what constitutes an adequate response to each question to satisfy

the purposes of the research project and to improve their use of the questionnaire. This information must be discussed as a part of interviewer training and reviewed before data collection begins.

3. *Ask the questions exactly as they are written.* Even small changes in question wording can alter the meaning of a question and a participant's response. Researchers must assume each participant has answered the exact same question. The consistent, unbiased wording of each question provides a strong foundation for the accuracy and reliability of study results. Neutral comments such as, "There are no right or wrong answers; we just want to know your opinion," should be used sparingly and only when interviewer feedback is required.

4. *Follow the order indicated in the questionnaire.* The order of questions in a survey instrument has been purposefully determined and carefully pretested. Arbitrary changes made in question order reduce the comparability of interviews and potentially introduce bias into questions that are sensitive to question order.

5. *Ask every question.* Interviewers need to ask every question, even when participants have answered a previous question or make comments that seem to answer a later question. Respondents' answers to questions often change as a result of small changes in question wording. In addition, the intent of questions that seem similar often are different. Researchers develop and pretest question wording carefully and with a specific purpose in mind. Unless respondents terminate an interview early, each question must be asked of each respondent.

6. *Do not suggest answers.* Interviewers must never assume to know a respondents' answer to a question, even after a respondent has answered seemingly similar questions in a consistent manner. All answers must be provided by the respondent.

7. *Provide transitions when needed.* A well-written questionnaire needs to contain transitional phrases that help the respondent understand changes in topics, question types, or question response categories. Interviewers use these transitional phrases to help guide a respondent through a questionnaire.

8. *Do not leave any question blank.* Interviewers need to make every effort to have participants answer every question, except those intentionally left blank because of skip patterns. Although researchers may choose to use a questionnaire even if questions are left blank, omitted questions reduce the reliability and external validity of survey results. It is best if each respondent answers every applicable question.

CALL SHEETS

Telephone surveys typically use call sheets that have lists of numbers and places to record the outcome of each call attempt (Fig. 12.1). Sometimes call sheets include special information such as giving history for donors or an

Telephone Call Sheet
UNIVERSITY FOUNDATION, FALL 1998

CALLER'S NAME: _____ DATE: _____

1st CALLBACK CALLER'S NAME: _____ DATE: _____

2nd CALLBACK CALLER'S NAME: _____ DATE: _____

Phone #	Name										When?	
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____
_____	_____	Coll: ____	Year: ____	Last Gift To: _____	CM	RF	NA	BZ	DIS	CB:	_____	OTH: _____

Colleges: 1) CAHE 2) Bus & Econ 3) Educ 4) El Eng 5) Eng & Arch 6) Lib Arts 7) Phar 8) Sci 9) Vet Sc 0) Other/None
Last Gift Categories: 1) Athletics 2) Academics/Colleges 3) Students/Residence/Scholarships 4) Other 5) Alumni Assn

FIG. 12.1. Sample telephone call sheet for a survey of university donors. A call sheet provides researchers with a detailed record of every call attempt. In this survey, specific individuals are being called, and records exist regarding the individual's giving history to the university. Coll = college, with numbered options provided at the bottom of the calling sheet; year of graduation from the university; Last gift to = the designation applied to the donor's most recent gift, with codes provided at the bottom of the calling sheet; CM = completed interview; RF = refused interview; NA = no answer; BZ = busy signal; DIS = disconnected number; CB = call-back appointment has been made; OTH = provide an opportunity for callers to jot down other observations or difficulties associated with the call attempt.

individual's code number if an organization is calling a different individual—such as a spouse or a parent—to obtain more complete information from a household. The basic information appearing on most call sheets includes the following:

CM = Completed interview
RF = Refused
NA = No answer (sometimes AM appears separately to indicate answering machine)
BUS = Business/beeper/fax/modem
BZ = Busy signal
TM = Terminated
DIS = Disconnected/out of service
LB = Language barrier
NE = No eligible respondents at this number
CB = Call back appointment (interviewers fill in the details
OTH = Other (interviewers fill in the details)

TIMING OF TELEPHONE SURVEYS

The optimum times of day and days of the week for doing telephone surveys vary depending on geographic region and time of the year. Although calling on Sunday evenings usually works well, a manager will not want to plan a survey for Super Bowl Sunday, for example. Generally, however, weeknights between 6:00 p.m. and 9:00 p.m. are considered reasonable hours for calling.

RESPONSE RATES

Research reports must include information regarding how many people approached for the survey actually completed it. Because probability sampling and inferential statistics assume 100% of eligible respondents will participate, people evaluate the validity of conclusions based partly on the response rate obtained. Because so many survey methods exist, the American Association for Public Opinion Research (AAPOR) has released a white paper discussing a myriad of specialized issues and methods for reporting outcome rates for surveys using mail surveys, random-digit dialing (RDD), in-person interviews, and other methods (AAPOR, 2004). Most practitioners generally will find the following several pieces of information sufficient to report outcomes in a way that complies with AAPOR's ethical guidelines:

1. *The total sample size.* This is the total number of people (or phone numbers or addresses) used in the study. For example, for an RDD survey, a manager might draw a sample of 2,500 random phone numbers to achieve

a final *n* of 800. Chapter 6 addresses how to draw the right size of total sample to achieve the desired valid sample.

2. *The valid sample size.* This is the number of sample elements (e.g., individuals, households, companies) in the total sample that remain after removing invalid phone numbers or addresses. These include nonworking phone numbers, respondents who do not fit the profile necessary to be included in the sample (e. g., not registered to vote), and respondents who do not speak English when surveyors do not speak other languages. In addition, research managers typically remove phone numbers and addresses from a sample if interviewers have made three to five attempts to reach someone at that location. The eliminated numbers are considered unreachable. Some market research firms only make one or two attempts, called *call-backs* in phone-survey lingo, but this practice is questionable scientifically. If an insufficient number of attempts have been made before the completion of the study, the location is considered a noncontact location and must remain in the valid sample. It is in the research manager's interest to minimize the number of noncontacts because they lower the response rate and can raise questions about the study's quality:

$$\text{Valid Sample Size} = \text{Total Sample Size} - \text{Ineligibles} - \text{Unreachables}$$

3. *The completion rate.* This is the number of people who complete a survey out of the *total* sample size. If the researcher can anticipate the completion rate, sometimes called the *minimum response rate* (AAPOR, 2004), this will determine how big a sample must be drawn to achieve the desired number of completed surveys. For example, if the completion rate is 31% and a sample of 384 is needed, the manager determines, by dividing 384 by .31, that 1249 numbers will be required to end up with 384 completed surveys:

$$\text{Total Sample Size} = \frac{\text{Target } n \text{ of Completed Surveys}}{\text{Completion Rate in Decimal Form}}$$

$$\text{TSS} = \frac{384}{.31} = 1{,}249$$

4. *The response rate.* This is the number of people who completed a survey out of the *valid* sample size. This number, also called a *cooperation* rate by AAPOR (2004), is the number most people want to see. Because the valid sample size is smaller than the total sample size, this number shows that a survey's quality is better than the completion rate makes it seem. Research managers strive to keep this number as high as possible. Unfortunately, phone surveys these days often have response rates as low as 45% to 55%, although some achieve rates as high as 80%. Mail surveys with a single

mailing (no reminders) usually garner only about a 30% response rate:

$$\text{Response Rate} = \frac{\text{Completed Questionnaires}}{\text{Valid Sample Size}}$$

5. *The refusal rate.* This is the number of people who declined to answer the survey out of the valid sample size. Research managers strive to keep this number low:

$$\text{Refusal Rate} = \frac{\text{Refusals}}{\text{Valid Sample Size}}$$

6. *The noncontact rate.* This is the number of people who could not be reached out of the total sample and, therefore, never had the opportunity to complete the survey or refuse to do so. They cannot be eliminated from the valid sample:

$$\text{Noncontact Rate} = \frac{\text{Noncontacts}}{\text{Total Sample Size}}$$

REPORTING UNIVARIATE RELATIONSHIPS

The minimum information usually required for a research report includes frequencies, percentages, means, and, for some clients, standard deviations. The frequencies tell the client how many people answered each question using each response. Frequency tables usually include percentages as well, so the reader can make informed comparisons across questions.

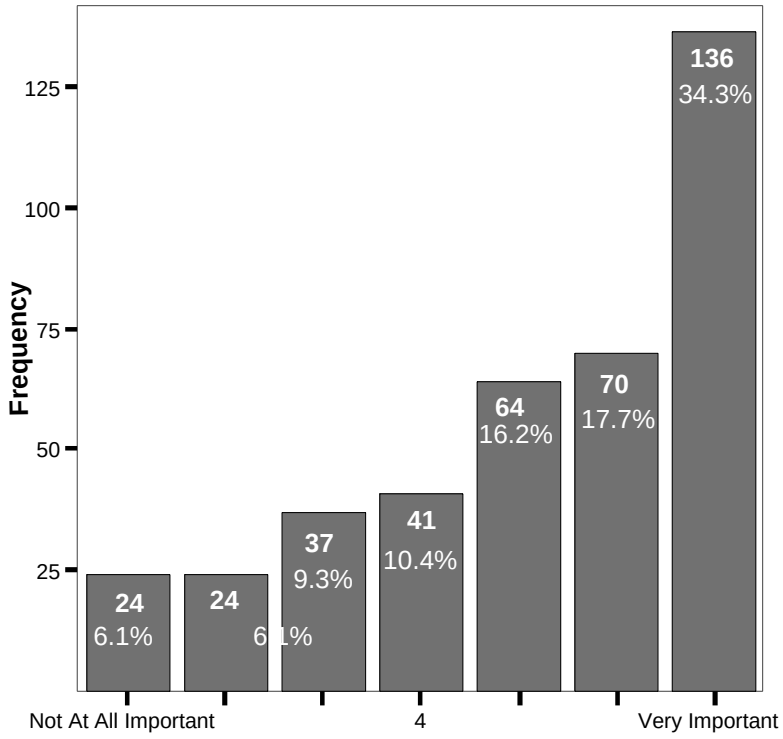
As shown in Table 12.1, a frequency table includes both the number of people who answered the question and the number who did not (called *missing*). The valid percentage is based on the number of people who answered the question (338), whereas the unadjusted percentage refers to the total sample size (400). In this example, the numbers used on the survey appear along with the descriptions of the responses, but most research reports include only the descriptions.

Researchers also can present frequencies visually, using tables, bar charts, histograms, or pie charts. Table 12.2 displays a bar chart, which often can give clients a better grasp of the range and strength of responses than they might get from just a recitation of mean and standard deviation. Pie charts, meanwhile, can communicate the contrasts and similarities between the usefulness of information sources, as shown in Figure 12.2. This particular example illustrates that registered voters tended to consider both daily news reports and interpersonal contacts as important sources of election information in 2004, but they were less likely to discount interpersonal sources than mediated ones.

TABLE 12.1
Sample Frequency Table: Purpose of Most Recent Visit to Convention Center

Valid	*Frequency*	*Percent*	*Valid Percent*
1.00 Attend a sporting event	121	30.3	35.8
2.00 Attend a musical or theater performance	54	13.5	16.0
3.00 Attend a fair or exhibit	51	12.8	15.1
4.00 Attend a rally or workshop	30	7.5	8.9
5.00 Attend a graduation ceremony	17	4.3	5.0
6.00 Attend a private social function	13	3.3	3.8
7.00 Get information	10	2.5	3.0
8.00 Other	42	10.5	12.4
Valid Total	**338**	**84.5**	**100.0**
Missing	**62**	**15.5**	
Total	**400**	**100.0**	

TABLE 12.2
Support for Tobacco Use Prevention Among Washington State Registered Voters in the 2004 Presidential Election



Note: Data excerpted from a random telephone survey of registered Washington State voters ($N = 397$) conducted October 24–28, 2004 (Pinkleton & Austin, 2005). Responses range from 1 (not at all important) to 7 (very important), with 4 representing the midpoint of the scale.

REPORTING RELATIONSHIPS AMONG VARIABLES

Often, a client wants to compare results for several groups, such as education levels, gender, ethnicity, or membership. The crosstab table usually provides the most intuitive presentation format. Crosstab tables usually present frequencies and column or row percentages. The layout of the table will encourage readers to interpret the data in a particular way. Note that in the first example, shown in Table 12.3, the table is arranged so that the row percentages display how many participants of each political orientation support tobacco use prevention efforts. In the second example, the row percentages display how many of the low, medium, and high supporters come from each political orientation. The first table shows more dispersed support among conservatives but strong support among moderates and liberals. This information can help communication managers understand which target groups need persuasion and which need reinforcement. The second table, shown in Table 12.4, shows that high supporters tend to cross the political spectrum fairly equally but that low support is disproportionately conservative. This information suggests that tobacco prevention represents an issue on which people from varied political orientations tend to agree. This could make tobacco use prevention a useful issue to emphasize when trying to unify the electorate.

TABLE 12.3
Political Orientation and Level of Support for Tobacco Prevention in the 2004 Presidential Election Among Washington State Registered Voters: Row Percentage Example

	Tobacco Use Prevention			
Political Orientation	*Low Support*	*Medium Support*	*High Support*	*Total*
Conservative				
Count	44	15	78	137
Row Percentage	32.1%	10.9%	56.9%	100.0%
Moderate				
Count	27	10	102	139
Row Percentage	19.4%	7.2%	73.4%	100.0%
Liberal				
Count	12	16	79	107
Row Percentage	11.2%	15.0%	73.8%	100.0%
Total				
Count	83	41	259	383
Row Percentage	21.7%	10.7%	67.6%	100.0%

Data from a random telephone survey of registered Washington State voters ($N = 397$) conducted during October 24–28, 2004 (Pinkleton & Austin, 2005).

TABLE 12.4
Political Orientation and the Level of Support for Tobacco Prevention: Column Percentage Example

	Tobacco Use Prevention			
Political Orientation	*Low Support*	*Medium Support*	*High Support*	*Total*
Conservative				
Count	44	15	78	137
Column Percentage	53.0%	36.6%	30.1%	35.8%
Moderate				
Count	27	10	102	139
Column Percentage	32.5%	24.4%	39.4%	36.3%
Liberal				
Count	12	16	79	107
Column Percentage	14.5%	39.0%	30.5%	27.9%
Total				
Count	83	41	259	383
Column Percentage	100.0%	100.0%	100.0%	100.0%

Data from a random telephone survey of registered Washington State voters ($N = 397$) conducted during October 24-28, 2004 (Pinkleton & Austin, 2005).

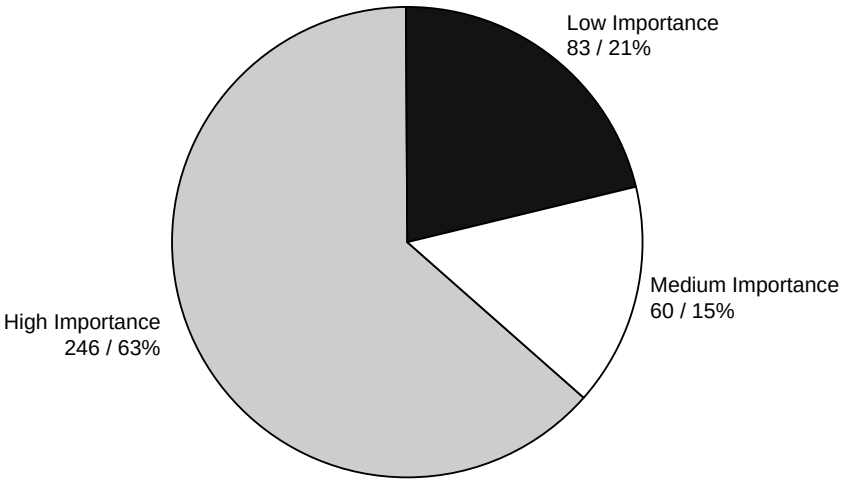
The research manager needs to determine which presentation provides the most helpful information to the program planner. Although it is possible to include both types of information in the same table, by displaying both row and column percentages and even total percentages, too much information can create confusion. Evaluate the communication programming needs carefully before creating the tables.

Crosstab tables only lend themselves to breakdowns across a limited number of categories, such as low/medium/high or male/female. For a variable such as age, which in a standard public opinion survey can range from 18 to 94 or more, managers need to condense it into decades or market-oriented categories (such as 18–34) to make the crosstab table interpretable. Strive to make tables readable at a glance.

Statistics such as the *Chi-square* can be used to highlight especially notable differences across groups. Many clients, however, do not want to see the statistics themselves and may find too much statistical information intimidating. Keep the presentation interpretable while ensuring statistical rigor behind the scenes. Meanwhile, always be prepared for the client who appreciates statistical depth.

Another way to present relationships is through a correlation coefficient. Correlations are useful for examining the strength and direction of relationships between two interval-level (e.g., "very interested" to "not at all interested") or ratio-level (e. g., "0 times to 7 times") variables. Correlations are less intuitive for a client without a statistical background, however. It can be useful, therefore, to run correlations as background while presenting

The Importance of Conversation With Family and Friends for the Washington State Registered Voters in the 2004 Presidential Election



The Importance of Daily News for the Washington State Registered Voters in the 2004 Presidential Election

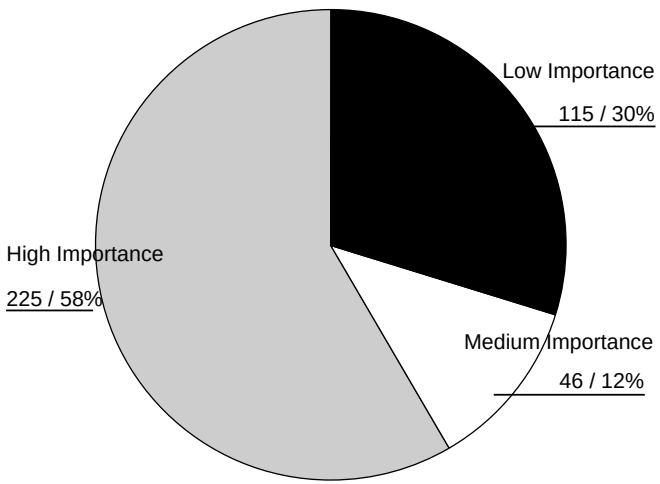


FIG. 12.2. Pie charts displaying registered voters' use of interpersonal and mass media information sources in the 2004 election. Data excerpted from a random telephone survey of registered Washington State voters ($N = 397$) conducted October 24–28, 2004 (Pinkleton & Austin, 2005).

condensed versions of the scales in a crosstab table. Condensing the data may mask subtle relationships; check the tables to make sure that they do not seem to tell a different story from the original statistics.

A correlation coefficient ranges between −1 and +1. A −1 indicates that the two variables are exact opposites. For example, if younger children always like an event featuring Cookie Monster more, people always dislike Cookie Monster more as they get older. A + 1 indicates the two variables are perfect matches, for example, if the price people will pay for a product increases exactly in proportion to how valuable they think the product is.

Of course, such perfect relationships do not exist. The statistician, therefore, looks to see whether the coefficient is "significantly" different from 0, which would indicate that two variables change with no relevance to each other. The closer the coefficient is to +1 or −1, the stronger the relationship is between the two variables. In the information-seeking example, examining the relationship between age and interest level, the correlation between the two original variables is about −.17, indicating a small and negative association between age and interest: Older people, in other words, have slightly less interest than younger people have, but the difference is not dramatic.

Research managers may encounter times when they need to consider complex relationships using sophisticated multivariate statistics. For the most part, this sort of analysis still needs to be translated into results interpretable by a statistical novice or math phobe. Keep in mind that even the most sophisticated analysis is useful only if it is understandable, and the most prescient research is helpful only if it gets used. Keep the presentation as simple and compelling as possible.

FINAL THOUGHTS

In the blur of program-planning deadlines, the communication manager can be tempted to plunge into a research project without thinking through the details of data entry or later presentation. The more communication managers think ahead, however, the more useful the final report is likely to be. The research plan can serve as an invaluable tool for determining what a research report should look like. Managers often map out the final report before doing the research, demonstrating—without the numbers, of course—what the answers to the questions raised in a situation analysis should look like. Planning to this level of detail can help ensure that the questions asked on a survey are designed to make it possible to create the tables desired. In addition, planning ahead for data entry and analysis helps focus the research manager's work and can save both time and money. Just as effective communication program plans focus on the final outcomes—the goals and objectives—from the start, the most effective research projects envision the final report well before the first survey responses are collected.

SIDEBAR 12.1
American Association for Public Opinion Research

III. Standards for Minimal Disclosure
Good professional practice imposes the obligation upon all public opinion researchers to include, in any report of research results, or to make available when that report is released, certain essential information about how the research was conducted. At a minimum, the following items should be disclosed:

1. Who sponsored the survey, and who conducted it.
2. The exact wording of questions asked, including the text of any preceding instruction or explanation to the interviewer or respondents that might reasonably be expected to affect the response.
3. A definition of the population under study, and a description of the sampling frame used to identify this population.
4. A description of the sample selection procedure, giving a clear indication of the method by which the respondents were selected by the researcher, or whether the respondents were entirely self-selected.
5. Sample sizes and, where appropriate, eligibility criteria, screening procedures, and response rates computed according to AAPOR Standard Definitions. At a minimum, a summary of disposition of sample cases should be provided so that response rates could be computed.
6. A discussion of the precision of the findings, including estimates of sampling error and a description of any weighting or estimating procedures used.
7. Which results are based on parts of the sample, rather than on the total sample, and the size of such parts.
8. Method, location, and dates of data collection.

From time to time, AAPOR Council may issue guidelines and recommendations on best practices with regard to the release, design and conduct of surveys.

Used by permission from the American Association for Public Opinion Research. For more information about the American Association for Public Opinion Research, go to www.AAPOR.org.