

HEADQUARTERS

US ARMY TEST AND EVALUATION COMMAND

ABERDEEN PROVING GROUND, MARYLAND

LEVEL 4



This document has been approved for public release and sale; its distribution is unlimited.

11 25 Jul 75

12 13 p.

6
Man-Materiel Systems
QUESTIONNAIRE AND INTERVIEW DESIGN
(SUBJECTIVE TESTING TECHNIQUES),
Volume I.

AD A O 6 0 4 9 9

DDC FILE COPY

041750
78 09 22 09 6 B

**Best
Available
Copy**

FOREWORD

TECOM Pamphlet 602-1, Volume I, Questionnaire and Interview Design, is the first of two volumes on subjective testing techniques. It presents techniques for the design and administration of questionnaires and interviews, as well as procedure for treatment of the data. Volume II will present techniques for the development and use of checklists (procedural, design, life support, and maintainability), observational records, and error reports.

The purpose of this pamphlet is to provide a text and reference materials of subjective testing techniques. It is designed to serve as a guide and to provide a structured approach to proper application of subjective techniques in planning, conducting, and reporting development tests of Army materiel. The pamphlet also includes procedures for the tabulation and analysis of the data obtained.

The techniques described in this pamphlet are vital tools for determining if military equipment and weapons systems are designed and constructed for effective operation and maintenance by qualified military personnel. They are particularly relevant for obtaining valid and reliable data required for assessment of the Soldier-materiel interface.

ADDRESS		
NO.	ISSUE	✓
DATE	ISSUE	
REMARKS		
<i>After m/ple</i>		
BY		
DISTRIBUTION APPROVAL		
DATE		EXPIRES
A		

DEPARTMENT OF THE ARMY
HEADQUARTERS, US ARMY TEST AND EVALUATION COMMAND
Aberdeen Proving Ground, Maryland 21005

TECOM Pamphlet
No. 602-1
Volume 1

25 July 1975

Man-Materiel Systems
QUESTIONNAIRE AND INTERVIEW DESIGN
(SUBJECTIVE TESTING TECHNIQUES)

		PARAGRAPH	PAGE
PART ONE	QUESTIONNAIRE DESIGN AND ADMINISTRATION		
CHAPTER 1	INTRODUCTION		
	GENERAL	1-1	1-1
	PURPOSE OF THE QUESTIONNAIRE	1-2	1-1
	PROBLEMS IN QUESTIONNAIRE DESIGN	1-3	1-2
CHAPTER 2	METHOD OF QUESTIONNAIRE DESIGN		
	INTRODUCTION	2-1	2-1
	PRELIMINARY PLANNING	2-2	2-1
	SELECTING THE QUESTION FORM	2-3	2-2
	PRELIMINARY WORDING OF QUESTIONS	2-4	2-2
	FORMULATING THE QUESTIONNAIRE	2-5	2-2
	PRETESTING THE QUESTIONNAIRE	2-6	2-2
CHAPTER 3	PRELIMINARY PLANNING		
	INTRODUCTION	3-1	3-1
	DEFINING THE OBJECTIVES	3-2	3-1
	LISTING THE INFORMATION REQUIRED	3-3	3-2
	CONSIDERING THE TEST GROUP	3-4	3-2
	SELECTING THE TYPE OF QUESTIONNAIRE	3-5	3-5
	REFERENCES	3-6	3-6
	SELECTED READINGS	3-7	3-6
CHAPTER 4	SELECTING THE QUESTION FORM		
	INTRODUCTION	4-1	4-1
	THE FREE-ANSWER QUESTION	4-2	4-1
	THE DICHOTOMOUS QUESTION	4-3	4-8
	THE MULTIPLE CHOICE QUESTION	4-4	4-12
	THE RATING SCALE	4-5	4-24
	REFERENCES	4-6	4-33
	SELECTED READINGS	4-7	4-33

		PARAGRAPH	PAGE
CHAPTER 5	QUESTION WORDING		
	INTRODUCTION	5-1	5-1
	VOCABULARY LEVEL	5-2	5-2
	AMBIGUITY	5-3	5-2
	BIAS	5-4	5-3
	OTHER CONSIDERATIONS	5-5	5-5
	REMINDERS FOR QUESTION WORDING	5-6	5-7
	REFERENCES	5-7	5-8
	SELECTED READINGS	5-8	5-8
CHAPTER 6	FORMULATING THE QUESTIONNAIRE		
	INTRODUCTION	6-1	6-1
	QUESTION SEQUENCE	6-2	6-1
	ESTABLISHING THE FORMAT	6-3	6-3
	CHECKING AGAINST THE OBJECTIVES	6-4	6-5
	USE OF PUNCH CARDS	6-5	6-5
	CODING AND CLASSIFYING RESPONSES	6-6	6-6
	CATEGORIZING RESPONSES	6-7	6-6
	SELECTED READINGS	6-8	6-6
CHAPTER 7	THE PRETEST		
	INTRODUCTION	7-1	7-1
	VALUE OF THE PRETEST	7-2	7-1
	THE METHOD	7-3	7-1
	ROLE OF THE INTERVIEWER	7-4	7-2
	PRETESTING THE QUESTIONNAIRE	7-5	7-2
	SELECTED READINGS	7-6	7-3
CHAPTER 8	QUANTIFICATION AND ANALYSIS OF DATA		
	INTRODUCTION	8-1	8-1
	QUANTIFICATION OF DATA	8-2	8-2
	TABULATION OF DATA	8-3	8-2
	ANALYSIS OF DATA	8-4	8-3
	REFERENCES	8-5	8-6
	SELECTED READINGS	8-6	8-6
PART TWO	PRINCIPLES OF INTERVIEWING		
CHAPTER 9	INTRODUCTION		
	GENERAL	9-1	9-1
	PURPOSE OF THE INTERVIEW	9-2	9-1
	SOURCES OF INTERVIEW BIAS	9-3	9-1
	TYPES OF INTERVIEWS	9-4	9-3

	PARAGRAPH	PAGE
CHAPTER 10	THE QUESTIONNAIRE INTERVIEW	
	INTRODUCTION	10-1 10-1
	ESTABLISHING RAPPORT	10-2 10-1
	ASKING THE INTERVIEW QUESTIONS	10-3 10-2
	OBTAINING THE RESPONSE	10-4 10-5
	REPORTING THE RESPONSES	10-5 10-6
	INTERVIEW TECHNIQUES	10-6 10-9
	SELECTED READINGS	10-7 10-11
CHAPTER 11	THE GROUP INTERVIEW	
	INTRODUCTION	11-1 11-1
	THE ROLE OF THE ADMINISTRATOR	11-2 11-1
	CONDUCTING THE GROUP INTERVIEW	11-3 11-3
	SELECTION OF INTERVIEWER	11-4 11-4
	QUANTIFICATION AND ANALYSIS OF RESULTS	11-5 11-5
APPENDIX A	SAMPLE INTERVIEW GUIDE	A-1
APPENDIX B	SAMPLE QUESTIONNAIRES (SELF-ADMINISTERED)	B-1
APPENDIX C	INTERVAL SCALE VALUES	C-1
APPENDIX D	SAMPLE ANCHORED INTERVAL SCALE	D-1
APPENDIX E	EXAMPLE OF QUESTIONNAIRE RATING INSTRUCTIONS	E-1
APPENDIX F	EXAMPLES OF SUMMARY TABLES	F-1
APPENDIX G	EXAMPLES OF STATISTICAL ANALYSIS TECHNIQUES	G-1
APPENDIX H	REFERENCES	H-1
APPENDIX I	SELECTED READINGS	I-1
APPENDIX J	LIST OF TABLES	J-1
INDEX		Index 1

PART ONE

QUESTIONNAIRE DESIGN AND ADMINISTRATION

CHAPTER 1

INTRODUCTION

1-1. GENERAL. The basic tool for obtaining subjective data is the questionnaire. It is the most frequently used and most difficult to construct of the subjective techniques. The questionnaire provides a structured method for asking a series of predetermined questions in order to obtain measurable expressions of attitudes, preferences, and opinions. The design of a questionnaire which will produce valid and reliable results requires a great deal of skill and experience. Unfortunately, questionnaire design and construction cannot be taught from books; the requirements for each test are somewhat different and present new and different problems. However, there are certain rules and principles of questionnaire design and administration which, when followed, eliminate some of the more common pitfalls which result in faulty questions and invalid results. The following chapters are intended to provide guidance, in an easy-to-understand manner, for planning, designing, and administering the questionnaire. The final chapter also includes guidance for the analysis and reporting of the results of the questionnaire. Appropriate examples of data collection and analysis techniques are given in the appendixes.

1-2. PURPOSE OF THE QUESTIONNAIRE.

a. The questionnaire is a subjective measurement tool for systematically obtaining attitudinal responses from a selected group of individuals. The function of the questionnaire is to communicate information. When properly formatted, it also aids in the tabulation of data and analysis of results. The questionnaire is used to assess a wide variety of qualitative variables such as comfort, Soldier acceptance, ease of use, and preference. It may be administered to small groups of technical personnel, such as those involved in highly controlled engineering tests, or to larger representative cross-sections of Army personnel.

b. Knowledge of individual or group attitudes provides valuable information regarding reactions, feelings, and preferences toward military systems. Since attitudes determine behavior, questionnaire responses of a representative sample of the Army population permit a reliable estimate of group reactions to military items or systems in actual field use. These results also may be used to resolve problems prior to type classification and to anticipate and thereby avoid future developmental problems.

c. The questionnaire is appropriate for use in all types of tests conducted by TECOM. It should be used to obtain subjective data when objective measurement is not feasible and when qualitative data are needed to supplement objective measurements.

1-1. PROBLEMS IN QUESTIONNAIRE DESIGN.

a. One of the most difficult problems to overcome in questionnaire design is the misunderstanding on the part of well-meaning individuals as to what a questionnaire is and how it should be used. Questionnaire methodology looks deceptively simple -- just ask people questions, add up the answers, and you know what they are thinking and why they feel that way. There are those who believe that anyone who can write well and use a little common sense can construct a good questionnaire. The seriousness of this faulty assumption is illustrated by the fact that an improperly designed and poorly worded questionnaire will still yield data in the form of numbers, frequencies, and percentages. These numbers are amenable to statistical analysis and may even produce statistically significant findings. The real tragedy is that these erroneous findings may be used to draw false conclusions which, in turn, contribute to faulty critical decisions regarding the military utility of an item and its suitability for issue to troops.

b. The questionnaire is not just a list of questions or a data form to be filled out. When properly designed, it is a scientific measurement instrument. Like all such instruments, however, it must be designed in accordance with rules and specifications and with specific aims regarding the item or subject to be measured. Like other measuring instruments, the questionnaire is subject to a variety of errors. These errors include errors in question wording, question sequence, sample selection, and the procedures employed in administering the questionnaire.

c. Use of the sound principles of questionnaire design will eliminate many of the problems cited above. The purpose of this pamphlet is to assist the designer of a questionnaire in avoiding these pitfalls and to provide guidance in constructing a questionnaire which will produce valid, reliable, and relevant test results.

CHAPTER 2

METHOD OF QUESTIONNAIRE DESIGN

2-1. INTRODUCTION. The method of questionnaire design applicable to the types of tests conducted by TECOM may be divided into five logical steps:

- a. Preliminary planning,
- b. Selection of the question form,
- c. Wording of the questions,
- d. Formulating the questionnaire, and
- e. Pretesting and administering the questionnaire.

Chapters 3 - 8, Part One, discuss these successive steps in detail. The essential elements of each of the five steps involved in questionnaire design and construction are summarized in this chapter.

2-2. PRELIMINARY PLANNING.

a. The preparation of a questionnaire requires great care and a background knowledge of the item or system to be tested. Knowledge also is required regarding the background of personnel to whom the questionnaire will be administered, the conditions under which it will be administered, and the type of analysis which will be made of the results. Too often a questionnaire is prepared on the basis of insufficient planning or no planning at all. The problems involved and the weaknesses in the design are frequently not recognized until such time as the data is analyzed or the results interpreted. Inadequate planning and poor design result in test reports which contain conclusions based upon faulty findings from insufficient or inaccurate data.

b. The planning and design for questionnaire construction and administration are closely related to the experimental design and plan for the entire test. Preliminary planning must include a review of all background documents to determine the test critical issues and objectives which must be answered by qualitative methods, determination of the best techniques to use, frequency of administration, and the analysis which will be used. This period of familiarization should help to clarify the objectives and scope of the questionnaire and provide a frame of reference within which to work.

2-3. SELECTING THE QUESTION FORM. There are three basic question forms:

- a. The open-end or free-answer,
- b. The dichotomous or two-way, and
- c. The multiple choice.

Each form has its merits and disadvantages which the questionnaire designer must be aware of and must weigh carefully before final selection. No one question form is superior to the others in all cases. In order to select one form over another, the designer must be aware of the advantages and disadvantages of each and choose that form which best meets the needs of the particular test situation.

2-4. PRELIMINARY WORDING OF QUESTIONS. The most important, and also the most difficult, aspect of questionnaire construction is the wording of the questions. Most authorities agree that faulty or improper wording of questions accounts for the greatest source of error in the questionnaire technique. Errors and distortions in the final data are often caused by misunderstanding and misinterpretation of questions due to use of an improper vocabulary level and ambiguous phrasing. Familiarity with the basic principles of good question wording will aid the designer in overcoming such errors.

2-5. FORMULATING THE QUESTIONNAIRE. In addition to selecting the question forms and wording the questions, it also is necessary to consider such factors as the sequence of the questions and the format for presentation and data collection. At this point, consideration must be given to the experimental design and statistical procedures to be employed. Finally, a check must be made of all questions to insure complete and accurate coverage of all data required by the test objectives and test critical issues.

2-6. PRETESTING THE QUESTIONNAIRE.

- a. A questionnaire is subject to many variables and must not be assumed to be perfected until it has been subjected to trial use. The trial, or pretest, should be conducted under conditions as representative of those of the actual test as possible. The pretest provides an opportunity to try the questionnaire out on a small sample of respondents. The results of this trial may then be used to make revisions and improvements as necessary before test administration. The pretest is the final and validating step in the method of questionnaire construction.

b. This five-step method must be followed carefully in the design and construction of each questionnaire. The procedure provides a systematic approach for the development of a valid and useful subjective test instrument. The chapters which follow elaborate on each of the five steps outlined above, concluding with a discussion of data quantification and analysis techniques.

CHAPTER 3

PRELIMINARY PLANNING

3-1. INTRODUCTION.

a. There are specific preliminary steps which must be followed carefully prior to the actual design of the questionnaire. These include defining the objectives to be answered, determining whether such objectives may best be met by quantitative or qualitative measurements, determining the kind and quantity of information required, the size and nature of the sample or group to be studied, and the type of questionnaire to be used.

b. The matter of questionnaire design is closely related to the overall test plan or experimental design. During the development of the experimental design it is necessary to obtain answers to such questions as: What variables should be measured? Which variables will be considered independent, dependent, controlled, and uncontrolled? What kind and what size sample will be used? The same questions must be answered in planning the questionnaire, therefore, the overall test requirements must be thoroughly reviewed to determine how, when, where, and how often the questionnaire will be employed.

c. It is helpful to begin planning the questionnaire by deciding what hypotheses to test; this will prove to be a constructive and time-saving approach. By deciding at the beginning the hypotheses to be tested, the designer is forced to consider what form the data will take - frequencies, percentages, ratios, or scale values. It is then possible to determine what tabulations and analyses of results are needed to draw valid conclusions. From this advanced consideration of the treatment of data it is possible to determine the type of questions that should be asked and how the data should be quantified, as well as the nature of the test sample required.

3-2. DEFINING THE OBJECTIVES.

a. Early efforts in preliminary planning must always be directed toward defining the objectives or purposes of the test and those sub-objectives which may best be answered by use of the questionnaire. The overall objectives of the test usually are stated in the test directive. Additional information and an outline of general topics to be covered are provided in the "background information" or other sections of the directive. Further guidance concerning what information is needed and why it is needed may be found in the requirements document (Required Operational Capability (ROC), Letter Requirement (LR), Independent Evaluation Plan (IEP), or Test Design Plan (TDP)). Analysis of these documents will help to identify the qualitative requirements of the new item or system and will provide the basis for the objectives or sub-objectives and the questions needed to answer them.

b. Some examples of the qualitative information required for the evaluation of an item or system are: suitability or acceptability for use and maintenance, comfort, preference, fit, ease of operation, compatibility, and appearance. A more complete listing of the important characteristics of clothing and equipment which may be assessed by the use of the questionnaire are listed at table 3-1. The original list (table 3-1) was later modified by rewriting the characteristics in "Soldier language" (paragraph H-6, appendix H). It is also advantageous in some instances to obtain qualitative data in support of quantitative measurements of such variables as durability, maintainability, transportability, and reliability.

3-3. LISTING THE INFORMATION REQUIRED.

a. After the requirements of the test have been determined and the role of the questionnaire clarified, the specific information to be obtained by the questionnaire must be listed. For example, in considering the comfort of an item of clothing, information will probably be required regarding its fit, restriction to freedom of movement, and warmth or coolness. Although the need for this information may be specified in the test directive or requirement document, it is often necessary to determine the need for specific data required to meet the general objectives. Care must be taken, therefore, in compiling a detailed listing of the information to be obtained by the questionnaire. It is important to insure that all necessary data are obtained; also it is important to avoid unnecessary collection of information which is of no benefit in answering the objectives.

b. The listing of information required is one of the greatest time-saving devices in questionnaire construction and its importance cannot be overemphasized. One of the reasons for its importance is the fact that questions will be designed later to cover each of the points listed. It must be decided at this time, early in the planning stages, which of the points should be covered by the questionnaire and which of the points could more adequately be covered by objective measurements. If this step is omitted or considered lightly, a great deal of time and effort may be expended in designing unnecessary questions resulting in useless data.

3-4. CONSIDERING THE TEST GROUP. Another aspect of preliminary planning concerns the test sample or group to whom the questionnaire is to be administered. The designer of the questionnaire must adapt the questions to the test participants who make up the test group. To accomplish this, there are several things that must be known or determined about the test group.

TABLE 3-1

IMPORTANT CHARACTERISTICS OF CLOTHING AND EQUIPMENT

a. <u>Original List</u>	b. <u>Modified List</u>
Appearance	Good Looks
Style	Stylish
Color	Right Color
Fit	Fits Well
Personal Comfort	Comfortable
Personal Protection	Protects Me
Item Quality (Durability)	Rugged
Item Usefulness	Many Uses
Item Effectiveness	Serves Purpose
Necessity	- - - - -
Maintenance	Easy to Care For
Ease of Handling and Carrying	Easy to Handle and Carry
Ease of Operation	- - - - -
Interference with Activities	Not Get in Way
Ease of Wearing	Convenient
Design Aspects	Well Designed
Physical Dimensions	Right Size
Weight	Right Weight
Protection for Clothing and Equipment	Protects Clothing and Equipment

a. Representativeness. First, it is important that the test group be representative of the eventual users in terms of the relevant factors which are likely to effect the test results. These factors may vary from test to test but generally include, as a minimum, training, MOS (job proficiency), skill level, intelligence, education, physical profile, and anthropometric dimensions. Unless care is taken to insure representativeness of the test group, valid inferences cannot be made from the results obtained.

b. Sample Size. The size of the selected sample is important in determining the type of questionnaire to be employed. A test with a sample size ranging from 1 to 20 participants should normally employ a personal or guided interview type questionnaire, while one involving 20 or more participants should normally employ a self-administered questionnaire. Determination of optimum sample size is dependent upon many factors, such as the type of sample required, the number of participants available, the number and types of items to be tested, the statistical confidence level, and precision of measurement. Entire books have been devoted to the subject of sampling. Sample selection is mentioned only to point out that it is a factor which should be considered during the planning stages of questionnaire design. It is recommended that individuals who are interested in a more comprehensive discussion of sample selection techniques consult one of the referenced texts on this subject (paragraphs H-4, H-7, and H-11).

c. Background. Knowledge of background information about test participants is important in determining the form and level of the questions to be used. Such information as the level of education of participants and ability to verbalize will aid in wording questions; words or terms which are beyond or excessively below the level of comprehension of the group must not be used. The efficacy of some rating scales, for example, depends to a great extent upon the ability of the group to understand the scale. The importance of background information will be more evident from the succeeding discussion of forms of questions.

d. Previous Training and Experience. Consideration also must be given to whether or not the test participants have had previous experience with the item under test or with similar standard items. If test participants have had experience, it is desirable to know whether the previous experience was obtained during training, controlled field exercises, normal usage, or in a combat environment. Such knowledge is helpful in constructing the questions to take advantage of the previous experience or, if appropriate, in designing the study to allow for the possible variable effect of different levels of previous training or experience among the test participants.

3-5. **SELECTING THE TYPE OF QUESTIONNAIRE.** Preliminary planning also must include consideration of the type of questionnaire to be used. The two types of questionnaires which meet the requirements of most test situations are the interview questionnaire and the self-administered questionnaire. Each type has characteristics and requirements which make it suitable for a particular use.

a. **The Interview Questionnaire.** The interview questionnaire is a guided type of questionnaire which is usually preferred for obtaining subjective data from a small group consisting of 20 or less test participants. The interview questionnaire, sometimes referred to as an interview guide, is designed to be administered by interviewing a respondent, asking the preworded questions exactly as written and directly from the questionnaire. The responses are recorded by the interviewer in the appropriate spaces provided on the questionnaire form. Since this technique provides for a face-to-face discussion with each respondent, it is possible to obtain detailed information of a specific nature. Also, it is possible with this method to follow up a general response by probing for details which could not be obtained otherwise. Another advantage is that this technique overcomes reading or comprehension difficulties, thereby minimizing misunderstandings and improving the reliability of the results. The disadvantages of the interview questionnaire are the time required for administration and the availability of qualified interviewers. An example of the interview questionnaire or interview guide is shown at appendix A.

b. **The Self-Administered Questionnaire.** The self-administered questionnaire is ideal for obtaining subjective data from large groups consisting of 20 or more test participants. This type of questionnaire may be used to obtain large quantities of data in a short time with little effort expended for administration of the questionnaire; it may be administered to individuals or groups. It does not require skilled interviewers and can be completed in less time than the interview questionnaire. However, it does require a great deal of care in design and preparation to insure readability and lack of ambiguity. The need for follow-up questions must be anticipated and included where appropriate to insure that responses are clarified. An example of the self-administered questionnaire is shown at appendix B.

c. **Factors Determining the Type of Questionnaire.** Factors which must be considered in determining whether to use the interview questionnaire or the self-administered questionnaire include the size of the test group, the time available for administration, the conditions under which the questionnaire must be administered, the availability of qualified interviewers, and the group to be tested. In a particular test situation, any one of these factors may outweigh all the others; however, all of these factors must be weighed when considering the type of questionnaire to be employed.

3-6. REFERENCES. See paragraphs H-4, H-6, H-7, and H-11, appendix H.

3-7. SELECTED READINGS. See paragraphs I-5 and I-6, appendix I.

CHAPTER 4

SELECTING THE QUESTION FORM

4-1. INTRODUCTION. There are several forms of questions available for use in developing a questionnaire. Authorities in the field of market research and public opinion polling favor the use of one form of question over another for specific purposes. No one form, however, has proved superior for all uses. Each question form has its own merits which make it preferable for use in a particular situation. Therefore, a well-planned questionnaire will use a variety of question forms to obtain the required data. The three principal forms of questions used in questionnaire construction are presented in this chapter. These are the free-answer, dichotomous, and multiple choice, along with their variations. The advantages and disadvantages of each type are discussed.

4-2. THE FREE-ANSWER QUESTION.

a. Free Expression. The free-answer, or open end, question is so named because with this question form the respondent freely expresses his/her feelings or ideas in his/her words. It is an easy question form to use and solicits a wide range of possible answers when all alternatives are not known. It is particularly useful in a pretest questionnaire to determine alternative responses prior to construction of the final test questionnaire. The response to the free-answer question is not restricted by predetermined alternatives from which the respondent must choose. This free-answer situation makes it a valuable question form since it elicits the respondent's opinions uninfluenced by preconceived ideas of the test designer. Unfortunately, this same characteristic makes it one of the most difficult forms to use since the response depends upon the knowledge of the respondent and his/her ability to verbalize.

b. Interview Questionnaire. When used in the interview questionnaire, the value of the free-answer question is largely dependent upon the ability and skill of the interviewer. It is often difficult for the interviewer to avoid expressing his/her feelings during the course of the conversation, thereby influencing the respondent. This is particularly true in a case where the respondent has difficulty expressing himself/herself. The validity of the results obtained from this form of question, when used in the interview questionnaire, is largely dependent upon the training and skill of the interviewer.

c. Self-Administered Questionnaire. Difficulties also may be encountered when this form is used in the self-administered questionnaire because a respondent may be reluctant to take the time to fully respond to the question. Another disadvantage of the free-answer form is the fact that summarizing and analyzing the wide variety of responses which are possible with this form is difficult and time-consuming. In tests involving a large number of respondents or a large number of questions which must be tabulated and analyzed, the data analysis problem can become a major consideration in determining the form of questions to employ. Examples of the free-answer form of question are:

- (1) What do you like best about the XM-69 grenade launcher?
- (2) How does the XM-69 launcher compare with the XM-38 with regard to ease of maintenance?
- (3) Do you have any additional comments or suggestions to make, either favorable or unfavorable, regarding the XM-69 launcher?

d. Free-Answer Forms. The free-answer question may be employed in several different forms to meet the needs of a specific questionnaire. The questionnaire designer must be familiar with each of the forms since one or more of these types are normally employed in a single questionnaire.

(1) Follow-up Questions.

(a) The free-answer form is most often used as a follow-up question to a previous question on the same topic. It is designed to obtain additional specific information. An example of the free-answer question used in this manner in an interview questionnaire is as follows:

"Did you have any difficulty in attaching the reserve parachute to the harness?" (check one) Yes _____ No _____
If yes, "What seemed to cause the difficulty?" _____

This type of question permits the use of one or more short follow-on questions regarding difficulties which the respondent may have encountered during the test. It provides a means of probing for more detailed information related to difficulties experienced by the respondent.

(b) The follow-up question also may be used to determine the reason for making a specific choice, as in the following example:

"Which do you like better, helmet A or helmet B?" (check one)
A _____ B _____

"Why do you prefer this type helmet?" _____

(c) The follow-up question should be worded as a complete interrogative sentence rather than "Why?" or "What?" A respondent usually has more difficulty understanding and answering a one-word follow-up question than a complete question. The difficulty is related to the fact that use of the one word "Why?" or "What?" tends to disrupt the continuity of thought. The respondent, upon first hearing the one-word question, may feel that his/her decision is being questioned. The likelihood of such a misunderstanding may be reduced during the interview questionnaire by exercising care in the tone of voice used to express the question. Constant repetition of the one-word query after each question also creates a certain amount of monotony or a "the heck with it" attitude on the part of the respondent. Whenever possible, a complete sentence should be used instead of the one-word question to overcome these difficulties.

(d) It is desirable to vary the wording of the follow-up questions. Some examples are:

1. "Please explain the difficulties that you had."
2. "What, in your opinion, makes that one the best?"
3. "What were your particular reasons for choosing this one?"

Even though a complete sentence is used to ask the follow-up "Why?" or "What?" question, the questionnaire designer and administrator must be aware of additional difficulties which exist.

(e) In the use of the interview questionnaire, the interviewer's job is to obtain a specific answer to the question, not one which avoids the issue. For example, the reply "I don't like the M-16 because I think the M-14 is better" does not directly answer the question; it is a restatement of the fact and not the reason for the preference. Such a restatement is often an indication that the respondent does not know or has not analyzed for himself/herself the reasons underlying his/her attitude or opinion. In such a case, the interviewer should probe to determine whether such a reply indicates a true lack of knowledge, a misunderstanding of the question, or an inability to verbalize the answer.

(f) Another difficulty which might arise in the use of the "Why?" or "What?" question is that of determining the validity of the reason given. The interviewer may sense that the respondent has a limited insight as to the reason for his/her opinion or actions. Unfortunately, there is no simple means of determining whether or not a response is valid. When there is a question regarding validity, the interviewer should probe for the answer while taking care not to "lead" or bias the respondent. Often, the respondent needs more time to think about the question. The interviewer, in this case, should be patient and repeat the question slowly and distinctly.

(2) The Argument Question. A second type of free-answer question is referred to as the "argument" type and consists of two parts. It attempts to obtain information as to the merits of the item as well as its faults. The "argument" form usually consists of a question such as, "What do you like best about the XM-21 system?" followed by, "What do you like least about the XM-21 system?" This is a useful form when it is necessary to obtain as much information as possible on one type of item, especially when the respondent is familiar with only one item. It may also be used to compare two or more items indirectly.

(3) The Information Question. The "information" type of question is a familiar form which is usually employed to obtain demographic data, such as name, job, age, rank, etc., required for completing the basic data section of the questionnaire. While the question form may appear simple, a great deal of care must be given to the use of the one-word question, since it may be misunderstood or misinterpreted if not clarified. For example, the word "job," without further clarification, could be interpreted by the respondent in any of the following four ways:

- (a) The respondent's present duty position; title;
- (b) Military occupational specialty (MOS);
- (c) The type of work he/she is actually doing; or
- (d) His/her civilian occupation.

Such difficulty in meaning often occurs in the interpretation of a one-word question because of a lack of adequate clarification. Care must be taken to describe exactly the information required, such as "MOS (primary)," "training (military schools only)," or "date (day, month, year)."

(4) The Suggestion Question. The "suggestion" type of question is a valuable device and should be employed at, or near, the end of a questionnaire. This free-answer type of question is used to obtain suggestions regarding the design, operation, maintenance, or training for an item under test. The respondent may be asked questions such as, "How could this item be improved?", or "What suggestions or criticisms do you have of this item?" By asking such a question of a participant with "hands-on" experience, it is often possible to obtain suggested improvements or modifications which the test planner or observer may have overlooked. Every respondent will not answer this type of question, but experience has shown that those who do answer offer valuable information which fully justifies its use.

(5) The Probing Question. A "probing" type of question may be added at the end of a questionnaire in an attempt to obtain additional opinions or ideas not covered by the preceding questions. The primary requirement of a probing question is that it must not influence the reply of the respondent in any way. The question "Do you have anything else, either favorable or unfavorable, that you would like to say about the item?" meets this requirement. It avoids suggesting an answer either for or against the item, while at the same time providing the respondent with an opportunity to present his/her ideas.

e. Care in Use of Free-Answer Form. Although the free-answer type of question generally provides a great deal of useful data, care must be taken not to use it indiscriminately. When used frequently in the same questionnaire, the respondent may lose interest due to the extra effort required on his/her part to answer the questions. He/she also may become apathetic and avoid giving complete or responsive answers. The free-answer type of question must be used with discretion and should be employed primarily in a case where probing is necessary in order to determine the underlying reason for a certain response or to avoid the possible suggestion of an answer.

f. Precoding the Response. In an instance when it is possible to anticipate in advance the most likely responses to a free-answer type of question, it may be advantageous to precode the question. Precoding involves the categorizing of anticipated responses into specific groups or categories. This simplifies the recording of the response(s) by the respondent or the interviewer, since all that is required is the placement of a mark "X" or "✓" opposite the selected category. This method is illustrated by the precoded answers to the following questions:

(1) "How many rounds did you fire today? (check one)"

(a) None

(b) 1-50

(c) 50-100

(d) More than 100

(2) "What kind of terrain did you drive over today? (check one or more, as applicable)"

- (a) Paved roads
- (b) Graveled roads
- (c) Dry dirt roads
- (d) Wet muddy roads
- (e) Cross-country trails
- (f) Other (describe) _____

This type of precoding saves time for the interviewer or respondent and the data analyst. Methods of categorizing and precoding the responses are discussed in greater detail in chapter 6. The questionnaire designer must be aware of the advantages and disadvantages of use of the free-answer type of question; the advantages and disadvantages are summarized in table 4-1.

TABLE 4-1

ADVANTAGES AND DISADVANTAGES OF USE OF THE FREE-ANSWER TYPE OF QUESTION

Advantages

1. Uninfluenced - allows a respondent to give his/her free answer, unguided by a preconceived idea or suggestion.
2. Elicits a wide variety of responses. (A good type of question to use in the pretest form of the questionnaire to insure that all possible forms of reply are included.)
3. A good introductory type of question for the interview questionnaire. It aids in establishing rapport by making a respondent feel that he/she is playing an important role in the interview and that his/her idea is of value.
4. Provides background information which is often valuable for interpreting other responses.
5. Most valuable form for soliciting ideas, suggestions, and reasons for opinions.

Disadvantages

1. Time-consuming - lack of uniformity in responses - requires a great deal of time to classify, code, and quantify for a meaningful interpretation. May be misinterpreted by the coder.
2. Poor or incomplete results may be obtained due to the inability of a respondent to verbalize his/her observation and opinion.
3. More effort required by respondent may result in resistance to completing answer or to responding at all.
4. Usefulness and validity of the response is more dependent upon the skill of the planner and interviewer than any other question form.
5. Difficult to obtain a specific, valid response without influencing the reply.

4-3. THE DICHOTOMOUS QUESTION.

a. Most Commonly Used. The dichotomous, or two-way, question is the most commonly used type of question because of its simplified form. It provides for two possible alternatives, such as yes-no, satisfactory-unsatisfactory, and approve-disapprove. From two directly opposed alternatives, a respondent is asked to choose one. In addition to the two choices, a "no opinion" or "no experience" option may be included for the respondent who is unable to select one of the two choices provided. The decision to use a "no opinion" or neutral point in the formulation of a question should be made only after careful consideration of the effects on the analysis and final results. Refer to paragraph H-5, appendix H, for a detailed discussion of the use of neutral, or mid-, points.

b. Choice of One Alternative. The dichotomous form does not offer the respondent an opportunity to express his/her personal view beyond the choice of one of the two alternatives. Therefore, this form is the easiest type to administer. The following are examples of dichotomous questions applicable to tests of military equipment:

(1) Is the digital altimeter satisfactory or unsatisfactory for use under nighttime flying conditions? (check one)

Satisfactory Unsatisfactory No opinion

(2) Did you have any difficulty in adjusting the harness of the T-19 parachute? (check one)

Yes No No Experience

(3) Were your feet warm or cold? (check one)

Warm Cold

(4) Were you issued the M-69 mask? (check one)

Yes No

(5) Do you consider this combat pack acceptable or unacceptable for military use? (check one)

Acceptable Unacceptable

c. Stating One or Two Alternatives. When formulating the dichotomous question, doubt often arises as to whether one or both alternatives should be stated. In questions 4-3b(1), (3), and (5), for example, both alternatives are stated, whereas in questions 4-3b(2) and (4), one choice is given and the other implied. In many instances, as in question 4-3b(2), it is awkward to state both alternatives; to state both alternatives in this question would result in an awkwardly worded and perhaps confusing question. There are many instances where, for the sake of clarity, it is desirable to state one alternative; however, it is generally considered to be good practice to state both alternatives. The reason for this is the fact that there is a tendency for the stated choice to be selected more often than the unstated or implied choice. By stating both alternatives a built-in bias in favor of the stated choice is avoided.

d. Forcing a Choice. The question of whether or not the respondent should be forced to make a choice between only two alternatives arises in the construction of a questionnaire. The truly dichotomous question, such as those shown in subparagraphs 4-3b(3), (4), and (5), above, is a forced-choice question. In each of these examples the respondent is provided only two alternatives, i.e., warm or cold, yes or no, and acceptable or unacceptable; he/she is forced to choose one of the stated categories. This kind of dichotomous, forced-choice question does not provide a middle ground nor does it provide for another level or degree of response between the two extreme categories. In many circumstances the forced-choice question is desirable since the respondent must crystallize his/her ideas and make a definite decision. The forced-choice question simplifies the analysis of the data since there are only two responses to be dealt with. Also, when a middle response is provided, there is a tendency for the respondent to pick this category as an easy way out without giving much thought to the question. On the other hand, when a genuine middle ground exists, forcing a response into a dichotomy may produce a distortion in results. The solution to the problem of the forced-choice may be found by assessing the degree to which respondents' opinions are likely to be crystallized and the requirements of the particular question or issue being evaluated.

e. The Unstated Middle Category.

(1) Another approach to the problem of the forced-choice question is the use of the unstated middle category. This involves providing a middle response among the possible answers, but not stating the middle ground in the question. In this way it is possible to use the dichotomous question with the advantage of forcing a choice, but allowing an escape for the respondent who is unable to accept one of the extremes without altering his/her true opinion. Use of the unstated middle ground and its effect on the results obtained are shown in the following example (paragraph H-9, appendix H):

"After the war do you think the Federal Government should regulate gas and electric companies more or less than it did before the war started?"

More 40% Same 23% Less 18% Don't Know 19%

"After the war, do you think the Federal Government should regulate gas and electric companies more, less, or about the same as it did before the war started?"

More 23% Same 54% Less 10% Don't Know 13%

Many more replies were received in the "same" category when it was stated than when it was not stated. This example serves to illustrate the advantage of using an unstated middle category in the alternative responses of a dichotomous question. The advantage diminishes, of course, if the same question is repeated at a later date to the same group of individuals since they may remember that the middle category is listed as a response alternative.

(2) The concept of the unstated middle category also may be applied to terms such as "no difference" and "equal." However, it must be remembered that the use of one of these middle categories within a question results in a tendency to select the middle ground as an easy way out. By omitting this choice from the question an attempt may be made to force a respondent to make a choice between the two alternatives which are stated. For example:

"Which altimeter was easier to read, the standard or the digital?"
(check one)

Standard
Digital
Both the same

f. Don't Know Category. Often, a dichotomous question provides a "don't know" or "no opinion" category within the possible answers, as shown in example questions 4-3b(1) and (2). This category serves the important purpose of allowing a respondent with insufficient experience on the item to provide an honest response without distorting the results. Whenever there is a possibility that some respondents may have had insufficient experience with the item at the time the questionnaire is to be administered to have formed a definite opinion, one of these categories should be included. However, such categories should be reserved only for this type of situation. Every attempt must be made when designing a question to discourage the "lazy" respondent from using a "middle" or "don't know" response. "Don't know," "no opinion," and other such neutral responses must be avoided if possible, particularly when the test sample size is small, i.e., 20 or less respondents. In this case, it is better to force a choice than to lose much needed data. A summary of the advantages and disadvantages of use of the dichotomous type of question is shown at table 4-2.

TABLE 4-2

ADVANTAGES AND DISADVANTAGES OF USE OF THE DICHOTOMOUS TYPE OF QUESTION

Advantages

1. The greatest advantage of use of the dichotomous question is that it is simple to administer and tabulate.
2. It forces a respondent to crystallize vague opinions and to make a decision.
3. Easy question for a respondent to answer without requiring verbalization or explanation.

Disadvantages

1. The primary disadvantage is that it provides for no degree of qualification of opinion, but forces a respondent to choose one of two alternatives given. Makes replies appear definite when they may not be.
2. A slight language difficulty or misunderstanding of one word can change the answer from one extreme to the other.
3. Danger of oversimplification by question designer in attempt to force the response into two arbitrary categories.

4-4. **THE MULTIPLE CHOICE QUESTION.** The multiple choice question is a compromise between the dichotomous question with its two alternatives and the open-end, or free-response, question with its infinite number of possible responses. The multiple choice question offers a respondent a specific number of alternative replies. This form of question is particularly useful when the issue cannot be divided into two mutually exclusive categories or when various levels or degrees of opinion are desired. The multiple choice responses are easily tabulated and scored so that data analysis is relatively quick and easy. It also requires less time per respondent to complete than the open-end question and, therefore, is more likely to produce the desired data. The multiple choice question, however, requires a great deal of care in the wording and selection of alternatives. The question designer must have sufficient knowledge of the item or system to select the most meaningful and valid alternatives or must obtain a listing of the possible alternatives from other qualified individuals by use of a pretest questionnaire. Such a questionnaire must employ open-end questions to obtain as much information as possible about the relative importance of various item characteristics. The multiple choice question is extremely versatile and may take many forms, as follows:

a. The Basic Multiple Choice Question.

(1) The basic multiple choice question provides a respondent with from three to five response alternatives. The respondent is instructed to select one of these response alternatives as representing his/her choice or opinion regarding a preferred item or item characteristic. As in all variations of this form of question, it serves the important function of presenting all possible responses so that each alternative has an equal chance of being selected. The following are examples of the basic multiple choice question:

(a) Which of the following colors do you prefer for the plastic serving trays? (check one)

Green Brown Yellow White

(b) Which type of sleeping bag closure did you like best?
(check one)

Snaps Velcro Zipper

(c) What do you consider the most important characteristic of the XM-70 weapon? (check one)

Weight

Ease of maintenance

Portability

Accuracy

Other (specify) _____

(2) The responses to a multiple choice question may be worded in a variety of ways; the responses may be single words, phrases, or complete sentences. In some forms of the multiple choice question, the responses may consist of a series of statements and the respondent asked to check the one which most closely agrees with his/her feelings. In the wording of question responses it is important that the responses be simple and as brief as possible. The various alternatives should contain approximately the same word length and degree of simplicity so that a respondent will be likely to give equal consideration to all alternatives. The advantages and disadvantages of use of the basic multiple choice question are summarized at table 4-3.

TABLE 4-3

ADVANTAGES AND DISADVANTAGES OF USE OF THE BASIC MULTIPLE CHOICE QUESTION

Advantages

1. The main advantage of use of the basic multiple choice question is the ability to obtain opinions of varying degrees.
2. It is simple to administer, tabulate, and analyze.
3. It is easy to complete and therefore, suitable for administration to a large group.

Disadvantages

1. Wording of question and alternatives requires great care to insure validity.
2. There is a danger of omitting some important responses.
3. By placing all alternatives on an equal basis for consideration, some points may be overemphasized which would not otherwise be considered by a respondent.

b. The Checklist Question.

(1) The checklist is a form of the multiple choice question. It consists of a statement followed by a list of three or more alternatives from which a respondent is asked to choose. The checklist question differs from the basic multiple choice question primarily in that a respondent is expected to choose more than one answer. The list of terms, or alternatives, for a checklist is usually derived from the administration of a pretest questionnaire to insure that applicable terms are not left out. It is important that every effort be made to include all applicable alternatives to avoid biasing the results. Examples of checklist questions are:

(a) Which of the following types of beverages did you drink with the ready-to-eat meal? (check one or more)

Coffee

Water

Tea

Milk

Hot chocolate

Other _____

(b) Which of the following items of footwear did you wear with the XM-19 skis? (check one or more)

Leather combat boots

Ski boots

Insulated boots

Mukluks

Other _____

(2) The advantage of the checklist question is that it serves as a reminder to a respondent who might otherwise fail to recall a relevant experience or situation. By listing all of the relevant characteristics, each characteristic has an equal chance of being remembered. The important point is that all possible characteristics must be listed or provision must be made for a respondent to add to the list by use of the "other" category.

(3) The usefulness of the checklist question, therefore, depends upon the validity and completeness of the list. Enumerating the items or situations for the respondent makes this type of question advantageous to use in place of the free-answer question under certain conditions. When a respondent is required to recall the specific items used, experiences, or characteristics of an item, fewer replies are obtained than when a checklist is used. This does not mean that the checklist should be used as a substitute for the free-answer question in all cases. It should be used as a substitute only when an accurate account of all experiences is desired or essential to the evaluation. Care must be exercised in the use and placement of the checklist question within the questionnaire to avoid suggesting replies to other questions which might not ordinarily be recalled. The advantages and disadvantages of use of the checklist question are listed at table 4-4.

TABLE 4-4

ADVANTAGES AND DISADVANTAGES OF USE OF THE CHECKLIST QUESTION

Advantages

1. Results in a higher proportion of replies than if the respondent were asked to recall the items or characteristics.
2. In some cases it may be more reliable than the open-end question since all responses are on an equal basis for selection.

Disadvantages

1. It tends to suggest answers to other questions or bring to mind events which otherwise would not have been important enough to remember.
2. All possible answers must be included in the listing for reliable results.

c. The Ranking Question.

(1) Arrangement. Ranking means arranging in order with regard to some specified common dimension such as comfort, fit, taste, or overall preference. The ranking question is used for ranking or comparing standard and experimental items or in comparing a number of experimental items among themselves. The question also may be used to compare or rank specific characteristics of one item against similar characteristics of another item. Adjectival rankings should always proceed from the "most preferred" to the "least preferred" and from "best" to "worst" characteristics. The order of placement of the rank responses also should remain the same for different questions or different characteristics to reduce the probability of marking error by a respondent. The order and sequence of questions and responses are discussed in more detail in chapter 6. The following are examples of ranking questions:

(a) How would you rank the four types of armor vests (A, B, C, and D) which you have worn with regard to comfort? (1 = most comfortable, 4 = least comfortable)

1. _____
2. _____
3. _____
4. _____

(b) How important are the following features of the XM-27 launcher for portability through the jungle? (Place 1 beside the most important, 2 beside next most important, etc.)

Position of carry _____
 Length of tube _____
 Weight _____

(c) Rank the following meat items of the ready-to-eat meal in the order of your preference. (1 = most preferred, 2 = next most preferred, etc.)

Hamburgers _____
 Sausage patties _____
 Corn beef hash _____
 Meat balls _____
 Beef stew _____
 Chicken & noodles _____

(2) Formulation. In formulating the ranking question, the question designer must be very clear about the results required. Ranking of preference or order of "merit" (like, dislike), for instance, may produce confusing results since the response does not reveal the reasons for the respondent's ranking. The response also may be misleading. For example, it may be assumed that a new experimental weapon is most preferred (ranked number one) for its greater accuracy when actually the higher rank may have been given because of its lighter weight or ease of maintenance. A ranking on a somewhat vague dimension may be used as a preliminary step in an exploratory series of questions. A follow-up, or "Why?" or "What?" type question should then be used to obtain the reason for the response. It is best to state the specific characteristic to be ranked, as in example 4-4c(1)(a), above, and make no assumptions regarding the basis for the ranking. This question should be followed up by a "Why?" question to determine the reason for the "comfort" or "discomfort" ranking if this is important to the evaluation of the item. The reasons given for the "comfort" or "discomfort" ranking often provide the basis for causal analysis - which aids the responsible developer in making modifications and improvements.

(3) Limitations. There are limits to the number of rankings that a respondent should be asked to make, dependent upon the type of characteristic or item to be ranked and the extent of the respondent's familiarity and experience with this or similar items. Under normal circumstances, a respondent should not be expected to reliably rank-order more than five items. It must be remembered that ranking reveals nothing about the differences between ranks. It should not be assumed, for instance, that the "distance" between ranks one and two is the same as the "distance" between ranks two and three. Ranking tells you the order or sequence, only, and nothing about the interval between the ranks. If information is needed regarding both the direction and intensity of differences between items or characteristics, a rating or scaling technique should be used. The use of rating scales is discussed in detail later in this chapter.

(4) Measurements. For purposes of analysis, the measurements obtained are classified as ordinal, or ranking, scales. The median is the most appropriate statistic for describing the central tendency of rank scores. Nonparametric statistics should be used for hypothesis testing (paragraph H-10, appendix H).

(5) Methods of Ranking. There are several methods of wording ranking, or comparison, questions, as shown in the examples above. The usual method is to require that the respondent rank the items from "most preferred" to "least preferred" or from "best" to "worst" as in question 4-4c(1)(c). It must be remembered that the reliability of the results obtained when the respondent is asked to rank more than four or five items is greatly dependent upon his/her comprehension and experience with the items. Therefore, under most circumstances, it is best to limit the number of items or characteristics to be ranked to four or five.

(a) Paired Comparisons. If it is necessary to rank more than five items, two alternate comparison methods may be used. The first of these methods is the "paired comparison method." In using this method, the respondent is requested to make a comparison between each item and every other item, but only comparing two items at a time. To obtain the desired rank-order from this procedure, it is necessary to include all possible combinations of pairs, i.e., A&B, A&C, A&D, A&E, B&C, B&D, C&D. For six items this method requires 15 comparisons ($1/2N(N-1)$), and the number of comparisons increases rapidly when the number of items is increased (for 10 items, this method requires 45 comparisons). The "paired comparison method" is involved and time-consuming for use in a questionnaire when the number of comparisons exceeds five or six. For a more detailed explanation of the uses of and analysis methods employed with the paired comparison technique, refer to the reference cited in paragraph H-3, appendix H.

(b) Numerical Ratings. The second method involves the use of a numerical rating. In this method the respondent is asked to rate each item by means of a numerical rating scale. This involves the selection or assignment of a numerical value, e.g., 1 through 5, to each item rated. The items are then ranked statistically by comparing the rating received for each item. These indirect ranking methods are considered to be more reliable than asking a respondent to rank a large number of items directly.

(6) Differences Between Items Being Compared. The differences between the items being ranked or compared is of extreme importance in the design of a ranking question. The following differences are those which must be considered, since they do vary and can affect the results obtained:

(a) Small observable differences between the standard and experimental items or between several different experimental items;

(b) Differences which may exist in characteristics other than the specific one being measured; and

(c) Differences within the standard group or experimental groups being evaluated.

One or all of these types of differences may be found in a given test sample and they create special problems which must be considered in the design of the question and response categories.

1. Difficulties When Differences are Slight. When the differences between items or groups of items being compared are slight, it is difficult for a respondent to rank the items. If forced to make a choice, he/she may carelessly select one over the other or choose the most familiar item. In this case, it is better to begin the series of questions by asking the respondent a question such as: "Did you notice any difference in the degree of protection afforded by items A, B, and C?" A "no" response indicates that the items are considered to be equal with regard to that characteristic. A follow-up determines whether the protection afforded was equally good or equally poor. A "yes" response is then followed up by the ranking question. This approach affords an opportunity to determine whether the respondent was able to discern differences in the items before asking for a ranking. As stated previously, the ranking provides no information regarding the degrees of differences which may exist.

2. Difficulties Caused by Differences Other Than Those Being Measured (The "Halo" Effect).

a. A problem may occur in tests when items differ noticeably in characteristics other than the one being measured. If, for example, two types of jackets are being compared for warmth and one fits better than the other, the respondent may unconsciously choose the best fitting jacket although the other is actually warmer. A respondent may have a great deal of difficulty separating one characteristic of an item from others so that his/her response is colored by overall feelings about the item. Similarly, a respondent's attitude toward any item of Army equipment may be influenced by his/her overall attitude toward the Army. This effect is called the "halo" effect since the overall impression influences the attitude toward specific items or characteristics of items. The questionnaire designer must remain acutely aware of the "halo" effect when planning the questionnaire. One method of minimizing the "halo" effect is to ask similar questions for each item, in turn, regarding the specific characteristics of interest such as comfort, warmth, fit, etc. A comparison of items is then made statistically without having the respondent make a direct comparison. This indirect method of ranking is made by obtaining an independent rating on specific characteristics and comparing the ratings obtained for the different items evaluated. For example, the question, "Is this type of parachute harness very easy, fairly easy, fairly difficult, or very difficult to adjust" may be asked independently for harness type A and harness type B. The following results may be obtained:

	<u>Harness A</u>	<u>Harness B</u>
Very easy	20%	5%
Fairly easy	60%	10%
Fairly difficult	10%	70%
Very difficult	10%	15%

b. From the results, a quantitative index may be derived for comparing the two types of parachute harnesses without asking a respondent to directly compare them. This indirect ranking method is one means of reducing the "halo" effect. The procedure resembles the rating method but differs in that no assumption is made about the intervals between ranks. Ranking by this method is less precise than by the use of a rating scale.

3. Difficulties When Items Differ Among Themselves.

a. Another source of difficulty is that differences may exist among items of one type with regard to a characteristic of interest. The differences may exist as a result of design or lack of adequate quality control; a clothing item (footwear, uniform, handwear, body armor, etc.) usually varies in weight according to size. Differences in size will more than likely affect the results of a question regarding differences in such characteristics as weight, fit, etc. Before constructing a ranking, or comparison, question, care must be exercised to insure that such within-group variations do not adversely influence the results.

b. Whenever a comparison, or ranking, is to be made between items, the items should differ only with respect to those characteristics being measured, with all other factors held constant. However, when this ideal situation does not exist, considerable thought must be given to the differences between the items, as well as to the situations in which they are to be used, operated, or maintained. When the items are not to be used under comparable conditions, or when the differences between them are such that the validity of the results may be compromised, a respondent should not be asked to make a direct ranking or comparison of the items. Instead, another direct or indirect measurement technique should be employed to obtain the desired data. Advantages and disadvantages of use of the ranking methods are shown at table 4-5.

TABLE 4-5

ADVANTAGES AND DISADVANTAGES OF USE OF THE RANKING METHODS

Advantages

1. Ranking permits the direct comparison of two or more items.
2. The direct ranking method is more reliable than other indirect methods for comparing a small number of items.

Disadvantages

1. The size of the intervals between ranks, or degree of differences between ranks, is unknown.
2. Items must be used under comparable conditions for valid results.
3. There is a practical limitation to the number of items (4 or 5) which a respondent may be expected to rank reliably.

4-5. THE RATING SCALE.

a. An Objective Method. The rating scale is a variation of the multiple choice question form. It may be used to give a numerical value to a judgment, and provides an objective method for rating attitudes and opinions. When properly constructed, the rating scale accurately reflects both the direction and intensity of differences in attitude or opinion. The rating scale is appropriate for use in the construction of both a questionnaire and an interview, and the results obtained are amenable to analysis by conventional statistical techniques (paragraph H-5, appendix H).

b. Clear Alternatives. With the rating scale, a respondent must select from a list of alternatives the one that most closely approximates his/her opinion. The respondent assumes the coding task, but the interviewer or questionnaire designer must be certain that the alternatives provided are clear and not ambiguous.

c. Examples. The following are examples of some of the various types of rating scales which may be used in the construction of a questionnaire:

(1) How would you rate the lightweight mortar for use in the handheld mode? (check one)

Good Fair Poor

(2) Reading the altimeter during free-fall was: (check one)

Very easy

Fairly easy

Fairly difficult

Very difficult

(3) How would you rate the fit of the CVC helmet? (check one)

(6) Fits extremely well

(5) Fit is quite satisfactory

(4) Fit is about average

(3) Fit needs improving

(2) Fit is not very satisfactory

(1) Fit is very poor

(4) How would you rate your like or dislike for the prefried bacon component of the Meal, Ready-to-Eat? (circle one)

9	8	7	6	5	4	3	2	1
Like Ex- tremely	Like Very Much	Like Moder- ately	Like Slight- ly	Neither Like Nor Dislike	Dis- like Slight- ly	Dis- like Moder- ately	Dis- like Very Much	Dis- like Extremely

d. Determining the Number of Steps.

(1) As shown in the examples above, the number of steps or alternative used in the rating scale may vary greatly. The number of steps in the scale may range from three, as in example 4-5(c)(1), to nine, as in example 4-5(c)(4). The number of steps or points to use should be determined on the basis of the degree of discrimination desired and on the number of different items or characteristics being compared or rated. For instance, on a characteristic such as comfort, do not expect a respondent to discriminate between more than six levels of comfort for most items. On the other hand, do not use a three-point scale to rate discrimination between three or more different types of items. It has been found through extensive testing that on a common characteristic such as taste, an individual may be expected to differentiate reliably between as many as nine alternatives (paragraph H-1, appendix H).

(2) The main advantage of increasing the number of points on the scale beyond three is that it enables the respondent to select an alternative which more closely approximates his/her opinion. A second advantage of an increased number of scale points is that it tends to encourage more response toward an extreme and away from the middle categories for a more valid representation of the true attitude or opinion. Generally, there is a tendency on the part of a respondent to avoid the extreme points at either end of the scale. The bias which could result from this tendency may be reduced by increasing the number of scale points while retaining the extreme alternatives. These extremes, either favorable or unfavorable, may be appropriate and applicable to the item being rated. On the other hand, a respondent may have difficulty in comprehending a scale with too many alternatives to choose from.

(3) A typical respondent is able to comprehend and reliably rate the characteristics of most military items or systems on the basis of a five- or six-point scale. The rating scale shown in example 4-5(c)(4) is a nine-point hedonic (pleasure) rating scale used in large scale food preference studies. The principal advantages of this type of scale are its reliability and the fact that it facilitates statistical analysis of results when employed with a large sample, usually in excess of 20 respondents. There is no specific rule regarding the number of scale points to be used for all groups under all circumstances. Therefore, it is necessary in each instance to consider the issue or characteristic to be measured and the individuals to whom it is to be administered prior to determining the number of steps to be used.

e. Selecting the Steps of the Rating Scale.

(1) Examples 4-5(c)(1) through (4) are semantic scales defined by a pair of polar (opposite) adjectives. In constructing the rating scale, a decision must be made concerning the number of alternatives or step intervals to be used, with the desired degree of precision weighed against the complexity of the respondent's rating task. The important characteristics of each item to be assessed must be determined and it must be decided whether a separate specific scale must be constructed for each of the relevant characteristics or whether a single general scale consisting of one set of alternatives can be used to rate all characteristics.

(2) First, the steps must be as evenly spaced as possible, since the use of unequal steps will bias the results as readily as a poorly worded question. That is, the steps must progress gradually; not in uneven steps such as "very good," "poor," "extremely poor." These alternatives encourage a negative response. Since a rating scale deals with two-sided issues and is bipolar (opposite extremes), it is essential that the steps progress evenly and that there be an equal number of positive and negative, or favorable and unfavorable, alternatives as shown in examples 4-5(c)(2) and (3). This principle of evenly spaced steps also applies when a middle or neutral category is used. Additionally, when a middle category is used, the number of steps on either side of the middle category must be the same, as shown in example 4-5(c)(4).

f. Use of Equal Appearing Intervals. A further refinement of the rating scale is the concept of equal appearing intervals, indicating that the alternatives are equidistantly spaced along the dimension being assessed. The term "interval" actually applies to the space between the points on the rating scale. As shown in example 4-5(c)(4), the distance between "like extremely" and "like very much" should appear equal to the distance between "like very much" and "like moderately." In addition to the equal-appearing adjectival terms, numbers or values of equal interval units are also assigned to points on the scale so that numerically equal distances are associated with equal appearing adjectival terms. For instance, numerical scale values of 1 through 6 were assigned in example 4-5(c)(3), and scale values of 1 through 9 in example 4-5(c)(4). In example 4-5(c)(3), the number 6 equates to the term "fits extremely well"; the number 5 equates to the term "fit is quite satisfactory," and so on. By assigning equal numerical values to the equal-appearing adjectival terms, a quantitative scale is developed. The scale is continuous from "best" to "worst" and the distances have the property of additivity and may be treated the same as measurements on the interval scale of measurement. Almost all common statistical procedures can be applied to interval scale values, including mean, standard deviation, Pearson correlations, t test, and F test (paragraphs H-5 and H-10, appendix H). The equal interval rating scale eliminates the idiosyncrasies of the observer in arriving at the final index of meaning, and this is the essence of objectivity.

(1) Standardized Rating Scales. One of the most important considerations in devising an equal interval rating scale is the selection and arrangement of the terms (adjectival statements) to be employed. Standardized rating scales were developed for measuring the acceptability of military clothing, equipment, and food items (paragraph H-8, appendix H). These standardized scales include specific scales for assessing item characteristics such as fit, comfort, protection, and ruggedness or durability. Examples of these specific rating scales are shown at table 4-6. A standardized list of general scale values also was developed for use as a guide in the preparation of equal interval rating scales; examples of six-point and nine-point general rating scales are shown at table 4-7. These scales were developed by the method of successive intervals (paragraph H-3, appendix H). The terms were validated by a large sample of Army personnel experienced in the use of Army equipment and systems; these personnel then rated and ranked the adjectival statements on a numerical scale from which mean values and standard deviations were derived. These standardized scales are a valuable guide for constructing equal interval scales for use in assessing military items. A complete listing of the scale values for both specific and general scales is included at appendix C.

TABLE 4-6

EXAMPLES OF SIX-POINT RATING SCALES FOR RATING SPECIFIC CHARACTERISTICS

<u>Comfort</u>	<u>Mean</u>	<u>Standard Deviation</u>
6 - Excellent comfort	5.42	0.69
5 - Comfort is very satisfactory	4.41	0.65
4 - About average in comfort	3.44	0.59
3 - Slightly uncomfortable	2.40	0.70
2 - Very uncomfortable	1.40	1.14
1 - So uncomfortable it can barely be worn	0.40	0.78
 <u>Protection</u>		
6 - Protects extremely well	5.05	0.77
5 - Protection is above average	4.12	0.77
4 - Protection is about average	3.23	0.85
3 - Protection needs improving	2.41	0.99
2 - Protection is below average	1.78	0.71
1 - Protection is poor	0.87	0.80
 <u>Fit</u>		
6 - Fits extremely well	5.49	0.92
5 - Fit is quite satisfactory	4.41	0.99
4 - Fit is above average	3.73	0.81
3 - Fit needs improving	2.98	0.97
2 - Fit is not very satisfactory	1.69	0.78
1 - Fit is very poor	0.80	1.07

TABLE 4-7

EXAMPLES OF NINE-POINT AND SIX-POINT GENERAL RATING SCALES
FOR RATING OVERALL ACCEPTABILITY AND UTILITY

Rate this item for: (Overall acceptability
(Ease of operation
(Ease of maintenance
(Compatibility

<u>Statement</u>	<u>Mean</u>	<u>Standard Deviation</u>
9 - Excellent	6.27	0.54
8 - Extremely good	5.74	0.81
7 - Very good in most respects	4.62	0.72
6 - Good	4.25	0.90
5 - Adequate	3.39	0.87
4 - Barely acceptable	2.40	0.85
3 - Not quite adequate	1.79	0.90
2 - Poor	1.06	1.11
1 - Extremely poor	0.36	0.76
6 - Extremely good	5.74	0.81
5 - Very good in most respects	4.62	0.72
4 - Moderately good	3.58	0.77
3 - Barely adequate	2.10	0.84
2 - Poor	1.06	1.11
1 - Extremely poor	0.36	0.76

(2) Selecting Scale Statements. First, determine the number of steps appropriate for the scale, then refer to the standardized scales shown at appendix C which provide an average and standard deviation for each statement listed. In selecting the statements, use the average values for determining the equality of the interval between terms. It is not necessary that the interval between average values be exactly equal, but the numerical difference between all statements should be approximately equal. For instance, if the statement "excellent comfort," with an average value of 5.42, is selected for the first alternative, the next selection may be "very comfortable," which has a value of 4.43, and is approximately one unit from 5.42. The next statement selected may be "about average comfort," with a value of 3.44, which also has an interval value of approximately one unit from 4.43. If more than one statement contains approximately the same average value, such as 4.43, and 4.41, select the statement with the smallest standard deviation. Whenever possible, the standardized terms with step intervals of known numerical value should be used to construct a rating scale. Never use more steps in the construction of the rating scale than will be used in the scoring and analysis of the results.

g. Formatting the Rating Scale. The format used for constructing a rating scale may vary in that the alternatives may be listed vertically as in example 4-5(c)(3) or horizontally as in example 4-5(c)(4). It is important, however, to list the "good" or "high" ends of the scale first, i.e., at the top of the vertical scale and at the extreme left of the horizontal scale. It is recognized that this placement violates the conventional order found in the mathematical coordinate system, however, a typical rater tends to think of good qualities first. The good or high end of each scale used in the questionnaire must be in the same direction as this is natural for the typical rater and helps to avoid possible errors which could occur if the order were reversed. The numerical values associated with the scale also should begin with the highest value and decrease down the scale. For instance, in example 4-5(c)(4), the first and most favorable term "like extremely" is assigned the value of 9, while the least favorable term "dislike extremely" is assigned the numerical value of 1. Assignment of the highest numbered value to the "best" or most favorable term and the lowest number to the "worst" or least favorable term also is natural for the typical rater.

h. Use of a Middle or Neutral Category. Another consideration in the development of a rating scale is whether or not a middle or neutral category should be used. As in the case of the dichotomous question, this depends upon whether or not a middle position actually exists in the issue to be assessed. For example, it may be difficult for an individual to make a valid response to the question in example 4-5(c)(4) on the acceptability of prefried bacon without the middle category. The respondent's taste may be such that he/she truly neither likes nor dislikes the prefried bacon, therefore, the middle category would be a logical choice. On the other hand, the middle category should be used only when it is a logical choice. In a majority of cases, the inclusion of a middle category merely serves to provide a respondent with an "out," an excuse for not thinking through the question and making a decision. The middle category should be omitted when a logical or obvious difference exists and, particularly, in those instances where the respondent sample size is small - less than 20. The middle category, or mid-point, on a bipolar scale is usually worded as "neither satisfactory nor unsatisfactory," "neither comfortable nor uncomfortable," etc. This type of response generally indicates that the respondent either was unable to distinguish a difference where a real difference was expected to exist or, as in many cases, he/she was too lazy or wished to avoid committing himself/herself, feeling that the response was "safe."

i. Wording of Terminal Categories. The wording of the terminal categories (extreme values of the scale) plays an important part in determining the response dispersion and the ability of the scale to discriminate or detect differences which may exist. It might appear that the two extreme categories in example 4-5(c)(4), "like extremely" and "dislike extremely" would be of little or no real value since a typical respondent would avoid these extremes. While it is best not to word terminal categories in such extreme terms that no one would use them, there are at least two good reasons for including them. One reason is that a respondent may use them to differentiate among several good or favorable characteristics or items. For instance, as a rater proceeds through the rating process he/she may discover a characteristic or item which is obviously more extreme than any he/she has previously judged to be in category 8, "like very much," or category 2, "dislike very much." If a more extreme category were not listed, a respondent would be forced to judge this characteristic or item equal to others which he/she feels are not equal. These extreme categories (9 and 1) serve as extensions that are occasionally needed and do not detract from the remainder of the scale or add to the difficulty of analysis. The other important reason is that the existence of the extreme categories will tend to spread out the responses on the scale and draw them toward those categories and away from the middle of the scale. This helps to overcome the general tendency of a rater to avoid terminal or extreme categories. In other words, if categories 1 and 9 were omitted from the scale, a rater would tend to avoid categories 2 and 8 and thereby decrease the dispersion of responses and increase the shift toward the middle category. Some of the advantages and disadvantages of use of the rating scale are shown at table 4-8.

TABLE 4-8

ADVANTAGES AND DISADVANTAGES OF USE OF THE RATING SCALE

Advantages

1. The rating scale provides a measure of the intensity of opinion or attitude.
2. The scaling process enables identification and elimination of ambiguous alternatives.
3. A scale of equal units is amenable to more powerful and efficient statistical analysis than is possible with data of unknown interval lengths.
4. An equal unit scale provides a more accurate estimate of the magnitude of differences between characteristics or items rated.
5. A standardized scale provides a basis for comparison of data from two different tests.
6. An equal interval scale provides a truer picture of the shape of the frequency distribution.
7. The rating scale is easier to complete and more interesting to the rater than most other measurement techniques.

Disadvantages

1. There is a tendency for a rater to avoid extreme terms and to use the middle category instead of giving the question adequate thought.
2. Difficulty in designing categories which are evenly spaced if other than a standardized scale is used.
3. Possible difference in interpretation of meaning of terms between rater and evaluator.

j. The Anchored Rating Scale. The anchored scale uses adjectives to anchor the extreme ends of a bipolar numerical scale. For example, the response scale to the question, "How would you rate the ease or difficulty experienced in loading the 81mm mortar?" may be structured as follows: (circle one)

Extremely difficult 1 2 3 4 5 6 7 8 9 Extremely easy

This type of rating scale is relatively easy to construct, but the results are more difficult to interpret than the equal interval rating scale which uses an adjectival term for each point on the scale. A respondent who circled the number "7" on the anchored scale may have intended this to be a "fairly easy" response. There is no way, however, that an analyst can be certain what term a respondent either consciously or unconsciously assigned to a particular number. While statistical analysis will produce an average score for a particular characteristic, it is difficult to accurately interpret the numerical value in an adjectival term. The anchored scale is valuable, however, when a large number of characteristics must be rated in a short period of time. The use of the anchored scale to evaluate the adequacy of New Equipment Training (NET) is illustrated in appendix D.

4-6. REFERENCES. See paragraphs H-1, H-3, H-5, H-8, H-9, and H-10, appendix H.

4-7. SELECTED READINGS. See paragraphs I-2, I-6, and I-10, appendix I.

CHAPTER 5

QUESTION WORDING

5-1. INTRODUCTION.

a. One of the most important aspects of questionnaire construction is the question wording. Question wording is the major cause of error in the questionnaire or interview technique. An improperly worded question may lead to misinterpretation and confusion on the part of a respondent and produce invalid results. The function of a question in the questionnaire or interview guide is to elicit communication on a specific subject. The goal is to word each question so that the desired information is furnished by the respondent with a minimum of distortion.

b. Before constructing the questionnaire, information should be available as to:

(1) How large the sample should be.

(2) Whether the questionnaire will be administered to a stratified sample of selected personnel or to a representative sample of the population.

(3) Whether or not the questionnaire will be administered to the same respondent group more than once.

(4) Whether or not responses will be influenced by experience factors, condition of the test item (new and worn), and seasonal or environmental changes.

These and other factors must be considered before question wording reaches the final stages.

c. Specific rules cannot be established to apply in each test situation and for each respondent group. The circumstances and purposes of each test vary as to test activities, experimental variables, location, and test participants. These factors influence the way each question for a particular test must be worded. This chapter sets forth some of the problems which may be encountered and offers suggestions for solving them. A few general rules are discussed. If the rules and suggestions are followed, such pitfalls as ambiguity, misinterpretation, and bias, which seriously affect the validity and reliability of results, may be avoided.

5-2. VOCABULARY LEVEL. A primary source of misunderstanding in question wording is the vocabulary level used. As previously mentioned, it is necessary to consider the group to whom the questionnaire will be administered early in the planning stages. It is important to determine the general vocabulary level of the test participants who will be responding to the questionnaire. The general level of education and formal training of the participants has a definite influence on the range of vocabulary which may be used in question phrasing. However, an important point to remember is that, whenever possible, the vocabulary level of the questionnaire must be established for the lowest vocabulary level rather than for the average vocabulary level. In other words, if a group consists mainly of personnel with college training and a high vocabulary level, but a few members have a lower vocabulary level, each question must be worded to conform to the verbal abilities of the minority. This assumes that adequate background information is available on the entire sample to make this determination. In those instances where such information is not available, or where a sample is too large to warrant the time to determine the vocabulary levels to be dealt with, a good rule is to use an eighth grade vocabulary level. Every effort must be made to use simple and familiar words to communicate with the respondent. Pilot testing, using respondents representative of those to whom the complete questionnaire will be administered, will aid in determining whether or not the questionnaire is easy to understand. A question must be reworded, if necessary, to insure that it is understood and interpreted by each respondent as intended.

5-3. AMBIGUITY.

a. A word or question which is ambiguous is capable of being understood or interpreted in two or more ways. It is essential that a question be worded so that it has exactly the same meaning to everyone who will be concerned with it, to include the individual who designs the questionnaire, the individual who administers it, the respondent who answers the question, the statistician who tabulates the results, and the report writer who analyzes and evaluates the results. A discrepancy in interpretation of a question by any of these individuals could have serious results. Above all, the respondent must understand since, in the final analysis, the results of the test depend upon his/her answers.

b. Ambiguity may result from the use of words which, although they may be in most individuals' vocabularies, mean one thing to one person and something else to another. For instance the question, "Under what conditions did you use the M-16?" could result in confusion to a respondent since he/she may interpret the word "conditions" to mean environmental conditions, daylight and nighttime conditions, or field activities. The question must be more specific to avoid the possible ambiguity which may arise from the use of the word "conditions." Another example of ambiguity, using common but easily misunderstood words, is the question, "What weapons have you used during the past week?" In this question, the word "weapons" is too general and could mean small arms, grenades, mortars, or even a tank. The word "used" also is subject to misinterpretation since it could be construed to mean carried, worn, or fired in training or in a testing situation. Less ambiguity would result if the question were worded more specifically such as, "What small arms have you fired during field test activities this week?" Some of the simplest words become problem words if they are too vague. Words such as "used," "conditions," "variety," "few," "any," and "all" are among the words which tend to be vague and confusing. In addition to being short and simple, the words chosen must be specific. It is suggested that a review be made of each question by asking: (paragraph H-9, appendix H)

- (1) Does the question mean what it is intended to mean?
- (2) Is it ambiguous; does it have any other meaning?
- (3) If so, does the context of the sentence make the intended meaning clear?
- (4) Could more simple words or phrases be used?

5-4. BIAS. One of the major pitfalls of the questionnaire technique is that the phrasing of a question may be biased or "leading" and, therefore, produce invalid results. A question is "leading" when it is worded in such a manner as to suggest an answer or indicate the questioner's point of view. It is recognized that bias in phrasing a question is probably the result of carelessness rather than intention; the fact is that the misuse or addition of one word may produce an invalid response. Bias may be introduced into a questionnaire by the ordering of questions so that an earlier question influences the response to a subsequent question and by the use of suggestive words, leading phrases, and "loaded" words or phrases. A "loaded" word or phrase is one which is emotionally colored and suggests an automatic sense of approval or disapproval so that a respondent reacts to the word or phrase rather than to the intent of the question.

a. Suggestive Words. Certain words, when used in a question, are likely to arouse a respondent's emotions or prejudices and, hence, have a suggestive effect. Public opinion pollers are aware of the power of suggestive words and sometimes use these words to arouse emotion and elicit a desired response. Words such as "Reds," "Communism," "Facist," "rich," "upper class," etc., are typical examples of suggestive words. This suggestive effect may occur in the development of a questionnaire used to assess military systems unless great care is exercised to avoid it. Words such as "new," "improved," "experimental," and "old" should be avoided. When the performance of a current standard item is to be compared with an experimental item, it is recommended that the two items not be labeled or referred to as "standard" and "experimental" in the questionnaire or in the interview situation. These terms may introduce a bias in the response which could invalidate the results. A respondent's answer must not be influenced by the wording of the question. The question, "Which do you prefer, the standard or the new, experimental uniform?" illustrates the type of suggestive wording which must be avoided. It would be better to phrase the question, "Did you like one uniform better than the other?" followed by the question, "If yes, which one did you prefer?" With this technique of questioning, the response is less likely to be affected by the wording of the question.

b. Loaded Words and Phrases.

(1) Both individual words and phrasing of an entire question can be "loaded." Some words, such as "new," "experimental," and "leaders," have a positive loading effect while others, such as "old," "standard," and "bosses," have a negative loading effect. The following are examples of questions obviously loaded toward a favorable response:

(a) "Don't you feel that the new boot would be the best for all around combat wear?"

(b) "Most troops have suggested adding a zipper to this item to replace the buttons; don't you think this is a good idea?"

(c) "Are you in favor of improving this item by adding additional insulation?"

(2) Loaded phrasing also may result from the question designer assuming too much when wording the question. The question, "What difficulties did you have using the night vision viewer?" assumes that the respondent experienced some difficulty using the item, which may or may not have been the case. In order to avoid such loading, a preliminary question could be asked, "Did you have any difficulty using the night vision device?" If answered in the affirmative, a follow-up question, "What seemed to cause the difficulty?" could be asked.

(3) In many instances, a leading question results from a failure to provide alternatives. For instance, the question "How much trouble did you have adjusting the sight on this weapon?" assumes that the respondent experienced some trouble adjusting the sight and the question does not provide for the possibility of a neutral or negative response. This question could be followed by alternatives such as:

- (a) "A lot of trouble,"
- (b) "Some trouble,"
- (c) "Little trouble,"
- (d) "No trouble."

This would provide the respondent with an obvious negative response although the question itself is still leading in the negative direction.

(4) Prestige or social bias also may be inadvertently included in a question such as a question prefaced by a phrase such as "AMC feels that this type of breach should reduce loading time on this weapon; do you agree?" or "Most NCO's agree that the Type A weapon is better than the Type B for quick-fire operation; which do you prefer?" These are exaggerations, but serve to illustrate another type of bias which must be avoided in question wording.

(5) Leading bias also may occur in the wording of a follow-up or "Why?" question. For instance, the question, "How often do you clean your weapon?" followed by, "Why don't you clean it more often?" implies that a respondent should clean his/her weapon more often. This kind of bias often occurs as a result of carelessness in question wording and may be avoided by a careful review or pretest of the questionnaire.

(6) There is a place for the deliberately loaded or leading question in the hands of a skillful question framer or interviewer where the respondent is intentionally led or put on the defensive. However, in designing a questionnaire for use in assessing a military system, each question must be as objectively worded as possible.

5-5. OTHER CONSIDERATIONS. In addition to vocabulary level, ambiguity, and bias, there are other factors which must be considered when wording a question for inclusion in the questionnaire. Careful consideration of each of these factors will help avoid faulty wording which can produce invalid results.

a. Thoroughness of Questions. As previously discussed, a questionnaire plan must be developed before proceeding with question wording. The plan outline must include a listing of all the major and minor issues to be covered. This listing must be checked carefully to insure that all objectives are included. The next step is to decide for each issue how much detail is required; consideration must be given to how much detail a respondent will remember or be able to provide. The following sample question goes into detail regarding the comparative tracking characteristics of weapon systems:

"Did you notice any difference in the two weapons systems with regard to tracking targets?" (check one) Yes _____ No _____

If "yes," which system was the easiest to operate? (check one)

Type A _____ Type B _____

Do you feel that operation of this system was: (check one)

Very much easier _____

Somewhat easier _____

Only slightly easier _____

Notice that this question proceeds from the general to the specific. Often a respondent will have had a wide variety of experiences over a long period of time with the system under test and he/she may have difficulty responding immediately to a specific question regarding detailed characteristics of the system. By structuring the question in steps from general, easy to answer questions to more detailed questions a respondent will be better able to remember his/her specific experiences with the system. The question framer must decide whether or not this degree of detail is necessary for the purposes of the test.

b. Double-barrelled Questions. The double-barrelled or double-negative question must be avoided since it is impossible to know what the respondent meant by the answer to this kind of question. For example, "Did you suffer from headaches or nausea while riding in the M-60 vehicle?" (check one) Yes _____ No _____. If the respondent answered "yes," did he suffer from headaches, nausea, or both? Another example of this type of question which is even more confusing, but often occurs, is "Did you have any difficulty lighting the Type A or Type B fuel tablet under windy conditions during daylight or nighttime operations?" (check one) Yes _____ No _____. An affirmative response in this case would reveal no useful information for test purposes. The question should be broken down into separate questions dealing with difficulty in lighting and the conditions under which difficulty was experienced. Care must be taken to avoid the double-barrelled question since it is not only confusing to a respondent but also produces invalid results.

c. Avoid "Trick" Questions. Some textbooks on market research and polling techniques recommend "trick" or "consistency check" questions to determine the reliability, sensitivity, or consistency of a respondent. One of the techniques suggested is to repeat a question later in the questionnaire, but to rephrase the question so that the respondent does not recognize it as the same question. Another "trick" technique is to reverse the order of the rating scale for every fourth or fifth question. In other words, change the normal method of listing the "favorable" or "best" alternative first and list the "least favorable" or "worst" condition first. The theory is that the respondent develops what is referred to as a "response set" in which he/she tends to check the same relative position on the scale without regard to the scale wording. These recommendations were made for framing questions for use in market research and opinion polls to be administered to the general public. Experience has shown, however, that frequently these techniques only "trick" the questioner and the data obtained are invalid or cannot be analyzed. It is best to avoid "trick" questions which may backfire and invalidate the results. A military test participant is usually aware of the fact that the questionnaire responses are important; with proper and thorough orientation of each participant regarding the purpose and importance of the test, such "tricks" are not necessary.

d. Instructions. Each questionnaire must contain instructions to the respondent. Instructions must always be included to direct the respondent in the way the answers should be recorded and to provide procedural direction. Examples are: "Check the one you prefer," "Check one," "Circle the answer which is closest to your opinion," and "If 'No,' continue with question 12." Instructions must also be brief and unambiguous.

e. Keep Questions Short. It is best to keep each question as short as possible, preferably not more than 20 words, so as to be less confusing to a respondent. In most cases, a short, precisely worded question gets the point across with clarity and directness. Every effort must be made to keep each question as short as possible without sacrificing clarity or accuracy.

5-6. REMINDERS FOR QUESTION WORDING. The following are a few reminders which will prove helpful when wording a question:

- a. Each question must be relevant to the issue at hand.
- b. Each question must have uniformity of meaning to all.
- c. Each question must be free of bias.
- d. A question should not ask for information which a respondent is not likely to be able to provide.

e. Each question must be clear and brief.

Reference to these reminders, as well as others mentioned in this chapter, should prove helpful in question wording which will produce accurate and consistent results.

5-7. REFERENCES. See paragraph H-9, appendix H.

5-8. SELECTED READINGS. See paragraphs I-2, I-5, and I-7, appendix I.

CHAPTER 6

FORMULATING THE QUESTIONNAIRE

6-1. INTRODUCTION. The success of a questionnaire is dependent not only upon its component parts - the individual questions - but also upon the manner in which the parts, or questions, are organized. The process of organization involves determining the sequence of the questions, establishing the format, checking the questions against the test objectives, and determining the method of analysis.

6-2. QUESTION SEQUENCE. The first step in formulating the questionnaire is to decide upon the order of the questions. Each question may be properly worded, with no bias or ambiguity, but if it is placed in the wrong sequence in the questionnaire, it could confuse a respondent or influence his/her answer as a result of a response made to a previous question. There are no hard-and-fast rules to follow in question arrangement, however, there are several guidelines which aid in establishing the proper sequence of questions. Intelligent and conscientious application of these guidelines serves to insure that a response will not be affected by the arrangement of the questions. As a general rule, it is best to begin the questioning with factual, easy-to-answer questions and then proceed to the more detailed and difficult questions which require more thought or consideration.

a. Logical Order.

(1) There must be a logical flow of thought throughout the questionnaire. In other words, one question should lead naturally to the next so that a train of thought is established in the mind of a respondent regarding the subject or issue being investigated. If the questions skip around a great deal, the respondent may become confused - and this could lead to inaccurate results. For example, in a test of acceptability of a field jacket, the criteria might include warmth, water resistance, and fit. Each of these criteria should be dealt with individually so that all questions relating to warmth are asked before questions regarding water resistance or fit. If the questions relating to these criteria are intermingled throughout the questionnaire, a respondent may have difficulty re-orienting his/her thinking for every question. Check over the questions to insure that they are in a logical order and natural and easy for a respondent to follow.

(2) It is essential, when reviewing the ordering of the questions, to think of the problems from the point of view of a respondent. Consider the test experiences the respondent may have had with the item prior to questioning and the conditions under which the questionnaire are to be administered. Remember, also, that what may seem perfectly logical may be extremely confusing to the respondent. Take extra care in assembling the questions in a sequence which is logical to the respondent; it will help a great deal in directing the respondent's train of thought and in eliciting accurate answers.

b. The Opinion Question. The first few questions in the questionnaire are extremely important since they set the tone for the entire questionnaire. The first question should be a relatively simple one that a respondent can and will answer. If a respondent has difficulty with or becomes confused by the first question, he/she may feel threatened, feel that doubt is being cast on his/her intelligence, or he/she may become discouraged and lose interest in the remainder of the questions. Even a self-administered questionnaire represents a "test" or threat to some individuals and the wording of the first question can instill confidence or confuse and discourage them. The questionnaire should begin with an easy, impersonal question and not ask detailed or complicated questions until later - when a kind of rapport has been established. On the self-administered questionnaire, a simple, dichotomous "yes-no" question should be used to put the respondent into the right frame of mind and help establish self-confidence for later more complicated questions. The first question also should be one which each respondent can answer easily, such as, "Did you wear the CVC helmet during this test period?" or "Did you fire the M-16 rifle while wearing the protective vest?" This type of question should elicit a definite response without too much effort on the part of the respondent. An important point to remember is that the opening series of questions should be factual and relatively easy to answer, followed by the attitudinal questions which require more thought and time to answer.

c. Affect of Preceding Questions. Question arrangement can have an influence upon a response. An inaccurate or biased response may be obtained as a result of preceding questions. The preceding questions can cause a mental "set" or pattern of thinking which influences the manner in which a respondent replies to the immediate question. For example, a questionnaire on the acceptability of a new fabric for the field uniform may include one question each regarding the criteria of fit, comfort, and crease-retention, followed by several questions concerning appearance of the garment. The series of appearance questions may be followed immediately by a question on the overall acceptability of the uniform. Such an arrangement of questions has a tendency to produce a mental "set" in which a respondent may forget the earlier criteria of fit, comfort, and crease-retention and base his/her idea of acceptability of the uniform solely on appearance. A bias also can be introduced by omitting one of the important criteria. For instance, if a question on appearance is omitted altogether, a respondent may forget this aspect of acceptability and base his/her answer solely on the criteria of fit, comfort, and crease-retention. The best solution to the problem of mental "set" is to maintain a balance among the questions involving each criteria. Every effort must be made to avoid over-emphasis of any criteria. The important point to remember in formulating the questionnaire is that the preceding questions can influence the answers obtained to subsequent questions.

6-3. ESTABLISHING THE FORMAT.

a. Simplicity. There is no single prescribed format which each questionnaire should follow. However, there are some basic principles which should be adhered to in designing the physical lay-out of the questionnaire. The first principle is that of simplicity. The questionnaire format, as well as the questions, must be easy for a respondent to understand and follow so that his/her determination of a response to a particular question cannot be confused with that for another question. In order to obtain comparable results, each respondent must understand and interpret instructions for responding in the same manner.

b. Instructions. Special attention must be given to the location and wording of instructions for completing the questionnaire and individual questions, as appropriate. In a self-administered questionnaire, a respondent in a group situation may be reluctant to ask questions. If each question is of the same general form, an example may be used to illustrate the mechanics of responding to the particular type of question. An example of instructions for a self-administered questionnaire is shown at appendix E. In the case of a questionnaire to be administered by personal interview, specific instructions on the administration of the questions, conditions under which they are to be administered, and the type of information expected will aid an interviewer or administrator in obtaining complete data. An example of instructions for an interviewer is shown at appendix A.

c. The Heading. The first section of each questionnaire must include the heading and basic data which are required for later identification, tabulation, and analysis of the results. The first item at the top of the first page of the questionnaire should be the title which indicates whether the form is designed to be used as a "questionnaire," "interview guide," or "checklist." The second item should be the test title, or item to be assessed, and also may include the specific criteria to be covered, such as comfort, acceptability, or maintainability. The next item in the heading should provide for recording the name or code number of the respondent and the date the questionnaire is administered. Other basic header data which may be desirable, and in some instances essential, include provisions for recording such information as rank or job title of respondent, test phase, test location, administrator or interviewer's name, exact time of administration, and general instructions which apply to the entire questionnaire. Examples of the basic data sections included in the headings of questionnaires are shown at appendixes A, B, and E.

d. Length of the Questionnaire. Another consideration in the design of a questionnaire is its length. The length of a questionnaire is dependent upon a number of factors, such as the type of questions used, the location and conditions under which the questionnaire is to be administered, the interest level of the test group, and the form of the questionnaire. A questionnaire designed for use in a personal interview situation may be longer than a self-administered questionnaire. Also, a questionnaire consisting primarily of short, dichotomous, or multiple choice questions may be longer than one containing numerous open-end questions. As a general rule, a questionnaire should be designed for completion by a typical respondent in 20-30 minutes. The questionnaire should be long enough to obtain required data, but not so long that the task of responding becomes boring or tiring to the respondent. The best rule to follow is to keep the questionnaire as short as possible.

e. Frequency of Administration. A prime consideration during the design of a questionnaire and plan of test is how often the questionnaire is to be administered. Several factors must be considered in making this determination. The first factor is whether or not the test group's attitudes or opinions will be influenced by time and experience with the system under test. A second factor which must be considered is whether or not the condition of the test item (new or used) is likely to influence the responses. Another very important factor which must be considered is the influence of varying climatological and environmental conditions on the responses to particular questions. For instance, the range of temperatures experienced, rainfall, snowfall, and daylight or nighttime conditions may affect the results obtained. Also, the types of terrain encountered (flat, hilly, wooded, sandy, muddy, etc.) may affect the responses obtained. If any of these conditions are expected to influence the results, the questionnaire should include the experiences or variables of interest. A general rule to follow is to plan to administer the questionnaire on at least two occasions, once when the test item or system is fairly new and again toward the end of the test period when maximum experience with the system is likely to have been obtained. If the test is to be conducted in specific phases, or activities are planned to be replicated, it may be desirable to repeat the questionnaire at the end of each use phase or after each replication. Care in planning must be exercised, however, to insure that the same questions are not asked frequently. Too much repetition results in boredom and indifference on the part of a respondent in providing answers. Each scheduled administration must be fully justified on the basis of need. The unwarranted or indiscriminate administration of a questionnaire may be prejudicial to the objectives of a test, wasteful in terms of time and effort expended for administration and analyses, and of little or no value in answering the test objectives.

6-4. CHECKING AGAINST THE OBJECTIVES.

a. Once a questionnaire has been drafted in its final form, and prior to pretesting, it must be checked against the objectives of the test and the list of issues or subobjectives formulated during the preliminary planning stages. This check must be performed very carefully to avoid the omission of necessary questions related to specific objectives. A small amount of time and effort are required to perform this check; it may help to avoid the necessity of rewriting the questionnaire at a later date. A close review of the completed questionnaire at this time may reveal an oversight or omission during the formulation stages and will help to insure that a follow-up "why" question, or "comment," is solicited wherever there is a possibility of a negative or unfavorable response. As previously stated, it is often the case that the reason for a negative response is of more value to the analysis than the simple frequency count of the number of such responses.

b. During this review, there is often a temptation to include extra questions to obtain "nice-to-have" information. It is true that valuable data sometimes results from such questions, however, more often than not the disadvantages of including the extra questions outweigh the advantages. Due to the addition of such questions, a respondent may not attach the desired degree of significance to the questionnaire and this could affect the answers obtained on vital questions. It is important, therefore, that only those questions which are expected to produce essential and highly relevant results be included in the questionnaire.

6-5. USE OF PUNCH CARDS.

a. For large scale tests or studies involving 100 or more questionnaires, the use of punch cards greatly facilitates the tabulation and analysis of data. Although no major difference exists between questionnaires which are tabulated by hand and those tabulated through the use of machine or manually punched cards, there are additional considerations which greatly simplify the task of transferring the data from the questionnaires to the cards. First, it is important that all possible answers on the questionnaire be assigned identifying numbers or letters and have corresponding numbers or letters on the cards. Also, it is important that the answers be located on the questionnaire so that they may be easily located and identified. The layout of the questionnaire should make maximum use of space and be designed to minimize page turning during transfer and tabulation.

b. Punch cards are rectangular shaped cards into which holes may be punched either by hand or machine. The punched holes may then be read or sensed electronically by machines called card readers. The holes in the cards are associated with code numbers of the questionnaire responses. Each punch card may be divided into as many as 80 columns and each column or position may contain up to 12 positions. It is possible, therefore, to record the results from 80 questions with as many as 12 levels of response to each question. One card per respondent is generally sufficient to record the results of a single administration of a questionnaire. Since punch card analysis requires a number of machines and operations, it is advisable to determine the availability of the machines prior to making a decision to use punch cards, and to confer with specialists in automatic data processing systems to obtain expert advice and assistance in setting up codes and punching cards.

6-6. CODING AND CLASSIFYING RESPONSES. The task of coding or assigning numbers or letters to responses to dichotomous, or multiple choice, questions is a relatively simple matter since the number of expected responses is known in advance. When open-end or free response questions are used, it is possible to anticipate most, if not all, of the types of responses which will be obtained, especially if the questionnaire is pretested prior to administration of the actual test. Some examples of simple codes which may be used to classify possible responses to questions are shown at table 6-1.

6-7. CATEGORIZING RESPONSES. In coding responses to open-end questions, it is usually necessary to perform a grouping, or classification, of responses because of the difficulty in tabulating, analyzing, and reporting the results of an extremely large number of responses. Many responses will be similar and may be grouped without a serious loss of useful information. In other cases, sacrifice of detail may be necessary to simplify the analysis and presentation of the results. The classification, or grouping, of responses to open-end questions is illustrated at table 6-2. The process of grouping, or categorizing, responses may be accomplished based upon the pretest results or upon the basis of anticipated responses; this helps to reduce the number of codes required to accommodate answers to open-end questions.

6-8. SELECTED READINGS. See paragraphs I-6 and I-9, appendix I.

TABLE 6-1

EXAMPLES OF CLASSIFIED AND CODED RESPONSES

a. An individual's service time:

- X Don't know; can't remember
- Y No answer
- 0 Less than 6 months
- 1 6 to 12 months
- 2 12 to 18 months
- 3 18 to 24 months
- 4 More than 24 months

b. Why boots were rated as uncomfortable:

- X Don't know; can't remember, etc.
- Y No answer
- 0 Rubbed blisters
- 1 Poor fit
- 2 Hurt arch
- 3 Too hot
- 4 Too tight
- 5 Too loose

TABLE 6-2

GROUPING OF RESPONSES

(Suggested Improvements to Combat Boots)

<u>Initial Responses</u>	<u>Frequency</u>	<u>Grouped Responses</u>	<u>Frequency</u>
Add non-slip soles	6	Improve traction	10
Soles that don't slip	2		
Soles that grip better	2		
Use material that won't leak	3	Improve water resistance	8
Waterproofing on the leather	2		
Better resistance to leaking	2		
Use water resistant material	1		

CHAPTER 7

THE PRETEST

7-1. INTRODUCTION. Thorough knowledge and understanding of the principles of questionnaire design and construction on the part of the questionnaire designer do not insure development of an effective measuring instrument. A questionnaire may not be assumed to be effective until it has been tried under conditions similar to those which are expected to exist in the actual test. A questionnaire designer cannot possibly anticipate all of the various problems that may arise in the actual test administration, therefore trial is necessary. The trial, or pretest, is simply the use of a proposed questionnaire on a small sample of respondents in order to detect weaknesses in the questionnaire and to obtain responses to free-answer questions which will aid in developing coded responses. Pretesting is of major importance in the development of a valid and reliable questionnaire and must not be omitted unless the urgency of the test is such that pretesting is not possible.

7-2. VALUE OF THE PRETEST.

a. The pretest is the best check on the quality of questionnaire design and adequacy of question wording. As previously mentioned, the questionnaire is a measurement tool whose value depends to a large extent upon its ability to communicate effectively. The pretest affords an excellent opportunity to determine whether or not the questionnaire communicates as planned. It provides a check on question sequence, vocabulary level, and ambiguity. The adequacy of the entire questionnaire design and format may be appraised by the pretest method to insure that the questionnaire is as good as human judgment can make it.

b. In addition, the pretest may provide information which has been overlooked previously. The participants in the pretest may suggest important questions which should be included in the final questionnaire. Other advantages are that the adequacy of instructions to respondents may be verified and the pretest results may be used as a guide to set up final tabulation and analysis procedures. The pretest has repeatedly proved to be of practical importance in the improvement of the questionnaire technique and must be employed whenever possible.

7-3. THE METHOD.

a. The pretest is the first step in validating a questionnaire. A draft form of the questionnaire is used as an interview guide. A number of individuals representative of the final sample are questioned using the questionnaire. During these interviews, the interviewer analyzes each question, the responses given, and the questionnaire as a whole for possible defects. Problems or deficiencies in question design, as well as recommendations for improvements, are recorded in detail. These results are then used by the questionnaire designer to make the necessary revisions and corrections.

b. The sample of pretest respondents need not be large and elaborate methods for selection are not required. It is advisable, however, that an effort be made to include individuals whose experience, training, and general level of intelligence are similar to those expected to participate in the final test sample. When the first pretest reveals need for revision or rephrasing of questions, it is necessary to conduct another pretest of the revised questions. The follow-up pretest, however, may be conducted with a smaller sample of respondents than the initial pretest.

7-4. **ROLE OF THE INTERVIEWER.** Regardless of the form of the questionnaire, whether self-administered or individual interview, the first pretest is conducted as a personal interview. Using this procedure, false, misleading, or confused answers may be detected immediately and followed up by the personal approach. Usually, several different interviewers are used in the pretest and it is important that individuals other than the designer of the questionnaire conduct some of the interviews. The author of the questionnaire may be so familiar with the questions that he/she may fail to detect weaknesses in his/her work. It is important that pretest interviewers be critical; they must be ready and free to criticize every question, as well as the sequencing and format of the questions. The interviewers also must be capable of critically evaluating the adequacy of each response during an interview to insure that each question is eliciting the type of reply for which it was designed. The interviewers must be sufficiently familiar with the questionnaire and interview concepts to enable them to perform adequate analyses of each question and to suggest improvements.

7-5. **PRETESTING THE QUESTIONNAIRE.** Pretesting a questionnaire may present several problems, however, the technique can be applied with a minimum of difficulty. For instance, the extremely limited sample size of a typical test presents a real problem. The planner of a large-scale civilian consumer-type survey may work with a sample size of hundreds, or even thousands, of respondents; in a military test, the sample generally numbers between 10 and 30 participants. Another problem with a questionnaire may be the fact that the questions assume experience with a specific item under particular experimental conditions, and the only individuals suitable to participate in the pretest are the actual test participants. If either of these unique conditions exist, the pretester may find it necessary to question individuals from among the test participant group.

a. Timing of the Pretest. Timing is an important factor in a pretest. In some instances, a pretest can be administered only after the participants have had sufficient experience with the test item to intelligently answer questions concerning it. On the other hand, the pretest must be given at a time which allows revisions, if any, to be made and incorporated into the questionnaire before the initial test administration. The orientation periods or break-in phases for a test system usually provide the best opportunity to conduct the pretest. Timing of the pretest, to allow for subsequent authorization for changes by the project director or test manager, requires close coordination among the test team administrator, test manager, and questionnaire designer. If time is limited and changes in the questionnaire design are necessary, recommendations must be forwarded to the responsible test manager or questionnaire designer by the fastest possible means. Changes to the questionnaire must be processed and approved in the same manner as required for the original questionnaire since such changes could affect the results obtained.

b. Pretest Sample. The pretesting of a draft questionnaire is usually a simple process. A cross-section of approximately twenty percent of the total number of test participants should be selected to constitute the pretest sample. The participants must not be informed that they are participating in a pretest of the questionnaire since this knowledge may influence the responses. The selected participants should be interviewed individually and apart from one another. The interviewers must not comment to the pretest respondents regarding deficiencies or omissions in the questionnaire. Recommendations for changes or improvements must be noted by the interviewers during the interviews and fully described and recorded immediately following the interviews. Written recommendations, together with oral comments or suggestions, must be provided to the individual responsible for making the revisions as soon as possible after all interviews are completed.

7-6. SELECTED READINGS. See paragraphs I-6 and I-8, appendix I.

CHAPTER 8

QUANTIFICATION AND ANALYSIS OF DATA

8-1. INTRODUCTION.

a. Consideration must be given to the quantification and analysis of data during the early planning stages of questionnaire development. Plans for the design and administration of the questionnaire, as well as for data quantification and analysis, must be developed in conjunction with the overall experimental design and plan of test. This is essential in order to determine such factors as:

- (1) The kind of sample required,
- (2) The number of respondents to be included in the sample,
- (3) The frequency and schedule of administration,
- (4) The number and type of questions to be used, and
- (5) The type of analysis which will be made.

b. It is unlikely that the questionnaire designer also will be a qualified statistician or knowledgeable of statistical analysis techniques. Therefore, it is essential that the individual responsible for the design of the questionnaire seek the assistance and advice of a qualified statistician, in addition to coordinating administrative requirements with the test supervisor. In order to insure that the questionnaire produces the desired data, it is suggested that the planning task be approached by first deciding what hypotheses will be tested. By this process it is possible to determine what analysis techniques can be employed to test the hypotheses and something about the tabulations which will be required to summarize the results. From these tabulations, it is possible to determine the type of questions needed and some characteristics of the sample required to produce the results. Again, it is emphasized that early consideration must be given to the form which the data will take and the kind of analysis which will be made of the data when it is obtained. If this step is ignored, a great deal of time and effort may be wasted in testing the initial hypotheses.

8-2. QUANTIFICATION OF DATA.

a. As stated at the beginning of the chapter, the questionnaire is used for measurement purposes. The product obtained from administration of the questionnaire consists of subjective words or phrases. This information must be quantified - converted to figures or numbers that can be tabulated and analyzed. The end product of the questionnaire may be a simple frequency distribution of responses to each question summarized in terms of numbers, proportions, or percentages. The data may be further summarized to include averages, standard deviations, or correlations. The summaries also may include statistical analyses showing the statistical significance of differences or correlations obtained. These quantified data must then be tabulated and analyzed. The results usually are summarized in tabular form for inclusion in the final report.

b. Data obtained from rating scales, dichotomous responses, and precoded answers are not difficult to quantify since numerical values or precoded numbers are assigned ahead of time. The written responses to open-end or "Why?" type questions require the additional steps of coding and classification prior to tabulation and analysis. The coding and classification process is the same as that described in detail in chapter 6.

8-3. TABULATION OF DATA.

a. After the data have been quantified, they must be organized into a form which will aid in the analysis and presentation in the report. The construction of a table is an effective means of organizing the data obtained from a questionnaire. A table serves several purposes. It may be used to list the raw data in terms of frequencies of response to each question; the response frequencies may be further organized by test group, test phase, date of administration, etc. A table may sometimes be lengthy, however, the importance of the results may justify including a raw data table as an appendix to the report. A table also may serve to summarize the findings by presentation of an organized and concise picture of the findings which support the conclusions and recommendations. Examples of the types of tables which may be used to summarize and present data are shown at appendix F.

b. A table should be designed and prepared so that it tells the complete story and stands alone without additional narrative explanation other than the heading and concisely worded footnotes. The heading should tell what the table is about. Headings of columns and rows should be descriptive of the contents and lines should be drawn to clearly distinguish the columns and rows. The table should be clearly labeled so that there is no need to refer to the text in order to interpret the contents.

c. In the organization of a table, careful consideration must be given to the most important information contained in the table and the main points it is intended to bring out. The table must then be designed so that the main points are the easiest to observe. In this regard, it must be remembered that people read from left to right across the page and, therefore, headings of columns must be more prominent and easier to follow than headings of rows. If the information contained in the columns and the rows is of equal importance, consideration should be given to the fact that long lists of data look better in columns and short lists in rows. In some instances, results may be effectively summarized and presented in a graph or chart. There are many types of graphic techniques, including the bar graph, pictograph, pie diagram, and trend chart, which may be used instead of a summary table or to summarize the principal findings from several tables. Examples of these and other graphic techniques are discussed in the reference cited at paragraph H-4, appendix H.

8-4. ANALYSIS OF DATA.

a. The dictionary definition of analysis is ". . . an examination of a complex, the elements and their relations." More explicitly, analysis consists of a review of the data, in terms of the objectives and criteria, to identify significant facts and relationships among the variables of interest.

b. Analysis must include a careful examination of the data to provide answers to questions such as:

- (1) "Are specific objectives met, not met, or exceeded?"
- (2) "What are the causes and effects of not meeting objectives?"
- (3) "Are there differences between standard and experimental items?"
- (4) "Are the differences obtained meaningful?"
- (5) "How do the findings relate to the conclusions and recommendations?" and
- (6) "Are there unexpected findings or new facts which warrant further investigation?"

c. The terms "analysis" and "statistical analysis" are often used synonymously. While it is possible to perform an analysis of results without the use of statistics, normally the analysis will include the use of some statistical techniques. No attempt will be made to cover all of the statistical methods available and applicable to the kinds of data which may be obtained from a questionnaire. However, a brief overview of the subject is considered to be warranted to familiarize the reader with those techniques which are used most frequently in the analysis of questionnaire data. There are many excellent texts readily available covering tests of significance, sampling, and correlation techniques (paragraphs H-2, H-4, H-7, and H-10, appendix H).

d. Statistics are important tools of the analyst, and an analysis of questionnaire data usually involves the use of either descriptive or sampling statistics, or both. Descriptive statistics are numerical descriptions of situations or conditions. Averages such as the mean, the median, and the mode are descriptive statistics; standard deviations and correlations are also descriptive statistics. Sampling statistics are used to determine how well the sample from which the measurements are taken actually represents the population from which the sample is drawn. Since it is seldom possible to employ the entire population of an item, or classification of personnel, in a test, sampling statistics must be used in order to make a judgment or inference about the population. For example, to determine which of three types of body armor is preferred for use by infantry Soldiers, obtain a sample of 30 each of the three types of armor and select a representative sample of 30 infantry Soldiers to alternately wear each of the three types of armor for the same period of time and under the same conditions. Then administer a questionnaire to obtain preferences for the three types of armor worn. Based on the analysis of the results, an inference may be made that the type with the highest preference is most preferred by the entire population of infantry Soldiers. The validity of such an inference, however, depends upon:

- (1) The margin of preference of the most preferred type of armor,
- (2) The representativeness of the sample of body armor used, and
- (3) The representativeness of the group of infantry Soldiers questioned.

There are statistical techniques which may be used to determine whether or not the observed differences are large enough to state with confidence that the differences obtained represent real differences in the larger population. A summary of some of the various types of statistical methods which may be used in the analysis of questionnaire data, together with references to texts which describe each of the techniques, is shown at table 8-1. Examples of statistical techniques, such as chi-square, t tests, and F tests, which are applicable to the analysis of the various types of data obtained from questionnaires and interviews are shown at appendix G.

TABLE 8-1
 STATISTICAL METHODS APPROPRIATE FOR ANALYSIS OF QUESTIONNAIRE RESULTS

CONDITIONS OF ANALYSIS	TYPE OF ANALYSIS			REFERENCES
	PARAMETRIC	NONPARAMETRIC	REFERENCES	
Two Sets of Measures	t Test	McNemar test Sign test Wilcoxon matched pairs signed rank	Para H-2 (pg 124)	Para H-10 (pg 63) Para H-2 (pg 250) Para H-10 (pg 75)
	F Test (Analysis of variance)	Chi Square test (two ind. samples) Mann-Whitney U-test Kolmogorov-Smirnov two samples	Para H-2 (ch 10)	Para H-10 (pg 104) Para H-10 (pg 116) Para H-10 (pg 127)
More Than Two Sets of Measures	F test	Cochran Q test	Para H-2 (ch 10)	Para H-10 (pg 161)
	Analysis of variance (ANOVA)	Chi Square (k ind. samples)	Para H-2 (ch 10)	Para H-10 (pg 175)
Two or More Sets of Measures	Analysis of Covariance	Friedman two-way ANOVA Kruskal-Wallis one-way ANOVA	Para H-2 (ch 12)	Para H-10 (pg 166) Para H-10 (pg 184)
	Product moment correlation (r)	Kendall rank correlation	Para H-2 (pg 193)	Para H-10 (pg 213)
MEASURES OF ASSOCIATION AND PREDICTION	Tetrachoric correlation (r _t)	Kendall coefficient of concordance	Para H-4 (pg 126)	Para H-10 (pg 229)
	Multiple correlation	Spearman rank correlation	Para H-4 (pg 392)	Para H-10 (pg 202)
SIGNIFICANCE OF DIFFERENCES				

These examples are not exhaustive, however, they do cover the most commonly used statistical methods. Other parametric and nonparametric techniques are given, together with their derivations, in the referenced texts on statistics.

e. Some of the advantages of statistical analysis techniques are summarized as follows: (paragraph H-4, appendix H)

(1) Statistics permit exact descriptions.

(2) They force personnel to be definite and exact in their planning and procedures.

(3) Statistics enable personnel to summarize and present results in a meaningful and convenient form.

(4) They enable personnel to draw general conclusions and to make predictions.

(5) They enable personnel to analyze some of the causal factors out of complex and otherwise confusing events.

8-5. REFERENCES. See paragraphs H-2, H-4, H-7, and H-10, appendix H.

8-6. SELECTED READINGS. See paragraphs I-3, I-4, and I-7, appendix I.

PART TWO
PRINCIPLES OF INTERVIEWING

CHAPTER 9

INTRODUCTION

9-1. GENERAL. When properly used, the personal interview is the most effective means of obtaining Soldier operator-maintainer information. The personal interview is a technique by which an individual is questioned by a skilled interviewer, who records all replies of the respondent. The data from many such interviews may be collected, tabulated, and analyzed in a manner similar to that for questionnaires. Proficiency in personal interviewing can be acquired; it consists of a combination of specific habits, skills, and techniques, rather than one general ability. This section sets forth some general rules and procedures for conducting an interview which should help the interviewer avoid unnecessary mistakes and obtain reliable data.

9-2. PURPOSE OF THE INTERVIEW.

a. The interview has been defined as a "conversation with a purpose." The purpose of an interview is to find out either objective facts related to the system about which the interviewee has some knowledge, or subjective facts, attitudes, or opinions about how he/she feels about something. The interview must be designed to obtain these facts with as much clarity and accuracy as possible.

b. The interview, used to obtain information from the Soldier operator-maintainer, attains its greatest value from the relationship which is established between the interviewer and the respondent. The inquiry is on a personal level, a conversational give-and-take, which encourages a respondent to give any information he/she can to the interviewer. In a properly conducted interview, where a genuine rapport is established between the interviewer and the interviewee, it is possible to obtain more detailed and reliable data than from the self-administered questionnaire.

9-3. SOURCES OF INTERVIEW BIAS. Bias in the interview situation can be defined as some attitude or prejudice on the part of either of the participants which affects the response by distorting or slanting it from the truth. Ideally, the interview results in the interviewee supplying accurate information to the interviewer. However, the influence of bias can alter the results to such an extent that the answers are of little or no value in the final analysis. The interview situation is highly susceptible to bias of many kinds. The interviewer may bias the interview by tone of voice, the way in which the questions are phrased, or even by facial expressions. He/she may unwittingly influence the respondent by pausing at certain points or by thoughtlessly agreeing with the respondent on a particular response in order to maintain rapport. The interviewer also may unconsciously bias the answers by communicating his/her ideas or feelings regarding a question. These, and other, sources of bias can be greatly reduced - first, through recognition of the problem and, then, by training and experience.

a. Interviewer Bias.

(1) In order to maintain objectivity during the interview, and to avoid the most damaging errors, the interviewer must be aware of the fact that he/she is potentially the greatest source of error and misunderstanding in the interview situation. The interviewer must be aware of the fact that he/she has prejudices, likes and dislikes, pride of opinion, and perhaps even a personal preference for a particular item being tested. These feelings, while normal, are potentially dangerous if they are permitted to influence the interview in any way. Careless interviewing, without careful consideration of the many factors which may affect a respondent's replies, almost certainly produces distorted and invalid results.

(2) It is assumed that the interviewer will not consciously influence the respondent's answers, however, great care must be exercised at all times to insure that the interviewer's ideas or opinions are not communicated to the respondent, either during the actual interview or at other times during the test. Often, an unconscious mannerism or casual remark is sufficient to influence a respondent. Bias is frequently communicated through some unconscious influence such as the interviewer's tone of voice, his/her inflection, gestures, or facial expression. A respondent is usually anxious to please the interviewer and to provide the desired answers, particularly when he/she does not have strong feelings on the subject. Therefore, any sign or cue from the interviewer is likely to influence the response. Although bias on the part of the interviewer is difficult to control, a knowledge of the sources of interviewer bias, together with a conscious effort to eliminate these influences, should succeed in keeping such bias to a minimum.

(3) One source of bias, which is often overlooked in the military interview situation, is the difference in rank or grade of the military interviewer and the interviewee. If the rank of the interviewer is higher than that of the interviewee, and if the rank is evident or known to the interviewee, this fact may bias and invalidate the results obtained. The higher rank of the interviewer may intimidate the respondent and cause him/her to be reluctant to make detrimental comments about the system under test. On the other hand, a respondent may exaggerate or over-emphasize the desirable features of a system in an effort to please his/her superior. The most desirable situation is the absence of any indication of rank or knowledge on the part of a respondent of the rank of a military interviewer. If the rank of the interviewer cannot be concealed, every effort must be made to select a qualified interviewer of a rank or grade no higher than that of the individual being interviewed.

b. Bias and the Interviewee.

(1) Another source of bias is the interviewee. The interviewee, or respondent, brings his/her bias with him/her to the test situation in the form of attitudes and opinions about the Army in general and, possibly, about the specific system under test. For example, he/she may be bitter about his/her particular job or current assignment, or about being sent to a particular test site to participate in a test; in many cases, such a negative attitude is transferred to the test system. On the other hand, the interviewee's past experience with the standard system, which the experimental system is designed to replace, may influence his/her attitude toward the test system. Depending upon his/her specific past experience, this could result in either a positive or negative bias toward the test system.

(2) The conditions under which the interview is conducted also may influence the interview. Such factors as the time of day, climatological and environmental conditions, and degree of privacy afforded are very important considerations. For example, the interview must be planned at a time of day which does not conflict with a scheduled meal time or immediately before the interviewee is scheduled for leave. In such instances, the interviewee may give incomplete and unresponsive answers in order to complete the interview as quickly as possible. The interviewee's responses also may be biased by discomfort or personal inconvenience caused by the interview environment. If the interview is conducted out of doors or in a room which is extremely cold or extremely hot, or if the interviewee is not afforded a comfortable seat, the data obtained are likely to be incomplete. The lack of quiet and privacy, away from the hearing of others, also may bias the responses obtained. Therefore, considerable thought must be given to scheduling and planning the interview in order to eliminate such sources of bias.

9-4. TYPES OF INTERVIEWS. The types of interviews which are most useful in tests of military systems are the individual or questionnaire type of interview and the group interview. In some test situations it may be desirable to use a combination of these two types of interviews to obtain the required results.

a. The Questionnaire Interview. In the questionnaire interview, the individual being interviewed is asked a series of questions which were constructed prior to the interview. The document on which these questions are listed is referred to as an interview guide. The questions for the interview guide are prepared following the same procedures for wording, sequencing, and formatting as previously described in Part One for the questionnaire. The interview guide differs from the questionnaire only in that more space is provided for recording responses to open-end questions and responses to follow-up questions. An example of the interview guide is shown at appendix A.

b. The Group Interview. The term "group interview" is used to denote the group administration of a self-administered questionnaire. Although the group interview, as referred to herein, is not an interview in the technical sense, it does rely on many of the same techniques as are employed in the personal interview. The results of a group interview also can be influenced by the administrator in much the same way as during the personal interview. The administrator's primary task during a group interview is to establish rapport with the group, motivating them so that they will be cooperative and responsive. The administrator also must give the necessary instructions and orientation on the questionnaire itself, answering any question which may arise in the interpretation of the instructions.

c. Special Applications.

(1) Each type of interview lends itself more readily to obtaining certain kinds of information.

(2) The questionnaire type of interview, for example, is valuable for obtaining specific facts, ratings, opinions, and observations. Through the use of this method, a standard set of questions may be developed and administered to each participant in the test sample. The data from such a questionnaire are easy to tabulate since all questions are exactly the same and they are administered to each individual in the same manner.

(3) The group interview is used when large quantities of data are required. It is one of the most efficient means of obtaining responses from large groups of individuals (20 or more) since all may be questioned simultaneously and under the same conditions. However, the group interview does not provide the means for obtaining the degree of detailed information possible with the guided questionnaire interview.

CHAPTER 10

THE QUESTIONNAIRE INTERVIEW

10-1. INTRODUCTION. The questionnaire interview is a technique whereby each participant is interviewed individually, apart from all others. It is a structured interview situation in that the interviewer asks questions exactly as they are written on a prepared questionnaire. When the questionnaire is used in this manner, it is referred to and is entitled "interview guide." The format of the interview guide is similar to that for a questionnaire, except more space is provided for responses and comments. The interview situation provides greater flexibility in obtaining the required information, however, it also involves greater risk of bias. The effectiveness of the individual interview method depends to a large extent upon the rapport established between the interviewer and the respondent.

10-2. ESTABLISHING RAPPORT.

a. Rapport is a word which has been adopted by users of the interview technique to signify the jointness of purpose and desire for cooperation which must be present during any interview. The word, taken from the French, means a sympathetic relationship between two or more people. The first essential for a successful interview is the establishment of this relationship between the interviewer and interviewee.

b. At the beginning of an interview, an interviewer can usually expect a certain amount of reluctance on the part of the interviewee. The degree of reluctance encountered depends largely upon the immediate situation and the attitude of the individual being interviewed. His/her personality, understanding of the situation, or any stress he/she may have recently experienced, can effect his/her feelings toward the interview. The first, and perhaps the most important, aspect of the interviewer's job is to overcome any negative attitude on the part of the interviewee toward the test or the interview situation. He/she must put the interviewee at ease and try to develop an atmosphere conducive to wholehearted acceptance and cooperation.

c. There are several ways of establishing rapport, and the approach used in each situation must be based upon the immediate conditions, including the attitude and needs of the individual being interviewed. The interviewer should begin by introducing himself/herself to the interviewee, followed by a brief but clear explanation of the purpose of the interview. During the course of this conversation, the interviewer should convey the fact that the questions are in no way a test of the interviewee's personal abilities or intelligence. To many individuals, the interview situation is similar to a "test" and represents a threat. The idea of threat should be overcome as early as possible during the interview, otherwise the reliability of the interviewee's responses may be jeopardized. The respondent must be assured that his/her statements will not be used in any way which could be harmful to him/her or his/her interests. The use of phrases such as "not interested in names" and "there's nothing personal in this study, we are only interested in statistics," may help to convey the desired impression of interest in statistical data which would not harm the interests of the respondent. There are several ways of producing and maintaining a positive and cooperative attitude in a respondent. A respondent must never be put on the defensive or made to feel that he/she is wrong. A little politeness, demonstrated by the use of words or phrases such as "please," and "would you mind," goes a long way in maintaining rapport. The interviewer must never talk down to a respondent or talk over his/her head. The use of technical terms and abbreviations should be avoided.

d. Another important factor in establishing rapport is the attitude of the interviewer. He/she must appear to be relaxed and at ease in the interview situation. This helps to put the respondent at ease and to reduce any tensions or inhibitions on the part of the interviewee. The attitude of the interviewer must convey sincerity, encouragement, patience, and understanding. His/her attitude should encourage frankness and honesty on the part of the respondent. A respondent will usually be open and candid when he/she is made to feel that his/her point of view is appreciated and respected. Successful interviewing requires that both participants make every effort to understand each other. The key to genuine mutual understanding is a display of sincerity and interest on the part of the interviewer.

10-3. ASKING THE INTERVIEW QUESTIONS. There are three important points to remember when asking the interview questions: the sequence of the questions; ask every question; and wording of the questions.

a. Question Sequence. During the construction of the questionnaire or interview guide, particular attention is given to the sequence of the questions. The order or arrangement of the questions is based upon a logical order of thought processes regarding the system or activities that occur during the test. Consideration also is given to possible "order effects," that is, the effect that a preceding question may have on the response to the question under consideration. It is essential, therefore, that the interviewer follow the order of questions as they appear on the interview guide and that the sequence not be altered without prior approval. The interviewer must control the interview at all times so that the respondent answers the questions in the sequence asked and does not anticipate the questions by his/her responses. The question sequence must be maintained since considerable thought and pretesting have usually been done to establish the best sequence for producing reliable results.

b. Asking Every Question. A common mistake on the part of an interviewer is to assume that he/she knows what the answer will be to a certain question based upon the response received to a previous related question and fail to ask it. The interviewer should never assume knowledge of a response to a particular question, but should ask every question as it is written. Even highly skilled and experienced interviewers are susceptible to this kind of assumption, which can result in a loss of valuable data. An example of when this kind of assumption may be tempting to an interviewer is when a rating scale precedes a preference question. The respondent may rate one of several items tested much higher than the others with regard to one or more characteristics. Based on these higher ratings, the interviewer may assume that this highly rated item also is preferred for overall use. In fact, however, the open preference question may produce a surprising response in that the respondent may prefer a less highly rated item for some reason or characteristic not covered by the ratings. For instance, on a test of experimental raincoats, the respondent may rate coat "A" better than coat "B" with regard to comfort, warmth, and appearance, but may state on overall preference for coat "B" based upon its superior resistance to water penetration. In this case, the importance of waterproofness would have been overlooked if the interviewer had assumed a preference based on the previous ratings. It is important, therefore, to ask every question and assume no responses.

c. Question Wording.

(1) A great deal of time and thought are given to the wording of every question before it appears in final form on the questionnaire or interview guide. Each question is carefully worded and usually pretested to insure that it:

- (a) Expresses the exact meaning desired,
- (b) Is as free from bias as possible, and
- (c) Has the same meaning to each respondent.

Questions are often rewritten many times so that they meet these standards.

(2) Experiments have shown that even a slight change in the wording of a question can distort the results. Any change in wording could influence the response, and there is no assurance that a change will be consistent with the original intent of the question. At the very least, the answers obtained will probably not be comparable to those obtained by other interviewers using the original wording. Whenever more than one interviewer is involved, it is essential that every interviewer ask each question in the same way during all interviews. If a problem develops during an interview as a result of the wording, the interview must be continued using the wording provided. A note about the faulty wording should be made on the interview guide or questionnaire and the matter brought to the attention of the test supervisor as soon as possible. Even in a small scale test involving only one interviewer, the policy of asking every question exactly as it appears on the interview guide must be adhered to in order to insure consistency.

(3) Although every effort is made to word each question so that it may be understood by each respondent, the interviewer may encounter a situation where he/she receives a blank stare or baffled expression in response to a question. Under such a circumstance, the interviewer should repeat the exact wording of the question slowly and distinctly, emphasizing the key words. Usually, the repetition is sufficient to elicit a response. If, after repeating the question, there is still no response, the interviewer should record "no response" for the question and proceed with the interview; this is the only way to insure comparability and reliability of the results. If every interviewer provided his/her explanation or interpretation of the questions, the results obtained would not be comparable or reliable.

(4) The interviewer also must avoid the usual tendency to rephrase or elaborate on the wording of a question. When a respondent appears confused or "stumped" by a question, it is often tempting to suggest a possible answer; this temptation must be overcome to avoid biasing the results. If a respondent seems to be having difficulty, the interviewer must limit his/her comments to harmless statements such as "take your time," "think the question over carefully," or "give me the best answer you can."

10-4. **OBTAINING THE RESPONSE.** The primary purpose of the interview is to obtain data in the form of answers to specific questions. It is the job of the interviewer to insure that the answers obtained are specific and usable. The interviewer's attitude and manner greatly facilitate the responses. The interviewer must be thoroughly familiar with the questions and ask them in a conversational manner - which helps to create an informal atmosphere rather than a sterile test situation. Too serious an attitude on the part of an interviewer may discourage free expression by the respondent.

a. Getting Specific Answers.

(1) The interviewer should not go to the next question until the respondent's reply to the current question has been obtained. The following are some examples of replies which are not specific:

Question: "Which would you prefer on the field jacket, button or zipper front closure"

Answer: "I can't decide."

(This question should be repeated by asking "Well, which would you prefer on the field jacket, button or zipper front closure?")

Question: "What was the main thing you liked about the variable protection body armor?"

Answer: "I liked it fine."

(The respondent has missed the point completely. The question should be restated stressing "the main thing.")

(2) Each of the above responses is an example of "an answer" which is not a specific, usable answer. The interviewer would fail if he/she accepted these replies. It is the job of the interviewer to make every effort to get a specific answer to the particular question asked, without influencing direction or intensity of the respondent's answer. Sometimes a simple repetition of a question, stressing key words, elicits an acceptable answer. In some instances, the interviewer may have to spend several minutes on one question, repeating it two or three times, before a satisfactory response is obtained. Repetition of the question is the basic method of probing when using the questionnaire interview or interview guide. It is the safest and most effective way to secure consistent and comparable replies to a specific question printed on the guide. By repeating the question, the interviewer encourages the respondent to develop his/her thoughts and to express them in his/her words. In many instances, repeating the question gives the respondent time to think and directs his/her attention to the real purpose of the question - encouraging a specific answer.

(3) In an instance when the interviewer does not fully understand the response, or a respondent's answer is not clear or complete, it may be necessary to ask the respondent to clarify his/her answer. It is unacceptable to ask the respondent a question such as "Do you mean this . . . ?" Such a question immediately suggests an answer, which would produce biased results. Neutral probing questions may be used if repeating the question does not work, such as "Well, what in particular did you have in mind?"; "I'm not sure I got that last point you made. Could you please expand it a little further?" or "Let's see now, you said . . . Just how did you mean that?" These and similar neutral phrases may be used to encourage a respondent to clarify a response or to provide a specific answer to a question without biasing the results. The interviewer must not skip a question and return to it later in the interview. The questions must be asked in the sequence in which they are written to avoid biasing the responses.

b. Don't Know Responses. There will be occasions when a respondent cannot answer a question as it is stated. In some cases the "don't know" response is legitimate because the respondent has not had particular experience with an item, misunderstands the question, or simply because he/she can't remember. In other instances, however, the initial "don't know" response is merely a coverup for uncertainty, misunderstanding, or even mental laziness. A preliminary "don't know" response may be to hide genuine feelings; it is an interviewer's job to reveal them. The interviewer must be aware of this problem and probe for a specific answer until he/she is certain that a respondent is unable to commit himself/herself on the question.

10-5. REPORTING THE RESPONSES.

a. Proper Recording. All of the effort expended in putting the respondent at ease, carefully asking the questions, and obtaining specific answers may be wasted unless each response is recorded properly.

b. Mistakes. If an interviewer forgets to check an answer, the result is the same as if he/she never asked the question; if he/she makes a mistake and marks the wrong category, the result will be invalid; if he/she fails to record a free answer exactly as the respondent gives it, the analyst may get a distorted view of the respondent's opinion. The interviewer's function in this regard is to accurately record and communicate the responses and opinions of each test participant to the analyst who must analyze, evaluate, and report the results. Failure to do this will result in unreliable data and invalid test results.

c. Effort Involved. The interviewer must not only make every effort to obtain complete and specific responses, but must exercise great care to insure that the responses obtained are recorded completely and accurately. An unfavorable interview condition, or frequent interruptions, during an interview causes confusion and may result in errors in recording the responses. The interviewer must be familiar with the questions and the entire questionnaire to avoid asking questions not included on the guide and to insure that all questions which are listed are asked and the responses recorded.

d. Types of Errors.

(1) Some of the most common types of errors in reporting interview responses are as follows:

(a) Omission of response. The most frequent error in response reporting is omission - the failure to check or write down the answer to a particular question.

(b) Asking a question which should be omitted. In some instances, the instructions for a question state that it should be asked only under certain conditions, such as "If a 'yes' response is given then ask . . ." In the case of a "no" response the question should not be asked. Asking a question which should be omitted is confusing to the respondent as well as to the analyst.

(c) Circling or checking the wrong answer. This type of error is usually committed as a result of carelessness. When the reply checked is different than the one given, it will probably not be detected as an error and will distort the results.

(d) Marking more than one answer. This error occurs most often as a result of the respondent changing his/her mind. The interviewer marks both answers and fails to erase or eliminate the original response. It is important to check one, and only one, answer for each question. If both answers remain checked, neither answer can be used since the analyst does not know which of the two represents the respondent's final opinion.

(e) Inserting a dash to indicate a "Don't Know" or "Refused Answer." The use of a dash or line instead of writing out the correct response cannot be accurately interpreted or placed in a coded category by the analyst.

(2) The most experienced interviewer makes an occasional error or omission in spite of rigorous checking, however, the number of errors can be minimized by adhering to the following basic principles:

(a) Become thoroughly familiar with the interview instructions, wording, and layout of the interview guide before conducting the interview.

(b) Inspect each completed questionnaire carefully before going to a new respondent. Check the interview guide for errors and omissions immediately after completing each interview. Correct any error and finish recording any free-answer response which isn't written out in full.

(c) Avoid unfavorable interviewing conditions. Error is more likely to occur when there are frequent interruptions or when the presence of others creates distractions. Interview each respondent individually, apart from all others, and in quiet, comfortable surroundings.

e. Reporting A Change in Meaning. If a respondent either qualifies a question or interprets a question in a manner different from that intended, his/her interpretation, or a qualifying statement, must be recorded on the questionnaire. For example, if the purpose of the interview is to ask the same question of 10 individuals, it is important to know about any deviation from the intended meaning of that question. By recording a misinterpretation, the analyst can void or treat the answer separately if he/she feels the respondent completely changed or missed the intended meaning. The following are examples of qualifications which should be recorded, in addition to the initial answer:

Question: "Do you think these gloves are suitable for combat use in the Arctic?"

Answer: "Yes, unless the temperature gets below about 15 degrees."

Question: "What do you consider to be the single most important characteristic of a good helmet?"

Answer: "Comfort, except in a combat zone where protection is the most important."

Question: "Based on your experience, would you say the armor vest is satisfactory or unsatisfactory?"

Answer: "I'd say satisfactory, but only if it has a higher collar."

f. Reporting A Free-Answer Response.

(1) The free-answer question poses a special problem in recording a respondent's opinion. For a free-answer response, it is necessary that the entire response be recorded verbatim. It is important that the interviewer record every word the respondent says since the significance of all statements must be determined by the analyst. For the multiple choice or dichotomous question, where an answer is checked, tabulation of the results simply involves adding up the numbers. For the free-answer question, however, it is necessary to devise a set of code categories and to determine into which coded category each answer falls. Deciding whether a particular reply should be classified in one category or another is often a difficult decision. The coding process may be particularly difficult, or even impossible, if the interviewer is careless in recording the exact response. If the interviewer fails to record the full response, or only writes a summary of the response, the analyst may be unable to accurately code the answer for inclusion in the data analysis. On the other hand, if the interviewer faithfully records the answer clearly and completely in the respondent's exact words, the answer may be confidently and accurately grouped into one of the coded categories.

(2) The free-answer response also must be recorded in the respondent's language. The interviewer should listen attentively to what the respondent says, the words used, the way the response is phrased, and then quote him/her directly. The interviewer should not summarize, but should record the respondent's words, including bad grammar, slang, or profanity. In this way, the respondent's emphasis and true feelings will be conveyed to the analyst.

(3) A knowledge of shorthand is not required in order to record an answer verbatim. First, it would be unusual for more than a few answers to be more than one or two sentences long, and the average individual speaks slowly enough that it is not difficult to keep up with him/her. Secondly, a respondent usually is glad to give an interviewer time to write down all that is said since the respondent is interested in having his/her opinions recorded completely and accurately. A respondent also may be requested to repeat an answer in order to insure that a response is complete and accurate.

10-6. INTERVIEW TECHNIQUES.

a. The following interview techniques should prove helpful to the interviewer in becoming adept at speedy and accurate recording of responses:

(1) Be prepared to write as soon as the question is answered. It may sound contradictory to previous statements about verbatim recording, but it is not always necessary to begin writing as soon as the interviewee begins to talk. Often the respondent's first few words are just a "warm-up" and serve as a pause while he/she organizes his/her thoughts. It is a waste of time and effort to record this part of a conversation and a waste of the analyst's time to try to interpret and clarify it. Irrelevant remarks such as, "Well now, I was talking to a friend about that just the other day . . ." or "I'm not sure whether this is what you mean, but I would say that . . ." need not be recorded, but when a respondent actually starts to answer the question, be prepared to write it out as fully and as accurately as possible.

(2) Write quickly, but legibly. In addition to recording all of the words of the respondent, it is also essential that the handwritten material be readable so that the answers may be accurately interpreted and classified by the analyst. In writing a response, the interviewer should take particular care to distinguish between n's, m's, u's, v's, and w's, which often appear as a series of indistinguishable loops; o's and a's, which often are not closed; and t's and l's and other looped letters. Breaking a word in the middle also makes translation of the written response extremely difficult. Illegible handwriting should be corrected by the interviewer immediately after each interview. This should be done concurrently with the review for errors and omissions.

(3) Use common abbreviations. The interviewer should make use of common abbreviations whenever possible. Examples of some common abbreviations are "equip." for "equipment"; "DK" for "don't know"; "exp." for "experimental"; and "std," for "standard." The important concern is to record each response in full.

(4) Don't erase. If a mistake is made when recording a response, cross out the error rather than take the time to erase. Crossing out takes far less time than erasing and is the preferred means for correcting an interview response.

b. The most rapid writer may fall behind in recording a response if the interviewee talks fast and does not pause between sentences. In such an instance, take a moment or two before asking the next question. Usually, if the essential rapport of the interview has been maintained, the respondent will be favorably impressed that care is being taken to accurately record all that he/she has said. If the respondent shows signs of impatience, a comment such as, "This is very interesting, I want to be sure I have recorded it exactly the way you expressed it," is appropriate.

TECOM Pam 602-1, Vol I

c. Complete and accurate recording of each free answer response usually has a positive affect on an interviewee. When an interviewee observes that his/her opinions are considered important enough to be recorded faithfully, he/she is more likely to give full and open answers.

10-7. **SELECTED READINGS.** See paragraphs I-1 and I-5, appendix I.

CHAPTER 11

THE GROUP INTERVIEW

11-1. INTRODUCTION.

a. The group interview is a variation of the interview technique whereby a questionnaire is simultaneously presented to a group of individuals assembled in one place. The procedure is relatively simple in that a questionnaire is distributed to each of those present and each respondent records his/her answers. This type of interview also is referred to as "self-administered" since the individual in charge of the interview only supervises the group while each interviewee fills out his/her own questionnaire.

b. The primary advantage of the group interview is the ability to obtain large quantities of information with a minimum expenditure of time and effort. In an instance where a large sample is required, as in a food preference or clothing study, the group interview may be the most efficient means of obtaining Soldier reaction. Using this method, as many as a hundred or more questionnaires may be administered at one session.

c. The gain in efficiency achieved by the group interview technique is offset by loss in reliability of the results. In a group interview situation, it is not possible to control all of the various influences which are known to affect individual responses. If the respondents are seated close to one another they are likely to talk, exchange ideas and opinions, or, simply by gestures or informal comments, influence the responses of others. Another notable disadvantage of the group questionnaire is that it presupposes the ability of each respondent to read and comprehend the written questions, however, misunderstanding or misinterpretation can invalidate the responses.

11-2. THE ROLE OF THE ADMINISTRATOR. In the section on the personal interview it was stated that the interviewer himself/herself is perhaps the greatest source of bias. It might appear that the administrator of the group interview has little or no influence on the results, but this is not true. Although interviewer-induced bias may be less evident in the group situation than in the personal interview, it is of considerable importance. The interviewer's, or administrator's, opinions of or attitudes toward a system under test can be conveyed to a group by the method of introducing the questionnaire or in informal remarks and thereby influence or bias the results. This is particularly true in a case where an administrator has a much higher grade or rank than the respondents.

a. Orientation and Rapport. The main functions of an administrator are to explain the purpose of the interview and to explain to the respondents what is expected of them. The object is to inform the group of the need for and importance of the data and to engender their full cooperation; motivation of the group to perform the required tasks and establishment of rapport are important. The quality of the data obtained is greatly affected by the degree of motivation and rapport achieved.

b. Administrative Functions.

(1) The administrator also is responsible for numerous administrative details in connection with the group interview and may require one or more assistants, depending upon the size of the group. The administrative arrangements which must be made for interviewing a large group of respondents simultaneously are more numerous and considerably more complicated than those for interviewing one individual. The administrator usually is responsible for scheduling the interview at a time convenient to all participants, arranging for a building or room of sufficient size and with adequate facilities for seating and writing, insuring an adequate supply of questionnaire forms and pencils, arranging transportation, if required, and monitoring the interview.

(2) As mentioned in the chapter concerned with bias, the conditions under which the interview is administered are extremely important. A time should be selected for the administration that is agreeable to both the supervisors of the test participants and the test participants themselves. If the time of administration is inconvenient for the respondents, such as during off-duty hours or during a meal hour, the motivation of the participants and the quality of data obtained will be adversely affected. Every effort must be made to schedule an interview at a time convenient to all concerned, while still timely with regard to relevant activities or experiences.

(3) Often, it is the responsibility of an administrator to make arrangements for a suitable room in which to hold a group interview. In this case, consideration must be given to selecting a centralized location convenient to most of the group (to minimize transportation requirements) and to the availability of seating and writing facilities. Consideration also must be given to environmental factors such as adequate heating or cooling, and lighting, and a lack of distracting noises or activities which would adversely affect the results. In addition to insuring an adequate supply of questionnaire forms and pencils, the administrator also should make arrangements for visual displays which may be required for the orientation; these may include actual samples of the test items to be assessed during the interview or photographs of the items in various configurations. Such displays are often helpful to the respondents in identifying the items discussed in the questionnaire.

c. Overall Supervision. Finally, it is the responsibility of an administrator to supervise and monitor the entire interview. He/she should make written notes of difficulties encountered by the respondents with the questionnaire wording or format so that this information may be considered in evaluating the results and used to modify the questionnaire, if necessary, for future administrations. The administrator also must maintain order and minimize conversation among the respondents to reduce the possible affects of bias on the results.

11-3. CONDUCTING THE GROUP INTERVIEW. After the interview participants are assembled and all materials are passed out, it is the administrator's task to explain the purpose of the interview and the administrative procedures to be followed. During this initial orientation, the administrator must motivate and establish rapport with the group. He/she must impress the group with the importance of the information being sought and the importance of the roles of all individuals participating in the test. The administrator should encourage the group to answer each question completely and honestly, assuring them that their replies will be treated as test data, will be strictly confidential, and that there will not be any personal consequences as a result of the answers given.

a. Interview Instructions. Following this informal orientation, the administrator should read aloud the instructions provided, making sure that all participants understand them. The following is an example of instructions for a group interview:

PLEASE READ ALOUD AND CAREFULLY

"The purpose of this questionnaire is to find out what Soldiers think about the Army and Army life. Soldiers all over the country will fill this out just as you are doing today. We are not trying to check up on you as an individual, but you can help make conditions better for yourself and others in the future by filling the form out properly. **DO NOT WRITE YOUR NAME OR ORGANIZATION ANYWHERE ON THE QUESTIONNAIRE.** Please answer each question as frankly and as truthfully as you can. You will find that most questions may be answered by a simple check mark opposite your choice of answer."

"If you have a question, or if you are not sure about how to answer, raise your hand and one of the supervisors will help you. Remember, we are interested in your opinions, so please don't look at your neighbors' answers and don't discuss the questions or possible answers with others during the interview. We sincerely appreciate your cooperation."

If there is no question regarding the instructions or the questionnaire, the group should be instructed to begin work.

b. Assistance During the Interview. If, during the course of an interview, questions are asked regarding the intent of questions or the meaning of words, the administrator must not attempt to explain or alter the wording of questions. This is the only way to avoid bias, to be consistent, and to insure that all respondents are working within the same frame of reference during a particular interview. If a respondent persists in asking for help, the administrator should only repeat the question, emphasizing the key words, as in the personal interview. If the respondent still doesn't understand or has questions, it is often helpful to simply state "Just do the best you can." This may give the encouragement needed to continue without affecting the response. Such techniques are referred to as "nondirective" assistance and have been found to be very effective.

c. Review for Completeness. As in the case of the personal interview, the administrator must review each completed questionnaire as soon as possible after completion of the interview. The review should include a check for omissions, illegible responses, and open-end responses or comments which require clarification. In those instances when the respondents' names or other coded means of identification are not used on the questionnaires, the completed forms should be checked as they are turned in or before the group is dismissed. This review will help to insure receipt of maximum usable data from each questionnaire administration.

11-4. SELECTION OF INTERVIEWER.

a. The demands of an interview, whether personal or group, are such that the individual selected for the job should be of above average ability. The primary qualifications to be considered in the selection of an interviewer are:

- (1) Above average intelligence,
- (2) Ability to meet and talk to many different types of people,
- (3) Interest in the job, and
- (4) Rank no higher than respondents unless evidence and knowledge of rank can be concealed.

While it is desirable to select an individual with previous interviewing experience, this is not always possible. It must be stressed once again that interviewing is a skill which can be mastered through conscientious effort and the application of the principles described herein.

b. Whether the individual selected to be an interviewer is experienced or not, a certain amount of training and orientation are necessary for each specific test project. The individual selected should familiarize himself/herself with the test plan, to include the test requirements, objectives, procedures, and schedules. He/she also must be thoroughly familiar with the questionnaire or interview guides to be used in the conduct of the interview. If possible, he/she should participate in the preparation of these documents and any pretesting which may be required. The pretest administration provides a means for checking for errors in the questionnaire, and also provides an opportunity for the interviewer to receive training and experience.

c. It must be remembered that the primary advantage of the interview is the flexibility it provides. This flexibility derives from the ability of the interviewer to communicate with the respondent. The interviewer is thereby able to make sure that the respondent understands the purpose of the test and the questions asked. He/she also is able to probe for additional information or to clarify responses. Most important, perhaps, is the ability of the interviewer to establish and maintain rapport with the respondent, providing the encouragement and motivation necessary to obtain complete and useful information. The skill with which the interviewer is able to perform his/her tasks has an important influence on the results obtained.

11-5. **QUANTIFICATION AND ANALYSIS OF RESULTS.** The quantification and analysis of data obtained from interviews are identical to the techniques described in chapter 8 and appendix G for analysis of questionnaire data. Interview data consist of nominal, ordinal, or ranking, and interval data, and appropriate statistical methods include both parametric and nonparametric techniques.

APPENDIX A

SAMPLE INTERVIEW GUIDE

TEST OF COMPOSITE BODY ARMOR

NAME _____ DATE _____
(First) (M.I.) (Last) (Day) (Mo.) (Yr.)

TEST PHASE: (circle one) I II III IV

Instructions: Interview each test participant separately and apart from all others. Have a sample of each type of body armor within view of the participant. Provide a copy of the interview guide to the participant so that he/she can follow the questions and see the choice of responses during the interview.

1. Which type of body armor did you wear during the past 3-day wear phase?
(check one)

Type S

Type E

2. What type of test activities did you participate in during this wear phase? (check)

Close order drill

Road march

Cross-country march

Night patrol

Weapons firing

3. a. Did you experience any difficulty in performing any of the test activities while wearing this type of body armor? (check one)

Yes No

- b. If yes, then ask, "What seemed to cause the trouble?" (describe in detail)

4. a. How would you rate this type of body armor with regard to comfort? (check one)

- 6. Excellent comfort
- 5. Comfort is very satisfactory
- 4. About average in comfort
- 3. Slightly uncomfortable
- 2. Very uncomfortable at times
- 1. So uncomfortable it can barely be worn

b. If any "uncomfortable" category is selected, then ask, "What do you feel caused it to be uncomfortable?" (describe)

5. a. How would you rate the fit of this type of body armor? (check one)

- 6. Fits extremely well
- 5. Fit is quite satisfactory
- 4. Fit is about average
- 3. Fit needs improving
- 2. Fit is not very satisfactory
- 1. Fit is very poor

b. If fit is considered to be less than "about average," then ask "What seemed to be the problem with the fit of the armor?" (describe)

6. a. How would you rate the degree of freedom of movement afforded by this type of body armor? (check one)

- 6. Excellent
- 5. Very good
- 4. Adequate
- 3. Not quite adequate
- 2. Poor
- 1. Extremely poor

b. If less than "adequate," then ask, "What seemed to be the problem?"

NOTE: Ask the following questions only after Phases II and IV:

7. a. Now that you have experienced wear of both the Type N and Type S body armor, which type would you prefer to wear in a combat situation? (check one)

- Type N
- Type S
- Either Type

b. If Type N or Type S is checked, then ask, "Why would you prefer this type?"

APPENDIX B

SAMPLE QUESTIONNAIRES (SELF-ADMINISTERED)

LIGHTWEIGHT COMPANY MORTAR SYSTEM (LWCMS)

NAME _____ RANK _____ DATE _____
 (First) (M.I.) (Last)

MOS _____ TIME IN MOS _____

The purpose of this questionnaire is to obtain your opinions of the Light-weight Company Mortar System (LWCMS). Since you have participated in the firing of the LWCMS your opinions and comments are extremely important. Please answer all of the following questions as honestly and accurately as you can. Each question provides space for you to give any additional comments or information which you feel may be helpful.

1. In your opinion, is the LWCMS a safe weapon system to handle and fire?
 (check one)

Yes

No

Comments: _____

2. Did you ever have malfunctions which caused a mission delay with the LWCMS? (check one)

Yes

No

Comments: _____

TECOM Pam 602-1, Vol I

3. Compared to the 81mm Mortar, how easy or difficult was it to emplace the LWCMS? (check one)

- a. Much easier
- b. Slightly easier
- c. About the same
- d. Slightly more difficult
- e. Much more difficult

Comments: _____

4. Did you detect any residual flour remaining after the round was fired with the LWCMS? (check one)

- Yes
- No

Comments: _____

5. Compared to the 81mm Mortar, how easy or difficult was it to load the round into the tube of the LWCMS? (check one)

- a. Much easier
- b. Slightly easier
- c. About the same
- d. Slightly more difficult
- e. Much more difficult

Comments: _____

6. Did you have any problems when firing the LWCMS in the hand-held mode?
(check one)

Yes

No

Comments: _____

7. Overall, how would you rate the Lightweight Company Mortar System?
(check one)

a. Very good

b. Good

c. Poor

d. Very poor

Comments: _____

8. Which mortar system would you prefer to see in the Army inventory?
(check one)

a. 81mm Mortar

b. Lightweight Company Mortar

c. No difference

Comments: _____

QUESTIONNAIRE

DH-I FOOD TEST

NAME _____ GRADE _____ DATE _____
 (First) (M.I.) (Last) (Day) (Mo.) (Yr.)

ORGANIZATION _____ TEST PHASE _____

Instructions: The Army would like to have your opinion of certain food items served to you during this meal. Your opinion, along with the opinions of other Army personnel will help in determining which foods are more acceptable to the Soldier.

The names of the foods to be rated during this meal are shown below. Under the name of each food item is a rating scale. Please eat some or all of the foods served to you. After you have eaten some, or all, of a particular food, rate it by drawing a circle around the words in the proper rating scale which best describe how much you liked or disliked that food. If after rating one food, you wish to make any additional comments about that item, you may do so in the spaces provided:

Baked Ham

9	Like
	Extremely
8	Like
	Very Much
7	Like
	Moderately
6	Like
	Slightly
5	Neither Like
	Nor Dislike
4	Dislike
	Slightly
3	Dislike
	Moderately
2	Dislike
	Very Much
1	Dislike
	Extremely

Comments: _____

Green Peas

9	Like
	Extremely
8	Like
	Very Much
7	Like
	Moderately
6	Like
	Slightly
5	Neither Like
	Nor Dislike
4	Dislike
	Slightly
3	Dislike
	Moderately
2	Dislike
	Very Much
1	Dislike
	Extremely

Comments: _____

Applesauce

9	Like
	Extremely
8	Like
	Very Much
7	Like
	Moderately
6	Like
	Slightly
5	Neither Like
	Nor Dislike
4	Dislike
	Slightly
3	Dislike
	Moderately
2	Dislike
	Very Much
1	Dislike
	Extremely

Comments: _____

APPENDIX C

INTERVAL SCALE VALUES

C-1. MEAN AND STANDARD DEVIATION OF STATEMENTS FOR USE IN DEVELOPING SPECIFIC RATING SCALES.

COMFORT

<u>Mean</u>	<u>Standard Deviation</u>	<u>Statement</u>
5.50	0.84	Superior comfort
5.44	0.75	Superior in comfort
5.42	0.69	Excellent comfort
5.04	1.01	Exceptionally comfortable
4.99	1.45	Extremely comfortable
4.93	0.93	Perfectly comfortable
4.43	0.99	Very comfortable
4.41	0.65	Comfort is very satisfactory
4.13	1.38	Unusually comfortable
3.72	0.78	Comfort is satisfactory
3.51	0.87	Generally quite comfortable
3.44	0.59	About average comfort
3.39	0.83	Fairly comfortable
3.24	0.80	Moderately comfortable
2.58	0.81	Not too uncomfortable
2.57	0.86	Usually comfortable, but some- times uncomfortable
2.40	0.70	Slightly uncomfortable
2.32	1.00	Uncomfortable at times
2.30	1.00	Comfort is not quite adequate
2.05	1.09	Comfort is not very satisfactory
1.84	1.05	Comfort is barely adequate
1.75	0.78	Below average in comfort
1.68	0.91	Somewhat uncomfortable most of the time
1.66	0.86	Slightly uncomfortable all the time
1.47	1.42	Quite uncomfortable
1.40	1.14	Very uncomfortable at times
1.12	2.02	Extremely uncomfortable
1.00	0.85	Much below average
0.81	0.78	So uncomfortable it can only be worn for a short time
0.49	0.87	Very uncomfortable
0.40	0.78	So uncomfortable it can barely be worn
0.26	0.94	So uncomfortable it can't be worn

PROTECTION

<u>Mean</u>	<u>Standard Deviation</u>	<u>Statement</u>
5.65	0.78	Protection is perfect in every respect
5.55	0.54	Excellent protection
5.47	0.92	Protection is superior
5.05	0.77	Protects extremely well
4.94	1.03	Protection is ideal
4.49	0.70	Protection is very satisfactory
4.26	1.02	Protects unusually well
4.12	0.77	Protection is above average
3.99	0.74	Protection is good
3.98	1.00	Protection is very good in most respects
3.50	0.87	Protection is satisfactory
3.31	0.92	Protection is adequate
3.26	0.82	Protection is moderately good
3.23	0.85	Protection is about average
3.20	0.94	Protects about as well as most equipment of its type
3.10	0.89	Protection could be better in some ways
2.90	0.76	Protection is fair but could stand improvement
2.60	0.83	Protection is fair
2.58	1.11	Protection could be improved
2.41	0.99	Protection needs improving
2.37	1.17	Protection is not adequate under extreme conditions
1.94	0.93	Protection is not quite adequate
1.83	0.94	Protection is barely adequate under moderate conditions
1.81	1.06	Protection is barely adequate
1.78	0.92	Protection is not very satisfactory
1.78	0.71	Protection is below average
1.44	1.23	Protection is hardly noticeable
1.13	0.98	Protection is slightly better than nothing at all
0.92	0.83	Protection is much below average
0.87	0.80	Protection is poor
0.66	1.42	Protection is completely inadequate
0.21	0.70	Protection is so poor item serves no purpose

RUGGEDNESS

<u>Mean</u>	<u>Standard Deviation</u>	<u>Statement</u>
5.24	0.74	Superior for rough usage
5.09	0.70	Extremely rugged and well made
4.96	0.09	Exceptionally rugged
4.90	0.81	Excellent for rough usage
4.65	0.73	Very rugged
4.48	1.00	Durability is ideal
4.33	1.15	Unusually rugged
4.32	0.88	Durability is very satisfactory
4.14	0.92	Above average in ruggedness and durability
3.93	0.68	Very rugged in most respects
3.85	0.82	Quite rugged
3.71	0.85	Adequate durability for rough usage
3.59	0.81	Durability is quite satisfactory
3.25	0.98	As rugged as most equipment of this type
3.24	0.95	Quite rugged but needs some improvement
3.23	0.73	Durability is satisfactory
3.08	0.77	Moderately rugged
2.98	0.82	Durability could be improved
2.94	0.62	Average durability
2.66	1.05	Not quite rugged enough
2.37	0.97	Durability is not quite adequate
1.92	1.20	Shows excessive wear after moderate usage
1.80	0.78	Not very rugged
1.56	1.01	Won't stand up under rough usage
1.53	0.84	Below average durability
1.46	0.88	Durability is not very satisfactory
1.42	1.08	Durability is barely adequate
1.40	1.35	Easily damaged, shows excessive wear
1.21	1.07	Flimsy material and/or construction
1.15	1.12	Won't stand up under normal usage
0.63	0.77	Very poor durability
0.55	0.67	Poorly made, low durability

FIT

<u>Mean</u>	<u>Standard Deviation</u>	<u>Statement</u>
5.81	0.71	Fits perfectly in every respect
5.66	0.79	Fit is excellent
5.61	0.88	Fit is superior
5.49	0.92	Fits extremely well
5.18	1.08	Fit is ideal
4.89	0.79	Fit is very good
4.71	1.17	Fits unusually well
4.41	0.99	Fit is quite satisfactory
4.29	0.79	Fit is very good in most respects
4.25	0.83	Fits comfortably
4.00	0.83	Fit is above average
3.73	0.81	Fit is about average
3.71	0.90	Fit is satisfactory
3.44	0.93	Fit could be improved
3.40	0.86	Fit is adequate
3.27	0.77	Fit could be better in some ways
3.25	0.74	Fit is moderately good
3.07	0.62	Fit is fair
2.98	0.97	Fit needs improving
2.87	1.00	Fit needs some adjustment
2.35	0.86	Fit is not quite adequate
2.00	0.74	Fit is poor but item is wearable
1.94	0.74	Fit is below average
1.79	1.01	Fit is barely adequate
1.69	0.78	Fit is not very satisfactory
1.25	1.23	Fit is poor
1.19	0.91	Fit is much below average
1.11	0.94	Fit is so poor item can only be worn for short periods
0.96	0.90	Fit is so poor item can only be used under limited conditions
0.80	1.07	Fit is very poor
0.76	1.10	Fit is so poor item can't be worn comfortably
0.30	0.86	Fit is so poor item is unusable

C-II. MEAN AND STANDARD DEVIATION OF STATEMENTS FOR USE IN DEVELOPING OVERALL ACCEPTABILITY AND GENERAL RATING SCALES.

<u>Mean</u>	<u>Standard Deviation</u>	<u>Statement</u>
6.27	0.54	Excellent
6.22	0.86	Perfect in every respect
5.74	0.81	Extremely good
5.19	0.75	Very good
5.03	0.98	Unusually good
4.62	0.72	Very good in most respects
4.56	0.75	Above average
4.35	0.95	Quite satisfactory
4.25	0.90	Good
4.13	1.11	More than adequate
3.77	0.85	About average
3.69	0.87	Satisfactory
3.58	0.77	Moderately good
3.39	0.87	Adequate
3.28	1.09	Could use some minor changes
3.10	1.30	Not good enough for extreme conditions
2.72	1.15	Not good for rough use
2.11	0.76	Not very satisfactory
2.10	0.84	Barely adequate
2.10	0.85	Not very good
2.03	0.79	Below average
2.00	0.87	Unsatisfactory but usable
1.97	1.12	Needs major changes
1.83	0.98	Barely acceptable
1.79	0.90	Not adequate
1.76	1.21	Not good enough for general use
1.22	1.08	Better than nothing
1.06	1.11	Poor
0.76	0.95	Very poor
0.69	1.32	Very unsatisfactory
0.36	0.76	Extremely poor

APPENDIX D
SAMPLE ANCHORED INTERVAL SCALE
NET EVALUATION QUESTIONNAIRE

Course: Operator	(1)	1
		2
ORG Main		

Evaluation Period (Week) (Use 99 for Final)	(2)	(3)

Evaluated by: (Last Name, Initials)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	

Date _____

Instructions: Circle a number between the adjectives which best represent your opinion of the instruction you have received during this evaluation period.

A. Instructors:

- 1. Used jargon or confusing terms Never 1 2 3 4 5 6 7 8 9 Always
- 2. Speaking ability (enunciation, volume, etc.) Poor 1 2 3 4 5 6 7 8 9 Excellent
- 3. Subject knowledge Poor 1 2 3 4 5 6 7 8 9 Excellent
- 4. Treatment of students Discourteous 1 2 3 4 5 6 7 8 9 Courteous
- 5. Aware of student understanding of subject material Never 1 2 3 4 5 6 7 8 9 Always
- 6. Preparation of instruction Poor 1 2 3 4 5 6 7 8 9 Excellent
- 7. Response to student questions Poor 1 2 3 4 5 6 7 8 9 Excellent
- 8. Overall rating Unsatisfactory 1 2 3 4 5 6 7 8 9 Outstanding

B. Instruction:

- 1. Basic concepts were made clear at beginning of block of instruction Never 1 2 3 4 5 6 7 8 9 Always
- 2. Basic concepts were developed logically Never 1 2 3 4 5 6 7 8 9 Always

3. Presentation of material was Boring 1 2 3 4 5 6 7 8 9 Interesting
4. Classroom discussions were Waste of time 1 2 3 4 5 6 7 8 9 Valuable
5. Material was presented Too slowly 1 2 3 4 5 6 7 8 9 Too rapidly
6. Coverage of material was Too superficial 1 2 3 4 5 6 7 8 9 Too technical
7. Training aids were Poor 1 2 3 4 5 6 7 8 9 Excellent
8. Training aids were used Too seldom 1 2 3 4 5 6 7 8 9 Too often
9. Lectures/conferences led into practical exercises Never 1 2 3 4 5 6 7 8 9 Always
- C. Practical Exercises (PE):
1. Time scheduled for PE's was Inadequate 1 2 3 4 5 6 7 8 9 Adequate
2. PE's were conducted on actual hardware Never 1 2 3 4 5 6 7 8 9 Always
3. All students participated in PE's Never 1 2 3 4 5 6 7 8 9 Always
4. PE's were conducted as scheduled Never 1 2 3 4 5 6 7 8 9 Always
5. What percentage of the instruction time was "hands on" for you? 10-20-30-40-50-60-70-80-90
- D. Lesson Assignments and References:
1. Assignments were necessary Never 1 2 3 4 5 6 7 8 9 Always
2. Assignments were Too simple 1 2 3 4 5 6 7 8 9 Too difficult
3. Manuals and reference materials were Too elementary 1 2 3 4 5 6 7 8 9 Too difficult
4. Manuals and reference materials were designed for easy use Never 1 2 3 4 5 6 7 8 9 Always
- E. Examinations:
1. Material covered in exams was presented during instruction/PE Never 1 2 3 4 5 6 7 8 9 Always
2. Exams were Too short 1 2 3 4 5 6 7 8 9 Too long
3. Exams were Too simple 1 2 3 4 5 6 7 8 9 Too difficult

APPENDIX E

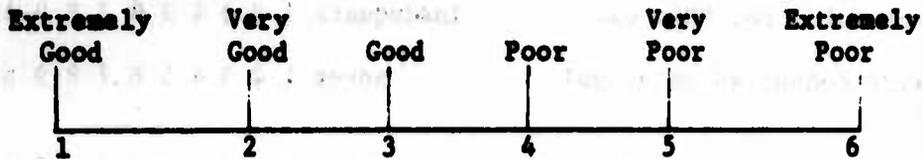
EXAMPLE OF QUESTIONNAIRE RATING INSTRUCTIONS

NIGHT VISION SIGHT
CREW SERVED WEAPONS
FIRING ACCURACY - ZERO AND BORESIGHTING (106mm RIFLE)

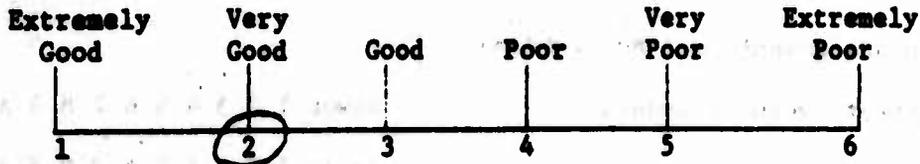
The Night Vision Sight Crew Served Weapons (AN/TVS-2B) (AN/TVS-5) are currently undergoing test. Since you have been involved in the testing and evaluation of this night vision sight, your opinions are extremely important. Therefore, please answer all of the questions as accurately and honestly as you can.

Many of the questions are of a rating scale type. Below is a sample item:

How would you rate the current Army pay scale?



Circle the number that best reflects your opinion of the Army pay scale. Do not make any marks between the numbers. Therefore if you rated the Army pay scale as "very good," your scale would look like this:



On all other questions, circle the appropriate letter or check the appropriate box.

All of these questions pertain to your use of the AN/TVS-5 and AN/TVS-2B on the 106mm rifle during the firing accuracy and zero and boresighting subtests. Therefore please answer the questions as they apply to these areas.

Name _____

Age _____

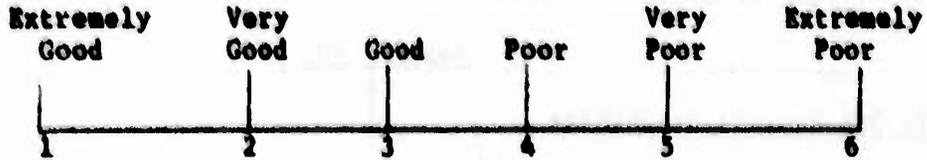
Rank _____

Unit _____

Time in Service _____

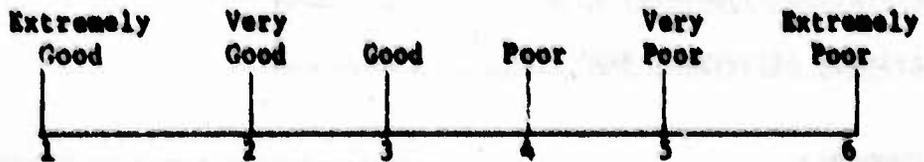
MOS _____

1. How would you rate the overall ease of using the AN/TVS-5?



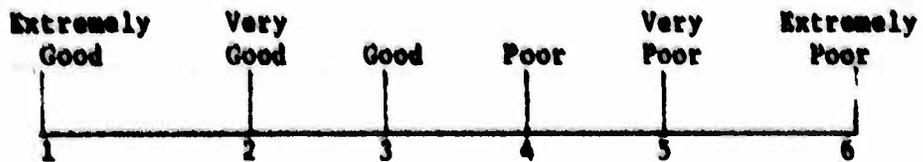
Comments _____

2. How would you rate the overall ease of using the AN/TVS-2B?



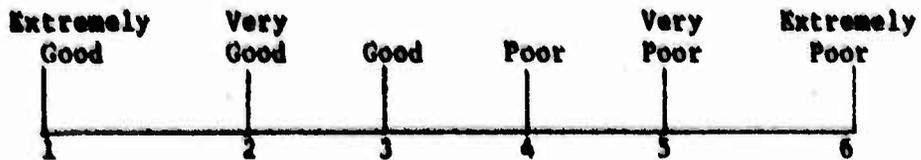
Comments _____

3. How would you rate the comfort of your body position while using the AN/TVS-5 mounted on the weapon?



Comments _____

4. How would you rate the comfort of your body position while using the AN/TVS-2B mounted on the weapon?



Comments _____

5. Are the controls on the AN/TVS-5 listed below easy to adjust while you are in your normal firing position?

	Yes	No
On/Off Reticle Brightness		
On/Off Tube Brightness		
Range Focusing Ring		
Diopter Ring		
Elevation Adjustment Knob		
Azimuth Adjustment Knob		

Comments _____

6. Are you able to utilize the AN/TVS-5 while wearing the M17A1 protective mask?

Yes

No

Comments _____

APPENDIX F

EXAMPLES OF SUMMARY TABLES

SUMMARY TABLE
 SUMMARY OF TEST PARTICIPANT'S RESPONSES REGARDING EASE OF DONNING AND DOFFING
 TEST OF FT KNOX V COMBAT BOOTS - ARMY AND USMC
 Phase II - Comparative Interview

Question:	Fort Benning-Army (211)				Camp Lejeune-USMC (164)				Combined (375)									
	Army-V		MC-V		Army-V		MC-V		Army-V		MC-V		N.D.					
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%				
Ease of Donning and Doffing?	2	1.0	187	88.6 ^a	22	10.4	15	9.1	101	61.6 ^a	48	29.3	17	4.5	288	76.8 ^a	70	18.7
Reasons:																		
Because of hooks (or eyelets)	1	0.5	168	79.6			4	2.4	91	55.5			1	0.3	259	69.1		
Easy to lace			13	6.2			12	7.3	7	4.3			4	1.1	20	5.3		
Smooth leather									2	1.2			12	3.2	2	0.5		
Boot strap			6	2.8											6	1.6		

NOTE: N.D. = No Difference
 a. Significant at 5% level

SUMMARY TABLE

USER'S RANKINGS OF MEAL-READY-TO-EAT MENUS^a

Meals Heated Using Delrin Fuel Tablet Without Grid

Menu No.	Mean Rating ^b	No. of Responses
1	7.3	38
3	7.6	31
4	7.7	40
2	7.7	30
12	7.8	32
13	7.8	38
7	7.8	39
9	7.9	29
6	8.0	35
8	8.0	32
14	8.1	33
10	8.1	37
5	8.2	26
11	8.2	31
15	8.2	38
	<u>7.9</u>	<u>509</u>

- a. Using Kramer's extension of DMRT for unequal sample sizes.
- b. Any two means not bracketed by the same line are significantly different, and any two means bracketed by the same line are not significantly different. Ratings based on 9-point hedonic rating scale.

APPENDIX G

EXAMPLES OF STATISTICAL ANALYSIS TECHNIQUES

The purpose of this appendix is to illustrate the application of some of the more commonly employed statistical techniques to the analysis of subjective test data. The types of analyses shown are only representative and are not intended to be exhaustive. In order to make the examples complete, summaries of the actual interview data obtained during a test of standard (std.) and experimental (exp.) armor vests are included for computational and reference purposes. A more detailed explanation of each of these techniques, together with the underlying rationale for their employment, are provided in the reference cited at paragraph H-2, appendix H.

t-test

A summary of distributions of test participant ratings for four characteristics is shown at table G-1. Twenty-two test participants wore both standard and experimental vests alternately for two weeks. At the end of each week, each participant was asked to rate both vests for each characteristic. The raw data for the second characteristic, comfort, are shown at table G-2, along with some tabular computations necessary for the following examples. Since the questionnaire employed an equal interval rating scale, parametric statistical procedures may be applied to the ratings if other assumptions appropriate for each method are met.

a. The first example involves a comparison of the means of the subjective comfort ratings for two vest types. The hypothesis to be tested is that there is no preference for one vest over the other. Since the test procedure involved each test participant rating both vests, it is appropriate to use the paired t-test to compare vest ratings. This procedure is outlined in section 9-4 of the reference cited at paragraph H-2, appendix H. From the difference column (table G-2) calculate:

$$N = 22$$

$$\sum d = 14$$

$$\sum d^2 = 32$$

then calculate

$$\bar{d} = 14/22 = .636$$

$$s_d^2 = \{32 - (14)^2/22\}/21 = 1.110$$

$$s_d = \sqrt{1.110} = 1.049$$

$$t = \frac{.636 - 0}{1.049 / \sqrt{22}} = 2.844$$

Reject if $t < -2.080$ or if $t > 2.080$ (where 2.080 is the critical t value for a 2-sided test with a significance level of .05 and 21 ($=N-1$) degrees of freedom). Since 2.844 is greater than 2.080, the hypothesis that comfort ratings during the first week are equal is rejected.

The same test procedure applied to the other characteristics summarized at table G-1 yields similar results, i.e., for all characteristics and for each week within a characteristic, the hypothesis of equality of preference is rejected at the .05 significance level.

b. The above analysis is appropriate for the actual test. It would often be true, however, that each individual wore only one type of vest. Under this test design, there is no basis for pairing, and a different t -test must be used, as outlined in section 9-3 of the reference cited at paragraph H-2, appendix H. In the example, the same data as above will be used; it will be assumed for illustration that each individual wore only one type of vest. The hypothesis to be tested is that there is no preference for one vest over the other. There is also a requirement on this test that the variability of ratings is equal for both vests.

For ratings of each vest (table G-2) calculate:

N	$\frac{EXP.}{22}$	$\frac{STD.}{22}$
$\sum x$	103	89
$\sum x^2$	503	381

Then calculate

\bar{x}	4.682	4.046
$\sum x^2 - (\sum x)^2/N$	20.773	20.955

$$s_p^2 = \frac{20.773 + 20.955}{22 + 22 - 2} = .9945$$

$$s_p = \sqrt{.9945} = .997$$

$$t = \frac{4.682 - 4.046}{.997 \sqrt{1/22 + 1/22}} = 2.116$$

and reject if $t < -2.018$ or if $t > 2.018$ (where 2.018 is the (interpolated) critical t value for a 2-sided test with a significance level of .05 and 42 ($N_1 + N_2 - 2$) degrees of freedom). Since 2.116 is greater than the critical value of 2.018, the hypothesis that the two vests are equally comfortable is rejected.

Analysis of Variance

The same comfort data also may be used to illustrate one of the many types of analysis of variance. In this example we ask whether there is a difference in preference for vests, averaged over both wear periods; whether there is a difference in comfort from week to week, averaged over both types of vests; and finally, whether differences in comfort vary independently with respect to vest type and period of wear. The first question, difference in preference for vests, is the same one answered by the previous t -test examples; the other questions are new.

Computational procedures are found in section 10-5 of the reference cited at paragraph H-2, appendix H, and are not repeated here. Only the final Analysis of Variance (ANOVA) table is shown.

ANOVA TABLE

SOURCE	SS	Df	M.S	F	Critical (5%)F*
Vest	10.22	1	10.22	13.63	3.96
Weeks	0.18	1	0.18	0.24	3.96
Vest X Weeks	0.05	1	0.05	0.07	3.96
Error	62.82	84	0.75		
Total	73.27	87			

*Interpolated from table A-7a of reference cited at paragraph H-2, appendix H.

The calculated F statistic for differences in comfort ratings between vest types (13.63) far exceeds the critical value (3.96) for a significance level .05 and with 1 and 84 degrees of freedom. Hence it is concluded, as before, that there is a real difference in preference between the two types of vests. There is no evidence whatsoever that there is a change in preference from week to week. Nor is there any evidence of any interaction between the two kinds of data classification, vest type, and period of wear.

Chi-square test

The results of daily interviews involving four questions are summarized at table G-3. The questions were phrased so that a simple yes - no answer was required. Visual inspection of the responses for all questions and each of the three wear periods shows a great deal of similarity in responses about both vests. Only for question 2 on the first wearing would the viewer even question whether the respondents felt the vests were not equal in the matter of restricting body movements.

The appropriate test to ascertain whether the proportion of yes answers is the same except for sampling variation is a contingency test using the chi-square statistic. This is discussed in section 13-3 and table 13-6 of the reference cited at paragraph H-2, appendix H. The data are set up as follows:

		Vest Type		
		STD	EXP	
Movement Restricted	Yes	9	5	14
	No	13	17	30
		22	22	44

and the test statistic is calculated by the formula in the reference cited at paragraph H-2, appendix H. The degrees of freedom for a contingency table are found by $(R-1)(C-1)$, where R and C are the number of rows and columns, respectively. For a 2 X 2 table, the degree of freedom is unity.

The test statistic, χ^2 , is 0.94; the critical value for the .05 significance level is 3.84. Accordingly, there is no reason to believe the vest types are different as regards restricting movement.

TABLE G-1
ARMOR VEST
SUMMARY OF INTERVIEWS ON FOUR CHARACTERISTICS

App	Characteristic	Rating	Numerical Score	1st Week Std	1st Week Exp	2nd Week Std	2nd Week Exp
D At 5% level of significance experimental vest rated better than standard vest on all characteristics for both weeks.	1. Freedom of Movement	Extremely good	(6)	0	1	0	1
		More than adequate	(5)	2	13	1	17
		Just adequate	(4)	19	7	17	3
		Not quite adequate	(3)	1	1	4	1
		Inadequate	(2)	0	0	0	0
		Extremely poor	(1)	0	0	0	0
		Average rating		4.0	4.6	3.9	4.8
	2. Comfort	Extremely comfortable	(6)	0	2	0	2
		Moderately comfortable	(5)	7	14	5	16
		Barely comfortable	(4)	12	5	15	2
		Slightly uncomfortable	(3)	1	0	1	2
		Moderately uncomfortable	(2)	1	0	1	0
		Extremely uncomfortable	(1)	1	1	0	0
		Average rating		4.0	4.7	4.1	4.8
	3. Fit	Extremely good	(6)	2	5	3	7
		More than adequate	(5)	15	17	17	13
		Just adequate	(4)	5	0	2	2
		Not quite adequate	(3)	0	0	0	0
		Inadequate	(2)	0	0	0	0
		Extremely poor	(1)	0	0	0	0
		Average rating		4.9	5.2	5.0	5.2
	4. Maneuverable efficiency	Extremely good	(6)	0	2	0	2
		More than adequate	(5)	4	11	4	16
		Just adequate	(4)	15	8	16	4
		Not quite adequate	(3)	3	1	2	0
		Inadequate	(2)	0	0	0	0
		Extremely poor	(1)	0	0	0	0
		Average rating		4.2	4.6	4.1	4.9

TABLE G-2

SUMMARY OF COMFORT RATINGS

Test Participant	First Week		Diff.
	Exp	Std	
1	5	4	1
2	5	4	1
3	5	4	1
4	1	1	0
5	6	4	2
6	5	4	1
7	4	5	-1
8	5	4	1
9	5	4	1
10	5	5	0
11	5	4	1
12	4	5	-1
13	5	4	1
14	5	5	0
15	5	4	1
16	4	5	-1
17	6	5	1
18	5	4	1
19	4	5	-1
20	5	2	3
21	5	3	2
22	4	4	0
SUM	103	89	14

TABLE G-3
ARMOR VEST
SUMMARY OF DAILY INTERVIEWS

Questions	Responses	1st Wearing		2nd Wearing		3rd Wearing	
		Std. (22)	Exp. (22)	Std. (22)	Exp. (19)	Std. (22)	Exp. (22)
1. Have you been able to satisfactorily perform all your duties?	Yes	22	20	18	18	20	22
	No	0	2	4	1	2	0
2. Were your body movements restricted in any way by the vest?	Yes	9	5	3	1	1	2
	No	13	17	19	18	21	20
3. Do you feel that the fit of this vest is adequate?	Yes	21	22	22	19	22	22
	No	1	0	0	0	0	0
4. Have you found that this vest interferes with the clothing or equipment you are wearing?	Yes	1	1	1	0	0	0
	No	21	21	21	19	22	22

APPENDIX H

REFERENCES

- H-1. Amerine, Maynard, et. al., Principles of Sensory Evaluation of Food, Academic Press, New York, 1965.
- H-2. Dixon, Wilfred J., and Frank J. Massey, Jr., Introduction to Statistical Analysis, McGraw-Hill Book Company, Inc., 1957.
- H-3. Edwards, Allen L., Techniques of Attitude Scale Construction, Appleton-Century-Crofts, Inc., New York, 1957.
- H-4. Guilford, J. P., Fundamental Statistics in Psychology and Education, McGraw-Hill Book Company, New York, 1950.
- H-5. Guilford, J. P., Psychometric Methods, McGraw-Hill Book Company, New York, 1954.
- H-6. Hembree, Howard W., Dimensionality of Soldier Acceptance: An Approach to Criterion Research, QM Contract DA-44-109-QM-129, University of Maryland, Technical Report No. 10, 1950.
- H-7. Natrella, M. G., Experimental Statistics, NBS Handbook No. 91, Bureau of Standards, 1963.
- H-8. Paul, Lee E., The Construction of Interval Scales for Measuring the Acceptability of Clothing and Equipment in Field Tests, QM R&E Field Evaluation Agency Technical Report R-4, 1960.
- H-9. Payne, Stanley L., The Art of Asking Questions, Princeton University Press, 1951.
- H-10. Siegel, Sidney, Nonparametric Statistics for the Behavioral Sciences, McGraw-Hill Book Company, New York, 1956.
- H-11. Winer, B. J., Statistical Principles in Experimental Design, McGraw-Hill Book Company, New York, 1971.

APPENDIX I

SELECTED READINGS

- I-1. Bradford, Ernest S., Marketing Research, McGraw-Hill Book Company, Inc., New York, 1951.
- I-2. Cantril, H., Gauging Public Opinion, Princeton University Press, 1947.
- I-3. Edwards, Allen A., Techniques of Attitude Scale Construction, Appleton-Century-Crofts, New York, 1957.
- I-4. Guilford, J. P., Psychometric Methods, McGraw-Hill Book Company, Inc., New York, 1954.
- I-5. Kahn, Robert L., and Charles F. Connell, The Dynamics of Interviewing, Wiley, New York, 1957.
- I-6. Oppenheim, A. N., Questionnaire Design and Attitude Measurement, Basic Books, Inc., New York, 1966.
- I-7. Osgood, Charles E., George J. Suci, and Percy H. Tannenbaum, The Measurement of Meaning, University of Illinois Press, Urbana, Illinois.
- I-8. Parton, Mildred, Surveys, Polls and Samples, Harper and Bros., New York, 1950.
- I-9. Payne, Stanley L., The Art of Asking Questions, Princeton University Press, 1951.
- I-10. Powell, Norman J., Anatomy of Public Opinion, Prentice Hall, Inc., New York, 1951.

APPENDIX J

LIST OF TABLES

	PAGE	
TABLE 3-1	IMPORTANT CHARACTERISTICS OF CLOTHING AND EQUIPMENT	3-3
TABLE 4-1	ADVANTAGES AND DISADVANTAGES OF USE OF THE FREE-ANSWER TYPE OF QUESTION	4-7
TABLE 4-2	ADVANTAGES AND DISADVANTAGES OF USE OF THE DICHOTOMOUS TYPE OF QUESTION	4-11
TABLE 4-3	ADVANTAGES AND DISADVANTAGES OF USE OF THE BASIC MULTIPLE CHOICE QUESTION	4-14
TABLE 4-4	ADVANTAGES AND DISADVANTAGES OF USE OF THE CHECKLIST QUESTION	4-17
TABLE 4-5	ADVANTAGES AND DISADVANTAGES OF USE OF THE RANKING METHODS	4-23
TABLE 4-6	EXAMPLES OF SIX-POINT RATING SCALES FOR RATING SPECIFIC CHARACTERISTICS	4-28
TABLE 4-7	EXAMPLES OF NINE-POINT AND SIX-POINT GENERAL RATING SCALES FOR RATING OVERALL ACCEPTABILITY AND UTILITY	4-29
TABLE 4-8	ADVANTAGES AND DISADVANTAGES OF USE OF THE RATING SCALE	4-32
TABLE 6-1	EXAMPLES OF CLASSIFIED AND CODED RESPONSES	6-7
TABLE 6-2	GROUPING OF RESPONSES	6-8
TABLE 8-1	STATISTICAL METHODS APPROPRIATE FOR ANALYSIS OF QUESTIONNAIRE RESULTS	8-5

INDEX

	PAGE
Ambiguity in wording	5-2
Analysis of data	8-3
Anchored rating scale	4-33
Argument question	4-4
Asking every question	10-3
Asking interview questions	10-2
Avoiding trick questions	5-7
Background of participants	3-4
Basic multiple choice question	4-12
Bias	5-3, 9-1, 9-2, 9-3
Leading questions	5-3, 5-5
Loaded words	5-4
Suggestive words	5-4
Bias and the interviewee	9-3
Bipolar rating scale	4-33
Care in use of free-answer	4-5
Categorizing responses	6-6
Checklist question	4-15
Examples	4-15
Use of	4-16
Circling wrong answers	10-7
Coding and classifying responses	6-6
Conducting the group interview	7-3
Considering the test group	3-2
Defining objectives, questionnaire	3-1
Descriptive statistics	8-4
Design of questionnaire	2-1
Dichotomous questions	4-8
Don't know category	4-10
Forcing a choice	4-9
Middle categories	4-9
Stating one or two alternatives	4-9
Double-barrelled questions	5-6
Effect of preceding questions	6-2
Equal interval scales	4-27
Establishing format	6-3
Establishing rapport	10-1

	PAGE
Factors determining type questionnaire	3-5
Factors in test participant selection	3-2
Follow-up questions	4-2
Forced choice	4-9
Format, questionnaire	6-3
Formatting the rating scale	4-30
Formulating the question	2-2
Formulating the questionnaire	6-1
Checking against objectives	6-5
Affect of preceding question	6-2
Establishing format	6-3
Frequency of administration	6-4
Heading	6-3
Length of questionnaire	6-4
Logical order	6-1
Punch cards	6-5
Question sequence	6-1
Free-answer question	4-1
Care in use	4-5
Examples	4-2, 4-3
Forms of	4-2, 4-3, 4-4, 4-5
Precoding	4-5
Getting specific answers	10-5
Group interview	11-1
Conduct of	11-3
Purpose	11-1
Quantification and analysis	11-5
Role of administrator	11-1
Selection of interviewers	11-4
Grouping responses	6-6
Halo effect	4-21
Heading, questionnaire	6-3
Hedonic rating scale	4-26
Indirect ranking	4-21
Information question	4-4
Instructions, interview	11-3
Instructions, respondent	5-7
Interview guide	3-5, 10-1
Interview questionnaire	3-5
Interview, types of	9-3
Interviewee bias	9-3
Interviewer bias	9-2

	PAGE
Interviewers, selection of	11-4
Interviewing	9-1
Purpose of	9-1
Sources of bias	9-1
Special applications	9-4
Types of interviews	9-3
 Keeping questions short	 5-7
 Leading questions	 5-5
Length of questionnaire	6-4
Listing information required	3-2
Loaded words or phrases	5-4
Location of interview	11-2
Logical order of questions	6-1
 Marking more than one answer	 10-7
Method of questionnaire design	2-1
Methods of ranking	4-20
Middle categories	4-9, 4-31
Multiple choice questions	4-12
Basic multiple choice	4-12
Examples	4-12
 Neutral category	 4-9, 4-31
Numerical ratings	4-20
 Objectives, questionnaire	 3-1
Obtaining the response	10-5
Omission of response	10-7
Opinion question	6-2
Orientation, interview	11-2
Overall supervision, interview	11-3
 Paired comparisons	 4-20
Precoding the response	4-5
Preliminary planning, questionnaire	3-1
Preliminary wording of questions	2-2
Pretesting the questionnaire	7-1
Method	7-1
Role of interviewer	7-2
Sample required	7-3
Timing of	7-3
Value of	7-1
Previous training and experience	3-4
Probing question	4-5
Problems in questionnaire design	1-2
Purpose of questionnaire	1-1

	PAGE
Quantification and analysis	8-1
Question form, selection of	4-1
Checklist question	4-15
Free-answer	4-1
Free-answer, forms of	4-2
Ranking question	4-18
Rating scale	4-24
Question sequence	6-1
Affect of preceding questions	6-2
Logical order	6-1
Opinion question	6-2
Question wording	5-1
Ambiguity	5-2
Bias	5-3
Instructions	5-7
Loaded words	5-4
Other considerations	5-5
Reminders	5-7
Vocabulary level	5-2
Questionnaire design	1-1, 2-1
Considering test group	3-2
Defining the objectives	3-1
Listing information required	3-2
Preliminary planning	2-1, 3-1
Problems in design	1-2
Purpose of questionnaire	1-1
Questionnaire format	6-3
Frequency of administration	6-4
Heading	6-3
Length	6-4
Questionnaire interview	9-3
Asking interview questions	10-2
Establishing rapport	10-1
Obtaining the response	10-5
Purpose of	9-3
Reporting responses	10-6
Questionnaire planning	2-1
Questionnaire, pretest of	7-1

	PAGE
Ranking question	4-18
Definition and use	4-19
Differences between items	4-20
Difficulties	4-21, 4-22
Examples	4-18
Methods	4-20
Numerical ratings	4-20
Paired comparisons	4-20
Rapport	10-1
Rating scales	4-24
Anchored scale	4-33
Determining number of steps	4-25
Examples	4-24, 4-25
Formatting scales	4-30
Neutral category	4-31
Selecting scale statements	4-30
Selecting the steps	4-26
Standardized scales	4-27
Reminders for wording questions	5-7
Reporting change in meaning	10-8
Reporting free-answer responses	10-9
Reporting the responses, interview	10-6
Response classification	6-6
Response coding	6-6
Review for completeness	11-4
Sample size	3-4, 7-2, 7-3
Sampling statistics	8-4
Selecting scale statements	4-30
Selecting the test group	3-2
Background	3-4
Previous training	3-4
Representativeness	3-4
Sample size	3-4
Selecting type of questionnaire	3-5
Factors determining	3-5
Interview questionnaire	3-5
Self-administered questionnaire	3-5
Selection of question form	2-2
Self-administered questionnaire	3-5
Semantic rating scales	4-26
Standardized rating scales	4-27
Stating one or two alternatives	4-9
Statistical analysis	8-3
Statistical procedures	4-19, 4-24, 4-27, 8-3
Suggestion question	4-4
Suggestive words	5-4

	PAGE
Tabulation of data	8-2
Terminal categories	4-31
Thoroughness of questions	5-6
Timing of pretest	7-3
Trick questions	5-7
Types of errors	10-7
Types of interviews	9-3
Value of pretest	7-1
Vocabulary level	5-2
Wording terminal categories	4-31

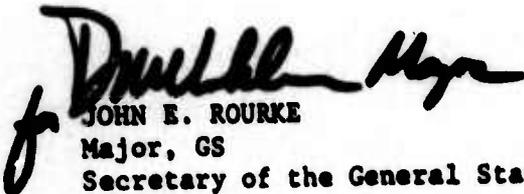
TECOM Pam 602-1, Vol I

The proponent office of this pamphlet is the Systems Analysis and Evaluation Directorate. Users are invited to send comments to the Commander, TECOM, ATTN: AMSTE-SY, Aberdeen Proving Ground, Maryland 21005.

FOR THE COMMANDER:

OFFICIAL:

WILLIAM H. TUCKER, JR.
Colonel, GS
Chief of Staff


JOHN E. ROURKE
Major, GS
Secretary of the General Staff

DISTRIBUTION:
A1 and B1