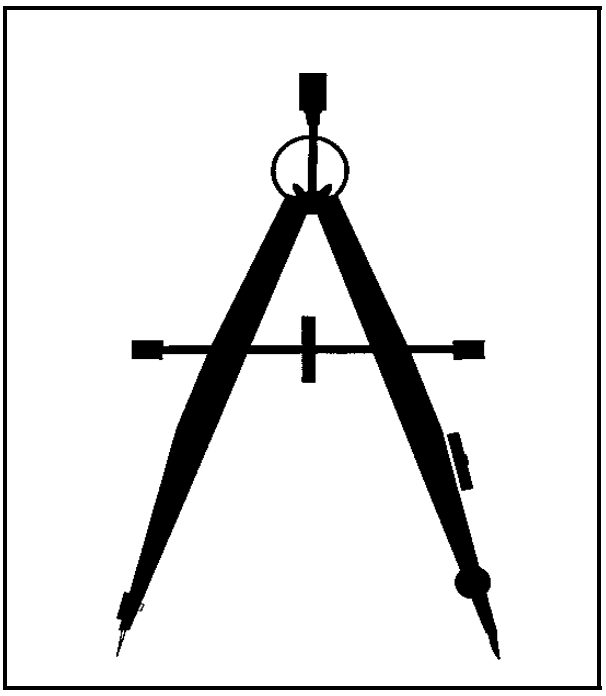
Air University
Sampling and Surveying
Handbook



Guidelines for planning, organizing,
and conducting surveys

Revised (PDF) Edition: May 2002

Previous Editions: 1996, 1993, 1988, 1985

FOREWORD

This handbook contains guidelines for planning, organizing, and conducting surveys. It should be useful to anyone embarking on a project requiring the gathering of data through the medium of the questionnaire. The text is designed to be easily readable, even for someone with a limited background on the subject.

The book is the product of the efforts of several people. Major Keith C. Ross did the majority of the work in fulfillment of his research requirements while a student at the Air Command and Staff College, Air University, Maxwell Air Force Base, Alabama in 1977. Lt Col Lawrence D. Clark was the research advisor for the project, and Major Thomas C. Padgett did the final editing and assembling on the original edition. Dr. Thomas R. Renckly, Air University Curriculum Coordinator, edited the 1988 reprint of the first edition, the 1993 second edition, and the 1996 Internet edition, in addition to providing supplemental information on bias in survey research (Chapter 5) and common statistical analysis errors (Chapter 6). If you have any questions about the book, Dr. Renckly can be reached at:

HQ AU/CFA
55 LeMay Plaza South
Maxwell AFB, AL 36112-6335
(334) 953-2989 or DSN 493-2989
thomas.renckly@maxwell.af.mil

PREFACE

To say that people's opinions and attitudes are more important now than ever before is very nearly an understatement. More and more we are seeing individuals and groups relying heavily on the opinions and attitudes of customers, constituents, concerned citizens, focus groups, etc., to provide information for decision making. It is also obvious even to the casual observer that surveys (including paper-based questionnaires, personal interviews, and telephone polls) play a crucial role in gathering these opinions and attitudes.

Surveys are also used as evaluation and control devices. They can be used to measure the effectiveness of an ongoing project, such as an information program, for example. By surveying the participants in a program, the effectiveness of the program can be determined. Also, management can use surveying as an aid to control, by finding new problem areas and insuring that old problem areas have been corrected (which is, for instance, one of the fundamental premises of total quality management).

The need for accurate information to fuel the decision-making process exists at all levels of management. This has created a trend for surveys to be generated at lower management levels by staff officers, many of whom are not experienced in survey development or administration. The growing necessity to survey and the relative lack of knowledge on surveying methodology leads to a significant demand for information on the subject. The primary purpose of this guide is to supply this information in simple, non-technical language.

An equally important purpose of this guide is to identify problems that may arise during development of a survey and to provide techniques and guidance for solving these problems. The procedures presented in this guide are designed to help you develop valid and useful surveys.

The steps in surveying are varied and complex. Therefore, this guide only highlights the major information, techniques, and procedures available to the surveyor. References offering more detailed treatments of these subjects are provided in the bibliography and appendices. Since the

surveyor frequently is unable to reach the entire group in which he is interested, this guide explains sampling techniques. To use these techniques necessitates only a rudimentary knowledge of statistics. Finally, although many of the techniques and procedures covered here apply equally well to the personal or telephone interview survey, the primary focus is on the self-administered and group-administered surveys.

One word of caution: Because the steps in survey preparation are closely interrelated, you should study this entire guide before beginning an initial survey effort.

TABLE OF CONTENTS

CHAPTER 1

Introduction to Surveying

Webster defines a survey as "the action of ascertaining facts regarding conditions or the condition of something to provide exact information especially to persons responsible or interested" and as "a systematic collection and analysis of data on some aspect of an area or group." A survey, then, is much more than the mere compiling of data. The data must be analyzed, interpreted, and evaluated. Only after this processing can data become information. The "exactness" of the information is determined by the surveyor's methods. Unless he makes a systematic collection of data, followed by a careful analysis and evaluation with predefined objectives, his collection of data cannot become "exact" information.

TYPES OF SURVEYS

Surveys can be divided into two general categories on the basis of their extensiveness. A complete survey is called a "census." It involves contacting the entire group you are interested in -- the total population or "universe." The other category is more common; it is a sample survey. A sample is a representative part of a whole group (universe). Thus a sample survey involves examining only a portion of the total group in which you are interested, and from it, inferring information about the group as a whole.

ADVANTAGES & DISADVANTAGES OF THE TWO TYPES OF SURVEYS

One of the decisions to be made in surveying is whether or not to sample. Parten (1950, p 109) presents a list of advantages and disadvantages of the sample survey. (These, in turn, imply the advantages and disadvantages of the census.) The three most important

considerations for the surveyor are: speed, low cost, and increased accuracy and analysis of the data.

By sampling only a small portion of a large population, it is possible to collect data in far less time than would be required to survey the entire group. Not only is data collection quicker, but data processing and analysis also require less time because fewer pieces of data need to be handled. Rapidly changing conditions and the short turn-around time imposed in many surveys make the efficient use of time a critical variable. If an accurate snapshot of the attitudes of a particular group is desired, currency is of paramount importance. Professional political pollsters make their living by providing quick snapshots of the "political climate." Results of such polls lose their accuracy very quickly (sometimes in as little as 24 hours--particularly in the days preceding a major election). So, for these pollsters, time is truly of the essence. It's probably a safe bet that those of you reading this guide will not need that degree of speed. Nevertheless, speed is essential to ensure the data are "fresh," especially when it comes to assessing public opinion in a volatile or contentious area before they change appreciably.

The smaller amount of data gathered by sampling as opposed to surveying an entire population can mean large cost savings. By limiting the group to be surveyed, less time, hence less cost, are involved in collecting, formatting, and analyzing the data. In addition, if automated data processing (ADP) equipment is being used to analyze data, your overall time investment will be even less, as will be the overall cost. Sampling allows you to do a credible job for a smaller investment of time and money.

Parten (1950) also notes that sampling enables the surveyor "to give more attention to each return received and to make certain that the data are as accurate as possible" (p 110). This attention may lead to more precise information than would a less careful collection of data from the entire population. Nothing more than a rudimentary quality control is possible for the great volume of raw data gathered in a census. The more data collected, the greater the potential for making "accounting" errors.

The disadvantages of sampling are few, but important. The main disadvantages stem from risk, lack of representativeness, and insufficient sample size, each of which can cause errors. Inattention to any of these potential flaws will invalidate survey results.

It is important to realize that using a sample from a population to infer something about the entire population involves a risk. The risk results from dealing with partial information. If risk is not acceptable in seeking the solution to a problem or the answer to a question, then a complete survey or census, rather than a sample survey, must be conducted.

Determining the representativeness of the sample is the surveyor's greatest problem when sampling. By definition, "sample" means a representative part of an entire group. To avoid the charge of using "biased data," it is necessary to obtain a sample that meets the requirement of representativeness, and this is not an easy task. Without a representative sample, a survey will, at best, produce results that are misleading and potentially dangerous. Procedures for minimizing the possibility of using a nonrepresentative sample are covered in Chapter 4.

The final major problem in sampling is to determine the size of the sample. The size of the sample you need for a valid survey depends on many variables including the risk you are willing to accept and the characteristics of the population itself. The determination of sample size is discussed in Chapter 4. Here, it is sufficient to say that if sampling becomes too complicated, or the required sample size becomes too large, the easiest solution may be to survey the entire population.

The decision as to whether to survey the entire population or only a sample of it is not based on the above advantages and disadvantages alone. It is affected by many other variables that are covered later in this guide.

TO SURVEY OR NOT TO SURVEY

Before attempting a survey, you should investigate some basic facts and answer some pertinent questions. The result of this investigation will be a greater realization of the work involved in producing a survey. Perhaps it will lead to a decision not to survey.

Surveys demand time(maybe more time than you have available. The exact amount of time varies greatly from survey to survey depending on the number of people to be surveyed and the content of the survey. A survey of a few questions administered to the people in your office may take only a day or so, whereas a larger survey administered to a great number of people located worldwide can take over three months from the

time the survey is delivered to the printer (see Appendix B). And this does not include the time needed to design the survey and construct the questionnaire. Moreover, coordination with officials and the customers of the survey takes additional time. If your estimate of the time needed to produce the survey exceeds your deadline date, you are likely to decide you do not have the time to conduct a survey. A hurried survey wastes both your time and that of your respondents. The results of a hurried effort are questionable at best.

Surveys are expensive to produce. The solution to the problem or the answer to the question may not be worth the cost to produce it. Even if it would be worth the price, you may not be able to obtain the needed funds, either from your own pocket or from your organizational budget. Although no standard estimates of survey cost are available, some of the items of expense can be examined. The primary expense is in time and effort; the time you spend producing the survey could be spent on other tasks. If other personnel are needed, they will have to be paid. Access to typewriters, word processors, and calculating machines (or computer resources) is a must. If you expect to gather a great deal of data, the cost of renting ADP time and of purchasing ADP scanner sheets should be examined. Surveys of more than 150-200 respondents cannot feasibly be tabulated by hand. The same is true for groups of less than 150 respondents if the survey questionnaire is lengthy. The final cost involves supplies. At a minimum you will need paper and envelopes. You may also have to pay either the cost of printing the survey questionnaire or the postage or both. Each of the above costs that applies to your survey should be estimated and the total cost measured against the survey requirement.

Since surveys are being used more and more, the information you want may have already been gathered. A search of some of the survey data sources listed in Appendix C might yield a solution to your problem or at least provide examples of how others have approached similar problems. So before you undertake a survey, first make sure the answer to your problem does not already exist. Next, evaluate the time you will need and determine the cost involved to produce the survey results, and then weigh these findings against the importance of the survey. Undertake a survey only if it is worth the time, effort, and cost to make it a good one.

GUIDE OUTLINE

The remaining chapters of this guide cover the various steps in surveying. Chapter 2 outlines the official policies and procedures within the Air Force for conducting surveys. Chapter 3 covers the determination of the purpose, hypotheses, and survey plan. Chapter 4 deals with the design of a sample survey and the technique for determining the required sample size. The concepts presented in Chapter 4 will not apply to you, of course, if you are conducting a full census rather than a sample survey. Chapter 5 outlines the construction of the survey questionnaire. Chapter 6 discusses some of the more common statistical errors committed by novice researchers and ways to avoid them. The various appendices contain checklists and data sources useful in surveying. The bibliography lists informative references on surveying.

CHAPTER 2

The Air Force Personnel Survey Program

The purpose of the Air Force Personnel Survey Program is to foster the development of compatible and effective surveys, and to minimize exposure of Air Force personnel to repeated or unwarranted survey solicitations (Air Force Instruction (AFI) 36-2601, 1993; p 1). This instruction also describes the survey policy responsibilities and explains how the survey program is conducted. This chapter will highlight some of the important points covered in AFI 36-2601, but Air Force personnel who plan to conduct official surveys within the Air Force should become familiar with the entire instruction.

The instruction designates the Military Personnel Survey Branch, Air Force Military Personnel Center (AFMPC/DPMYAS), Randolph AFB, Texas 78150, as the controlling and approving agency for Air Force military personnel surveys. Any member of the Air Force wanting to conduct a survey covered by this instruction must submit a written request through channels to AFMPC/DPMYAS for approval. "Any survey of Air Force civilian personnel must conform to the Air Force Labor Relations Program described in AFI 36-701." (AFI 36-3601, paragraph 7.3). Surveys going to non-DOD civilians (e.g., dependents of military personnel, government contractors, general public, etc.) are a special concern. These must be approved through the Office of Management and Budget (OMB), specifically the Office of the Administrative Assistant for the Secretary of the Air Force, Information Management Policy Division (SAF/AAIA), 1600 Air Force Pentagon, Washington, DC 20330-1600.

The types of surveys requiring approval are defined in detail in paragraphs 2 and 3 of AFI 36-2601. In general, surveys covered by this instruction must be personnel surveys and not occupational surveys. The latter type, referred to as a job inventory, is used to identify the duties and tasks that comprise an Air Force career field.

It is especially important to note that this instruction does not apply to surveys that concern aspects of base activities that the base

commander is authorized to change (e.g., surveys of base conditions, commissary or hospital services, etc.). In such cases, the survey does not require approval under AFI 36-2601 (paragraph 7.8). These are the major exclusions, but paragraphs 7.8 and 7.9 identify other less common exclusions.

Paragraph 7.2 of AFI 36-2601 specifies in detail the information that must accompany the request for approval to conduct a survey. You should be sure that:

- available information is inadequate to satisfy your needs
- currently programmed surveys cannot produce the required information
- the need for the data justifies the cost to obtain it
- a survey will produce the best data with the minimum inconvenience to the respondents.

The Air Force Personnel Survey Program specifies that all data collected must be treated as privileged information and that respondents will in no way suffer adverse actions as a result of their participation (or non-participation). The introductory paragraphs of AFI 36-2601 specify that all surveys subject to the provisions of AFI 37-132 (Air Force Privacy Act Program) must contain a Privacy Act Statement. This requires all respondents be advised of:

- the Federal statute or executive order that authorizes the solicitation of the information
- the principal purpose(s) for which the data are to be used
- the routine uses to be made of it
- whether furnishing the information is mandatory or voluntary
- the effects (if any) on the individual of not providing all or part of the requested information.

Finally, AFI 36-2601, paragraph 9, specifies conditions under which release of survey results must be coordinated with HQ AFMPC/DPMYAS. Every member of the Air Force who administers a survey should be familiar with and follow the guidelines established by this instruction and appropriate command/unit operating and implementing instructions.

CHAPTER 3

Developing the Purpose, Hypotheses, & Survey Plan

The first steps in producing a survey are the most important. They determine where you are going (the purpose), how you will know when your are there - or what you expect to find (the hypotheses, objectives, or research questions), and by what route you will go (the survey plan). If these steps are not well planned, all the remaining steps will be wasted effort.

THE PURPOSE

The first step in producing a survey is to define the purpose or objective of the survey. "A clear statement of purpose is necessary not only as a justification/explanation of the project, but also as a guideline to determine if future actions in the project are in support of the original purpose" A Guide for Development..., 1974; p 2). Without knowledge of the exact nature of the problem (objective), you cannot decide exactly what kind of data to collect or what to do with it once you have it. Usually a staff officer is given a problem or objective; it seldom originates with him. But this does not relieve the individual of responsibility for insuring that:

- the problem is well stated
- the surveyor understands exactly what the problem is
- the stated problem is the real problem

The survey should be designed to answer only the stated problem. Adding additional *interesting* objectives will lengthen and complicate the survey while clouding the real issue.

THE HYPOTHESIS, OBJECTIVE, OR RESEARCH QUESTION

Once the problem has been clearly stated, the next step is to form one or more hypotheses. The hypothesis is actually your educated guess about the answer to the problem. It should not be a capricious guess, however. It ought to be based on your prior experience related to the problem, or perhaps any knowledge you may have of previous research done on the topic. Without such a framework in which to make an educated guess, you really have no basis for making a guess at all. If you do not have a clear basis for formulating an hypothesis, you should instead develop one or more objectives or questions to frame the scope of your questionnaire.

For example, if a problem is identified on the base as declining use of the Officers' Club, an immediately obvious question comes to mind: "Are the officers on this base satisfied with the Officers' Club facilities?" This would be suitable as a research question. It is possible, though doubtful, if you could come up with a supportable hypothesis, or educated guess, as to the answer to the problem. You may, for instance, have gathered some anecdotal evidence (overhearing colleagues talking) of dissatisfaction with the club's facilities. But, this may not be sufficient for making an educated guess that this is the real reason for the decline in club use. The problem could be seasonal; it could be related to a decline in the officer population on the base; or a number of other possibilities. The point is that without some credible evidence to support an hypothesis, you should probably not formulate one.

If you formulated an hypothesis for the current example on the basis of the anecdotal evidence available to you, you would naturally construct a questionnaire to survey the opinions of officers regarding their use, or lack thereof, of the Officers' Club and the reasons for it. You might never think to gather data from the base military personnel office to see if the officer population is lower now than usual or if there are seasonal (cyclic) trends in the size of the officer population on the base. In other words, establishing the hypothesis may blind you to collecting data on other possible causes of the problem. This is why all researchers are cautioned not to formulate hypotheses unless they have a solid base in theory or previously gathered evidence that suggests the hypothesis is, in fact, probable.

Hypotheses must be carefully written. They should not contain moral judgments or biased statements such as "All pilots are good leaders." There are many ideas on what constitutes a good leader and your idea may not be the same as those of the people you will contact. Avoid words like should, best, good, bad, and ought.

Hypotheses should be as specific as possible. Avoid words such as most and some. If by most you mean a majority, then say majority. A survey can more easily be designed to test whether "more than 75 percent approve" than whether "most approve."

A well-formulated hypothesis, objective, or research question translates the purpose into a statement that can be investigated scientifically. The level of difficulty you will face in producing a valid survey will increase dramatically if they are not well formulated. Take care in doing this step, and it will save you much effort later in the survey development process.

THE SURVEY PLAN

The next step after determining the purpose and hypotheses is constructing the survey plan. The purpose of the survey plan is to ensure that the survey results will provide sufficient data to provide an answer (solution) to the problem you are investigating. The survey plan is comprised of three different parts:

- data collection plan
- data reduction and reformatting plan
- analysis plan

None of these plans stands on its own. Decisions you make on how you will analyze your data will affect your data collection plan. The type of data reduction you do will affect not only the types of analyses you can do, but also the amount and types of data you need to collect. Because these plans are closely interrelated, they should be developed concurrently.

THE DATA COLLECTION PLAN

The purpose of the data collection plan is to ensure that proper data are collected in the right amounts. Your hypothesis and your data analysis

plan determine the appropriateness of the data. For example, if you plan to analyze your results by age group to test a hypothesis, then you must collect data from each age group whose opinions you want to know. The right amount applies to sample data. As pointed out earlier, the use of sample data involves risk, and the amount of that risk is determined by the size of your sample. The amount of risk you are willing or able to accept should be stated in your analysis plan. Proper and right come together when your analysis plan involves both sampling and analyzing data by groups. You not only have to collect data from some members of each group you plan to analyze, but you also have to see that each group provides a response rate that is high enough to ensure your meeting your minimum risk level. The concept of the proper sample size is covered in greater detail in Chapter 4.

THE DATA REDUCTION AND REFORMATTING PLAN

The purpose of the data reduction and reformatting plan is to identify up front and to decrease as much as possible the amount of data handling (reduction and reformatting) you will have to do. This plan is highly dependent on the other two plans. As previously mentioned, if your collection plan calls for a great deal of data, you should plan to use a computer to analyze the data. If ADP scanner sheets are to be used to record respondents' answers, include the sheets with the questionnaire so the respondent can fill out the scanner sheet. This will save a great deal of time that you would have to spend if you transferred the survey data to the scanner sheets yourself. It also eliminates the possibility of your making errors in transferring data. You should coordinate in advance with the ADP personnel to make sure they will be able to scan your answer sheets and, if necessary, analyze your data within your timeframe. ADP shops are busy places. The prudent surveyor will “book” the scanning and analysis jobs well in advance with ADP personnel to ensure their resources are available when needed.

A strong potential for error and tedious corrective work lies in data reduction and reformatting. Proper care in developing this plan can save a great deal of time later and preclude error.

OPEN- AND CLOSED-ENDED QUESTIONS

The use of Automatic Data Processing (ADP) necessitates the use of closed-end questions -- a type of question you should consider even if you are hand-tabulating your data. A closed-end question lists possible answers from which the respondent picks the one he/she likes best. An example is the common multiple-choice question. The open-end question is one to which the respondents write the answer out in their own words. At first glance, the open-end question seems superior since respondents supply their answers rather than ones from your list of answers. But the wide variety of answers respondents generally provide to open-end questions turns out to be a great handicap later. For every open-end question, there are virtually an infinite number of possible answers. Since you cannot analyze an infinite number of answers, you must devise some means of categorizing this diversity of answers into a smaller, more manageable group. You will find yourself spending a tremendous amount of time reading, comparing, categorizing, and recording each answer. Much of this time can be saved if you use care in developing the questionnaire and constructing your own categories in advance. Construct each question so that every possible major category of response is contained in the answer list.

Then, later, all the computer will have to do is count the number of answers in each category. By having the survey respondent, not you, categorize the answer, you will collect data that is more valid, reliably, and accurate than if you did the categorizing yourself. Additional information on closed-end questions is provided in Chapter 5.

THE ANALYSIS PLAN

Finally, an analysis plan ensures that the information produced by the analysis will adequately address the originally stated hypotheses, objectives, or questions. It also ensures an analysis that is compatible with the data collected during the survey. In the analysis plan, you determine which statistics you will use and how much risk you can take in stating your conclusions. Each of these decisions will affect the amount and type of data you collect and how you will reduce it. Novice researchers often misuse statistical analyses out of ignorance of the assumptions on which the statistics are based. The most often committed error in statistical analysis by novices is using a statistical technique with inappropriate data. The results of such analyses appear to be legitimate, but are actually impossible to interpret correctly. We will discuss some of these common errors and how to avoid them in Chapter 6.

CONCLUDING THOUGHTS

Oppenheim (1966) suggests that to make sure all these parts of the survey plan are correctly interlocked, you can simply approach the natural sequence of survey operations in reverse order. First determine what conclusions you are interested in; then decide what statistics and results will be needed to draw these conclusions. From this, the type of questions needed and the nature of the sample can be determined.

A conscientious survey plan will help you produce a well-designed survey. The *proper* data will be processed correctly and efficiently to produce the information required to shed light on, and hopefully provide a solution to, the original problem.

CHAPTER 4

Sampling Techniques and Related Statistical Concepts

Chapter 1 identified some of the problems associated with sampling:

- acceptance of a risk of error
- choice of a representative sample
- determination of the size of the sample

This chapter outlines procedures for dealing with these challenges. First, different techniques designed to produce a representative sample from different types of populations are explained. Next, the relationship between risk and sample size is investigated. Finally, techniques are discussed for quantifying the amount of risk present in your results and for determining the sample size necessary to achieve the confidence and reliability specified in your analysis plan.

SAMPLING METHODS

Your overarching goal in doing a survey is to determine what some group thinks or feels about some issue. If money, time, or other resources were not a concern, the most accurate data you could get would come from surveying the entire population of interest. Since limited resources are a reality we all have to deal with, however, we are often forced to survey the views of only a few members of the population. But never lose sight of the fact that the real purpose is to discover the views of the entire population. Obviously, then, we want to be able to say with as much confidence as possible that the views of the group we surveyed represents the views of the entire population. Using a combination of powerful statistical tools, known as inferential statistics, and unbiased sampling

techniques, any surveyor can collect data that actually represent the views of the entire population from which the sample was taken. Two things are absolutely necessary, however, to ensure a high level of confidence that the sample represents the population:

- an unbiased sample
- a sufficiently large sample

Bias as a statistical term means error. To say that you want an unbiased sample may sound like you're trying to get a sample that is error-free. As appealing as this notion may be, it is impossible to achieve! Error always occurs -- even when using the most unbiased sampling techniques. One source of error is caused by the act of sampling itself. To understand it, consider the following example.

Let's say you have a bowl containing ten slips of paper. On each slip is printed a number, one through ten. This is your "population." Now you are going to select a sample. We will use a random method for drawing the sample, which can be done easily by closing your eyes and reaching into the bowl and choosing one slip of paper. After choosing it, check the number on it and place it in the sample pile.

Now to determine if the sample is representative of the population, we must know what attribute(s) we wish to make representative. Since there are an infinite number of human attributes, we must precisely determine the one(s) we are interested in before choosing the sample.

In our example, the attribute of interest will be the average numerical value on the slips of paper. Since the "population" contained ten slips numbered consecutively from one to ten, the average numerical value in the population is:

$$\frac{1+2+3+4+5+6+7+8+9+10}{10} = 5.5$$

As you can see, no matter what slip of paper we draw as our first sample selection, it's value will be either lower or higher than the population average. Let's say the slip we choose first has a 9 on it. The difference between our sample (9) and the population (5.5) averages is +3.5 (plus signifies the sample average is larger than the population average). The difference between the sample average and the population average is known as sampling error. That is, the sample mean (average)

plus (or minus) the total amount of sampling error equals the population mean.

On our second pick, we choose a slip that has a 1 on it. Now the average of sample values is:

$$\frac{9+1}{2} = 5.0$$

The sampling error has shrunk from its previous value of + 3.5 to its new value of - 0.5 (minus signifies the sample mean is now smaller than the population mean). Each time we choose a slip from the population to include in the sample, one of three mutually exclusive things can occur -- the sample mean will become:

- larger than the population mean
- smaller than the population mean
- equal to the population mean

On average, each sampling brings the sample mean a bit closer to the population mean. Ultimately, if we sampled everyone from the population, the sample mean and the population mean would be equal. This is why a complete census is completely accurate - there is no sampling error. Yet, if we are forced to use only a sample from the population, the larger the sample the less sampling error we will have, generally speaking.

Equally important to the size of the sample is the determination of the type of sampling to be done. In our example, we randomly (blindly) chose from the population. Random sampling always produces the smallest possible sampling error. In a very real sense, the size of the sampling error in a random sample is affected only by random chance. The two most useful random sampling techniques are simple random and stratified random sampling methods. These will discussed shortly.

Because a random sample contains the least amount of sampling error, we may say that it is an unbiased sample. Note that we are not saying the sample contains no error, but rather the minimum possible amount of error.

Nonrandom sampling techniques also exist, and are used more frequently than you might imagine. As you can probably guess from our previous discussion, nonrandom sampling techniques will always produce

larger sampling errors (for the same sample size) than random techniques. The reason for this is that nonrandom techniques generate the expected random sampling error on each selection plus additional error related to the nonrandom nature of the selection process. To explain this, let's extend our sampling example from above.

Let's say we want to sample from a "population" of 1000 consecutively numbered slips of paper. Because numbering these slips is time consuming, we have 10 people each number 100 slips and place all 100 of them into our bowl when they finish. Let's also say that the last person to finish has slips numbered from 901 to 1000, and these are laid on top of all the other slips in the bowl. Now we are ready to select them.

If we wanted to make this a truly random sampling process, we would have to mix the slips in the bowl thoroughly before selecting. Furthermore, we would want to reach into the bowl to different depths on subsequent picks to make sure every slip had a fair chance of being picked.

But, let us say in this example that we forget to mix the slips in the bowl. Let's also say we only pick from the top layer of slips. It should be obvious what will occur. Because the top layer of slips is numbered 901 through 1000, the mean of any sample (of 100 or less) we select will hover around 950.5 (the true mean of the numbers 901 through 1000). Clearly, this is not even close to the true population mean (500.5 -- the mean of the numbers from 1 to 1000). Sampling error amounts to the difference between the true population mean and the sample mean. In this example, the sampling error can as large as 450 (950.5 - 500.5).

This was a simple, and somewhat absurd, example of nonrandom sampling. But, it makes the point. Nonrandom sampling methods usually do not produce samples that are representative of the general population from which they are drawn. The greatest error occurs when the surveyor attempts to generalize the results of the survey obtained from the sample to the entire population. Such an error is insidious because it is not at all obvious from merely looking at the data, or even from looking at the sample. The easiest way to recognize whether a sample is representative or not is to determine if the sample was selected randomly. To be a random sampling method, two conditions must be met. If both are met, the resulting sample is random. If not, it is a nonrandom sampling technique:

- every member in the population must have an equal opportunity of being selected,
- the selection of any member of the population must have no influence on the selection of any other member

All nonrandom sampling methods violate one or both of these criteria. The most commonly used nonrandom methods are:

- systematic sampling (selecting every nth person from a group)
- cluster sampling (selecting groups of members rather than single members)
- convenience or incidental sampling (selecting only readily available members)
- judgment or purposive sampling (selecting members who are judged to be appropriate for the study)

SIMPLE RANDOM SAMPLING

A simple random sample is one in which each member (person) in the total population has an equal chance of being picked for the sample. In addition, the selection of one member should in no way influence the selection of another. Simple random sampling should be used with a homogeneous population, that is, one composed of members who all possess the same attribute you are interested in measuring. In identifying the population to be surveyed, homogeneity can be determined by asking the question, "What is (are) the common characteristic(s) that are of interest?" These may include such characteristics as age, sex, rank/grade, position, income, religious or political affiliation, etc. -- whatever you are interested in measuring.

The best way to choose a simple random sample is to use a random number table (or let a computer generate a series of random numbers automatically). In either case, you would assign each member of the population a unique number (or perhaps use a number already assigned to them such as SSAN, telephone number, zip code, etc.). The members of the population chosen for the sample will be those whose numbers are identical to the ones extracted from the random number table (or computer) in succession until the desired sample size is reached. An example of a random number table and instructions for its use appear in Appendix D. Many statistical texts or mathematical tables treat random number generation. A less rigorous procedure for determining randomness is to write the name of each member of the population on a separate card,

and with continuous mixing, draw out cards until the sample size is reached.

The simple random sample requires less knowledge about the population than other techniques, but it does have two major drawbacks. One is if the population is large, a great deal of time must be spent listing and numbering the members. The other is the fact that a simple random sample will not adequately represent many population attributes (characteristics) unless the sample is relatively large. That is, if you are interested if choosing a sample to be representative of a population on the basis of the distribution in the population of gender, age, and economic status, a simple random sample will need to be very large to ensure all these distributions are equivalent to (or representative of) the population. To obtain a representative sample across multiple population attributes, you should use the technique of stratified random sampling.

We made this point earlier in this chapter, but it's such an important concept that it bears repeating. To determine if the sampling method you use is random or not, remember that true random sampling methods must meet two criteria:

- Every member in the population must have an equal opportunity of being chosen for the sample (*equality*)
- The selection of one member is not affected by the selection of previous members (*independence*)

Both simple random and stratified random sampling methods meet these two criteria. Nonrandom sampling methods lack one or both of these criteria. We discuss stratified random sampling next.

STRATIFIED RANDOM SAMPLING

This method is used when the population is heterogeneous rather than homogeneous (or as discussed above, when you want to obtain a representative sample across many population attributes). A heterogeneous population is composed of unlike elements; such as, officers of different ranks, civilians and military personnel, or the patrons of a discount store (differing by gender or age).

A stratified random sample is defined as a combination of independent samples selected in proper proportions from homogeneous groups within a heterogeneous population. The procedure calls for categorizing the heterogeneous population into groups that are

homogeneous in themselves. If one group is proportionally larger than another, its sample size should also be proportionally larger. The number of groups to be considered is determined by the characteristics of the population. Many times the survey plan will determine some or all of the groups. For example, if you are comparing enlisted and officer segments on your base, each of these will be a separate group.

After dividing the population into groups, you then sample each homogeneous group. Different sampling techniques can be used in each of the different groups, but keep in mind that random techniques produce the minimum amount of sampling error. Finally, you should calculate the sample statistics for each group to determine how many members you need from each subgroup.

We will discuss the calculations involved in determining the size of your sample later in this chapter. These calculations are designed to determine the size of a simple random sample. Since the stratified sampling technique requires you to create simple, homogeneous subgroups from a large heterogeneous group, think of the calculations for a stratified sample as a series of simple random sample size calculations for each homogeneous subgroup. The only other information you must know is the proportion of the population possessing the attribute contained in each homogeneous subgroup.

For example, let's say we want to draw a random sample from a population of military personnel to assess their opinions on some issue. In addition, we would like to determine if the opinions differ by officer-enlisted affiliation and gender of the individuals surveyed. We recognize that the population we want to draw our sample from is heterogeneous with respect to the two attributes of interest to us. So, we have to create homogeneous subgroups (four to be exact):

- Enlisted, male
- Enlisted, female
- Officer, male
- Officer, female

Now, each group is homogeneous on both attributes. To ensure each subgroup in the sample will represent its counterpart subgroup in the population, we must ensure each subgroup is represented in the sample in the *same proportion* to the other subgroups as they are in the population. Let's assume that we know (or can estimate) the population of Air Force military personnel to be distributed as follows: 70 percent male, 30

percent female and 65 percent enlisted, 35 percent officer. With that, we can determine the approximate proportions of our four homogeneous subgroups in the population:

- Enlisted, male .65 x .70 = .455
- Enlisted, female .65 x .30 = .195
- Officer, male .35 x .70 = .245
- Officer, female .35 x .30 = .105

Thus, a representative sample of the Air Force population (by race and enlisted-officer affiliation) would be composed of 45.5 percent enlisted males, 19.5 percent enlisted females, 24.5 percent officer males, and 10.5 percent officer females. Each percentage should be multiplied by the total sample size needed to arrive at that actual number of personnel required from each subgroup or stratum.

As this example illustrates, stratified random sampling requires a detailed knowledge of the distribution of attributes or characteristics of interest in the population to determine the homogeneous groups that lie within it. A stratified random sample is superior to a simple random sample since the population is divided into smaller homogeneous groups before sampling, and this yields less variation within the sample. This makes possible the desired degree of accuracy with a smaller sample size. But, if you cannot accurately identify the homogeneous groups, you are better off using the simple random sample since improper stratification can lead to serious error.

SYSTEMATIC SAMPLING

Sometimes it is more expeditious to collect a sample of survey participants systematically. This is frequently done, for instance, in exit polling of voters or store customers. It is a nonrandom sampling technique, but is used primarily for its ease and speed of identifying participants.

To use the systematic approach, simply choose every K^{th} member in the population where K is equal to the population size divided by the required sample size. If this quotient has a remainder, ignore it (round down). For example, if you need 100 members in your sample and the population consists of 1000 people, you need to sample every 1000/100 (or 10^{th}) member of the population. When using this method, some

suggest you should choose your starting point at random by choosing a random number from 1 to K.

If you recall the characteristic requirements for a random sample discussed above (equality and independence), you can see that systematic sampling methods lack both characteristics. Every member from the population does not have a equal chance of being selected, and the selection of members for the sample depends on the initial selection. Regardless of how you select your starting point, once selected, every subsequent member of the sample is automatically determined. This method is clearly nonrandom.

Some suggest that by mixing the population well you can turn this into a random sampling technique. They are wrong. Regardless of how much you mix the population before selecting a starting point, the fact remains that once that point is chosen, further selection of members for the sample is nonrandom (no independence).

Recognize the limitation of this type of sampling. Since it is nonrandom, the resulting sample will not necessarily be representative of the population from which it was drawn. This will affect your ability to confidently generalize results of the survey since you may not be sure to which segment of the population the results will apply. As a word of advice, unless you have experience in systematic sampling techniques, and have full knowledge of the population to be sampled, you should avoid using this method.

JUDGMENT SAMPLING

Judgment sampling involves asking an expert on an issue being investigated to define the members that should comprise the sample. The representativeness of the sample is determined solely by the judgment of the researcher. Since each member in the population does not have an equal chance of being chosen, a judgment sample is also a nonrandom sampling method. Since the sample does not meet the criterion of randomness - the basis for many statistical sampling applications (a judgment sample should never be used in a statistical evaluation effort.

PURPOSIVE SAMPLING

As the name implies, purposive sampling involves selecting members from the population to comprise a sample because they specifically meet some prescribed purpose of possess specific attributes of interest that address the purpose of a particular research problem under investigation. Purposive sampling is used primarily in causal-comparative (*ex post facto*) research where the researcher is interested in finding a possible cause-and-effect link between two variables, one of which has already occurred. The researcher intentionally selects the samples in such a way that one possesses the causal (independent) variable and one does not. The purpose of the research governs the selection of the sample and, thus, excludes members of the population who do not contribute to that purpose.

REPRESENTATIVE SAMPLING

The types of sampling methods discussed above are only a few of the many available. You will find others in the references listed in the bibliography. Each type is designed to obtain the most representative sample possible from different kinds of populations. Before using any sampling method yourself, first think about the population to which you want to generalize the results of your survey (which population do you want to represent). If generalizing results is not your aim, any sampling method will do. If generalizing results is important, use a sampling method that will ensure your sample is representative of the population from which you draw it. Random sampling methods typically ensure a high degree of confidence that the results do, in fact, represent those of the whole population.

FACTORS INFLUENCING SAMPLE SIZE

As pointed out in Chapter 1, when you sample you are dealing with only partial information. And you must accept a risk of being wrong when inferring something about the population on the basis of sample information. In the analysis portion of your survey plan, you identify the amount of risk you are willing (or allowed) to take. This amount of risk relates directly to the size of your sample. Simply stated, the less risk you are willing to take, the larger your sample must be. If you cannot accept any risk, you should survey the entire population (take a census) and you need not study this chapter any further.

When determining your risk level, keep in mind the time and cost involved in obtaining the sample size sufficient to achieve the risk level you can accept. You may find it impossible to produce a sample large enough to meet that risk level.

Another factor bearing on sample size is also obtained from your analysis plan. It is the number of groups you are planning to examine within the population. For example, if you are planning to compare two groups (enlisted and officer) on a base (your population), each of the groups must be sampled and each of the samples must be large enough to ensure satisfying your risk level.

CONFIDENCE LEVEL AND PRECISION

Risk, as it relates to sample size determination, is specified by two interrelated factors:

- the confidence level
- the precision (or reliability) range.

To minimize risk, you should have a high confidence (say 95 percent) that the true value you seek (the actual value in the population) lies somewhere within a small interval (say + or - 5 percent) around your sample value (your precision). Sawyer (1971; p 49) uses a baseball game analogy to explain confidence level, precision range, and their relationship. A baseball pitcher may feel that he can get very few of his pitches (perhaps 10 percent) over the exact center (small precision range) of home plate. But since home plate is 17 inches wide, he may feel that he can get 95 percent of his pitches over the center of the plate with a precision of plus or minus 8 1/2 inches (a 95 percent confidence level). If the plate is widened to 30 inches, he may feel 99 percent confident. So when we widen the range of precision (or reliability), we increase our confidence level. Likewise, if we reduce the range, we reduce our confidence level. Most surveying organizations use a 95 percent confidence level and a ± 5 percent precision level as the absolute minimum.

DETERMINING THE SIZE OF THE SAMPLE

Once you determine your desired degree of precision and your confidence level, there are several formulas you can use to determine

sample size depending on how you plan to report the results of your study. We'll discuss three of them here. If you will be reporting results as percentages (proportions) of the sample responding, use the following formula:

$$n = \frac{P(1-P)}{\frac{A^2}{Z^2} + \frac{P(1-P)}{N}}$$

Where n = sample size required
N = number of people in the population
P = estimated percentage of the population possessing attribute of interest
A = Accuracy desired, expressed as a decimal (i.e., .01, .03, .05, etc)
Z = number of standard deviation units of the sampling distribution corresponding to the desired confidence level (see Appendix E)

If you will report results as means (averages) of the sample responding, use the following formula:

$$n = \frac{P^2}{\frac{A^2}{Z^2} + \frac{P^2}{N}}$$

Where n = sample size required
N = number of people in the population
P = estimated standard deviation of the attribute of interest in the population
A = Accuracy desired, expressed as a decimal (i.e., .01, .03, .05, etc)
Z = number of standard deviation units of the sampling distribution corresponding to the desired confidence level (see Appendix E)

If you plan to report results in a variety of ways, or if you have difficulty estimating percentage or standard deviation of the attribute of interest, the following formula may be more suitable for use:

$$n = \frac{NZ^2 * .25}{[d^2 * (N-1)] + [Z^2 * .25]}$$

Where n = sample size required
N = total population size (known or estimated)
d = precision level (usually .05 or .10)
Z = number of standard deviation units of the sampling distribution corresponding to the desired confidence level (see Appendix E)

We illustrate this formula with the following example. If the total population (N) is 10,000, and you wish a 95% confidence level and ± 5 percent precision level (d = .05, Z = 1.96 from Appendix E), then:

$$n = \frac{10{,}000 * 1.96^2 * .25}{[.05^2 * 9999] + [1.96^2 * .25]} = 369.98$$

So, a representative sample of 370 (369.98 rounded up) would be sufficient to satisfy your risk level. Inspection of the formula shows that the required sample size will increase most rapidly if:

- the confidence level (Z factor) is increased, or
- the precision level (d) is made smaller.

If you have stratified your population into more than one group, the size of each group will be its proportion (percentage) in the population times the total sample size as computed above. To illustrate, recall our earlier example of four stratified groups. Using the n of 370 calculated above, each of these strata should have the following sample sizes:

- Enlisted, male 370 x .455 = 168.35 = 168
- Enlisted, female 370 x .195 = 72.15 = 72
- Officer, male 370 x .245 = 90.65 = 91
- Officer, female 370 x .105 = 38.85 = 39

Finally, you should adjust the computed sample size (n) by dividing n by the expected response rate. For instance, if you expect 75 percent response rate, you should make your sample size equal $\frac{n}{0.75}$. If you can't anticipate a response rate, assume a 50 percent response rate (i.e., double the n value). This sort of adjustment should ensure you get a sufficient number of responses regardless of return rate.

CHAPTER 5

The Questionnaire

The final step in preparing the survey is developing the data collection instrument. The most common means of collecting data are the interview and the self- or group-administered questionnaire. In the past, the interview has been the most popular data-collecting instrument. Recently, the questionnaire has surpassed the interview in popularity, especially in the military. Due to this popularity, this chapter concentrates on the development of the questionnaire.

THE QUESTIONNAIRE PROS AND CONS

It is important to understand the advantages and disadvantages of the questionnaire as opposed to the personal interview. This knowledge will allow you to maximize the strengths of the questionnaire while minimizing its weaknesses. The advantages of administering a questionnaire instead of conducting an interview are:

- lower costs
- better samples
- standardization
- respondent privacy (anonymity)

The primary advantage is lower cost, in time as well as money. Not having to train interviewers eliminates a lengthy and expensive requirement of interviewing. The questionnaire can be administered simultaneously to large groups whereas an interview requires each individual to be questioned separately. This allows the questions to reach a given number of respondents more efficiently than is possible with the interview. Finally, the cost of postage should be less than that of travel or telephone expenses.

Recent developments in the science of surveying have led to incorporating computers into the interview process, yielding what is commonly known as computer automated telephone interview (or CATI) surveys. Advances in using this survey technique have dramatically reshaped our traditional views on the time-intensive nature and inherent unreliability of the interview technique. Yet, despite resurgence in the viability of survey interviews, instruction in the development and use of the CATI technique is well beyond the scope of this handbook.

Many surveys are constrained by a limited budget. Since a typical questionnaire usually has a lower cost per respondent, it can reach more people within a given budget (or time) limit. This can enhance the conduct of a larger and more representative sample.

The questionnaire provides a standardized data-gathering procedure. Using a well-constructed questionnaire can minimize the effects of potential human errors (for example, altering the pattern of question asking, calling at inconvenient times, and biasing by "explaining". The use of a questionnaire also eliminates any bias introduced by the feelings of the respondents towards the interviewer (or vice versa).

Although the point is debatable, most surveyors believe the respondent will answer a questionnaire more frankly than he would answer an interviewer, because of a greater feeling of anonymity. The respondent has no one to impress with his/her answers and need have no fear of anyone hearing them. To maximize this feeling of privacy, it is important to guard, and emphasize, the respondent's anonymity.

The primary disadvantages of the questionnaire are nonreturns, misinterpretation, and validity problems. Nonreturns are questionnaires or individual questions that are not answered by the people to whom they were sent. Oppenheim (1966) emphasizes that "the important point about these low response rates is not the reduced size of the sample, which could easily be overcome by sending out more questionnaires, but the possibility of bias. Nonresponse is not a random process; it has its own determinants, which vary from survey to survey" (p 34).

For example, you may be surveying to determine the attitude of a group about a new policy. Some of those opposed to it might be afraid to speak out, and they might comprise the majority of the nonreturns. This would introduce non-random (or systematic) bias into your survey results, especially if you found only a small number of the returns were in favor of

the policy. Nonreturns cannot be overcome entirely. What we can do is try to minimize them. Techniques to accomplish this are covered later in this chapter.

Misinterpretation occurs when the respondent does not understand either the survey instructions or the survey questions. If respondents become confused, they will either give up on the survey (becoming a nonreturn) or answer questions in terms of the way they understand it, but not necessarily the way you meant it. Some view the latter problem as a more dangerous occurrence than merely nonresponding. The questionnaire instructions and questions must be able to stand on their own and must use terms that have commonly understood meanings throughout the population under study. If novel terms must be used, be sure to define them so all respondents understand your meaning.

The third disadvantage of using a questionnaire is inability to check on the validity of the answer. Did the person you wanted to survey give the questionnaire to a friend or complete it personally? Did the individual respond indiscriminately? Did the respondent deliberately choose answers to mislead the surveyor? Without observing the respondent's reactions (as would be the case with an interview) while completing the questionnaire, you have no way of knowing the true answers to these questions.

The secret in preparing a survey questionnaire is to take advantage of the strengths of questionnaires (lower costs, more representative samples, standardization, and privacy) while minimizing the number of nonreturns, misinterpretations, and validity problems. This is not always as easy as it sounds. But an inventive surveyor can very often find legitimate ways of overcoming the disadvantages. We provide some suggestions below to help.

THE CONTENTS

The key to minimizing the disadvantages of the survey questionnaire lies in the construction of the questionnaire itself. A poorly developed questionnaire contains the seeds of its own destruction. Each of the three portions of the questionnaire - the cover letter, the instructions, and the questions - must work together to have a positive impact on the success of the survey.

The cover letter should explain to the respondent the purpose of the survey and motivate him to reply truthfully and quickly. If possible, it should explain why the survey is important to him, how he was chosen to participate, and who is sponsoring the survey (the higher the level of sponsorship the better). Also the confidentiality of the results should be strongly stressed. A well written cover letter can help minimize both nonreturn and validity problems. An example is given in Appendix F. In support of the statement above regarding level of sponsorship, the signature block on the letter should be as high level as you can get commensurate with the topic being investigated. For instance, a survey about Air Force medical issues or policy should be signed by the Air Force Surgeon General or higher, a survey on religious issues by the Air Force Chief of Chaplains, etc. Another tip that seems to help improve response rate is to identify the survey as official. Even though the letter is on government stationery and is signed by an military official, it may help to mark the survey itself with an OFFICIAL stamp of some sort. In general, the more official the survey appears, the less likely it is to be disregarded.

The cover letter should be followed by a clear set of instructions explaining how to complete the survey and where to return it. If the respondents do not understand the mechanical procedures necessary to respond to the questions, their answers will be meaningless. The instructions substitute for your presence, so you must anticipate any questions or problems that may arise and attempt to prevent them from occurring. If you are using ADP scanner sheets, explain how you want the respondent to fill it in - what portions to use and what portions to leave blank. Remember anonymity! If you do not want respondents to provide their names or SSANs, say so explicitly in the instructions, and tell them to leave the NAME and SSAN portions of the scan sheets blank.

If you need respondents' SSAN and/or name included on the survey for tracking or analysis purposes, you will need to put a Privacy Act Statement somewhere on the survey (refer to Chapter 2). The "Instructions" page is usually a good place for this statement. It places it in a prominent place where all respondents will see it, but does not clutter the instrument itself or the cover letter.

The third and final part of the questionnaire is the set of questions. Since the questions are the means by which you are going to collect your data, they should be consistent with your survey plan. They should not be ambiguous or encourage feelings of frustration or anger that will lead to nonreturns or validity problems.

TYPES OF QUESTIONS

Before investigating the art of question writing, it will be useful to examine the various types of questions. Cantelou (1964; p 57) identifies four types of questions used in surveying. The classifier or background question is used to obtain demographic characteristics of the group being studied, such as age, sex, grade, level of assignment, and so forth. This information is used when you are categorizing your results by various subdivisions such as age or grade. Therefore, these questions should be consistent with your data analysis plan. The second and most common type of question

is the multiple choice or closed-end question. It is used to determine feelings or opinions on certain issues by allowing the respondent to choose an answer from a list you have provided (see Chapter 3). The intensity question, a special form of the multiple-choice question, is used to measure the intensity of the respondent's feelings on a subject. These questions provide answers that cover a range of feelings.

The intensity question is covered in greater detail later in this chapter. The final type of question is the free response or open-end question. This type requires respondents to answer the question in their own words (see Chapter 3). It can be used to gather opinions or to measure the intensity of feelings. Multiple-choice questions are the most frequently used types of questions in surveying today. It is prudent, therefore, to concentrate primarily on factors relating to their application.

QUESTIONNAIRE CONSTRUCTION

Many researchers have investigated the complex art of question writing. From their experiences, they offer valuable advice. Below are some helpful hints typical of those that appear most often in texts on question construction.

1. Keep the language simple. Analyze your audience and write on their level. Parten (1950; p 201) suggests that writing at the sixth-grade level may be appropriate. Avoid the use of technical terms or jargon. An appropriate corollary to Murphy's Law in this case would be: If someone can misunderstand something, they will.

2. Keep the questions short. Long questions tend to become ambiguous and confusing. A respondent, in trying to comprehend a long question, may leave out a clause and thus change the meaning of the question.

3. Keep the number of questions to a minimum. There is no commonly agreed on maximum number of questions that should be asked, but research suggests higher return rates correlate highly with shorter surveys. Ask only questions that will contribute to your survey. Apply the "So what?" and "Who cares?" tests to each question. "Nice-to-know" questions only add to the size of the questionnaire. Having said this, keep in mind that you should not leave out questions that would yield necessary data simply because it will shorten your survey. If the information is necessary, ask the question. With the availability of desktop publishing (DTP) software, it is often possible to give the perception of a smaller survey (using smaller point/pitch typefaces, etc.) even though many questions are asked. A three-page type written survey can easily be reduced to a single page using DTP techniques.

4. Limit each question to one idea or concept. A question consisting of more than one idea may confuse the respondent and lead to a meaningless answer. Consider this question: "Are you in favor of raising pay and lowering benefits?" What would a yes (or no) answer mean?

5. Do not ask leading questions. These questions are worded in a manner that suggests an answer. Some respondents may give the answer you are looking for whether or not they think it is right. Such questions can alienate the respondent and may open your questionnaire to criticism. A properly worded question gives no clue as to which answer you may believe to be the correct one.

6. Use subjective terms such as good, fair, and bad sparingly, if at all. These terms mean different things to different people. One person's "fair" may be another person's "bad." How much is "often" and how little is "seldom?"

7. Allow for all possible answers. Respondents who cannot find their answer among your list will be forced to give an invalid reply or, possibly, become frustrated and refuse to complete the survey. Wording the question to reduce the number of possible answers is the first step. Avoid dichotomous (two-answer) questions (except for obvious demographic questions such as gender). If you cannot avoid them, add a third option, such as **no opinion**, **don't know**, or **other**. These may not

get the answers you need but they will minimize the number of invalid responses. A great number of "don't know" answers to a question in a fact-finding survey can be a useful piece of information. But a majority of "other" answers may mean you have a poor question, and perhaps should be cautious when analyzing the results.

8. Avoid emotional or morally charged questions. The respondent may feel your survey is getting a bit too personal!

9. Understand the should-would question. Selltiz, et al. (1963, p 251) note that respondents answer "should" questions

from a social or moral point of view while answering "would" questions in terms of personal preference.

10. Formulate your questions and answers to obtain exact information and to minimize confusion. For example, does "How old are you?" mean on your last or your nearest birthday? Does "What is your (military) grade?" mean permanent or temporary grade? As of what date? By including instructions like "Answer all questions as of (a certain date)", you can alleviate many such conflicts. (Refer to hint 13 below.)

11. Include a few questions that can serve as checks on the accuracy and consistency of the answers as a whole. Have some questions that are worded differently, but are soliciting the same information, in different parts of the questionnaire. These questions should be designed to identify the respondents who are just marking answers randomly or who are trying to game the survey (giving answers they think you want to hear). If you find a respondent who answers these questions differently, you have reason to doubt the validity of their entire set of responses. For this reason, you may decide to exclude their response sheet(s) from the analysis.

12. Organize the pattern of the questions:

- Place demographic questions at the end of the questionnaire.
- Have your opening questions arouse interest.
- Ask easier questions first.
- To minimize conditioning, have general questions precede specific ones.
- Group similar questions together.
- If you must use personal or emotional questions, place them at the end of the questionnaire.

Note: The next two hints apply to the entire questionnaire including the cover letter, instructions, and question.

13. Pretest (pilot test) the questionnaire. This is the most important step in preparing your questionnaire. The purpose of the pretest is to see just how well your cover letter motivates your respondents and how clear your instructions, questions, and answers are. You should choose a small group of people (from three to ten should be sufficient) you feel are representative of the group you plan to survey. After explaining the purpose of the pretest, let them read and answer the questions without interruption. When they are through, ask them to critique the cover letter, instructions, and each of the questions and answers. Don't be satisfied with learning only what confused or alienated them. Question them to make sure that what they thought something meant was really what you intended it to mean. Use the above 12 hints as a checklist, and go through them with your pilot test group to get their reactions on how well the questionnaire satisfies these points. Finally, redo any parts of the questionnaire that are weak.

14. Have your questionnaire neatly produced on quality paper. A professional looking product will increase your return rate. As mentioned earlier, desktop publishing software can be used to add a very professional touch to your questionnaire and improve the likelihood of its being completed. But always remember the adage "You can't make a silk purse out of a sow's ear." A poorly designed survey that contains poorly written questions will yield useless data regardless of how "pretty" it looks.

15. Finally, make your survey interesting!

INTENSITY QUESTIONS AND THE LIKERT SCALE

As mentioned previously, the intensity question is used to measure the strength of a respondent's feeling or attitude on a particular topic. Such questions allow you to obtain more quantitative information about the survey subject. Instead of a finding that 80 percent of the respondents favor a particular proposal or issue, you can obtain results that show 5 percent of them are strongly in favor whereas 75 percent are mildly in favor. These findings are similar, but the second type of response supplies more useful information.

The most common and easily used intensity (or scaled) question involves the use of the Likert-type answer scale. It allows the respondent to choose one of several (usually five) degrees of feeling about a statement from strong approval to strong disapproval. The "questions" are in the form of statements that seem either definitely favorable or definitely unfavorable toward the matter under consideration. The answers are given scores (or weights) ranging from one to the number of available answers, with the highest weight going to the answer showing the most favorable attitude toward the subject of the survey. The following questions from the Minnesota Survey of Opinions designed to measure the amount of "anti-US law" feelings illustrate this procedure:

1. Almost anything can be fixed up in the courts if you have enough money.

Strongly Disagree (1)	**Disagree** (2)	**Undecided** (3)	**Agree** (4)	**Strongly Agree** (5)

2. On the whole, judges are honest.

Strongly Disagree (1)	**Disagree** (2)	**Undecided** (3)	**Agree** (4)	**Strongly Agree** (5)

The weights (shown by the numbers below the answers) are not shown on the actual questionnaire and, therefore, are not seen by the respondents. A person who feels that US laws are unjust would score lower than one who feels that they are just. The stronger the feeling, the higher (or lower) the score. The scoring is consistent with the attitude

being measured. Whether “agree” or “disagree” gets the higher weight actually makes no difference. But for ease in interpreting the results of the questionnaire, the weighting scheme should remain consistent throughout the survey.

One procedure for constructing Likert-type questions is as follows (adapted from Selltiz, et al., 1963; pp 367-368):

1. The investigator collects a large number of definitive statements relevant to the attitude being investigated.

2. Conduct and score a pretest of your survey. The most favorable response to the attitude gets the highest score for each question. The respondent's total score is the sum of the scores on all questions.

3. If you are investigating more than one attitude on your survey, intermix the questions for each attitude. In this manner, the respondent will be less able to guess what you are doing and thus more likely to answer honestly.

4. Randomly select some questions and flip-flop the *Strongly Agree* -- *Strongly Disagree* scale to prevent the respondent from getting into a pattern of answering (often called a response set).

The intensity question, with its scaled answers and average scores, can supply quantitative information about your respondents' attitudes toward the subject of your survey. The interested reader is encouraged to learn and use other scales, such as the Thurstone, Guttman, and Semantic Differential scales, by studying some of the references in the bibliography.

A number of studies have been conducted over the years attempting to determine the limits of a person's ability to discriminate between words typically found on rating or intensity scales. The results of this research can be of considerable value when trying to decide on the right set of phrases to use in your rating or intensity scale. When selecting phrases for a 4-, 5-, 7-, or 9-point Likert scale, you should choose phrases that are far enough apart from one another to be easily discriminated, while, at the same time, keeping them close enough that you don't lose potential information. You should also try to gauge whether the phrases you are using are commonly understood so that different respondents will interpret the meaning of the phrases in the same way. An obvious

example is shown with the following 3 phrases: *Strongly Agree*, *Neutral*, *Strongly Disagree*

These are easily discriminated, but the gap between each choice is very large. How would a person respond on this three-point scale if they only ***agreed*** with the question being asked? There is no middle ground between *Strongly Agree* and *Neutral*. The same thing is true for someone who wants to respond with a mere ***disagree***. Your scales must have enough choices to allow respondents to express a reasonable range of attitudes on the topic in question, but there must not be so many choices that most respondents will be unable to consistently discriminate between them. Appendix H provides several tables containing lists of phrases commonly used in opinion surveys with associated "scale values" and standard deviations (or inter-quartile range values). Also provided is a short introduction describing how these lists can be used in selecting response alternatives for your opinion surveys. The information in that appendix is derived from research done for the U.S. Army Research Institute for the Behavioral and Social Services at Fort Hood, Texas.

BIAS AND HOW TO COMBAT IT

Like any scientist or experimenter, surveyors must be aware of ways their surveys might become biased and of the available means for combating bias. The main sources of bias in a questionnaire are:

- a nonrepresentative sample
- leading questions
- question misinterpretation
- untruthful answers

Surveyors can expose themselves to possible nonrepresentative sample bias in two ways. The first is to actually choose a nonrepresentative sample. This bias can be eliminated by careful choice of the sample as discussed earlier in Chapter 4. The second way is to have a large number of nonreturns.

The nonreturn bias (also called non-respondent bias) can affect both the sample survey and the complete survey. The bias stems from the fact that the returned questionnaires are not necessarily evenly distributed throughout the sample. The opinions or attitudes expressed by those who returned the survey may or may not represent the attitudes or opinions of those who did not return the survey. It is impossible to determine which is

true since the non-respondents remain an unknown quantity. Say, for example, a survey shows that 60 percent of those returning questionnaires favor a certain policy. If the survey had a 70 percent response rate (a fairly high rate as voluntary surveys go), then the favorable replies are actually only 42 percent of those questioned (60 percent of the 70 percent who replied), which is less than 50 percent! a minority response in terms of the whole sample.

Since little can be done to estimate the feelings of the nonreturnees, especially in a confidential survey, the only solution is to minimize the number of nonreturns. Miller (1970; p 81) and Selltiz et al. (1963; p 241) offer the following techniques to get people to reply to surveys. Some of these have already been mentioned in earlier sections of this chapter.

1. Use follow-up letters. These letters are sent to the non-respondents after a period of a couple of weeks asking them again to fill out and return the questionnaire. The content of this letter is similar to that of the cover letter. If you are conducting a volunteer survey, you should anticipate the need for following up with non-respondents and code the survey in some unobtrusive way to tell who has and who has not yet responded. If you don't do that, but still need to get in touch with non-respondents, consider placing ads in local papers or base bulletins, announcements at commander's call, or notices posted in public places. If at all possible, provide a fresh copy of the survey with the follow- up letter. This often increases return rate over simply sending out a letter alone.

2. Use high-level sponsorship. This hint was mentioned in an earlier section. People tend to reply to surveys sponsored by organizations they know or respect. If you are running a military survey, obtain the highest-ranking sponsorship you can. Effort spent in doing this will result in a higher percentage of returns. If possible, use the letterhead of the sponsor on your cover letter.

3. Make your questionnaire attractive, simple to fill out, and easy to read. A professional product usually gets professional results.

4. Keep the questionnaire as short as possible. You are asking for a person's time, so make your request as small as possible.

5. Use your cover letter to motivate the person to return the questionnaire. One form of motivation is the have the letter signed by an individual known to be respected by the target audience for your questionnaire. In addition, make sure the individual will be perceived by the audience as having a vested interest in the information needed.

6. Use inducements to encourage a reply. These can range from a small amount of money attached to the survey to an enclosed stamped envelope. A promise to report the results to each respondent can be helpful. If you do promise a report, be sure to send it.

Proper use of these techniques can lower the nonreturn rate to acceptable levels. Keep in mind, though, that no matter what you do, there will always be non-respondents to your surveys. Make sure the effort and resources you spend are in proportion with the return you expect to get.

The second source of bias is misinterpretations of questions. We have seen that these can be limited by clear instructions, well-constructed questions, and through judicious pilot testing of the survey. Biased questions can also be eliminated by constructing the questions properly and by using a pilot test. Finally, internal checks and a good motivational cover letter can control bias introduced by untruthful answers. Although bias cannot be eliminated totally, proper construction of the questionnaire, a well-chosen sample, follow- up letters, and inducements can help control it.

BIAS IN VOLUNTEER SAMPLES

This section illustrates the many diverse, and sometimes powerful factors that influence survey findings as a result of using volunteers in a survey. The conclusions expressed here regarding volunteer samples are provided to make the surveyor aware of the often profound effects of non-respondent bias on survey data.

The exclusive use of volunteers in survey research represents another major source of bias to the surveyor -- especially the novice. Although it may not be immediately evident, it is nonetheless empirically true that volunteers, as a group, possess characteristics quite different from those who do not generally volunteer. Unless the surveyor takes these differences into consideration before choosing to use an exclusively volunteer sample, the bias introduced into the data may be so great that the

surveyor can no longer confidently generalize the survey's findings to the population at large, which is usually the goal of the survey.

Fortunately, research findings exist which describe several unique characteristics of the volunteer subject. By using these characteristics appropriately, the surveyor may avoid inadvertent biases and pitfalls usually associated with using and interpreting results from volunteer samples. The following list provides 22 conclusions about unique characteristics of the volunteer. The list is subdivided into categories representing the level of confidence to be placed in the findings. Within each category, the conclusions are listed in order starting with those having the strongest evidence supporting them. (from Rosenthall and Rosnow, The Volunteer Subject, 1975; pp 195-196):

Conclusions Warranting Maximum Confidence

1. Volunteers tend to be better educated than nonvolunteers, especially when personal contact between investigator and respondent is not required.

2. Volunteers tend to have higher social-class status than nonvolunteers, especially when social class is defined by respondents' own status rather than by parental status.

3. Volunteers tend to be more intelligent than nonvolunteers when volunteering is for research in general, but not when volunteering is for somewhat less typical types of research such as hypnosis, sensory isolation, sex research, small-group and personality research.

4. Volunteers tend to be higher in need for social approval than nonvolunteers.

5. Volunteers tend to be more sociable than nonvolunteers.

Conclusions Warranting Considerable Confidence

6. Volunteers tend to be more arousal seeking than nonvolunteers, especially when volunteering is for studies of stress, sensory isolation, and hypnosis.

7. Volunteers tend to be more unconventional than nonvolunteers, especially when volunteering is for studies of sex behavior.

8. Females are more likely than males to volunteer for research in general, more likely than males to volunteer for physically and emotionally stressful research (e.g., electric shock, high temperature, sensory deprivation, interviews about sex behavior).

9. Volunteers tend to be less authoritarian than nonvolunteers.

10. Jews are more likely to volunteer than Protestants, and Protestants are more likely to volunteer than Roman Catholics.

11. Volunteers tend to be less conforming than nonvolunteers when volunteering is for research in general, but not when subjects are female and the task is relatively "clinical" (e.g., hypnosis, sleep, or counseling research).

Conclusions Warranting Some Confidence

12. Volunteers tend to be from smaller towns than nonvolunteers, especially when volunteering is for questionnaire studies.

13. Volunteers tend to be more interested in religion than nonvolunteers, especially when volunteering is for questionnaire studies.

14. Volunteers tend to be more altruistic than nonvolunteers.

15. Volunteers tend to be more self-disclosing than nonvolunteers.

16. Volunteers tend to be more maladjusted than nonvolunteers, especially when volunteering is for potentially unusual situations (e.g., drugs, hypnosis, high temperature, or vaguely described experiments) or for medical research employing clinical rather than psychometric definitions of psychopathology.

17. Volunteers tend to be younger than nonvolunteers, especially when volunteering is for laboratory research and especially if they are female.

Conclusions Warranting Minimum Confidence

18. Volunteers tend to be higher in need for achievement than non-volunteers, especially among American samples.

19. Volunteers are more likely to be married than nonvolunteers, especially when volunteering is for studies requiring no personal contact between investigator and respondent.

20. Firstborns are more likely than laterborns to volunteer, especially when recruitment is personal and when the research requires group interaction and a low level of stress.

21. Volunteers tend to be more anxious than nonvolunteers, especially when volunteering is for standard, nonstressful tasks and especially if they are college students.

22. Volunteers tend to be more extroverted than nonvolunteers when interaction with others is required by the nature of the research.

Borg and Gall (1979) have suggested how surveyors might use this listing to combat the effects of bias in survey research. For example, they suggest that:

> The degree to which these characteristics of volunteer samples affect research results depends on the specific nature of the investigation. For example, a study of the level of intelligence of successful workers in different occupations would probably yield spuriously high results if volunteer subjects were studied, since volunteers tend to be more intelligent than nonvolunteers. On the other hand, in a study concerned with the cooperative behavior of adults in work-group situations, the tendency for volunteers to be more intelligent may have no effect on the results, but the tendency for volunteers to be more sociable could have a significant effect. It is apparent that the use of volunteers in research greatly complicates the interpretation of research results and their generalizability to the target

population, which includes many individuals who would not volunteer. (pp 190-191)

SUMMARY

The questionnaire is the means for collecting your survey data. It should be designed with your data collection plan in mind. Each of its three parts should take advantage of the strengths of questionnaires while minimizing their weaknesses. Each of the different kinds of questions is useful for eliciting different types of data, but each should be constructed carefully with well- developed construction guidelines in mind. Properly constructed questions and well-followed survey procedures will allow you to obtain the data needed to check your hypothesis and, at the same time, minimize the chance that one of the many types of bias will invalidate your survey results.

CHAPTER 6

Common Statistical Analysis Errors

By far the most common analysis error committed by novices is the use of the wrong type of statistical tests with survey data. Novice surveyors most frequently use intensity scale questions that make use of a Likert-type scale. In the section of Chapter 5 entitled **Intensity Questions and the Likert Scale**, we discussed the fact that these scales are typically assigned numerical weights to each adjective in the response set.

Professional surveying organizations empirically *anchor* their instruments. This is done to ensure that each adjective in the response set is an equivalent distance from its adjacent neighbors in the set. Anchoring is a labor intensive and complicated mathematical process whose explanation is beyond the scope of this handbook. Suffice it to say that most surveyors do not follow the process with the surveys they develop. The advantage of anchoring is that it creates a weighted scale along the entire response set of adjectives in which each adjective is a (mathematically) uniform distance from its neighboring adjectives. This creates what is known as an *interval weighted scale*. Without anchoring, one cannot be sure of the distance between responses in the set. Likert, or any type of multiple-choice scale, that is not anchored will produce either *nominal* or *ordinal* data. For example, consider the following response set:

Very Important	Somewhat Important	Not Important	Quite Unimportant

Is the distance between **Very Important** and **Somewhat Important** the same as that between **Not Important and Quite Unimportant**? How about the distance between any other pair of neighboring adjectives. Of course, you could claim that the distances are

equal, but how do you know all or even most respondents will see them as equal. All responses will be based their perceptions of the distances between each adjective.

So, you can see that the numerical weights you assign to each adjective is arbitrary. You could just as well weight them 5, 4, 3, 2, 1; or 2, 1, 0, -1, -2; or 100, 50, 25, 0. Each scale is as arbitrary as the next. The reason this is important is because whether the data generated by the survey are on an interval scale or not determines the particular kind of statistical tests you should use to analyze your data.

Most surveyors use descriptive statistics to provide general analysis of the response data. The most common descriptive statistics are the mean, variance, standard deviation, range, frequency counts, and percentage distribution. Some of these (such as the mean, variance, and standard deviation) require interval data be used to make correct interpretations of results. Percentages and frequency counts will work on any type of data (nominal, ordinal, or interval). The mathematical computations for these statistics will work regardless of the kind of data you input, however. So, by using nominal or ordinal data in computing the mean will yield a result, but it may not necessarily be a true, or meaningful result. Consider the following example. You send out a survey containing one question to 100 people. The question is:

How do you feel about the President's economic policy?

Like it a lot	**Neutral**	**Hate it**
(1)	**(2)**	**(3)**

All 100 people respond to the question. Fifty say they like it a lot and 50 say they hate it. So, we weight each of the responses accordingly and get a total weight of:

$$(50 \times 3) + (50 \times 1) = 150 + 50 = 200.$$

Dividing this total weight by the number of respondents (100) yields the average or mean response for the survey:

$$\frac{200}{100} = 2.0 \text{ -- (equivalent to a neutral rating).}$$

We interpret this to say that on average, people we surveyed are neutral toward the President's economic policy. Obviously, this is an erroneous interpretation of the actual results. Not one of the respondents was neutral to the question, yet the *average* response is neutral.

The underlying problem is that the original data were not based on an interval scale. That is, our one-question survey did not have an anchored scale. Anchored scales display the desirable quality of having equal intervals between each point along the scale. Without anchoring, you cannot be sure that scale points are equidistant from each other. In our current example, interval data were not generated because the scale was not anchored. Consequently, we should not have computed the mean in the first place, because we could not reliably interpret it. Of course, if the survey contained more questions and/or a diverse set of response scales, the problem would only have been magnified.

The proper descriptive analysis for nominal or ordinal data is to report frequencies (or percentages) of responses per category. In our example, it would have been most correct to simply report that 50 percent of those responding indicated they like the President's economic policy and 50 percent indicated they hate it. Such a report is very easy to interpret, and provides accurate, useful data for decision-makers.

Some surveyors are also interested in determining if responses from different groups of respondents are statistically different or not. Similarly, some are interested to know if respondents' answers to certain questions of the survey are related somehow either to their answers to other questions on the same survey or to some demographic characteristic (their gender, rank, age, race, etc.). To answer these types of questions, surveyors must use a class of statistics known as *inferential statistics*. As with the descriptive statistics discussed above, there are different inferential statistics for use with interval data and with nominal or ordinal data. The former are called *parametric statistics*, while the latter are called *non-parametric statistics*.

It is enough here to mention the names of the most basic statistical tests used to answer questions about differences between respondent groups and relationships between responses. On the parametric side, the t-test is a common test to determine if a statistically significant difference exists between two (and only two) groups of respondents. To test for significant differences between three or more groups, the most common parametric test used is Analysis of Variance (ANOVA). On the non-parametric side, one should use a *Chi-Square test* if the data are in the

form of frequencies or counts within categories, or a *Mann-Whitney (U) test* if the data are in the form of ranks. The *Chi-Square* (χ^2) test works regardless of how many groups (categories) there are.

To determine relationships between responses, a useful parametric test is the *Pearson Product Moment Correlation Coefficient* (also known as the Pearson correlation coefficient or, simply, the Pearson r), while on the non-parametric side, there is the *Contingency Coefficient* (C). A book written by Bruning and Kintz (1973), entitled Computational Handbook of Statistics, provides step-by-step procedures for manually calculating these and many other useful statistics with the use of just a hand calculator. Their handbook is highly recommended as a basic resource text. You should be able to obtain a copy through your local library, or purchase one through a local bookstore.

APPENDIX A

Steps in Surveying

1. Define the purpose. Be specific!
2. Review existing data. Is a survey needed?
3. Read applicable regulations.
4. Define the hypothesis.
5. Define the population.
6. Develop the survey (& sampling) plan.
7. Develop cover letter, instructions, & Privacy Act Statement.
8. Develop survey questions.
9. Pretest instrument.
10. Edit and revise questionnaire.
11. Obtain approvals as required.
12. Survey (gather data).
13. Quality control/data reduction.
14. Analysis and interpretation of results.
15. Prepare report for customer(s).

APPENDIX B

Survey Development Timetables

Timetable for Major Surveys

	Calendar Day
1. Deliver to printer	1
2. Printing completed	14
3. Receipt of questionnaire by local surveying activities	24
4. Receipt of survey questionnaires by respondents	30
5. Six-week administration period ends	75
6. Answer sheets or completed questionnaires returned by respondents	80
7. Receive completed answer sheets by data reduction activity	90
8. Initial results available	100

Source: Guide for the Development of the Attitude and Opinion Survey, October 1974; pp. 16-17

Timetable for Surveys With Expedited
Printing and Direct Mailing to Respondents

	Calendar Day
1. Printing completed	1
2. Receipt of survey questionnaires by respondents	17
3. First of questionnaires returned to surveyor	27
4. Follow-up letters sent	37
5. Final set of questionnaires returned to surveyor	50

Source: Guide for the Development of the Attitude and Opinion Survey, October 1974; pp. 16-17

APPENDIX C

Survey Data Sources

- Air University
  HQ AU/CFA
  55 LeMay Plaza South
  Maxwell AFB AL 36112-6335
  Attn: Survey Control Officer

  Approval source for all Air University surveys.

- National Technical Information Service (NTIS)
  US Department of Commerce
  5285 Port Royal Road
  Springfield VA 22161

  Good source of published reports of Government agencies.

APPENDIX D

How to Use a Random Number Table

1. Number each member of the population.
2. Determine population size (N).
3. Determine sample size (M).
4. Determine starting point in table by randomly picking a page and dropping your finger on the page with your eyes closed.
5. Choose a direction in which to read (up to down, left to right, or right to left).
6. Select the first M numbers read from the table whose last X digits are between 0 and N. (If N is a two digit number, then X would be 2; if it is a four digit number, X would be 4; etc.).
7. Once a number is chosen, do not use it again.
8. If you reach the end of the table before obtaining your M numbers, pick another starting point, read in a different direction, use the first X digits, and continue until done.

Example: N = 300; M = 50; starting point is column 3, row 2 on Random Number Table (next page); read down. You would select population numbers 43, 13, 122, 169, etc., until you had 50 unique numbers.

59468
99699
14**043**
15**013**
12600
33**122**
94**169**
etc.....

TABLE of RANDOM NUMBERS

	1	2	3	4	5	6	7	8	9	10
1	96268	11860	83699	38631	90045	69696	48572	05917	51905	10052
2	03550	59144	59468	37984	77892	89766	86489	46619	50236	91136
3	22188	81205	99699	84260	19693	36701	43233	62719	53117	71153
4	63759	61429	14043	44095	84746	22018	19014	76781	61086	90216
5	55006	17765	15013	77707	54317	48862	53823	52905	70754	68212
6	81972	45644	12600	01951	72166	52682	37598	11955	73018	23528
7	06344	50136	33122	31794	86723	58037	36065	32190	31367	96007
8	92363	99784	94169	03652	80824	33407	40837	97749	18361	72666
9	96083	16943	89916	55159	62184	86206	09764	20244	88388	98675
10	92993	10747	08985	44999	35785	65036	05933	77378	92339	96151
11	95083	70292	50394	61947	65591	09774	16216	63561	59751	78771
12	77308	60721	96057	86031	83148	34970	30892	53489	44999	18021
13	11913	49624	28519	27311	61586	28576	43092	69971	44220	80410
14	70648	47484	05095	92335	55299	27161	64486	71307	85883	69610
15	92771	99203	37786	81142	44271	36433	31726	74879	89384	76886
16	78816	20975	13043	55921	82774	62745	48338	88348	61211	88074
17	79934	35392	56097	87613	94627	63622	08110	16611	88599	02890
18	64698	83376	87527	36897	17215	74339	69856	43622	22567	11518
19	44212	12995	03581	37618	94851	63020	65348	55857	91742	79508
20	89292	00204	00579	70630	37136	50922	83387	15014	51838	81760
21	08692	87237	87879	01629	72184	33853	95144	67943	19345	03469
22	67927	76855	50702	78555	97442	78809	40575	79714	06201	34576
23	62167	94213	52971	85794	68067	78814	40103	70759	92129	46716
24	45828	45441	74220	84157	23241	49332	23646	09390	13031	51569
25	01164	35307	26526	80335	58090	85871	07205	31749	40571	51755
26	29283	31581	04359	45538	41435	61103	32428	94042	39971	63678
27	19868	49978	81699	84904	50163	22652	07845	71308	00859	87984
28	14292	93587	55960	23159	07370	65065	06580	46285	07884	83928
29	77410	52135	29495	23032	83242	89938	40516	27252	55565	64714
30	36580	06921	35675	81645	60479	71035	99380	59759	42161	93440
31	07780	18093	31258	78156	07871	20369	53977	08534	39433	57216
32	07548	08454	36674	46255	80541	42903	37366	21164	97516	66181
33	22023	60448	69344	44260	90570	01632	21002	24413	04671	05665
34	20827	37210	57797	34660	32510	71558	78228	42304	77197	79168
35	47802	79270	48805	59480	88092	11441	96016	76091	51823	94442
36	76730	86591	18978	25479	77684	88439	34112	26052	57112	91653
37	26439	02903	20935	76297	15290	84688	74002	09467	41111	19194
38	32927	83426	07848	59372	44422	53372	27823	25417	27150	21750
39	51484	05286	77103	47284	00578	88774	15293	50740	07932	87633
40	45142	96804	92834	26886	70002	96643	36008	02239	93563	66429

APPENDIX E

Table of Z Values

Confidence Level	Z Factor
99.9	3.2905
99.7	3.0000
99.5	2.8070
99.0	2.5758
98.0	2.3263
95.5	2.0000
95.0	1.9600
90.0	1.6449
85.0	1.4395
80.0	1.2816

APPENDIX F

Sample Cover Letter, Privacy Act Statement, And Instruction Sheet

SAMPLE COVER LETTER

(Letterhead)

FROM: EXO (Major Ross, 3-2044) *{Date}*

SUBJ: Survey On-Base Facilities

TO:

1. Periodically, this command surveys its personnel to determine the effectiveness of base facilities and the desires of members concerning the requirements for any additional facilities. The attached questionnaire was developed to obtain this information. Results from this survey will be used to improve current facilities and to help plan for new ones.

2. This is your chance to express your opinions on current base facilities and to identify the need for additional ones. Please answer the questions as candidly as possible to provide us a valid assessment regarding facility improvements and additions. Participation in this survey is voluntary, and no attempt will be made to attribute the answers with specific respondents. I solicit your prompt cooperation in this project and thank you for your time.

JOHN J. JONES, Col, USAF
Base Commander

Attachment:
Questionnaire

SAMPLE PRIVACY ACT STATEMENT

PRIVACY ACT STATEMENT FOR USAF-SCN XX-XXX

In accordance with AFR 12-35, paragraph 30, the following information is provided as required by the Privacy Act of 1974:

a. Authority:

(1) 5 USC 301, Departmental Regulations; and/or

(2) 10 USC 8012, Secretary of the Air Force, Powers, Duties, Delegation by Compensation.

b. Principal Purpose: To sample Air Force officer opinion and attitudes concerning base facilities.

c. Routine Uses: To provide data as part of a base facilities study.

d. Participation in this survey is voluntary and respondents will not be identified.

e. No adverse action of any kind may be taken against any individual who elects not to participate in any or all parts of this survey.

SAMPLE SURVEY INSTRUCTIONS
(Designed for ADP Scanner Sheets)

SURVEY INSTRUCTIONS

1. Do not write your name or social security account number (SSAN) on the answer sheet.

2. There are no right or wrong answers to the questions on this survey. Select the or most appropriate response for each question.

3. Use a No. 2 pencil when marking your answers on the answer sheet. DO NOT use pen or marker.

4. Be sure your answer marks blacken the entire rectangle on the answer sheet.

5. Be sure to mark your answers carefully so that you enter them opposite the same answer sheet number as survey question number.

6. Upon completion, please place your answer sheet in the attached envelope and place the envelope in base distribution.

7. To help us ensure we meet our suspenses, please try to return your completed answer sheet by (return date).

8. Thank you for your time and cooperation.

APPENDIX G

Rating or Intensity Scales

Excerpted from ARI Technical Report #P-77-1, U.S. Army Research Institute for the Behavioral Social Sciences, Fort Hood Field Unit, July 1976, pp VIII-E-1 through VIII-E-24. (DTIC No. ADA037815)

Selection of Response Alternatives Using Scale Values and Standard Deviations

Using scale values and standard deviations to select response alternatives will give a more refined set of phrases than using an order of merit list. In general, response alternatives selected from lists of phrases with scale values should usually have the following characteristics:

- The scale values of the terms should be as far apart as possible.
- The scale values of the terms should be as equally distant as possible.
- The terms should have small variability (small standard deviations or interquartile ranges).
- Other things being equal, the terms should have parallel wording.

Tables VIII-E-1 through VIII-E-24 give lists of phrases which have scale values and, when possible, standard deviations or interquartile range. They are based on empirical evidence, and may be used to select response alternatives. Bibliographic source information supporting the citations in the heading of each table can be found in DTIC document ADA037815.

Table VIII-E-1 -- Acceptability Phrases (from: U.S. Army, 1973)

Phrase	*Average*	*Std. Dev.*
Excellent	6.27	0.54
Perfect in every respect	6.22	0.86
Extremely good	5.74	0.81
Very good	5.19	0.75
Unusually good	5.03	0.98
Very good in most respects	4.62	0.72
Good	4.25	0.90
Moderately good	3.58	0.77
Could use some minor changes	3.28	1.09
Not good enough for extreme conditions	3.10	1.30
Not good for rough use	2.72	1.15
Not very good	2.10	0.95
Needs major changes	1.97	1.12
Barely acceptable	1.79	0.90
Not good enough for general use	1.76	1.21
Better than nothing	1.22	1.08
Poor	1.06	1.11
Very poor	0.76	0.95
Extremely poor	0.36	0.76

Table VIII-E-2 -- Degrees of Excellence: First Set
(from: Myers and Warner, 1968))

Phrase	*Scale Value*	*Std. Dev.*
Superior	20.12	1.17
Fantastic	20.12	0.83
Tremendous	19.84	1.31
Superb	19.80	1.19
Excellent	19.40	1.73
Terrific	19.00	2.45
Outstanding	18.96	1.99
Wonderful	17.32	2.30
Delightful	16.92	1.85
Fine	14.80	2.12
Good	14.32	2.08
Pleasant	13.44	2.06
Nice	12.56	2.14
Acceptable	11.12	2.59
Average	10.84	1.55
All right	10.76	1.42
OK	10.28	1.67
Neutral	9.80	1.50
Fair	9.52	2.06
Mediocre	9.44	1.80
Unpleasant	5.04	2.82
Bad	3.88	2.19
Very Bad	3.20	2.19
Unacceptable	2.64	2.04
Awful	1.92	1.50
Terrible	1.76	0.77
Horrible	1.48	0.87

Table VIII-E-3 -- Degrees of Excellence: Second Set
(from: Jones and Thurstone, 1955)

Phrase	*Scale Value*	*Std. Dev.*
Best of all	6.15	2.43
Excellent	3.71	1.01
Wonderful	3.51	0.97
Mighty fine	2.88	0.67
Especially good	2.86	0.82
Very good	2.55	0.87
Good	1.91	0.76
Pleasing	1.58	0.65
OK	0.87	1.24
Fair	0.78	0.85
Only fair	0.71	0.64
Not pleasing	-0.83	0.67
Poor	-1.55	0.87
Bad	-2.02	0.80
Very Bad	-2.53	0.64
Terrible	-3.09	0.98

Table VIII-E-4 -- Degrees of Like and Dislike
(from: Jones and Thurstone, 1955)

Phrase	*Scale Value*	*Std. Dev.*
Like extremely	4.16	1.62
Like intensely	4.05	1.59
Strongly like	2.96	0.69
Like very much	2.91	0.60
Like very well	2.60	0.78
Like quite a bit	2.32	0.52
Like fairly well	1.51	0.59
Like	1.35	0.77
Like moderately	1.12	0.61
Mildly like	0.85	0.47
Like slightly	0.69	0.32
Neutral	0.02	0.18
Like not so well	-0.30	1.07
Like not so much	-0.41	0.94
Dislike slightly	-0.59	0.27
Mildly dislike	-0.74	0.35
Dislike moderately	-1.20	0.41
Dislike	-1.58	0.94
Don’t like	-1.81	0.97
Strongly dislike	-2.37	0.53
Dislike very much	-2.49	0.64
Dislike intensely	-3.33	1.39
Dislike extremely	-4.32	1.86

Table VIII-E-5 -- Degrees of Good and Poor
(from: Myers and Warner, 1968)

Phrase	*Scale Value*	*Std. Dev.*
Exceptionally good	18.55	2.36
Extremely good	18.44	1.61
Unusually good	17.08	2.43
Remarkably good	16.68	2.19
Very good	15.44	2.75
Quite good	14.44	2.76
Good	14.32	2.08
Moderately good	13.44	2.23
Reasonably good	12.92	2.93
Fairly good	11.96	2.42
Slightly good	11.84	2.19
So-so	10.08	1.87
Not very much	6.72	2.82
Moderately poor	6.44	1.64
Reasonably poor	6.32	2.46
Slightly poor	5.92	1.36
Poor	5.72	2.09
Fairly poor	5.64	1.68
Quite poor	4.80	1.44
Unusually poor	3.20	1.44
Very poor	3.12	1.17
Remarkably poor	2.88	1.74
Exceptionally poor	2.52	1.19
Extremely poor	2.08	1.19

Table VIII-E-6 -- Degrees of Good and Bad (from: Cliff, 1959)

Phrase	*Scale Value*
Extremely good	3.449
Very good	3.250
Unusually good	3.243
Decidedly good	3.024
Quite good	2.880
Rather good	2.755
Good	2.712
Pretty good	2.622
Somewhat good	2.462
Slightly good	2.417
Slightly bad	1.497
Somewhat bad	1.323
Rather bad	1.232
Bad	1.024
Pretty bad	1.018
Quite bad	0.924
Decidedly bad	0.797
Unusually bad	0.662
Very bad	0.639
Extremely bad	0.470

Table VIII-E-7 -- Degrees of Agree and Disagree
(from: Altemeyer, 1970)

Phrase	*Mean*	*Std. Dev.*
Decidedly agree	2.77	.41
Quite agree	2.37	.49
Considerably agree	2.21	.42
Substantially agree	2.10	.50
Moderately agree	1.47	.41
Somewhat agree	.94	.41
Slightly agree	.67	.36
Perhaps agree	.52	.46
Perhaps disagree	-.43	.45
Slightly disagree	-.64	.38
Somewhat disagree	-.98	.47
Moderately disagree	-1.35	.42
Quite disagree	-2.16	.57
Substantially disagree	-2.17	.51
Considerably disagree	-2.17	.45
Decidedly disagree	-2.76	.43

Table VIII-E-8 -- Degrees of More and Less
(from: Dodd and Gerberick, 1960)

Phrase	*Scale Value*	*Interquartile Range* [a]
Very much more	8.02	0.61
Much more	7.67	1.04
A lot more	7.50	1.06
A good deal more	7.29	0.98
More	6.33	1.01
Somewhat more	6.25	0.98
A little more	6.00	0.58
Slightly more	5.99	0.57
Slightly less	3.97	0.56
A little less	3.96	0.54
Less	3.64	1.04
Much less	2.55	1.06
A good deal less	2.44	1.11
A lot less	2.36	1.03
Very much less	1.96	0.52

[a] Minimum = 0.05

Table VIII-E-9 -- Degrees of Adequate and Inadequate
(from: Matthews, Wright, and Yudowitch, 1975)

Phrase	*Mean*	*Std. Dev.*
Totally adequate	4.620	.846
Absolutely adequate	4.540	.921
Completely adequate	4.490	.825
Extremely adequate	4.412	.719
Exceptionally adequate	4.380	.869
Entirely adequate	4.340	.863
Wholly adequate	4.314	1.038
Fully adequate	4.294	.914
Very, very adequate	4.063	.876
Perfectly adequate	3.922	1.026
Highly adequate	3.843	.606
Most adequate	3.843	.978
Very adequate	3.420	.851
Decidedly adequate	3.140	1.536
Considerably adequate	3.020	.874
Quite adequate	2.980	.979
Largely adequate	2.863	.991
Substantially adequate	2.608	1.030
Reasonably adequate	2.412	.771
Pretty adequate	2.306	.862
Rather adequate	1.755	.893
Mildly adequate	1.571	.670
Somewhat adequate	1.327	.793
Slightly adequate	1.200	.566
Barely adequate	0.627	.928

Phrase	*Mean*	*Std. Dev.*
Neutral	0.000	.000
Borderline	-0.020	.316
Barely inadequate	-1.157	.638
Mildly inadequate	-1.353	.621
Slightly inadequate	-1.380	.772
Somewhat inadequate	-1.882	.732
Rather inadequate	-2.102	.974
Moderately inadequate	-2.157	1.017
Fairly inadequate	-2.216	.800
Pretty inadequate	-2.347	.959
Considerably inadequate	-3.600	.680
Very inadequate	-3.735	.777
Decidedly inadequate	-3.780	.944
Most inadequate	-3.980	1.545
Highly inadequate	-4.196	.741
Very, very inadequate	-4.460	.537
Extremely inadequate	-4.608	.527
Fully inadequate	-4.667	.676
Exceptionally inadequate	-4.680	.508
Wholly inadequate	-4.784	.498
Entirely inadequate	-4.792	.644
Completely inadequate	-4.800	.529
Absolutely inadequate	-4.800	.431
Totally inadequate	-4.900	.412

Table E VIII-E-10 -- Degrees of Acceptable and Unacceptable
(from: Matthews, Wright, and Yudowitch (1975)

Phrase	*Mean*	*Std. Dev.*
Wholly acceptable	4.725	.563
Completely acceptable	4.686	.610
Fully acceptable	4.412	.867
Extremely acceptable	4.392	.716
Most acceptable	4.157	.915
Very, very acceptable	4.157	.825
Highly acceptable	4.040	.631
Quite acceptable	3.216	.956
Largely acceptable	3.137	.991
Acceptable	2.392	1.456
Reasonably acceptable	2.294	.722
Moderately acceptable	2.280	.722
Pretty acceptable	2.000	1.125
Rather acceptable	1.939	.818
Fairly acceptable	1.840	.924
Mildly acceptable	1.686	.700
Somewhat acceptable	1.458	1.241
Barely acceptable	1.078	.518
Slightly acceptable	1.039	.522
Sort of acceptable	0.940	.645
Borderline	0.000	.200
Neutral	0.000	0.000
Marginal	-0.120	.515
Barely unacceptable	-1.100	.300

Phrase	*Mean*	*Std. Dev.*
Slightly unacceptable	-1.255	.589
Somewhat unacceptable	-1.765	.674
Rather unacceptable	-2.020	.836
Fairly unacceptable	-2.160	.880
Moderately unacceptable	-2.340	.681
Pretty unacceptable	-2.412	.662
Reasonably unacceptable	-2.440	.753
Unacceptable	-2.667	1.381
Substantially unacceptable	-3.235	.899
Quite unacceptable	-3.388	1.066
Largely unacceptable	-3.392	.818
Considerably unacceptable	-3.440	.779
Notably unacceptable	-3.500	1.044
Decidedly unacceptable	-3.837	1.017
Highly unacceptable	-4.294	.535
Most unacceptable	-4.420	.724
Very, very unacceptable	-4.490	.500
Exceptionally unacceptable	-4.540	.607
Extremely unacceptable	-4.686	.464
Completely unacceptable	-4.900	.361
Entirely unacceptable	-4.900	.361
Wholly unacceptable	-4.922	.269
Absolutely unacceptable	-4.922	.334
Totally unacceptable	-4.941	.235

Table VIII-E-11 -- Comparison Phrases
(from Matthews, Wright, and Yudowitch, 1975)

Phrase	*Mean*	*Std. Dev.*
Best of all	4.896	.510
Absolutely best	4.843	.459
Truly best	4.600	.721
Undoubtedly best	t4.569	.823
Decidedly best	4.373	.839
Best	4.216	1.459
Absolutely better	4.060	.988
Extremely better	3.922	.882
Substantially better	3.700	.922
Decidedly better	3.412	.933
Conspicuously better	3.059	.802
Moderately better	2.255	.737
Somewhat better	1.834	.801
Rather better	1.816	.719
Slightly better	1.157	.776
Barely better	0.961	.656
Absolutely alike	0.538	1.623
Alike	0.216	.847
The same	0.157	.801
Neutral	0.000	0.000

Phrase	*Mean*	*Std. Dev.*
Borderline	-0.061	.314
Marginal	-0.184	.919
Barely worse	-1.039	.816
Slightly worse	-1.216	.498
Somewhat worse	-2.078	.860
Moderately worse	-2.220	.944
Noticeably worse	-2.529	1.030
Worse	-2.667	1.423
Notably worse	-3.020	1.038
Largely worse	-3.216	1.108
Considerably worse	-3.275	1.206
Conspicuously worse	-3.275	.887
Much worse	-3.286	.808
Substantially worse	-3.460	.899
Decidedly worse	-3.760	.907
Very much worse	-3.941	.752
Absolutely worse	-4.431	.823
Decidedly worst	-4.431	.748
Undoubtedly worst	-4.510	.872
Absolutely worst	-4.686	1.291
Worst of all	-4.776	1.298

Table VII-E-12 -- Degrees of Satisfactory and Unsatisfactory (from: U.S. Army, 1973)

Phrase	*Scale Value*	*Std. Dev.*
Quite satisfactory	4.35	.95
Satisfactory	3.69	.87
Not very satisfactory	2.11	.76
Unsatisfactory but usable	2.00	.87
Very unsatisfactory	0.69	1.32

Table VIII-E-13 -- Degrees of Unsatisfactory (from: Mosier, 1941)

Phrase	*Scale Value*
Unsatisfactory	1.47
Quite unsatisfactory	1.00
Very unsatisfactory	0.75
Unusually unsatisfactory	0.75
Highly unsatisfactory	0.71
Very, very unsatisfactory	0.25
Extremely unsatisfactory	0.10
Completely unsatisfactory	0.00

Table VIII-E-14 -- Degrees of Pleasant (from: Cliff, 1959)

Phrase	*Scale Value*
Extremely pleasant	3.490
Very pleasant	3.174
Unusually pleasant	3.107
Decidedly pleasant	3.028
Quite Pleasant	2.849
Pleasant	2.770
Rather pleasant	2.743
Pretty pleasant	2.738
Somewhat pleasant	2.505
Slightly pleasant	2.440

Table VIII-E-15 -- Degrees of Agreeable (from: Mosier, 1941)

Phrase	*Scale Value*
Very, very agreeable	5.34
Extremely agreeable	5.10
Highly agreeable	5.02
Completely agreeable	4.96
Unusually agreeable	4.86
Very agreeable	4.82
Quite agreeable	4.45
Agreeable	4.19

Table VIII-E-16 -- Degrees of Desirable (from: Mosier, 1941)

Phrase	*Scale Value*
Very, very desirable	5.66
Extremely desirable	5.42
Completely desirable	5.38
Unusually desirable	5.23
Highly desirable	5.25
Very desirable	4.96
Quite desirable	4.76
Desirable	4.50

Table VIII-E-17 -- Degrees of Nice

Phrase	*Scale Value*
Extremely nice	3.351
Unusually nice	3.155
Very nice	3.016
Decidedly nice	2.969
Pretty nice	2.767
Quite nice	2.738
Nice	2.636
Rather nice	2.568
Somewhat nice	2.488
Slightly nice	2.286

Table VIII-E-18 -- Degrees of Adequate (from: U.S. Army, 1973)

Phrase	*Scale Value*	*Std. Dev.*
More than adequate	4.13	1.11
Adequate	3.39	.87
Not quite adequate	2.40	.85
Barely adequate	2.10	.84
Not adequate	1.83	.98

Table VIII-E-19 -- Degrees of Ordinary (from: Cliff, 1959)

Phrase	*Scale Value*
Ordinary	2.074
Very ordinary	2.073
Somewhat ordinary	2.038
Rather ordinary	2.934
Pretty ordinary	2.026
Slightly ordinary	1.980
Decidedly ordinary	1.949
Extremely ordinary	1.936
Unusually ordinary	1.875

Table VIII-E-20 -- Degrees of Average (from: Cliff, 1959)

Phrase	*Scale Value*
Rather average	2.172
Average	2.145
Quite average	2.101
Pretty average	2.094
Somewhat average	2.080
Unusually average	2.062
Extremely average	2.052
Very average	2.039
Slightly average	2.023
Decidedly average	2.020

Table VIII-E-21 -- Degrees of Hesitation (from: Dodd and Gerberick, 1960)

Phrase	*Scale Value*	*Interquartile Range* [a]
Without hesitation	7.50	6.54
With little hesitation	5.83	3.40
Hesitant	4.77	1.06
With some hesitation	4.38	1.60
With considerable hesitation	3.29	3.39
With much hesitation	3.20	5.25
With great hesitation	2.41	6.00

[a] Minimum = 0.5

Table VIII-E-22 -- Degrees of Inferior (from: Cliff, 1959)

Phrase	*Scale Value*
Slightly inferior	1.520
Somewhat inferior	1.516
Inferior	1.323
Rather inferior	1.295
Pretty inferior	1.180
Quite inferior	1.127
Decidedly inferior	1.013
Unusually inferior	0.963
Very inferior	0.927
Extremely inferior	0.705

Table VIII-E-23 -- Degrees of Poor (from: Mosier, 1941)

Phrase	*Scale Value*
Poor	1.60
Quite poor	1.30
Very poor	1.18
Unusually poor	0.95
Extremely poor	0.95
Completely poor	0.92
Very, very poor	0.55

Table VIII-E-24 -- Descriptive Phrases
(from: Dodd and Gerberick, 1960)

Phrase	*Scale Value*	*Interquartile Range* [a]
Complete	8.85	.65
Extremely vital	8.79	.84
Very certain	8.55	1.05
Very strongly	8.40	1.04
Very critical	8.29	1.12
Very important	8.22	1.16
Very sure	8.15	.95
Almost complete	8.06	.58
Of great importance	8.05	.91
Very urgent	8.00	.90
Feel strongly toward	7.80	1.60
Essential	7.58	1.85
Very vital	7.55	1.05
Certain	7.13	1.44
Strongly	7.07	.67
Important	6.83	1.14
Good	6.72	1.20
Urgent	6.41	1.53
Crucial	6.39	1.73
Sure	5.93	1.87
Vital	5.92	1.63
Moderately	5.24	.99
Now	5.03	.53
As at present	5.00	.50
Fair	4.96	.77
Don't know	4.82	.82
Undecided	4.73	1.06
Don't care	4.63	2.00
Somewhat	3.79	.94
Indifferent	3.70	2.20
Object strongly to	3.50	6.07
Not important	3.09	1.33
Unimportant	2.94	1.42
Bad	2.89	.93
Uncertain	2.83	2.50
Doesn't make any difference	2.83	3.13
Not sure	2.82	1.24
Not certain	2.64	2.62
Non-essential	2.58	1.67
Doesn't mean anything	2.50	2.71
Insignificant	2.12	1.14
Very little	2.08	.64
Almost none	2.04	.57
Very unimportant	1.75	1.25
Only as a last resort	1.70	7.30
Very bad	1.50	1.13
None	1.11	.59

[a] Minimum = 0.5

APPENDIX H

Sample Sets of Response Alternatives

Excerpted from Questionnaire Construction Manual, ARI Technical Report #P-77-1, U.S. Army Research Institute for the Behavioral and Social Sciences, Fort Hood Field Unit, July 1976, ppVIII-F-1 through VIII-F-4. (DTIC No. ADA037815)

It is sometimes valuable and is a time saver to have lists of response alternatives available to use. The tables in this section give some examples of response alternatives that have been selected on different bases. These sets do not exhaust all possibilities.

The sets of response alternatives that appear in Table VIII-F-1 were selected so that the phrases in each set would have means at least one standard deviation away from each other and have parallel wording. Some of the sets of response alternatives have extreme end points; some do not. The sets of response alternatives shown in Table VIII-F-2 were selected so that the phrases in each set would be as nearly equally distant from each other as possible without regard to parallel wording. Table VIII-F-3 contains sets of response alternatives selected from lists of descriptors with only scale values given. The phrases were selected on the bases of equal appearing intervals. Table VIII-F-4 has sets of response alternatives selected from order of merit lists of descriptors.

Table VIII-F-1

Sets Selected So Phrases Are at Least One Standard Deviation Apart and Have Parallel Wording (18 Sets)

Set #	Response Alternatives
1	Completely acceptable Reasonably acceptable Barely acceptable Borderline Barely unacceptable Reasonably unacceptable Completely unacceptable
2	Wholly acceptable Largely acceptable Borderline Largely unacceptable Wholly unacceptable
3	Largely acceptable Barely acceptable Borderline Barely unacceptable Largely unacceptable
4	Reasonably acceptable Slightly acceptable Borderline Slightly unacceptable Reasonably unacceptable
5	Totally adequate Very adequate Barely adequate Borderline Barely inadequate Very inadequate Totally inadequate
6	Completely adequate Considerably adequate Borderline Considerably inadequate Completely inadequate
7	Very adequate Slightly adequate Borderline Slightly inadequate Very inadequate
8	Highly adequate Mildly adequate Borderline Mildly inadequate Highly inadequate
9	Decidedly agree Substantially agree Slightly agree Slightly disagree Substantially disagree Decidedly disagree
10	Moderately agree Perhaps agree Neutral Perhaps disagree Moderately disagree
11	Undoubtedly best Conspicuously better Moderately better Alike Moderately worse Conspicuously worse Undoubtedly worst
12	Moderately better Barely better The same Barely worse Moderately worse

Table VIII-F-1 (Cont'd)

Sets Selected So Phrases Are at Least One Standard Deviation Apart and Have Parallel Wording (18 Sets)

Set #	Response Alternatives
13	Extremely good Remarkably good Good So-so Poor Remarkably poor Extremely poor
14	Exceptionally good Reasonably good So-so Reasonably poor Exceptionally poor
15	Very important Important Not important Very unimportant

Set #	Response Alternatives
16	Like extremely Like moderately Neutral Dislike moderately Dislike extremely
17	Strongly like Like Neutral Don't like Strongly dislike
18	Very much more A good deal more A little more A little less A good deal less Very much less

Table VIII-F-2

Sets Selected So That Intervals Between Phrases Are as Nearly Equal as Possible (15 Sets)

Set #	Response Alternatives
1	Completely acceptable Reasonably acceptable Borderline Moderately unacceptable Extremely unacceptable
2	Totally adequate Pretty adequate Borderline Pretty inadequate Extremely inadequate
3	Highly adequate Rather adequate Borderline Somewhat inadequate Decidedly inadequate
4	Quite agree Moderately agree Perhaps agree Perhaps disagree Moderately disagree Substantially disagree
5	Undoubtedly best Moderately better Borderline Noticeably worse Undoubtedly worst

Set #	Response Alternatives
6	Fantastic Delightful Nice Mediocre Unpleasant Horrible
7	Perfect in every respect Very good Good Could use minor changes Not very good Better than nothing Extremely poor
8	Excellent Good Only fair Poor Terrible
9	Extremely good Quite good So-so Slightly poor Extremely poor
10	Remarkably good Moderately good So-so Not very good Unusually poor

Set #	Response Alternatives
11	Without hesitation With little hesitation With some hesitation With great hesitation
12	Strongly like Like quite a bit Like Neutral Mildly dislike Dislike very much Dislike extremely
13	Like quite a bit Like Like slightly Borderline Dislike Dislike moderately Don't like
14	Like quite a bit Like fairly well Borderline Dislike moderately Dislike very much
15	Very much more A little more Slightly less Very much less

Table VIII-F-3

Sets Selected From Lists Giving Scale Values Only
So That Intervals Between Phrases Are as Nearly Equal as Possible (10 Sets)

Set #	Response Alternatives
1	Very, very agreeable Usually agreeable Quite agreeable Agreeable
2	Rather average Quite average Unusually average Decidedly average
3	Very, very desirable Completely desirable Very desirable Desirable
4	Extremely good Somewhat good Slightly good Extremely bad
5	Slightly inferior Rather inferior Unusually inferior Extremely inferior
6	Extremely nice Decidedly nice Nice Slightly nice
7	Ordinary Slightly ordinary Unusually ordinary
8	Extremely pleasant Decidedly pleasant Somewhat pleasant
9	Poor Very poor Very, very poor
10	Very, very agreeable Extremely agreeable Very agreeable Quite agreeable Agreeable

Table VIII-F-4

Sets Selected Using Order-of-Merit Lists of Descriptor Terms (4 Sets)

Set #	Response Alternatives
1	Very good Good Borderline Poor Very poor
2	Very satisfactory Satisfactory Borderline Unsatisfactory Very unsatisfactory
3	Very superior Superior Borderline Poor Very poor
4	Extremely useful Of considerable use Of use Not very useful Of no use

BIBLIOGRAPHY

Air Force Instruction 36-2601. The Air Force Sample Survey Program. Washington, D.C.: Department of the Air Force, Publication Pending (1993).

Altemeyer, R. A. (1970). *Adverbs and Intervals: A Study of Likert Scales*. American Psychological Association, Proceedings of the Annual Convention of 1970, S (pt. 1) pp 397-398.

Borg, Walter R. and Gall, Meredith Damien (1979). Educational Research. Third Edition. New York: Longman, Inc.

Bruning, James L. and Kintz, B. L. (1977). Computational Handbook of Statistics. Second Edition. Glenview, IL: Scott, Foresman and Company.

Cantelou, Louis W., Jr (1964). *Opinion Research as a Managerial Tool for Internal Public Relations: The USAF Sample Survey and Internal Information Program*. Unpublished master's thesis, School of Public Relations and Communications, Boston, Mass.: Boston University Press.

Cliff, N. (1959). *Adverbs As Multipliers*. Psychological Review, 66, pp 27-44.

Cochran, W. G. (1963) Sampling Techniques. New York: Wiley and Sons, Inc.

Deming, William E. (1950) Some Theory of Sampling. New York: John Wiley and Sons, Inc.

Department of the Air Force. (October 1974). A Guide for the Development of the Attitude Opinion Survey. Washington, D.C.

Dodd, S. C., & Gerberick, T. R. (1960*). Word Scales for Degrees of Opinion*. Language and Speech, 3, pp 18-31.

Dyer, Robert., Matthews, J. J., Stulac, J. F., Wright, C. E., Yudowitch, Kenneth (1976). Questionnaire Construction Manual. Palo Alto: Operations Research Associates (prepared for the U.S. Army Research Institute for the Behavioral and Social Sciences, Fort Hood, Texas under Contract DAHC19-74-C-0032)

Huff, D. (1954). How to Lie with Statistics. New York: W. W. Norton and Company, Inc.

Jones, L. V., & Thurstone, L. L. (1955). *The Psychophysics of Semantics: An Experimental Investigation*. Journal of Applied Psychology, 39, pp 31-36.

Kish, Leslie. (1965) Survey Sampling. New York: John Wiley and Sons, Inc.

Matthews, J. J., Wright, C. E., & Yudowitch, K. L. (1975) Analysis of the Results of the Administration of Three Sets of Descriptive Adjective Phrases. Palo Alto: Operations Research Associates (prepared for the U.S. Army Research Institute for the Behavioral and Social Sciences, Fort Hood, Texas under Contract DAHC19-74-C-0032)

Miller, Delbert C. (1970). Handbook of Research Design and Social Measurement. New York: David McKay Co., Inc.

Mosier, C. I. (1941). *A Psychometric Study of Meaning*. Journal of Social Psychology, 13, pp 123-140.

Myers, J. H., & Warner, W. G. (1968). *Semantic Properties of Selected Evaluation Adjectives*. Journal of Marketing Research, 5, pp 409-412.

Oppenheim, A. N. (1966). Questionnaire Design and Attitude Measurement. New York: Basic Books, Inc.

Parten, M. (1950). Surveys, Polls, and Samples: Practical Procedures. New York: Harper and Brothers.

Rosenthal, Robert and Rosnow, Ralph, L. (1975). The Volunteer Subject. New York: John Wiley.

Sawyer, Lawrence. (November-December 1971*) Statistics Confuse Me, Grandfather*, Internal Auditor, Vol. 28, No. 6, pp 49-52.

Selltiz, Claire, Marie Jahoda, Morton Deutsch, and Stuart W. Cook (1963). Research Methods in Social Relations. New York: Holt, Rinehart and Winston, Inc.

U.S. Army Test and Evaluation Command (1973). Development of a Guide and Checklist for Human Factors Evaluation of Army Equipment and Systems, U.S. Army Test and Evaluation Command (TECOM), Contract DAADO5-73-C-0388.

Young, Pauline V. (1959) Scientific Social Surveys and Research. New York: Prentice-Hall, Inc.