UNCLASSIFIED

AD 651 795

PSYCHOLOGICAL RESEARCH ON EQUIPMENT DESIGN

Paul M. Fitts

Army Air Forces
Washington, D.C.

1947

UNCLASSIFIED

Army Air Forces
Aviation Psychology Program
Research Reports

Psychological Research on Equipment Design

REPORT NO. 19

Edited by
PAUL M. FITTS
Chief, Psychology Branch
Aero-Medical Laboratory
Wright Field

1947

Preface

The research on equipment design reported in the present volume was carried out by the psychology staff of the Aero Medical Laboratory, Engineering Division, at Wright Field during the past year, and by members of the staff of the Department of Psychology, AAF School of Aviation Medicine during the past three years.

Chapters have been written by the individuals who carried out the investigations reported therein. Many aspects of the research program, however, cannot be credited to any one individual. Research plans, in particular, usually have been formulated through the joint efforts of many individuals.

The Psychology Branch at Wright Field owes its existence in large measure to the support of Col. W. R. Lovelace, who was at Wright Field, and to Cols. L. E. Griffis and J. C. Flanagan, who were in the Office of the Air Surgeon, Headquarters, Army Air Forces, when the branch was activated. After his assignment as Chief of the Research Division in the Air Surgeon's Office, Col. Otis O. Benson gave strong support to psychological research on equipment problems.

The research carried on in the Aero Medical Laboratory, and the preparation of the present volume, have been made possible by the support and guidance of Cols. W. R. Lovelace, L. E. Griffis, and E. J. Kendricks, each of whom has served as Chief of the Laboratory. Lt. Col. A. P. Gagge and Dr. J. W. Heim have been of invaluable assistance in initiating and carrying out research.

Many individuals within the AAF Aviation Psychology Program have contributed to the research on equipment design. Among these are Col. A. W. Melton and other psychologists at the School of Aviation Medicine who initiated equipment research in the AAF. Credit is due particularly to those men who transferred to Wright Field after other research units were closed at the end of the war, bringing with them a rich background of experience in human problems of aircrew selection and training.

Several individuals outside of the AAF contributed directly to the research reported in this volume. Prof. F. C. Bartlett of Cambridge University discussed equipment-design research with the editor on numerous occasions during the summer of 1945, and permitted study of research equipment and research techniques in his laboratory. Drs.

W. F. Hunter and C. W. Bray, and other psychologists who worked on equipment problems for the Applied Psychology Panel of N. D. R. C., did much to convince the AAF of the importance of psychological research on equipment. Lt. Col. R. V. Garrett and Maj. Richard Crane, both long-time advocates of standardization and simplification of the pilot cockpit, gave much encouragement to psychological study of aviation equipment. Dr. E. F. DuBois of the Committee on Aviation Medicine, and Mr. D. K. Morrison, who was in charge of one of the Committee's flight safety projects, stimulated a great deal of interest in the reduction of pilot error through functional cockpit design.

Recognition is due also to the Anthropology Unit of the Aero Medical Laboratory, which, through its excellent work on human body-size requirements, won wide acceptance of the fact that human considerations are important in airplane design and paved the way for the cordial reception that has been accorded other research workers in this field.

Miss Patricia Wall has supervised the typing of the report, Miss Mary Cowles has been in charge of the art work, Miss Sally Bedworth has prepared most of index, and Mrs. K. D. Young has proofread the manuscript. Dr. Glen Finch has read and criticized the report. Dr. W. F. Grether has been of constant assistance in planning the research carried out by the Psychology Branch during the past year and in preparing this report. This assistance is gratefully acknowledged.

Wright Field, 1 October 1946. PAUL M. FITTS

Contents

Chapter | Page

CHAPTER ONE

Introduction to Psychological Research on Equipment Design

PAUL M. FITTS

ENGINEERING PSYCHOLOGY AS A FIELD OF RESEARCH

The purpose of psychological research on equipment design, or engineering psychology as it may be called, is to collect data on the psychological capacities and limitations of individuals and to apply these data to engineering design problems to the end that the equipment which is finally produced will be adapted to the requirements of the man who is to use it. This is a relatively new field of psychological research. In the past most engineering effort has been devoted to improving equipment from the point of view of mechanical efficiency. Engineers have been aware of the desirability of designing equipment to meet the requirements of the human operator, but in most cases have lacked the scientific data necessary for accomplishing this aim. Psychologists, on the other hand, have centered their attention on the study of techniques for selecting and training men to use existing equipment rather than on the investigation of equipment design problems.

Experiences of AAF pilots furnish many striking examples of the need for designing equipment in relation to the psychological characteristics of the user. The following description by a pilot of the problems of making an instrument landing approach is an example in point: "Before you hit the cone of silence everything is strictly routine. Afterwards things really start popping. You start putting the landing gear down, changing prop pitch, dropping your flaps, making radio calls, and at the same time the most difficult thing is to fly the beam. You only have a few seconds from the cone of silence to the field, and you want to stay right on the beam. If you have a good copilot and you train him properly the two of you can handle it all. You only have to be a chauffeur and drive it down the beam. Nine times out of ten, though, you get your wheels and flaps down all right but wind up off the beam." This account is one of a series of recordings obtained during interviews with AAF pilots.

The following description, obtained in the same manner, illustrates further the need for simplifying the pilot's equipment. "Reaching back from all the flying experiences that I've had, I find that flying formation under combat conditions in instrument weather was really the most difficult. Not only was there mental strain and uncertainty because you didn't know whether you were going to get through to the target, but you didn't know whether there were other ships ahead of you, and you didn't know where the rest of the group or the rest of the wing was. You just knew where that ship on your wing was and that was all. And you knew positively that you were much safer to stay with him than to break away from him under any condition. You were determined before you got in the weather that you were going to stay right there, but when you did get in the overcast you would start experiencing all sorts of sensations due to the fact that you couldn't watch your instruments long enough to convince yourself you weren't doing acrobatics, that he wasn't in a steep turn, that you weren't losing altitude, and so forth. I found that it was a mental and physical strain all the time and it took a lot more out of me to fly a mission in which I went through even a few minutes of instrument flying in formation than it did to fly a much longer one under normal conditions."

Results achieved in the space of a few years by psychologists who have worked in the field of engineering psychology, and the growing demand on the part of engineers for the type of information which can be obtained only through psychological research, indicate the scope and importance of this field of specialization in psychology. Sufficient actual research has already been accomplished to show that improvements in operator efficiency which can be obtained from minor design changes frequently turn out to be much greater than improvements which can be obtained through months of intensive training or through careful screening of operators on the basis of aptitude. In other words, variations in operator efficiency resulting from design changes in equipment sometimes are far more important than variations in efficiency which are due to the aptitude or the level of training of operators. Almost all types of equipment, whether for military or civilian use, for business or recreation, can be improved through the application of psychological research techniques to problems of adapting the equipment to the user. However, psychological research on equipment design is especially important as regards those items of equipment that are difficult or dangerous to use, that can be operated only by carefully selected and trained men, or that place a premium on the final degree of proficiency attained by the operator.

Relation of Engineering Psychology to Other Fields of Psychology

Most research on equipment design belongs to the field of experimental psychology. The study of perceptual and motor-abilities problems demands the application of research methods that date back to early psychophysical experimentation in psychology. Problems in engineering psychology also involve the psychology of learning, of individual differences, and of the total reacting organism.

Research on equipment design overlaps to some extent the field of industrial psychology. The latter has in the past been concerned with such applied problems as the selection, training, and upgrading of workers for particular jobs, and with problems of efficiency which are defined by machinery now in use. Few industrial psychologists, however, have attempted to improve the basic machines and working tools used in industry.

Engineering psychology employs many research techniques used for personnel selection and training, especial procedures for measuring the aptitudes and abilities required for different jobs, and for studying the effectiveness of training methods. The same criteria of proficiency used in selection and training studies often can be used in equipment-design studies. In addition, personnel research often reveals characteristics of design which make it difficult for an individual to operate an item of equipment effectively, and thus identifies those aspects of the equipment that should be modified.

The principal difference between engineering psychology and other special fields of psychology is in point of view and final objectives. In most fields of applied psychology, and clinical psychology in particular, the interest is in changing the individual or in placing the individual in an environment or in a work situation where he can adapt successfully. Engineering psychology is concerned with adapting one important aspect of the environment, the machines of a technological society, to man's own requirements. Broadly conceived, the techniques of engineering psychology can be applied to many aspects of our present-day industrial civilization for the purpose of improving them in terms of human requirements.

PLACE OF PSYCHOLOGY IN THE ENGINEERING DEVELOPMENT PROGRAM OF THE AAF

Interest in psychological problems of equipment design developed rapidly during the recent war. The intense effort to produce new weapons, the race against time in industrial production, and the magnitude of the program required to train men to operate these new machines resulted inevitably in many instances in which the final man-machine combination failed to function effectively.

The earliest large-scale studies of equipment-design problems were made by the Applied Psychology Unit at Cambridge University, England. This unit, early in the war, began work on problems in the design of aviation and of armored-force equipment (1, 2).[1] Near the end of the war, the German Air Force also was beginning to conduct psychological research on equipment-design problems (3).

In the United States psychologists working for the National Defense Research Committee carried on some research on equipment-design problems during the war. The NDRC Applied Psychology Panel initially devoted most of its efforts to selection and training problems, but later gave much attention to problems of equipment design, especially the design of fire-control equipment (8). Division 7 of NDRC also worked on psychological problems in equipment design (5).

Aviation psychologists in the United States Navy and in the Army Air Forces became interested in equipment-design problems chiefly as an outgrowth of their work on selection and training of aircrew. In 1943 Dr. William McGehee, who had been working on pilot-training problems at the Naval Instrument Flying School in Atlanta, carried out an experimental study to determine the effect on flying proficiency of different arrangements of flight instruments (6). At about the same time the Department of Psychology of the AAF School of Aviation Medicine, with approval from Headquarters, AAF, initiated a series of studies on equipment-design problems. Much of the equipment research at this unit was carried out by Dr. Roger B. Loucks (see ch. 8), and by Dr. Joseph Weitz (see ch. 13). Dr. Arthur W. Melton, the Director, and other members of the Department of Psychology staff collaborated in planning the studies.

A number of other agencies, both civilian and military, have been active in equipment-design research during the past few years. These agencies include the Special Devices Division of the United States Navy, the Naval Medical Research Institute, and the Committee on Aviation Medicine of the National Research Council. Special mention also should be made of the work of Dr. S. S. Stevens and associates at the Harvard Psycho-Acoustic Laboratory (7).

On 29 May 1945, the Headquarters of the Army Air Forces issued a directive to the Air Matériel Command at Wright Field pointing out the need for establishment of a psychological-research facility at Wright Field to study equipment-design problems. As a result of this directive there was activated on 1 July 1945 a Psychology Branch of the Aero Medical Laboratory. This was the last unit within the AAF Aviation Psychology Program to be established during the war. The Psychology Branch is now part of the peacetime AAF Aviation Psychology Program and is the unit responsible for all aspects of engineering psychology for the Army Air Forces. A list of personnel assigned to the branch during its first year of operation is given in Appendix A of this report.

[1] Throughout this report, boldface numbers in parentheses are used to refer to the numbered references listed at the end of the chapter.

Organization of Psychological Facilities in the Air Matériel Command

The official mission of the Psychology Branch of the Aero Medical Laboratory is to conduct "psychological research to determine the capacities of individuals to operate new types of equipment as an aid in the designing of such equipment to the end that the final project will be best adapted to the man who must use it." This responsibility extends to contracts for psychological research on equipment design problems carried on outside of the command. Research in the Psychology Branch is conducted by a staff of professional psychologists assisted by psychological technicians, apparatus technicians, and statistical and clerical personnel. The staff is composed chiefly of civilians. A few rated flying personnel and a few enlisted men also are assigned to the staff. The predominance of civilian over military personnel is in keeping with the general organization of Wright Field, which in peacetime provides for a staff composed primarily of civilian engineers and research workers.

The Aero Medical Laboratory, of which the Psychology Branch is a part, is one of 15 laboratories making up the Engineering Division of the Air Matériel Command. The Aero Medical Laboratory was initially established in 1935. It is responsible for many aspects of human requirements in relation to equipment. Physiologists, biophysicists, anthropologists, biologists, physicians, nutrition experts, physicists, psychologists, and engineers in the several branches of the laboratory work on a coordinated program of biological research.

The Engineering Division in turn is responsible for development of all types of equipment peculiar to the needs of the Army Air Forces. This includes responsibility for development of new airplanes and equipment employed in aircraft or used in communicating with aircraft. The division is also responsible for development of special training equipment used by the AAF. The primary mission of the Engneering Division is equipment development. During peacetime, however, substantial emphasis is given to basic research.

It will be seen that engineering psychology in the Army Air Forces is centralized in the command that is responsible for all engineering development, and in the laboratory that houses the biological sciences. The Psychology Branch is also a part of the AAF Aviation Psychology program which is directed through a chief psychologist in the Office of the Air Surgeon in Headquarters, Army Air Forces.

The advantages of the present location of the Psychology Branch are numerous. The present centralized organization permits a desirable amount of specialization. Psychologists can work on problems that are common to many different types of equipment. For example, an individual can specialize in perceptual problems, in motor abilities problems, in fatigue problems, or in some other field of primary interest. A centralized organization prevents duplication of effort and lends itself to maximum utilization of limited research personnel through a jointly planned research program. Other advantages include close contact with engineers and with the laboratories that are working on all new engineering developments. It is also of advantage to be associated with other biological scientists and to be able to engage in cooperative research on problems that overlap several specialized fields.

Activities and Responsibilities of the Psychology Branch

Activities of the Psychology Branch fall into four areas: (1) research activities; (2) coordination of university research projects; (3) consultation with scientists and engineers from other Wright Field laboratories; and (4) liaison with outside agencies. Of these areas, research is considered most important and receives most attention. The research program in turn is concerned with problems of three levels of generality: practical problems that are specific to one or a few items of equipment; basic problems that are of importance for the design of many different items; and broad theoretical problems. In general, specific practical problems are not investigated unless the problem is of sufficient importance to justify the effort. The majority of research projects are designed to obtain answers to basic questions that are common to many engineering design problems. In the long run it is much more efficient to emphasize this type of research. A certain amount of time also is devoted to research that is of general theoretical interest and to the study of methodological problems. Regardless of the level of the problem, however, every effort is made to design experiments in such a way that findings will be of as wide application as possible.

Another activity of the Psychology Branch is consultation with engineers and scientists from other laboratories and from industry. Consultation services may involve interpretation of available information or giving of professional advice on design problems requiring an immediate answer. Staff members also participate in mock-up studies. Many questions that arise in connection with the planning of crew positions and layout of equipment and working space for these mock-ups are psychological in nature and can be answered on the basis of well-established psychological principles.

A further responsibility of the Psychology Branch is the coordination of research which is carried out by universities under contract with the Air Matériel Command. After a general problem has been assigned, it is the policy to allow universities a great amount of freedom in designing experiments and in carrying out research. However, the Psychology Branch offers assistance in defining problems and in supplying information regarding new engineering developments to be encountered in future aircraft. The branch also is responsible for maintaining liaison with other military and civilian organizations engaged in the psychological study of equipment-design problems as well as with the aircraft and equipment industry.

RESEARCH AREAS

Problems of equipment-design research can be classified in terms of types of equipment involved or in terms of the nature of the psychological questions which must be answered. Classification in terms of psychological problems is considered to be more meaningful. When this basis is used, problems fall logically into five areas: (1) design of display systems in relation to human perceptual abilities; (2) design of control systems in relation to human motor abilities; (3) determination of human limits in operating equipment; (4) study of the user's acceptance of his equipment; and (5) problems involving the design of complex systems of equipment. It is obvious that these various areas overlap to a considerable extent. The first two are treated in detail in the two following chapters.

Display Problems

The design of a display system involves the problem of presenting through instruments, necessary information that otherwise could not be perceived. Design of devices for providing this information is in many respects a psychological problem, involving selection of the sense modality to be utilized, selection of the specific cues to be provided the operator, and choice of a method of indication. The design of aircraft instruments, of radar scopes, of communication systems, and of warning devices are problems in this area.

An important equipment-design problem is the design of flight instruments that can be read quickly, yet accurately. The following account illustrates this particular problem: "An instructor had his ship and five cadets in AT-6's at 18,000 feet. He gave directions to break up the formation and descend to 10,000 feet, then practice three-turn spins and return to the field. This cadet didn't follow instructions. He started his spin at 18,000 feet, thinking he would practice a spin on the way down and break it at 10,000. Of course, spins don't break as fast at that altitude. The cadet thought the ship wasn't coming out of the spin. He had been instructed to bail out of a spinning ship when it got down to 3,000 feet, so he looked at his altimeter. It read 13,000 but he thought it read 3,000, so he bailed out. It took him a long time to get down to old mother earth. We lost one AT-6 on that deal." An account of another pilot's experience which did not turn out as well illustrates a similar difficulty with the same instrument: "A buddy of mine was flying at 1,500 feet. He looked up and there was a bunch of trees in front of him. The funny thing was, after it was all over and the plane had crashed, he felt and saw that his bones were all OK and said, 'Well, I guess I must have hit a tree.' What had happened was that he had misread his altimeter. He was actually flying at an indicated altitude of 500 feet, which was just the altitude of the terrain at that point."

Control Problems

Many different types of controls can be used with most machines. Problems for psychological research in this area include investigations to discover which design variables have an important influence on the effectiveness with which an operator can use a control. Psychologists also should determine the quantitative relationship between each important design variable and the effectiveness of over-all use, so that controls can be built especially for the kind of control operation required. Basic to the study of problems in this field is the study of human motor abilities.

Accounts of flight experiences frequently contain descriptions of mistakes in the use of controls which might have been prevented by better control design. The following experience of an AAF pilot is an example: "The mistake of which I am speaking was made on the way from Gander Lake, Newfoundland, to Marrakech, North Africa. Our B-29's were on their way over. We had been out from Gander Lake about 2 hours when we encountered fuel-pressure trouble in number-one engine. The gage was reading about 4 pounds per square inch and the motor was backfiring. I told the engineer to try to clear it out and bring the pressure up. In trying to do so in a hurry, he pulled off number-four engine, the wrong one. For a while we sat up there with just two engines while he was trying to get number four started again. Eventually we had to feather number one and go on into Marrakech on three engines. I believe that the reason why our engineer, who was a green man, made this mistake was because the engineer on the B-29 faces aft. In an emergency he got excited and pulled the engine control which, if he had been facing forward, would have been the number-one control."[3]

Study of Human Limitations

In the design of many items the question of the limits of human ability in operation of the equipment is of special importance. Regardless of how well an item of equipment is designed in relation to the human requirements of the operator, there are always finite limits to the speed and accuracy with which the operator can use it. Such limitations must be studied and defined in order to predict whether it will be possible for individuals to use proposed new items of equip-

[3] For an individual facing forward in an aircraft the engines are numbered from left to right. However, since the engineer, in this case, was riding backwards, the number-one engine actually was on his right. The controls, however, were arranged with the number-one control to his left and number four control to his right.

ment. The determination of human limitations is especially important in relation to the prevention of accidents. In order to be certain that equipment can be operated safely it is necessary to know not only the average performance of individuals who use it, but also what the extreme range of human variability in its operation will be.

Acceptability of Equipment

Studies of the acceptability of equipment involve such problems as pilot comfort, the confidence of the operator in his equipment, and the amount of pride and satisfaction involved in its use. Many cases are known of operators who have refused to use an item of equipment such as a crash helmet, a pair of goggles, or a new instrument because they did not like it, in spite of the fact that it was satisfactory from every other point of view.

It is important that the user have confidence in equipment, such as his blind-flying instruments, that is used during critical periods. Pride in his flying clothing, parachute, and other personal equipment is also desirable.

Comfort is a special problem in this field. Apart from any possible causal relationship between subjective feelings of discomfort and loss of efficiency, it is desirable that comfort be considered in its own right. Problems of comfort include the design of seats (to which anthropology has made a special contribution), the provision of adequate working space, of adequate lighting, of protection from undue noise and vibration, and the relief of monotony and boredom.

Equipment Systems Problems

Problems for study in this field include questions of the arrangement of controls and displays for motion and time economy in sequential use, and the arrangement of complex systems of equipment for efficient over-all operation. A special problem concerns the integration of systems of equipment to be used simultaneously or consecutively by a number of different operators.

Arrangement of equipment in the pilot cockpit recently has received much attention. In the past, equipment often has been located in the cockpit without particular regard to the over-all problems of pilot efficiency. Placement of the pilot seat and location and arrangement of the numerous instruments and controls in the cockpit so that the pilot can see out of the airplane, check his instruments, and operate controls without getting out of his seat, is a difficult problem. Because of their complexity, the study of these and other systems-engineering problems requires the use of special research methods and techniques.

Other equipment systems of importance, especially in new types of aircraft, are those used by the navigator, the radar operator, the bombardier, the gunner, and the radio operator. Development of remote-control systems for guiding airplanes and missiles from a distance is another systems-engineering problem which deserves psychological study.

The communication system between different members of an aircrew and between the ground and the airplane also presents many psychological problems. A combination of voice communication and radar equipment is coming into use for controlling airplanes from the ground. Control of aircraft traffic around an airport and control of aircraft from the ground during blind landings involves communication between the pilot, the ground controller, and often a number of special radar operators. This complicated system involves many psychological links and human reaction times. The entire system, as well as its components, presents a problem worthy of considerable psychological study.

Relation Between the Various Research Areas

The first two areas discussed above, display and control problems, lend themselves readily to experimental designs in which one variable at a time is studied or covariation of a few factors is allowed. The study of human limitations is in some respects a special case of the first two areas. Study of the acceptability of equipment to the user is clearly distinct from the first three areas in that the emphasis is upon a subjective rather than an objective criterion. The last area, that of systems problems, differs from the others chiefly as regards the complexity of the variables being studied. Systems-design problems do not lend themselves to systematic experimental evaluation of each variable in the system. In comparing different systems it is possible to arrive at an over-all quantitative determination of the efficiency of any two systems; however, the identification of the specific factors responsible for the superiority of one system over the other or the identification of features which, if modified, would lead to over-all improvement, requires the use of techniques such as those employed in motion and time analysis. In general, it is believed that the first two areas are the most important at this stage and should receive major attention.

METHODS AND TECHNIQUES

In the application of psychological research to problems in a new field, such as equipment design, it is to be expected that many problems of methodology will be encountered. Much of the work of the Psychology Branch during its first year of operation has involved consideration of techniques suitable for use in equipment-design research. Most of the following discussion concerns problems of methodology rather than the answers to these problems.

Methods for Clarifying Psychological Problems in Equipment Design

Initiation of a program of research requires consideration of problems to be studied and an evaluation of their relative importance.

The implications of different problems, the likelihood of obtaining psychological findings of wide application, and the techniques suitable for use in studying problems are factors to be considered.

Six different techniques have been employed in exploring equipment-design problems to the extent necessary to clarify the psychological questions involved. These techniques are the following: (1) analysis of experiences and opinions of operators who have used the equipment; (2) observation of operator behavior; (3) participation by psychologists in the use of equipment; (4) analysis of available records; (5) analysis of the opinions of engineers; and (6) review of reports of related research findings. Often several of these procedures are used. Each of these techniques will be discussed briefly.

1. Useful information often can be obtained from individuals who have been using existing types of equipment. Descriptions of personal experiences in operating equipment are valuable. An investigation is now under way, for example, in which a series of 100 pilots is being interviewed and descriptions obtained of specific experiences in taking off, flying on instruments, landing, using controls, and using instruments. These accounts of specific experiences are relatively free from most of the biases that influence statements of opinions. Following the collection and analysis of descriptions of concrete experiences, it is often worth while to investigate certain questions in greater detail through the use of questionnaires or other techniques for obtaining responses from large numbers of individuals.

2. Reports are sometimes received that operators are experiencing difficulty in using an item of equipment or in carrying out a particular operation. Often the exact cause of the difficulty is not known. In this case it may be desirable to observe or to obtain records of behavior on the job. In such preliminary observation no attempt is made to control the conditions of work or to introduce experimental variations in procedure.

3. A somewhat similar technique is involved when a trained psychologist learns to operate an item of equipment in order to become familiar with its use, and to observe his own difficulties as an operator.

4. In some cases records of difficulties or errors in the use of equipment are available and can be studied by the research worker. Accident reports are an example of records of this sort. An analysis of navigator logs by the Psychological Research Project (Navigator) served as a stimulus for the study of psychological factors in the design of air navigation plotters which is reported in chapter 5 of this volume.

5. Another valuable source of information in clarifying psychological aspects of equipment-design problems is the opinion of engineers. This source is particularly valuable in the case of items of equipment,

such as new-type radar sets, which have never been placed in operational use.

6. A review of the psychological literature sometimes results in the location of information of value in planning research on a new problem. In general, however, it has been found that little of the research reported in the literature gives pertinent answers to the questions asked by engineers.

Planning Experimental Procedures

Having clarified the question at hand, the next step is the selection of suitable experimental procedures for studying the problem. This requires decisions regarding (1) the level of behavior to be studied, (2) the conditions under which research is to be carried on, and (3) the experimental design to be employed.

Level of behavior studied.—A distinction can be made between different levels of behavior in terms of the degree of complexity of the response required of the individual and the complexity of the total situation employed by the experimenter. The lowest level of behavior which lends itself to equipment studies is that in which the experimenter is concerned primarily with perceptual or motor processes of the individual. Sensory or motor tasks often can be made quite specific and one variable studied at a time. The next higher level of behavior which lends itself to equipment design studies is that in which the individual is required to carry on some simple perceptual-motor activity. The study reported in chapter 17 of this report involving the use of a simple pursuit task furnishes an example of research at the second level of complexity. At a still more complex level it is possible to simulate in the laboratory the characteristics of the total job situation. Studies such as the one employing the Link Instrument Ground Trainer, reported in chapter 8 of this report, or the study involving simulation of the task of an aerial gunner, reported in chapter 18, are examples of research studies carried on at this level. The simulated job is a useful experimental situation provided the task can be standardized and suitable scoring devices developed. The final and most complex level of behavior that can be selected for study is the performance of actual tasks, for example, the study of pilot reactions in the air. Research conducted in the air is admittedly both difficult and expensive. However, it appears desirable at times to verify laboratory findings in the air. In general it is believed that flight testing should be employed as a final step after the completion of as much work as can be done efficiently in the laboratory.

Conditions of research.—Choice of the level of behavior to be studied often determines also the conditions under which research should be conducted. Experimental studies of equipment design can be carried on in a laboratory, in a specially equipped aircraft, or at a station where training or routine flying is in progress. The majority of the work of the Psychology Branch has been and will continue to be done in the laboratory. Flight testing is carried out on a limited scale. Field studies usually are limited to the collection of data needed to clarify psychological questions involved in equipment design. Occasionally, in addition, it is planned to utilize routine training activities or flight operations as a means of collecting data.

Number of subjects.—Equipment usually is designed for the average individual. From this point of view, therefore, it is important to obtain as representative a population as possible and thus to maximize the number of subjects. However, economy in the collection of data usually makes it desirable to secure considerable data on the same individual, once he has been scheduled for testing. In practice, experimental designs should call for a sufficient number of representative subjects to provide an adequate sample of the population and sufficient measures on each individual to provide the required amount of data.

In equipment-design studies individual subjects frequently can be used as their own controls. Where this is permissible it is much more efficient to use the same individuals under experimental and control conditions than to use different individuals.

Choice of experimental design.—In many equipment-design problems a large number of experimental variables can be thought of which it might be important to study. The use of experimental designs which make maximum use of small samples frequently is indicated as an initial step in selecting from this large number of variables those which are most important in determining operator efficiency. The variables selected in this manner then can be studied systematically with larger numbers of cases in order to work out more exact quantitative relationships between each variable and the criterion. The study of any selected group of variables demands also that the interaction between variables be investigated. Frequently it is desirable to determine which variables are subject to interaction effects and then to adopt an experimental design which permits systematic control or study of these interactions.

Choice of Appropriate Criteria

The problem of the criterion has been recognized in many fields of psychology. The choice of suitable criteria is an important step in planning equipment-design studies. The obvious criterion is proficiency in using the equipment in question. Choice of a criterion of proficiency, however, revolves around the purpose for which the equipment is used. In some cases, for example, speed of operation may be of major importance. In other cases precision or accuracy may be more important than speed. Special items of equipment place a premium on still other response characteristics of the operator. Lead-computing gun sights, for example, often require great smoothness of operation since irregularities in the rates of motion imparted to the sight are amplified by the computer.

In investigations where the acceptability of equipment is the chief interest special criteria are indicated. Since acceptability often is defined in terms of the subjective feelings of the individuals who are to use the equipment, techniques must be employed for measuring attitudes such as confidence in the equipment.

A criterion of considerable importance is the amount of learning time required to reach a satisfactory level of proficiency in equipment operation. Apart from the probable relationship between rate of learning and final level of proficiency, the amount of training required on a new item of equipment is an important practical consideration. For example, the design of a parachute opening device, or the operation of a fire-extinguisher button in the cockpit should, if possible, be so simple that the operator will be able to use it instantly after long periods of no practice. Economy in training also is important. For example, an over-all reduction in the amount of time required to learn to fly safely would be highly desirable, not only for the training of military personnel but for the training of civilian flyers.

Often it may be desirable to utilize several criteria. It is common practice, for example, to obtain both speed and error scores. Often research apparatus can be designed to provide simultaneously several different scores. If possible, criteria should be chosen that are subject to precise objective measurement and to the recording of quantitative total scores. In the final analysis, choice of criteria involves an evaluative judgment by the experimenter.

Development of Research Equipment

During the first year of its existence at Wright Field the Psychology Branch has directed a major portion of its work to the design and development of suitable research equipment.[3] Some problems have been studied without special equipment or with simple apparatus that was available.

Perhaps the most economical research medium is the printed test. The research projects reported in chapters 4, 6, and 11 of this volume employed printed testing materials. This technique is particularly well adapted to the study of perceptual problems.

The assembly or development of equipment for laboratory studies of perceptual and motor capacities is a relatively straightforward

[3] A number of individuals have contributed significantly to the development of research equipment used in the research studies reported in this volume. Special credit is due to J. Bakalas, J. F. Boory, J. R. Brick, M. M. Ducody, H. Muehlhauser, R. J. Roettele, and R. B. Smith.

problem. In many cases existing equipment or modifications of standard laboratory equipment can be used. Tests employed in the AAF aircrew selection program, for example, have been used in studying control problems.

The development of apparatus for simulating activities of a high level of complexity, such as instrument flying, however, presents a difficult problem. Many training devices, such as those for simulating on the ground performance of blind flying instruments and radar equipment, are already available. Few existing training devices, however, can be used in psychological research without considerable modification. Most synthetic trainers have no provision for quantitative scoring, and little attention has been given to the requirement of day to day stability.

It has been necessary, in several cases, to contract for the development and manufacture of special research equipment by outside engineering firms. Plans for the coming year call for the delivery of equipment for simulating instrument flying problems, and for measuring pilot behavior and aircraft performance in an air-borne laboratory.

Selection of Subjects

It is not possible at the present time to estimate accurately the extent to which research findings in the field of engineering psychology may be influenced by characteristics of the population from which the data are secured. The general aptitude level of the population employed as experimental subjects is probably an important factor. Pilots and other aircrew members are very highly selected as regards coordination, perceptual ability, mechanical aptitude, and similar characteristics that are important in learning to fly. It has been considered desirable for this reason to use only aircrew members or individuals of similar levels of aptitude as subjects in studies of aviation equipment.

Age, height, weight, visual acuity, and auditory acuity are other factors that may be of importance in selecting subjects. In all of these characteristics aircrew members represent a more homogeneous group than the general population. Thus far it has not been feasible to investigate the importance of each of these variables in a systematic manner, but such studies are planned.

Another question is the desirability of using trained subjects or of using inexperienced subjects. Where new equipment is to be used by individuals who have already had a large amount of specialized training with similar equipment, it is considered desirable to use as research subjects individuals with a similar level of training. In other instances it has been considered advisable to use inexperienced subjects. In studies such as the one reported in chapter 5 of this volume, in which the design of air navigation plotters was studied, high-school mathematics students were employed because it was considered advisable to use as subjects individuals with no previous specialized navigation training. An alternative procedure would have been to give a group of experienced navigators a considerable amount of practice in the use of the new plotters before making the experimental comparisons. In many cases it will be desirable to use both experienced and inexperienced individuals as subjects, in order to evaluate the design requirements of both groups.

RESEARCH OBJECTIVES

The objectives of research on equipment design can be summarized as follows: (1) identification of those variables in design which are most important in determining the ability of individuals to use equipment; (2) determination of quantitative functions defining the relationships between these design variables and operator efficiency; and (3) application of these findings to engineering design problems. Much careful research will be required in order to achieve the first two of these objectives. However, it is believed that the responsibilities of the research worker do not end with the completion of his research and writing of a research report. He has a further responsibility to see that his findings are reported in such a way that they can be used by the engineer, and that correct applications are made.

Identification of Significant Design Variables

As stated in the section on methodology, the psychological research worker first is confronted with the problem of clarifying the psychological problems involved in equipment design and in determining which design variables are important from the human point of view. The identification of these variables is the first objective of research. Characteristics of design that are of no particular importance to the human operator can be decided entirely upon the basis of engineering considerations. Variables that are found to be important for the human operator should be studied systematically so that when equipment is designed the pertinent data on human requirements will be available.

Defining the Relationships Between Specific Design Variables and Operator Effectiveness

Experimental psychologists will agree that much of the research in psychology has been concerned with establishing qualitative differences. The most commonly used statistics in psychological research are those employed in testing the null hypothesis. In the field of equipment design, however, it is usually not sufficient to be able to say that one design is preferable to another or that a statistically significant difference exists between two alternative designs. Most engineering problems require the expression of psychological finding in quantitative terms and the determination of functional relationships over a wide range of conditions. For example, it is not enough to know that the legibility of instrument dials is influenced by the size of the dial; it is necessary to know how accurately dials of various sizes can be read and to express quantitatively the speed and accuracy of reading for a continuous range of dial sizes. The engineer may want to know how small a particular dial can be made and a man still be able to discriminate a given number of differences in the position of the indicator hand. In this example, as in many other cases, the design of equipment frequently is a compromise between many conflicting demands, and quantitative data over a wide range of conditions are required if the most intelligent compromise is to be reached.

A further objective of psychological research on equipment problems is determination of the extent to which research findings can be generalized. It is important not only to determine quantitative relationships but to specify precisely the types of equipment or situations to which these relationships apply.

Application of Research Findings to Engineering Design Problems

Psychological findings can be applied at different stages in the design and production of equipment. The most immediate application comes in the modification of equipment that has already been built. Application at this level, however, not only is expensive but frequently requires retraining of individuals who have already learned to use the older equipment. Therefore, it is not feasible to apply psychological findings at this stage unless sufficiently important results will be achieved by the modification.

Application of psychological research can be made at the final assembly or mock-up stage where various finished items of equipment are being assembled into a complete system. Psychological principles relating to the layout and arrangement of complex systems of equipment can be applied with great benefit at this stage.

Psychological findings can be applied to the redesign of prototypes of specific equipment items. Frequently manufacturers make a few models of a new item and subject these prototypes to preliminary tests before mass production is started. During this stage it is possible to change the equipment as a result of psychological findings. However, only limited changes are possible since the basic design of the equipment has already been fixed.

Undoubtedly the most important point of application of psychological data is during the initial designing of new items of equipment. At this stage the engineer has much greater freedom to adopt designs that will meet psychological requirements, and the human factor can be given its rightful consideration. Therefore, the primary objective of psychological research on equipment design should be to provide information which can be used by the engineers who are engaged in the initial design of equipment. This objective often requires that psychologists anticipate problems that will arise in the development of new equipment and that they have the answers to these problems in time to influence the initial design.

An even more basic application of psychological findings is in determining what new types of equipment should be developed. Psychological data could be applied, for example, in determining the kind of information required by an aircraft pilot in carrying out specified operations. Psychologically one of the primary problems of the pilot is to maintain his orientation in three-dimensional space. Psychologists should be able to determine the minimum amount of information that a pilot must have in order to remain oriented constantly, and to determine also the simplest methods of displaying this information so that he can fly efficiently without any outside visual reference. Knowledge of human abilities and limitations is also needed in deciding what equipment should be operated by the pilot and what equipment should be made entirely automatic.

In the practical application of research findings to design problems a question frequently arises as to the relative importance that should be attached to different criteria. The following are examples of such questions: Should displays be designed for use during normal or during emergency flight conditions? What relative importance should be attached to speed and what importance to accuracy data? The answers to such questions more often involve judgment of relative values than an application of existing information. For a further discussion of objectives of psychological research on aviation equipment design the reader is referred to a previous article by the writer (4).

ORGANIZATION OF THE PRESENT VOLUME

It will be apparent to the reader that the present report introduces a program of research in a new field of psychology. The chapters which follow contain discussions of research problems and objectives and reports of 17 separate research projects. Fifteen of the chapters reporting specific research projects were prepared by present or past staff members of the Psychology Branch at the Aero Medical Laboratory and two by former staff members of the Department of Psychology at the AAF School of Aviation Medicine.

Chapters 2 and 3 contain comprehensive discussions of research problems and suggestions for programs of research in the two most important areas of equipment design. These chapters cover perceptual problems and human motor-abilities problems in relation to design of equipment. The research reports which make up the major portion of the volume will give the best introduction of all to problems of research on equipment design. Some of these reports are preliminary in nature and represent the results of an initial attack on a new problem. It is believed that all of the research studies, although oriented in terms of more or less practical problems of equipment design, make a contribution to the methodology and body of scientific knowledge of psychology.

REFERENCES

1. Bartlett, F. C. Instrument controls and display—efficient human manipulation. Flying Personnel Research Committee, British Air Ministry, *Report No. 565*, December 1943.
2. Craik, K. J. W. The fatigue apparatus (Cambridge cockpit). Flying Personnel Research Committee, British Air Ministry, *Report No. 119*, March 1940.
3. Fitts, P. M. German applied psychology during World War II. *Amer. Psychol.*, 1, 1946, 151–161.
4. Fitts, P. M. Psychological requirements in aviation equipment design. *Journal of Aviation Medicine*, 17, 1946, 270–275.
5. Franklin Institute. Final report on tracking studies. *O.S.R.D. Report No. 330-1706-232*, 15 November 1945.
6. McGehee, W. Comparative study of pilot fatigue resulting from extended instrument flight using the standard AAF and British instrument panels. *Final Report, Project TED No. ATL-R601*, Bureau of Medicine and Surgery, U. S. Navy Department (Restricted).
7. Stevens, S. S. Machines cannot fight alone. *American Scientist*, 34, 1946, 387–400.
8. Wolfle, Dael (ed.). Summary technical report of Applied Psychology Panel, NDRC, volume II. *Training and equipment*, Joint Research and Development Board, Washington, D. C.: 1946. (Restricted.)

BLANK PAGE

CHAPTER TWO

Survey of Display Problems in the Design of Aviation Equipment

WALTER F. GRETHER

INTRODUCTION

In the field of equipment design the term "display" has come into common usage as meaning any method of providing information which cannot be obtained directly through the sense organs. In flying, particularly under instrument conditions, very little of the essential information about the attitude, performance, and location of the airplane can be directly perceived, and that which can be perceived is often erroneous or inaccurate. The kinesthetic sensations, for example, are notoriously misleading as indicators of the attitude in space, and must be deliberately ignored by the pilot in favor of visual displays of the airplane's attitude on the flight instruments. Information regarding the direction of flight, air speed, condition of the engines, fuel supply, and many other things is likewise registered on visual indicators. Displays are by no means limited to the visual sense, however, since hearing and, to a lesser extent, the kinesthetic and tactual senses are also employed to transmit information to the aviator.

There are few, if any, situations where the human being is provided with such a variety of displays to which the appropriate reactions must be made as quickly and accurately as in the airplane. Likewise, there are probably no other common situations where failure to react correctly to the displayed information can lead to such serious consequences. Thus, the achievement of maximum efficiency in methods of display in aviation is a goal which should need no elaboration.

The actual display problems in aviation are as numerous as the devices and instruments used to convey information to human beings in any way working with aircraft, either in the air or on the ground. In this chapter, however, the discussion of problems will be limited primarily to the equipment used in actual flight or on the ground to control aircraft in flight. Such equipment may be grouped into several major categories: (1) instruments of all types used to present primarily quantitative data; (2) visual and auditory warning devices; (3) radar and television scopes; (4) gun and bomb sights; (5) tables, graphs, check lists, computers, plotters, and maps; (6) identification marks and operating instructions on controls and miscellaneous equipment; (7) radio navigation aids, radio voice communication and interphone systems; (8) signal flares, airway beacons, and runway lights; (9) miscellaneous cues, such as flight-control pressures, control-knob shapes, engine noise, and vibration.

Not all of the multitude of equipment items covered by this listing can be considered as problems for research in equipment design. In many cases excellent and successful display methods are already in use. In other cases, the nature of the equipment is such that even a crude method of display accomplishes the purpose. There are many types of equipment, on the other hand, about which there is considerable disagreement over the best method of presenting information, or concerning which records of operator errors indicate the need for improvement. It is toward basic problems in the design of equipment of this latter type that psychological research is being directed. In the psychological study of aviation display methods, with the aim toward their improvement, it is helpful to group the existing problems in terms of psychological research areas rather than types of equipment. The ensuing portions of this chapter are therefore classified into sections on sensory discrimination, attention-getting value, and similar psychological categories.

The discussion which follows is intended not as a complete cataloging of the psychological research problems in the field of aviation displays. It is intended, rather, to illustrate the nature of problems which exist and to point out the kinds of research investigations which should be most fruitful. Although no attempt is made to include a comprehensive review of the research literature which applies to the problems mentioned, occasional references are included to studies which are particularly pertinent.

SENSORY DISCRIMINATION

All types of displays must, of course, involve sensory discrimination in some form, but in most instances the stimuli to be differentiated greatly exceed the threshold requirements. Where difficulties exist, they are usually in the interpretation of the stimuli rather than their differentiation. Nevertheless, there are a considerable number of situations where sensory discrimination constitutes the basic problem.

Sensory Adaptation

During night operations it is usually necessary to maintain the best possible visibility of outside objects while, at the same time, preserving the effectiveness of visual displays within the airplane. This is the familiar problem of dark adaptation for which the major governing principles are already well known and for which solutions already are available in terms of cockpit lighting. Adaptation of the retinal rods can be preserved best by using dimly illuminated red markings in the cockpit. This can be done with red floodlighting, red indirect lighting, or with ultraviolet light projected upon reddish fluorescent markings. Although such systems are successful in preserving dark adaptation, their use brings in other problems of maintaining objects within the cockpit sufficiently above the threshold to be adequately visible.

A somewhat different problem of visual adaptation arises in the use of radar and television scopes on which the maximum brightness of the image is limited. In some situations the operator is required to shift his fixation from brightly illuminated objects to a dimly lit scope, and vice versa. Very often this shift must also be accompanied by a radical change in accommodation of the lens of the eye. The resulting eye strain, loss of time, and added chance for erroneous scope readings is obviously undesirable.

Color, Brightness, and Pattern Discrimination

Numerous situations exist in which discriminations of color, brightness, or pattern differences play a major role. Early in the war, for example, it was found that at a distance the insignia on American planes were easily confused with those on Japanese planes. This problem was solved by adding a bar to the American star and circle. Similar problems arise wherever identification marks or signals must be noticed and correctly perceived at great distances. Maps, control knobs, cargo parachutes, navigation lights, signal flares, and warning lights are other examples of the use of brightness, color, and pattern differences to provide essential information. To meet the problem of achieving maximum visual differentiation in such equipment, the large amount of available visual data can often be applied successfully, although actual field tests are desirable as final proof of the successful application of known general principles.

A somewhat related visual discrimination problem arises in the identification of targets or other data on radar and television scopes. Because of static and other imperfections in the scope image, the critical pip or blip may be so near the visual threshold that it will pass unnoticed, particularly if the operator is not maximally alert and prepared for the stimulus. In this case, the possible methods of increasing the discriminability in the image would seem to be (1) improvements in image definition, (2) selection of optimum image brightness and color, (3) selection of optimum scanning rates, and (4) selection of optimum scope size.

Auditory Discrimination

Auditory discrimination plays an important role in the use of some types of equipment, particularly in radio and interphone systems. The understanding of speech, code, or radio-beam signals is often difficult because of static, imperfect functioning of the equipment, and the high noise level within aircraft. Improvement of auditory discrimination can be brought about in a number of ways, such as (1) training in methods of speaking and in choice of words (2); (2) shielding of head sets; (3) increase in fidelity of sound reproduction; (4) use of certain kinds of electronic distortion (6); and (5) selection of code and radio range signals of maximum discriminability. The evaluation of such modifications in the equipment or operating techniques requires actual performance comparisons by means of standard psychological techniques.

Tactual and Kinesthetic Discrimination

Tactual and kinesthetic discriminations are involved in several types of aviation equipment. Pressure on the flight controls is used by the pilot as a cue in flying the airplane. Likewise, the shapes and locations of control knobs aid the pilot, radar operator, and bombardier in nonvisual identification of controls. Because of the close relation to motor performance, such discrimination is considered as a phase of control problems discussed in chapter 3. Some research in this area is reported in chapters 13, 14, 15, and 16.

ATTENTION-GETTING VALUE

Although closely related to sensory thresholds, the attention-getting value of aviation displays constitutes a fairly distinct area of research. The greater the extent to which a stimulus exceeds the sensory threshold, the more likely it is to be noticed. However, there are factors besides mere extent above threshold which determine the attention-getting value of a stimulus.

Attention-getting value is of primary concern in the design of various warning devices, although it is often difficult to distinguish between a purely warning device and an instrument. Readings on most instruments, if outside normal limits, constitute warnings of impending danger. For this reason the practice has arisen of placing colored marks on instrument faces to indicate tolerance ranges for several operating conditions. Other displays, however, serve purely as warnings of emergency conditions, and usually use lights or sounds as stimuli.

Although considerable data are available on the stimulus factors which determine attention-getting value, additional studies are required in simulated flight situations. The most helpful research studies in this field would seem to be the following: (1) relative attention-getting value of various changes in visual stimulation, presented with degree of contrast, area, and distance from the major fixation point equalized; (2) effectiveness of visual warning stimulation as a function of location of the stimulus with respect to the area of major fixation for the best types of stimulus change, other factors being held constant; (3) effectiveness of warning stimulation as a function of color differences for the best types of stimulus change, other factors being held constant; (4) relative attention-getting value of several changes in auditory stimulation presented to subjects wearing headphones (but not presented through headphones) under conditions simulating noise levels occurring during flight; (5) relative effectiveness of warning stimuli in several sensory modalities; (6) effectiveness of warning devices as a function of the number of similar devices present; (7) methods of differentiating among various types of warnings and indicating the response to be made; (8) comparison of tolerance marks on instrument dials with other types of warning signals.

LEGIBILITY

Legibility, like attention-getting value, is in many respects merely a special problem in sensory thresholds. But since it represents a rather distinct area of psychological research, namely, the speed and accuracy of reading printed materials, it is given separate treatment. A large number of research data on legibility are available in psychological literature (7). Some additional data are required, however, for the particular problems of legibility which are unique to aviation. Some of the more significant problems in this area for psychological research are as follows: (1) print-size requirements for a variety of specific aviation conditions; (2) style of letters and digits which are differentiated most easily and are least likely to be confused when rotated from the normal position; (3) speed and accuracy of dial readings as a function of dial size; (4) speed and accuracy of dial readings as a function of spacing of dial graduations; (5) speed and accuracy of dial readings as a function of shape and size of pointer and shape and size of scale graduations. For some preliminary research data on this problem, see chapter 7.

INTERPRETABILITY

Even though a display may be excellently designed with respect to sensory differentiation, attention-getting value, and legibility of the print and scale markings, the operator may still fail to react quickly and in a manner that is appropriate to the displayed information. This factor in the design of displays, for lack of a better term, has been labeled "interpretability." Of all the difficulties encountered in the designing of adequate aviation displays, this is probably the most serious and the most elusive. The psychological principles which govern interpretability appear to be relatively unknown, and this would seem to be one of the most challenging areas of psychological research in equipment design. As an aid to the discussion of problems in this area, pictures of some of the instruments to be discussed are shown in figure 2.1.

Graduated Scales

An interpretability problem which is common to many displays is the design of graduated scales used on instruments, computers, and other devices. There are many ways in which scales can be varied, among them (1) graduation intervals (usually 1, 2, or 5 units), (2) differentiation among graduations (usually by length or width of line), (3) intervals between numerals, and (4) method of indicating shift in graduation values on nonlinear scales. Suitable research on these variables should provide general principles which can be applied in designing a dial for any specific purpose.

A unique scale design problem has been noted in the Weems navigation plotter, where the protractor scale increases from right to left instead of in the conventional direction. For a more detailed discussion of this problem and research applied to it see chapter 5.

Linear vs. Circular Instrument Scales

The scales on aircraft instruments are almost universally circular. It has been suggested that many types of quantitative values would be more easily interpreted on linear scales in either the vertical or horizontal plane, depending on the type of data being presented. There has apparently been no research comparing the relative ease of interpreting circular and linear scales on aircraft instruments.

Dial vs. Counter Type Instruments

Another suggestion which has been made frequently for improvement of aircraft instruments is that the dial and pointer be replaced with a counter type of indicator from which the numerical value can be read directly. Although this should eliminate many errors, the reading of actual numbers may in some cases require more time than the mere checking of the position of a pointer on a dial. The writer is not aware of any research showing the relative merits of these two types of display as aviation instruments.

Graphs vs. Tables

Many items of information for use in cruise control, navigation, and bombing are supplied in the form of graphs or tables. Many errors are known to occur in the use of these, and the extraction of the necessary data may require excessive time. There are doubtless many improvements possible in the manner in which graphs and charts are made up. Furthermore, there is little known about the

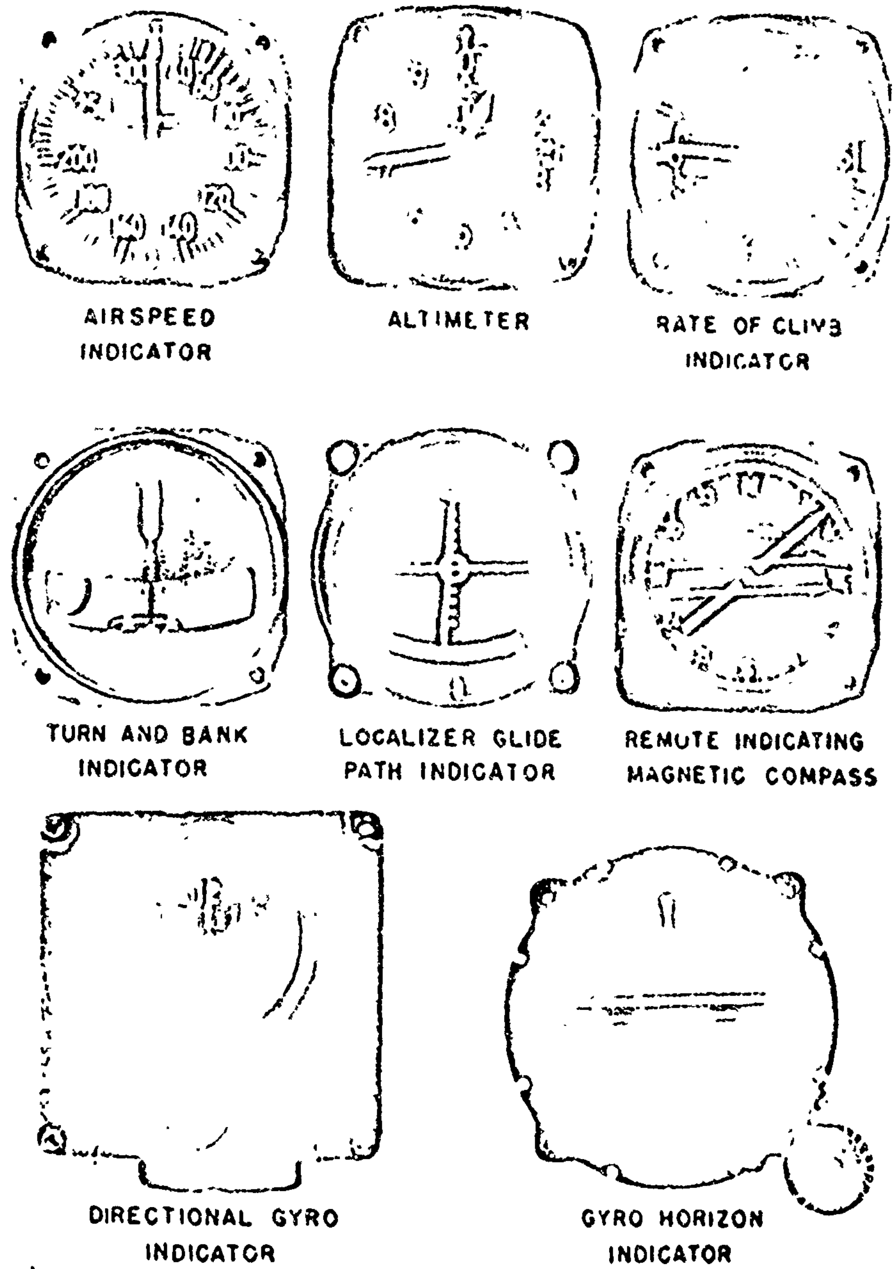


FIGURE 2.1.—Front view of eight standard aircraft flight instruments.

BLANK PAGE

relative merits of graphs versus tables for specific purposes. For further discussion of this problem and some experimental data, see chapter 4.

Labels

There are many items of aviation equipment on which there are printed labels and rudimentary operating instructions. In some cases the equipment is for emergency use only, and it is highly important that the instructions be brief and easily understood. The choice of words in such displays is very critical. Although it might be difficult to derive general principles to cover such labeling problems, psychological research could doubtless make some contributions.

Special Instrument Problems

Many instruments used in aircraft present unique problems of interpretation. In some of these instruments the primary problem is in the direction of movement of the moving element. These will be discussed in the following section on "Relation of Indicator to Control Movement." One instrument which presents a unique problem is the clock or watch used by the aviator. In military usage, time is expressed on a 24-hour basis, without a. m. and p. m. to differentiate between before and after noon. Thus, 3:47 p. m. becomes 1547 in military time. The reading of military time from a conventional clock thus involves a mental addition process. But clocks with 24-hour dials, designed for reading directly in military time, produce interference with firmly established clock-reading habits. For further discussion and research data on this particular problem, the reader is refered to chapter 6.

Another instrument which presents an interpretation problem is the altimeter (see fig. 2.1), which covers the altitude range in 20-foot steps. This is usually done with three pointers on a single scale. The first pointer indicates feet of altitude in hundreds, the second in thousands, and the third in ten-thousands. The synthesis of these three pointer readings is very confusing to the novice. Moreover, the altimeter includes another scale for setting in barometric pressure, which is expressed in inches of mercury rather than pressure altitude in feet. Some simplified method of indicating altitude and barometric pressure would be desirable.

Combining of Instruments

The whole array of instruments which the pilot faces is in itself a serious problem. Under some flight conditions it is virtually impossible to read and synthesize all the information displayed. To make the problem more pressing, new instruments are continually being developed which are added to the instrument panel without others being removed. To solve this dilemma it would seem desirable to determine which instrumental indications are most necessary and then to find the most simple, natural, and direct methods of displaying this information in the optimum area of the instrument panel. The instruments of lesser importance, if they cannot be eliminated entirely, might then be located in less conspicuous positions where they can be read if necessary.

There are a number of current trends toward the combining of displays into single instruments. This would seem desirable provided the indications are simple and natural. There is one danger, however, in carrying this trend too far. It is claimed that continued fixation on one instrument often induces a hypnotic state which may seriously impair the pilot's general alertness.

Radar Scopes

Problems of interpretation also arise in the use of radar scopes. The image on the scope has only slight resemblance to the actual landscape, or photograph of it, which is represents. The relative brightness of objects is often reversed, details are less clear, the proportions may be distorted, and the three planes of space may be differently presented. In some applications it is necessary to present three-dimensional space on a two-dimensional surface. All three dimensions can be represented at present with two separate scope images, but this introduces possible confusion of range with either azimuth or elevation. One type of radar scope presentation (B-scope) causes gross distortion of the terrain presented. For a report on the relative interpretability of this type of scope, see chapter 11.

RELATION OF INDICATOR TO CONTROL MOVEMENT

Simple Quantitative Instruments

A considerable proportion of the instruments used in aircraft respond directly or indirectly to control manipulations. Some of these, such as temperature gages, are unrelated to the attitudinal movements of the plane. It would seem desirable for the direction of movement on such instruments to have the most direct and natural relation to the movement of the control which affects it. Research is needed to determine what are the most natural or habitual relationships between control and indicator movements. Additional variables to be included in such research would be (1) the spatial location and plane of movement of the control with reference to the instrument; (2) the form of the control (whether an unstructured or pointed knob, a switch, or a lever); and (3) the meaning of the graduations on the instrument. Some experimental data on this problem are presented in chapters 9, 10 and 17.

Flight Instruments

In the so-called "flight instruments" there is an additional major variable in this direction-of-movement problem. That is the movement of the airplane which the display indicates. Here it is important to achieve the most natural combination of indicator movement and control movement in relation to the response of the airplane. This problem can be illustrated by reference to a number of current instruments (see fig. 2.1). The pointer on the rate-of-climb indicator moves down when the plane is nosed downward by pushing the stick or control wheel forward, and vice versa. The pointer on the turn-and-bank indicator moves to the left when the plane is turned to the left by application of pressure on the left rudder. In these examples the indicator and the airplane can be said to move in the direction in which the controls are displaced. Other instruments, however, move in the opposite direction. On the artificial (gyro) horizon, the horizon bar moves up (with reference to the instrument panel) when the plane is nosed down and rotates to the right when the plane banks to the left. The pilot director indicator, used in bombing, when displaced to the right, signals to the pilot that he is to turn to the right. But as he turns to the right, the indicator moves to the left. On the cross pointer, or localizer glide-path indicator, used for blind landings, the pointers indicate the direction of the correct flight path from the airplane. Thus, for example, as the pilot noses the plane downward the pointer moves upward, in a direction opposite to the control being made. It would seem desirable, where possible, to eliminate these apparent inconsistencies in direction of indicator movements. Before this is done, however, research must show which movement relationships are most satisfactory. In two studies of the artificial horizon a reversal of the existing movement relationships was found to be superior (3). See also chapter 8.

Frame of Reference

The problem of direction of movement is not actually as simple as the preceding discussion would imply. There are often special circumstances to be considered. On the pilot's magnetic compass and the directional gyro indicator, the cylindrical card bearing the scale and degree markings moves toward the same side as the plane is turning. But this is confusing since the true compass directions are displaced 180° on the card, as is necessary for viewing the card from the back. On the conventional compass and directional gyro, furthermore, it is possible to consider the lubber line as being the moving element rather than the card. On the artificial horizon the small reference plane, rather than the horizon bar, must be considered as being the moving element if correct sensing is to be achieved. Actually, in terms of the earth below, the lubber line on the compass and the reference airplane on the artificial horizon are the moving elements. In the visual bomb sight an image of the terrestrial target is moved under a pair of cross hairs which remain fixed in the visual field. Some bombardiers consider themselves to be moving the cross hairs, others to be moving the target.

The direction of movement problem is thus complicated considerably by the operator's frame of reference while performing the control task. To specify the optimum direction of movement in any given display will thus require an understanding of what the operator uses as his frame of reference. The Gestalt principles governing figure-ground relationships should be tested for application to this problem.

DIRECTIONAL ORIENTATION

The use of compass directions is basic to all forms of navigation, whether by the navigator, pilot, or radar operator. The bombardier must also use compass directions in identifying his target and controlling the bombing run. In the air most of the usual cues for maintaining directional orientation are lacking, and complete reliance must be placed on instruments. There is considerable uncertainty about the instrument designs for most effective presentation of directional information.

There is, for example, disagreement regarding the optimum design of the remote-indicating type of magnetic compass, where directions are presented on a dial. There are a number of possible arrangements: (1) the dial may be fixed, with North at the top; (2) there may be a fixed lubber line at the top, with the dial rotating behind it; (3) the dial may be adjustable by the pilot so that he can set his compass course at the top, with the pointer indicating his heading with reference to the dial; (4) the dial may be fixed, with one pointer used to indicate the desired course and another to indicate the actual heading (as in fig. 2.1). No doubt still other arrangements are possible. Actual performance measurements, with different methods of indicating direction, will be necessary to determine the relative merits of the various possible designs.

Similar directional orientation problems arise with equipment used by the navigator, bombardier, and radar operator. Some air-borne radar equipment is provided with azimuth stabilization which keeps North at the top of the scope image, regardless of the direction of flight. The same type of directional orientation of the visual field may be possible in periscopic bomb and gun sights. There are probably some situations in aviation where such azimuth stabilization is helpful, but there has been no adequate determination of which situations these might be. Probably the direction in which the operator faces with reference to the direction of flight is an important variable in this problem.

INFLUENCE OF ENVIRONMENTAL CONDITIONS

The aviator is exposed to a variety of atypical environmental conditions. Some of these, such as extremes of illumination and vibration, have a direct influence on the effectiveness of a display. Other environmental influences, such as lowered oxygen pressure at altitude, do not influence the display itself but may reduce the ability of the operator to use a display. This loss may be either in ability to interpret the display or in ability to perform the appropriate control movements. It is with these environmental effects on the operator, rather than on the display, that this section is concerned. It is important to identify and measure these effects of the environment so that suitable corrective steps may be undertaken. In the past, most of the work in this field has been carried out by physiologists, probably because the physiological effects of the environment are more easily identified and measured. During the war, however, a number of psychologists cooperated with physiologists in such studies, particularly on anoxia.

Anoxia

The deleterious effect on human efficiency of reduced oxygen pressure with increase of altitude has long been recognized. For moderate altitudes this problem has been solved by the use of oxygen masks and pressurized cabins. The prevention of anoxia at extreme altitudes to be flown in future planes is a problem not yet solved. One of the earliest and most serious effects of anoxia is known to be reduced effectiveness of visual discrimination and visual perception.

Acceleration (G)

As the accelerative or G forces encountered in maneuvering aircraft are increased, resulting in disturbance of the normal distribution of blood to the brain and other parts of the body, there is known to be first a narrowing of the visual field, then complete loss of vision, then loss of hearing, and finally loss of consciousness. Effects are somewhat different for positive G (head to foot) and for negative G (foot to head). Human tolerance to G can be increased somewhat by the G suit and by voluntary muscular contractions, both of which resist the flow of blood away from the upper part of the body. It is not known what, if any, losses occur in sensory and interpretive processes at G levels below those necessary to produce narrowing of the visual field. Research is needed to show whether or not such losses occur and whether there is a selective effect on different perceptual processes, so that equipment or tactics can be modified if necessary. Results also may have important implications for an understanding of cerebral functions. For some experimental results in this area see chapter 20.

Fatigue

The piloting of aircraft on long flights, while causing subjective feelings of fatigue, is believed also to cause an increased tendency toward pilot error. One of the most difficult operations in flying, landing under instrument conditions, is often required of the pilot at the end of a long flight when his proficiency is probably lowest.

Fatigue, whether in aviation or other situations, is a condition which is poorly understood. In the case of physical exertion the cause of fatigue can be reasonably well identified. The causes of the fatigue experienced by the pilot are much less specific. Probably the high degree of concentration required by the task, inability to move around, the inherent danger, the noise, vibration, and extremes of temperature, and the monotony are all contributing factors. The effects of long flights on the pilot's efficiency in performing his duties have likewise not been determined. The psychological effects of fatigue are very elusive when subjected to experimental study. When a fatigued subject is placed in a test situation he seems to be able temporarily to counteract the effects of the fatigue on the performance being measured. Thus, the objective record fails to reveal any loss of efficiency. Two British investigators, who carried out research on fatigue during the war (4, 5), suggested that the actual losses caused by fatigue are not so much in ability to perform the primary task as in more subtle aspects of behavior; namely: (1) a lowering of standards for performance of the task; (2) a tendency to respond to single instruments rather than the situation as displayed by the entire instrument panel; (3) failure to attend to instruments not related to the primary task; and (4) increase in irritability and tendency to blame difficulties on the test appartaus. This picture of fatigue is interesting and suggestive but should be subjected to further experimental tests. Adequate understanding of fatigue is necessary for achievement of maximum safety in long flights. Such knowledge should aid in designing equipment such as warning devices or blind-landing equipment which is much better suited to the fatigued pilot. Moreover, if the major causes of fatigue could be identified, it should be possible to reduce or eliminate them.

Emotional Stress

Stressful situations arise frequently in the air because of malfunctioning of equipment, human errors, or weather. It is at such times that the reactions of the human beings to their equipment become most critical. A number of psychologists have asserted that emotional stress results in reversion to more primitive reaction patterns. In aviation situations this would mean lapse of most recently learned skills in favor of older or more rudimentary forms of reaction. Thus an instrument which can be interpreted correctly only after considerable training may be reacted to incorrectly during stress. Conversely, an instrument which displays the information in the most simple, direct, and natural manner would be more likely to elicit the appropriate reaction. The cross pointer (see section on "Relation of Indicator to Control Movement") which is extremely difficult to use (1) provides an illustration of this principle. As the airplane nears the end of the runway the instrument becomes more and more sensitive. Many naive pilots, performing this increasingly difficult task, have been observed suddenly to reverse their interpretation of the instrument in the course of a last-minute desperate effort to bring the pointers back to center.

Other Environmental Factors

A number of environmental influences probably act upon aircrew personnel in ways which influence their reactions to displayed information. Those factors deserving experimental study would seem to be (1) temperature, (2) noise, (3) vibration, and (4) clothing and other encumbering equipment such as goggles, oxygen masks, pressure suits, and survival equipment worn on the body.

REFERENCES

1. Anonymous. *Report on the pilot and instrument approach system evaluation tests conducted at Indianapolis.* Washington, D. C., Army Air Forces, Headquarters, Air Transport Command, July 1946.
2. Black, J. W. Final report in summary of work on voice communication. Applied Psychology Panel, NDRC Project SC-67, *OSRD Report No. 5568*, 1945.
3. Browne, R. C. Comparative trial of two attitude indicators. Flying Personnel Research Committee, Great Britain, *Report No. 611*, 1945.
4. Davis, D. R. Experiments on mental fatigue on the Silloth Trainer. Flying Personnel Research Committee, Great Britain, *Report No. 486*, 1942.
5. Drew, G. C. Experimental study of mental fatigue. Flying Personnel Research Committee, Great Britain, *Report No. 227*, 1940.
6. Gross, N. B. and Licklider, J. C. R. *The effects of tilting and clipping upon the intelligibility of speech.* Office of Research and Inventions, United States Navy, Contract N5-ori-76, Psycho-Acoustic Laboratory, Harvard University, 1946.
7. Paterson, D. G. and Tinker, M. A. *How to make type readable.* New York: Harpers, 1940.

BLANK PAGE

CHAPTER THREE

An Analysis of Human Motor Abilities Related to the Design of Equipment and a Suggested Program of Research

JUDSON S. BROWN AND WILLIAM O. JENKINS

INTRODUCTION

The study of human motor abilities received a definite impetus during the war because designers of aircraft equipment, gunnery equipment, radar equipment, and similar complex devices discovered that the efficient operation of such equipment depends in part upon the degree to which the design of the device is adapted to the motor capacities of the operator. As a result it has become increasingly evident that there is a definite need for basic information regarding the capacities of human beings to make positional, rotary, translatory, and rhythmical movements of members of the body.

Although psychologists and physiologists have made extensive and elaborate studies of the sensory and perceptual capacities of human beings, relatively little time and effort have been expended in studying motor abilities. Thousands of pages have been devoted to vision, audition, gustation, olfaction, and touch, but few to fundamental motor capacity. A survey of current psychological texts and reference works reveals a somewhat surprising paucity of relevant material. A few pages are devoted to studies of the acquisition of complex motor skills, such as typewriting, and to the transfer of such skills from one member of the body to the other, but little is said regarding the accuracy with which an individual can position or move his limbs.

In the present chapter, an attempt has been made to classify motor reactions into several fairly distinct types and to suggest experimental procedures and significant variables in each case. In general, the scope of the treatment is restricted to the study of movements of the articulate members of the body. No space is devoted to complex motor activities, such as walking, running, and swimming, which involve more-or-less gross movements of the entire body, since such activities are, for the most part, of little significance for the design of equipment. Furthermore, the complex problems associated with the learning of skilled movements are not treated herein. In studying the acquisition of skilled movements, the customary procedure is to hold the task constant while varying the number of trials, the degree to which trials are massed, the amount and kind of interpolated activity, and other factors calculated to affect the degree and rate of learning. Here, however, the primary concern is not with the course of learning on a single task but with the basic initial abilities of individuals to perform a variety of tasks. Research of this type can, of course, lead directly into extensive studies of learning where it may be desirable to discover not only which movement or series of movements is most efficient initially, but also which leads to the highest level of proficiency after extensive practice.

It has been customary, in thinking about the psychological principles of equipment design, to divide the problems into two major groups: those related to design and operation of control mechanisms and those related to methods of displaying information to the operator. The research outlined here bears directly upon the problems of operating controls, since information about the capabilities of the operator is in most cases a necessary prerequisite to the design of a psychologically adequate control device. It is believed that the accumulation of accurate data on motor abilities will solve a great many of the most general control problems. Many questions about specific pieces of equipment cannot, of course, be answered without additional detailed research.

It is to be hoped that the material presented here will form the basis for a systematic series of integrated studies of basic motor abilities. In the concluding section, most of the existing publications on motor abilities that were available to the writers have been summarized.

TYPES OF MOTOR REACTIONS

Human motor reactions can be separated for convenience into three fairly distinct classes. These are (1) static reactions, (2) positioning reactions, and (3) movement reactions.

The term *static reactions* includes all of the instances where a bodily member is held for a time in a fixed position in space, the maintenance of that position being the central task imposed on the individual concerned. Although static reactions constitute, from the point of view of equipment design, the least important class of motor reaction, there is a definite need for sound empirical data on individual differences in static motor ability, on the relation of such ability to other more complex adjustments, and on the degree to which static reactions are affected by the position of the limb, fatigue, knowledge of results, and numerous other conditions. The two characteristics of static reactions which appear to deserve attention are (1) the relatively minute, high-frequency tremor movements and (2) the large, slow changes in static position.

Positioning reactions are those in which the members of the body are moved from a position of rest to a specified position in space. Here the terminal accuracy of the positioning movement is of primary significance. The ability to perform such reactions has never been carefully investigated, and it is believed that a systematic study of their variability and of their sensitivity to changes in experimental conditions would provide valuable information about human motor ability. The study of positioning reactions may be regarded as an investigation of motor localization—an analogue on the motor side of the well-known psychological studies of auditory and tactual localization. In experiments designed to investigate positioning reactions, the subjects might be required to move their limbs to a point straight ahead, to a point 90° to the left or right, or up or down, etc. Or the subjects could be instructed to bisect a spatial interval by moving a limb to a point halfway between two limiting stops, to draw a line of specified length, and so forth. Since the ability to make accurate positioning reactions seems to depend upon the ability to discriminate direction and extent (in motor, not perceptual terms), the experimental procedures reported in a later section have been divided accordingly.

Movement reactions are simply movements of the bodily members at given rates, in given directions, and along specific pathways or courses. Movement reactions are the most important, the most numerous, and the most complex of the motor reactions being considered. The number of possible studies of movement reactions is almost unlimited, since extensive variations may be introduced in the manner in which the stimulus is presented, in the kind of responses required, in the mechanism (if any) which is manipulated, and in the method of measuring the reactions. In the more detailed outline of movement reactions given later a distinction has been drawn between discrete movements, repetitive movements, serial movements, and continuous movements.

In the present report, each of the three types of movement is treated separately, and the suggested research program covers only the study of each in isolation from the others. The problem of the interrelations of the abilities involved in the three types of movement would constitute a desirable extension of the present program. A few references pertaining to this area are summarized in the concluding section.

GENERAL EXPERIMENTAL VARIABLES

A survey of the factors which can be varied experimentally in studies of motor ability reveals several which apply rather generally to all types of motor reactions. These *general variables* are the following:

1. Bodily member employed.
2. Limb position: This includes the position of the reacting limb with respect to the body; the positions of other limbs; the degree of flexion of the joints; the point of limb support; and the position of the extremities of the limb with respect to the more proximal joints, particularly the degree of supination, pronation, flexion, or extension of the wrist and hand.
3. Knowledge of results: This includes the manner in which the subject is informed of his success; the amount of information given him; and the particular sense modalities furnishing the motor-corrective cues, whether proprioceptive, tactual, visual, or a combination of these.
4. Environmental factors: Included here are such items as temperature, vibration and jolting, anoxia, drugs, acceleration, and deceleration.
5. Work output: This includes the degree to which the reacting member is artificially restricted by loading and/or frictional resistance.
6. Activity level: Included here are the amount, duration, and type of preceding, simultaneous, and/or interspersed motor activity of the reacting limb and other members, including the head and trunk. Fatigue, frequency of reacting, number of practice trials, and cooperative use of two limbs are subsumed under this heading.

STUDIES OF STATIC REACTIONS

As has been suggested above, the study of static reactions involves (1) the measurement of minute tremors, and (2) the measurement of gross drifts of the limb from null position. The experimental procedures for testing such reactions appear to be relatively simple. Verbal instructions, supplemented perhaps by a visually presented target, would serve to acquaint the subject with the experimental situation and the desired reaction. He would then initiate and maintain the response until notified that the trial was ended. The characteristics of the response would be measured by any of a number of graphic or photographic methods. In the case of positioning and movement reactions, however, the experimental procedures for defining the reaction to the subject and the methods of measuring his responses are much more complex. These will be discussed in detail in the appropriate sections.

From a practical standpoint, the study of tremor may be important for those motor operations in which extremely fine adjustments must be made of delicate mechanical parts or assemblies having insufficient inherent inertia to damp out the fundamental unsteadiness of the adjusting limb. Experiments in this category would be concerned with the factors which affect the amplitude, frequency, and plane of oscillation of limb tremors. As a secondary objective it would be desirable to study a sufficient number of subjects so that normative data on steadiness could be provided for each of a number of experimental conditions.

If the apparatus for recording tremor were suitably constructed, it would be possible to obtain simultaneous records of both tremor and the extent to which the limb gradually drifts away from its initial position. The study of slow changes in the basic limb position, especially as they are affected by the position or movement of other limbs and the head, is of practical importance in connection with the operation of high-speed aircraft at altitudes below about 100 feet. In such a situation, an unintentional forward or sideward movement of the stick may result in a sudden and disastrous loss of altitude. No studies of these slow changes in limb position have been made, so far as the writers are aware.

The general experimental variables listed above apply here. However, the use of vision in providing sensory cues applies to tremor only, since its use in the case of slow changes would eliminate the phenomenon being studied.

STUDIES OF POSITIONING REACTIONS

A preliminary inquiry into the psychological factors involved in the performance of positioning reactions indicates that the ability to discriminate the direction of a motor reaction, and the ability to discriminate its extent, are of primary importance. In studying such discriminatory abilities, the use of vision would, for the most part, be precluded, since its use in either case would reduce errors to an insignificant level. Because both directional and extentional judgments are involved simultaneously in the majority of positioning reactions, the relative accuracies of the two judgments can be determined only by the use of experimental techniques which isolate one from the other. Accuracy in performing directional reactions, uninfluenced by cues arising from the discrimination of extent, can be studied by holding extent constant and measuring errors in a plane perpendicular to the line of movement followed by the limb in arriving at the terminal position. In studying motor discrimination of extent, the effect of secondary cues due to direction can be eliminated by restricting the movement to a single dimension, errors being measured along the line connecting the starting and terminal points. If desirable, the over-all accuracy of positioning reactions, involving both directional and extentional cues, could be determined by measuring errors in three dimensions from the target position.

Data obtained in studies of discrimination of direction would have considerable practical application in the design of complex equipment. Information regarding the variability of such reactions as a function of direction and remoteness could be used in determining where control handles or switches should be located with respect to the operator and how widely they should be separated to prevent errors. Similarly, studies of the ability to discriminate extent would enable the equipment designer to achieve a better match between the range of excursion of controls or their degree of sensitivity and the capacities of the operator.

Discrimination of Direction

In experimental studies included under this heading, measurements would be made of the ability of an individual to move his limbs accurately to a specified terminal position. The location of that position would be varied in azimuth and in elevation, errors being measured radially about the target in two dimensions. The remoteness of the terminal position from the subject would be held constant for a particular group of directional reactions, but its effect upon such reactions could be studied systematically.

Experimental procedures for defining the standard reactions, along with relevant methods of measuring the subjects' responses are listed later. Also listed are those experimental variables which, in addition to the general variables given previously, apply specifically to the discrimination of direction. The question of which variables to combine with which procedures can, in all probability, be answered most efficiently by exploratory experiments.

Experimental procedures.—The desired reaction could be defined to the subject in purely verbal terms by telling him to point straight ahead, directly to the right, etc. His ability to translate such verbal instructions into movements could then be estimated by the method of production [1] or by the method of recognition.[2]

The terminal position might be defined by visual means by showing a subject the target at which he is to point. If, during the time he is shown the target, he is allowed to practice by moving his limb out to

[1] In the method of production, as used here and in succeeding pages, the subject makes a response following a verbal or visual definition of the desired reaction, without having previously made that same response in the experimental situation. Thus, he emits or produces the response without benefit of proprioceptive cues from a previous reaction.

[2] In the method of recognition, the subject makes a choice from amongst a group of stimuli or movements, one of which is the same as that presented originally. The choice would never be made by the use of vision alone, but would always involve either an active or passive motor reaction.

the target, either the method of reproduction[3] or the method of recognition could be employed in testing his ability. If the subject is not allowed to practice, the testing procedure is limited to the method of production.

The terminal position could be defined by passive movements of the subject's limbs to the desired position, with either the methods of recognition or reproduction being subsequently employed.

Active positioning movements of the subject himself might also serve to define a terminal position. Thus he could be instructed to move his limb out in any direction, note the point achieved, and then either reproduce or recognize the position. In general, this method would probably be unsatisfactory since the standard reaction is generated by the subject and is not, therefore, under the control of the experimenter.

No especially elaborate apparatus would be needed for the majority of studies of discrimination of direction. Paper targets could be mounted in various locations with the subject's reactions indicated automatically by a marker held in his hand (see ch. 15). The labor of record reading could be eliminated by having the subject hold a stylus to which a thread is attached. The thread would run through a minute hole in the center of the target and would have a small weight at the other end. The position of the weight on a vertically mounted scale would indicate the magnitude of the radial error, but not its direction.

Experimental variables.—Starting position and direction and remoteness of terminal position could be studied. With these factors defined, limb position and degree of limb flexion are automatically determined for a subject of a given size.

Temporal factors also could be studied, for example, the time between presentation of standard stimulus and the reaction, time between successive reintroduction of the standard stimulus, time between reactions, and time between initiation and completion of a reaction.

Discrimination of Extent

In the studies to be included under this heading, measurements would be made of the ability of an individual to move his limb a specified distance from one point to another. Errors would be measured along the line connecting the two points. The experimental procedures and variables for this area of research are listed in the following paragraphs.

Experimental procedures.—The desired reaction, in this case the distance to be covered between two points, could be defined to the

[3] In the method of reproduction, the subject attempts to duplicate a motor reaction which has been made previously. In a certain sense, every trial following an initial trial involves reproduction, although in that case it is not a reproduction of the standard reaction but of the subject's own relatively inaccurate attempt to duplicate that reaction.

subject in verbal terms by telling him to move his limb a certain number of inches from a starting position, or by telling him to move his limb to a point halfway between two points, etc. The methods of production and recognition would then be applied.

The correct distance could be specified by means of a visual display, the subject being shown the position of two points and the distance between them. If practice trials are allowed, the methods of reproduction and recognition could be employed, otherwise only the method of production would be applicable.

The distance could be defined by passive movements of the subject's limb between the two points marking the limits of the distance. Reproduction or recognition could then be used in estimating proficiency.

Active movement might be required of the subject in defining the standard distance. Here he would be asked to move his limb from left to right between two stops and then would be required either to reproduce or to recognize the reaction.

The apparatus could be extremely simple; perhaps a piece of paper and a pencil would suffice for the most part. However, if it were necessary to keep the factors of limb loading and friction down to an absolute minimum, more delicate photographic or graphic recording methods would be indicated.

Experimental variables.—In addition to the general variables, the effect of the following factors could be studied: (1) extent, starting position, direction, path, and rate of movement; and (2) temporal factors (see those listed under Discrimination of Direction).

STUDIES OF MOVEMENT REACTIONS

Experimental studies of movement reactions are undoubtedly of greater importance and of more general interest than studies of either static or positioning reactions. Although data on these latter types of motor activities are, as has been noted, of some significance for the design of complex equipment, such activities are far less frequent in the over-all operation of control mechanisms than are movement reactions. Moreover, although the act of reaching out to a given position in space and grasping a control lever (positioning reaction) is obviously the first step in operating that lever, the manner in which it is moved is ordinarily the most significant feature of the reaction in terms of the effects produced.

A preliminary examination of the many possible kinds of movement reactions suggests the feasibility of dividing them into four major groups: *discrete movements*, *repetitive movements*, *serial movements*, and *continuous movements*. The criteria for making these distinctions stem principally from the experimental procedures and operations employed in the study of movement reactions. Qualitative differences are useless as differentiating criteria because of the fact that movements in one group may have the same characteristics and involve the same musculatures as do segments of movements classed in another group. Although the four groups form a series in which complexity, loosely defined, increases from discrete movements to continuous movements, many inversions of the order could result from the experimental procedures employed in each case. In the material which follows, a detailed description of each of four classes of the movements is given, along with the operations and characteristics which appear to differentiate them.

Discrete Movements

For purposes of the present study, a discrete movement may be defined as a single unitary movement of any articulate member of the body from any position in space to any other. Studies of discrete movement could be concerned primarily with the rate at which the movement was made and its path through space. Thus a subject might be required to move his arm at a constant or varying rate, in a straight or curved line, in any direction. His accuracy in attaining the prescribed rate and in following the designated path would be determined. Because of the fact that the term *discrete movements*, like most topical headings, is not without ambiguity, it is essential that the laboratory operations and characteristics which differentiate these movements from others be specified. These are as follows: (1) The reaction occurs after the stimulus has ceased to act, the stimulus here being the words used by the experimenter or the visual display introduced in order to define the required movement to the subject. (2) The accuracy of the movement is defined in terms of the degree to which it deviates in rate or in form from that of the movement specified by the experimenter. (3) On successive repetitions, the same stimulating conditions are maintained, the same measurements are made, and approximately the same response is evoked.

Experimental procedures.—The desired movement could be defined verbally by telling the subject to move his limb at a constant rate from left to right, around in an arc, etc. The methods of production and recognition would then apply. The use of the method of recognition in this case would require that the subject's hand be moved passively at several rates over a given course. He would attempt to recognize the one movement in which a constant rate was maintained between two points.

The standard movement could be defined visually. Here a static model, such as a line drawn on paper, would be adequate to indicate the path, but a moving display would be needed to indicate rate. If, during the presentation of the visual display, the subject were allowed to practice by moving his limb along the path at the desired rate, his ability to perform the movement would be estimated by either the method of reproduction or by that of recognition. If he were not allowed to practice, but merely observed the display, the methods of production or recognition would be applicable.

The standard movement might also be defined by moving the subject's limb passively along a given path at a particular rate. Either reproduction or recognition could be used in evaluating proficiency.

The rate of the standard movement might be defined by allowing the subject to move his hand actively along a visually presented pathway at his own rate. On the following trial he would attempt to reproduce both his own self-generated rate and the correct path. It is probably not feasible to allow the subject to define his own path by an active movement.

The apparatus used in studying discrete movement should be capable of recording the rate of the movements and their course (preferably in three dimensions) without appreciably interfering with the movements being studied. Photographic techniques, in which a flashing light or spark gap is attached to the moving member, satisfy these requirements reasonably well. Much simpler graphic methods would, of course, suffice in some cases.

Experimental variables.—The following variables, in addition to the general variables already listed, are relevant to the study of discrete movements: (1) rate of movement; (2) magnitude, direction, location, and form of the path; (3) temporal factors (see those listed under Discrimination of Direction); and (4) characteristics of the mechanical system, if any, interposed between the limb movement and the observed movement of an indicator. These characteristics include the ratio between the amplitude of movement of a control and the amplitude of movement of the indicator; the relation between the plane of movement of the control and the plane of movement of the indicator; the relation between the type of movement of the control (rotary, linear, etc.) and the type of movement of the indicator; the relation of the direction of movement of the control to direction of movement of the pointer; magnitude of lag; and degree of backlash.

Repetitive Movements

The members of this class, which includes such movements as tapping and turning a crank handle, are characterized by the fact that they are, in essence, discrete movements which are performed a number of times in fairly rapid succession. In some instances, e. g., in tapping, each discrete movement is followed by a return movement functioning principally to bring the limb back into position for the next discrete movement. In others, e. g., turning a crank, the initial movement is such that its termination automatically leaves the limb in a position for the succeeding movement. Here the discreteness of the movements may become almost completely obscured, and the overall result can be distinguished from continuous movements only by reference to other characteristics of the situation such as the stimulus conditions.

For the most part, investigations of repetitive movements would be directed toward the study of the maximum rates obtainable, the nature of motor blockage which occurs as the rate is increased to the physiological limit, and the form of the movement elements as a function of rate, amplitude, and other factors.

Data obtained in studies of repetitive movements would relate directly to the problem of equipment design wherever crank-turning movements, tapping movements, or reciprocal pumping movements are required of the operator. Fewer movements of this type appear to be involved in the operation of aircraft equipment than in the operation of artillery pieces, tank guns, and machine tools.

As in the case of simple discrete reactions, certain characteristic conditions serve to distinguish repetitive movements from other types. These are as follows: (1) The type of repetitive reaction required is first specified by the experimenter; the subject then performs the reaction in the absence of corrective stimuli other than those which are self-generated. (2) Proficiency is determined by recording the number of movements made in a given time, by analyzing the shape of the components, by measuring the variability of timing and form within a series, by recording the occurrence of motor blockage, etc. (3) The component unit-movements are nearly identical, each involving approximately the same muscle groups as the others.

Experimental procedures.—The desired reaction could be elicited by verbal instructions alone or by a combination of such instructions and visual or auditory displays. The method of production would apply in the majority of instances since usually the subject would be told to tap as rapidly as possible, or to tap at a rate of, say, once a second, and so on. The methods of reproduction and recognition would clearly not apply if the subject were instructed to move at his maximum possible rate. Nor would they apply to the tasks in which the subject would be required to translate a verbally defined rate into motor terms.

The rate of a repetitive tapping or turning movement, as well as its amplitude, could be defined by attaching an appropriate device to the subject's limb which would move it passively in the desired manner. The methods of recognition and/or reproduction might then be employed.

The apparatus required for studies of repetitive movements would be extremely simple if only maximum rates of tapping or winding were to be measured. Electromagnetically operated counters and tachometers would meet most requirements. Were it desired to record the detailed characteristics of the component reactions, however, recourse would be had to more elaborate graphic or photographic methods.

Experimental variables.—In addition to most of the general experimental variables listed previously, the following apply directly to repetitive movements: (1) amplitude (of reciprocal movements) and radius (of turning movements)—this would include the determination of optimal amplitudes for various rates; (2) rate; and (3) kind of movement, for example, path traversed and direction.

Serial Movements

Serial movements may be defined provisionally as those which are composed of a number of discrete movements involving starting and stopping of the moving member, changes in direction, and the like, each component being made in response to a specific, relatively discrete change in the external stimulating conditions. Here the primary concern is with the ability of an individual to perform the entire series of reactions, with little or no significance being attached to the more minute characteristics of the components. The study of movements of this variety may have considerable practical importance, since there are a number of aircrew positions as well as industrial situations where it is desirable to be able to perform a series of knob-turning or switch-throwing reactions with speed and accuracy. One related problem which deserves considerable attention concerns the ways in which discrete movements can be combined into a series in the most efficient manner.

Serial movements may be distinguished from other movement reactions by the following properties: (1) The reactions ordinarily occur in the presence of the stimuli. (2) The over-all task may be such that the subject paces his own reactions (i. e., the completion of a reaction automatically results in the appearance of a new stimulus), or the rate of reaction may be predetermined (new stimuli appear at certain time intervals regardless of whether the reaction has occurred or whether it is correct). (3) Different, relatively discrete stimuli are associated with each of the different component movements. (4) Proficiency is measured primarily by recording the total time required to complete the series of movements and by noting the number of errors made. The detailed characteristics of each of the components could, of course, also be measured. (5) During the initial runs through the reaction sequence, each response will be relatively independent of every other, but on repeated trials as learning takes place, some responses will acquire the capacity to serve as stimuli for other responses, and some stimuli will acquire the capacity to evoke responses somewhat removed in the series.

In many serial-reaction tasks a rather clear distinction can be made between the movements of adjusting the individual control mechanisms and the connective movements which function solely to carry the adjusting limb from one control mechanism to another. For purposes of exposition, these movements will be referred to hereafter as *adjustive* and *connective* movements respectively. By varying the nature of the serial-reaction task, either adjustive or connective movements may be studied in relative isolation. Connective movements can be eliminated from the task by using a single control mechanism, such as a knob with a pointer attached, and requiring the subject to perform a series of adjustive movements with that control only. The elimination of multiple controls and hence the factor of their spatial separation excludes the connective movements. Likewise, the task may be set up so as to reduce adjustive movements to a minimum by requiring that the limb be moved as rapidly as possible from one to another series of points in space, pausing at each point only long enough to push a button or make a mark on a piece of paper with a pencil.

The Complex Coordination Test developed at the AAF School of Aviation Medicine is probably a good example of a serial-reaction task in which adjustive movements (as differentiable from the connective movements) have been eliminated. In that test, the subject moves a stick and a rudder bar to one after another of a series of positions indicated indirectly by means of rows of lights. Here the connective movements are, in essence, the only reactions required; once they have been made correctly, the light signals change to indicate the direction for the next connective movement. No qualitatively different movement reactions take place after each of the basic positioning or connective movements has been performed.

In serial-reaction tasks where it is desirable to include both adjustive and connective movements, experiments could be designed in which either type or reaction is held constant while the other is varied. (See Experimental Variables listed below.) This procedure is in contrast with that described previously where either of the reactions could be reduced to an unimportant minimum. Experiments conducted along these lines should yield data on how best to combine connective movements with adjustive movements in order to increase over-all performance efficiency.

Experimental procedures.—The procedures for defining the task to the subject are less complex in the case of serial-reaction situations than in the case of positioning reactions and discrete movements. One reason for this difference is that in serial-reaction tasks the subject is usually allowed almost complete freedom in determining the exact paths his movements will follow and the rates at which the movements are made. It is only the end result of the movements, i. e., the turning of a given knob through an angle of, say 180°, which is ordinarily specified in detail. In general, verbal instructions, supplemented by visual inspection of the apparatus and perhaps preliminary practice trials would suffice to acquaint the subject with the task.

The nature of serial-reaction tasks is such that the methods of production, reproduction, and recognition, as employed for positioning movements and discrete reactions cannot be applied without seriously distorting their original meanings. If interpreted rather loosely, the method of production is relevant to some extent, since the subject converts the verbal statements of the experimenter into action and hence produces the desired response. On the second presentation of the serial task, or on a trial following a practice trial, the subject may in a sense be reproducing his reactions of the previous trials. Since he is not instructed, however, to follow exactly the same movement paths on successive trials, but only to accomplish the over-all task as rapidly and efficiently as possible, the method of reproduction is not, strictly speaking, applicable. The method of recognition could be employed only if a rather large number of serial-reaction apparatuses were to be constructed to serve as comparison tasks, a procedure that is obviously impracticable because of the time, effort, and expense involved.

The type of apparatus required for studies of serial-reaction movements would depend entirely upon the variables being studied. In exploratory studies conducted in the Aero Medical Laboratory, use has been made of the Finger Dexterity Test employed in the AAF Aviation Psychology Program. This test consists of a board in which 48 square holes are cut. Round-headed pegs with square bases are inserted in the holes. The basic task is to lift the pegs from the holes, rotate them 180°, and reinsert them in the holes. By varying the spacing of the pegs and the required direction of rotation, a large number of serial-reaction tasks can be devised, in which either the connective or the adjustive movements, or both, are varied systematically (see ch. 16).

Experimental variables.—In addition to the general variables outlined in an earlier section, the following apply specifically to serial-reaction situations: (1) Adjustive movements. For example: Direction of movement (clockwise, counterclockwise); types of movements (twisting, pulling, pushing); amplitude of movement; homogeneity or heterogeneity of types of movements within the series, etc. (2) Connective movements. Here variations could be introduced in magnitude, degree of uniformity, direction, in the ways in which different connective movements are combined, etc. (3) Characteristic of the mechanical system. (See paragraph with this heading under the experimental variables listed for Discrete Movements.) (4) Type of display. This involves variations in the kinds of cues provided to the subject so that the adjustive movements can be performed in the right manner and in the correct order.

Continuous Movements

The movements which appear to fall logically into this class are those in which continuously changing motor adjustments are made in response to continuously changing stimulus configurations. The task of the aerial gunner in tracking a rapidly moving target by the manipulation of suitable controls in a turret is an excellent example of what is meant here by a continuous-movement reaction. In general, most of the tasks, which in the literature on psychomotor tests are referred to as pursuit tasks or eye-hand-coordination tasks, fall into this category. As was noted in the introduction to this report, complex motor activities (many of which are obviously continuous) involving gross movements of the whole body, are believed to be relatively unimportant with respect to the design of equipment and hence are not considered here.

Continuous-movement reactions may be distinguished from other movements by the following characteristics: (1) The reactions take place in the presence of the external stimuli which both evoke and guide the movements. (2) In some cases, the stimuli are completely independent of the reactions, whereas in others, the reactions alter the nature of the stimuli. (3) The stimuli are in a constant state of flux and demand therefore coordinate variations in the reactions. (4) The desired movements are not specified in detail by the experimenter prior to the test period, but are determined more or less continuously by the stimuli occurring during the course of the testing. (5) Proficiency is usually determined by obtaining a continuously integrated record of the disparity between the subject's control movements and the movements which would have resulted in a perfect performance.

For purposes of study, the continuous-movement reactions of importance for this report can be divided into two major classes which are differentiable in terms of the stimulating conditions presented to the subject and the reactions he is required to make. The two classes are *following-pursuit reactions* and *compensatory-pursuit reactions*.

In following-pursuit reaction situations, the movements of the stimulus are completely independent of any controlling actions performed by the subject. As a result, the physical aspects of the stimulus can be accurately specified and controlled. Here the operator manipulates certain control mechanisms in order to keep a pair of cross hairs or other indicator superimposed upon, or adjacent to, a moving target. The displacements between the indicator and the target, as well as the general pattern of movement of the target, serve as cues to the operator to enable him to follow the target.

The compensatory-pursuit task differs from the following-pursuit in that the subject's control movements act to keep the moving target or indicator from drifting away from a stationary null position rather than acting to keep a moving indicator in coincidence with a moving target. A pilot who is flying by reference to the artificial horizon alone is performing compensatory-pursuit movements of this type. Thus the movements of the control mechanisms compensate for movements of the indicator (or indicators) from the null position. Because of this characteristic, the stimuli received by the subject from the indicator change markedly as a function of his reactions. This means that the physical aspects of the stimulus situation differ from subject to subject and cannot be accurately controlled or easily measured. Furthermore, the magnitude of the stimulus deviations leading to the evocation of compensatory movements is inversely related to the level of proficiency attained; the better the subject's performance, the smaller the drifts of the pointer from the null position. In the case of following-pursuit tasks, an increase in proficiency does not reduce the over-all movement of the target, although it does result in a decrease in the average distance separating the target and the follower.

Probably both the following- and the compensatory-type of pursuit can, by changing the stimulus conditions, be converted into a third type which might be called planned pursuit. In the planned-pursuit task, the subject is able to obtain information about the kinds of reactions he will be required to make at some time in the future. If he can see the course he will be required to follow while he is reacting to some other portion of the course, he will be enabled to plan and to anticipate. The task of landing an airplane under contact conditions, where the pilot can see the airport and its runways some time before he must make his corrective movements is an example of what is meant here by a planned-pursuit task.

Experimental procedures.—As in the case of serial-movement reactions, the task set for the subject in the continuous-movement situation is defined by verbal instructions and perhaps by the use of preliminary practice trials in which errors may be corrected. In the following-pursuit and compensatory-pursuit tasks, the particular movements which must be performed at each moment are indicated by the position of the moving target with respect to the position of the follower or the location of the zero point. Usually no specific instructions are given regarding the kinds and rates of movements to be used, the subject being free to make whatever movements are most suitable for solving the problem at hand. As a consequence of these characteristics, the methods of production, reproduction, and recognition apply only to the same limited extent that has been indicated for serial movements.

Experimental variables.—The general variables listed above apply here, and in addition the following: (1) Stimulus characteristics.

Variations can be introduced in the speed of movement of the target, direction of movement, amplitude, path, conditions of illumination, design of indicating elements, etc. (2) Control mechanisms. This includes the radius of action of the control levers, the type of control (wheel, stick, rudder), kind and degree of damping, pressure required to operate controls (absolute magnitude, degree of linearity), number of controls (whether in a task involving several dimensions of movement several controls are used, or only one), etc. (3) Characteristics of the mechanical system. In addition to factors already listed in paragraph 4 under Experimental Variables for Discrete Movements, the mechanical system can be varied to provide either velocity control, direct control, or aided control. (4) Task characteristics. The overall task can be changed by increasing the number of pursuit tasks which must be performed simultaneously, by comparing foot-operated controls with hand-operated controls, by comparing the performance of two men in operating a two-dimensional tracking device with that of a single individual, etc. (5) Because of the wide variety of pursuit tasks which could be designed and constructed, no general recommendations can be made regarding the kind of apparatus which would be suitable for the study of continuous movements.

SUMMARY OF AVAILABLE LITERATURE PERTAINING TO MOTOR ABILITIES

As mentioned in an earlier section of this report, many pages of the psychological literature are devoted to studies of the sensory and perceptual abilities of human beings, but few to fundamental motor ability as described here (see Boring (7), Woodworth (65), and Troland (59) in this regard).

Historically, research related to the area of motor abilities is of interest. During the latter half of the nineteenth century, research on motor activity dealt principally with the idea of the muscle sense or kinesthesis. Studies conducted at that time emphasized the introspective and histological methods and were directed almost solely toward an analysis of the various types of sensations arising from movements of the limbs and toward an attempt to discover the specific receptors in the muscles, tendons, and joints which might be capable of mediating the various sensations. Reference may be made here to the work of Bell, Fechner, Goldscheider, Sherrington, Titchener, and others reported by Boring (7) and Troland (59).

In the early part of the twentieth century, the work of efficiency experts such as Taylor and Gilbreth led to numerous studies of movement efficiency in industrial situations. This work has been ably summarized and criticized by Farmer (18). These studies were aimed almost without exception toward the discovery of techniques for the elimination of waste motions and the consequent increase in work output per unit time. Considerable emphasis was laid upon proper posture and the use of smooth, circular, and rhythmical motions. Little information, however, was gathered about the basic abilities involved.

Stetson and his students (26, 32, 51, 52, 53, 54, 55) and others, e. g., Davis (15), have, in recent years, made extensive studies of skilled movements. Through the use of action-current recording devices these investigators have been able to specify accurately the particular muscle groups involved and the phasing of their contractions with respect to the total movement cycle. They have not, unfortunately, provided us with data at the macroscopic level on the gross characteristics of simple human movements.

A few items of literature are available concerning basic motor abilities as related to equipment design. Among these may be mentioned the work of Craik and his associates (3, 12, 13, 14, 28, 29, 30, 60, 61), the details of which are given below, and the work of German investigators reported by Fitts (19). A recent memorandum by Jenkins (35) summarizes a number of the problems on the design and use of controls in the equipment field.

In the sections which follow, publications accessible to the authors have been summarized under the same organizational plan followed in the earlier part of this report.

Static Reactions

An examination of the available literature indicates that a number of studies have been made of tremor-type reactions, but that a majority of these have been somewhat unsystematic in their approach to the problem. No differentiation appears to have been made between tremor and drift reactions in this area and the latter type of behavior has not been investigated in the case of limbs, although investigations of body sway are plentiful in the literature. Measures of tremor have been mostly in terms of amount of movement in inches or millimeters, or frequency of tremor reactions per second.

Whipple (64) describes typical apparatus for measuring steadiness-type static reactions and summarizes the literature available in 1924.

Two examples of apparatus for studying tremor may be cited. In a recent article, Edwards (17) describes a work-adder type of device for recording finger tremors in three dimensions. By means of three threads attached to the finger, small riders were caused to move up inclined planes, one for each dimension. The positions of the riders on the planes indicated the summated tremor. Fossler (20) studied frequency of finger tremor by means of a small movable coil suspended in the field of a strong electromagnet, a device developed by Travis and Hunter (58). The current induced in the coil was recorded by a sensitive oscillograph on motion-picture film.

The results of a number of studies of the effects of various factors on steadiness may be summarized as follows. It might be noted that many of these findings should be checked with additional experimentation.

Great variations from individual to individual and from day to day are reported by a number of investigators (17, 20, 64, 66), although Paulsen (43) reports odd-even reliabilities of 0.98 and test-retest correlations of 0.73.

Tremor appears to decrease as age increases from childhood to early maturity (17, 64).

Men may exhibit greater tremor than women, but the evidence is somewhat equivocal (16, 17, 64).

Steadiness in maintaining a fixed position with a lever is greatly increased by friction (12).

Most studies of the use of visual cues in tremor reactions indicate that steadiness is improved with visual reference (12, 17, 44). Young (66), however, found that the use of visual cues did not decrease tremor.

Young (66) reports that the greater the effort exerted in maintaining the fixation, the less the success and, presumably, the greater the tremor.

Praise and reproof have ambiguous effects (16), with praise increasing the steadiness of women and decreasing that of men and reproof acting in the opposite fashion.

Fatigue decreases steadiness; the rate, amplitude, and irregularity of tremor oscillations varies directly with fatigue (8).

Exercise is reported (44) to decrease steadiness, but according to other investigators (16, 17), exercise has small or ambiguous effects. Static and mental work apparently decrease steadiness (44).

The amount of tremor decreased in Edwards' study (17) as the point of limb support was shifted from none, to elbow, to wrist, and to palm; whereas Fossler (20) reports that the point of support has only a slight effect.

Less tremor is found when the subject is seated than when he is standing, and position of the limb relative to the body is reported to have appreciable effects on steadiness (12, 17). Craik (12) found that with visual reference, subjects were able, on the average, to keep a pointer on a line, though oscillations still occurred. Tremor increased as the hand was moved more than 8 inches above or below the heart.

Reaction-time is significantly shorter and variability smaller when the stimulus is introduced at the top or bottom phase of the tremor cycle (56).

Fossler (20) found tremor frequencies ranging from 5 to 500 per second, with roughly 70 percent lying in the range from 8 to 80.

Subjecting the hand to heat (20) is reported to decrease the low frequencies while cold tends to decrease the high ones.

Positioning Reactions

Excerpts from the available literature which pertain to discrimination of direction and discrimination of extent are summarized in the following paragraphs.

Discrimination of direction.—Literature concerned with the problem of how accurately subjects can attain terminal positions with the limbs in the absence of vision appears to be nonexistent. Studies involving the use of vision have been summarized by Whipple (64). A situation was employed in which the subject struck a target in time with the beat of a metronome employing a ballistic-type movement. Accuracy was measured radially about the target. The difficulties of controlling the type of reaction are obvious. The findings suggest the following conclusions: accuracy increases gradually with age, particularly in the range of 5 to 8 years; sex differences are not appreciable; normative data indicate that mean error for college subjects is about 4–6 mm., with a range of 3–9 mm.; the performance of the right and left hands correlates 0.54.

A study has been completed recently by Fitts which yields information concerning the ability of individuals to locate equidistant targets with their hands in the absence of direction vision (see ch. 15).

The work of McNeill (37) in France concerning the form, dispersion, and speed of trajectories of hitting movements, performed both with and without the aid of vision, appears to be relevant here, but the original material was not available at the time of writing.

Discrimination of extent.—The available literature with regard to this problem is somewhat more voluminous than that for discrimination of direction. The findings may be summarized as follows:

When visual corrective cues are not employed, short distances tend to be overestimated (45). The use of vision increases the accuracy of performance appreciably (1, 45).

The use of a pointer held in the hand results in a tendency for extents to be underestimated as compared with judgments of the same extents with the finger (9). Presumably vision was not employed.

A finding indirectly relevant is that size of objects judged by touch or manipulation is smaller than that judged by vision (4, 11). The judgment without visual corrective cues appears to depend on the estimation of the length of the sides or diameter involved (11).

Somewhat equivocal experimental evidence obtained by Weber (62) and Weber and Dallenbach (63), suggests that extents are underestimated when the subject's movements are impeded by frictional resistance.

Robinson and Richardson-Robinson (48) studied the ability of 81 women undergraduates to draw lines varying in length from 6 to 33 centimeters in steps of 3 centimeters. Corrected odd-even correlations were computed among the sets of reproduced lengths. The correlations were highest between lines of the same approximate length and decreased in a roughly linear fashion as the magnitude of the difference between the correlated lines increased. The range of correlations was from 0.96 to 0.35.

Movement Reactions

A few scattered studies are available in the literature which deal directly with movement reactions. The reports of investigations available to the writers are summarized below according to the categories of discrete movements, repetitive movements, serial movements, and continuous movements.

Discrete movements.—With regard to the rate with which discrete movements can be made and the variables affecting this rate, the following statements may be made.

Beeler (5) found that rate of movement with a control stick was greater for push than for pull; maximum rates increased both with a decrease in maximum stick force and an increase in maximum stick displacement; and maximum rates with zero force were 251 in./sec. for push and 140 in./sec. for pull. Hertel (27), however, reports considerably lower figures. He also states that the rate of movement of the feet is less than that of the hands.

An investigation of rate as a function of loading indicated that the former decreased directly with the latter (36). Maximum velocities of the hand averaged 139 in./sec.

Glanville and Kreezer (23) found that the left arm could be moved more rapidly than the right, flexion movements were slower than extension movements, and averaged velocity was about 0.45° per millisecond.

Researches by Montpellier (40) dealt with acceleration as a function of time. Subjects were required to move objects from one place to another with rapid and slow movements. Among other findings, it was noted that there was a continual variation in acceleration, i. e., the speed of the movement was not constant at any one point, and that the maximum acceleration was located usually in the first and last quarters of the movement.

With regard to force as a variable operating in discrete movements the following studies apply. The maximum push exertable on a foot pedal has been studied as a function of the position of the pedal relative to the seat (33). It was found that maximum push could be exerted when the position of the pedal was such that the angle of the lower leg with respect to the thigh was 165°. The maximum force exerted by 38 subjects when seated averaged 700 pounds, with 6 subjects pushing over 900 pounds. In addition it was found that with a knee angle of 165°, maximum push could be exerted when the foot pedal at the end of its travel was in such a position that the line of the subject's thigh was 20° above the horizontal.

Hertel (27) reports that the maximum forces which can be exerted on airplane controls are as follows: (1) Elevator operation of a stick: two-handed push, 220 pounds; two-handed pull, 185 pounds; one-handed pull or push, 150 pounds. (2) Aileron operation of a stick: two-handed push or pull, 50 pounds; one-handed push, 40 pounds; one-handed pull, 35 pounds. (3) Aileron operation of a wheel-type control: two-handed turning, 62 pounds. (4) Rudder pedal operation: pushing with 1 foot, 476 pounds. Gough and Beard (24) report comparable values for a stick-type control.

Hick (29) studied the ability to apply incremental and decremental forces to relatively constant basic forces in a nearly isometric situation involving a hand control. Increments, positive or negative, superadded to a steady muscular force had their means shifted in the direction of overshooting, as compared with the condition of no or small basic force. This overshooting was most marked with relaxations from a steady pull. The shift of the mean was of the order of 5 to 15 percent for a basic force of 4 pounds. Jenkins has found that Weber's law holds from about 10 to 40 pounds in a situation requiring pressure to be exerted on an approximately isometric stick-type control (see ch. 12). Below 10 pounds, relative accuracy decreased rapidly. Absolute accuracy increased in a roughly linear fashion from 1 to 40 pounds.

Findings of studies concerned with the path or form of discrete reactions may be summarized as follows: It has been found that counterclockwise rotation for the right hand is more accurate in a situation involving the drawing of circles, spirals, and the like (47). This finding held for both right- and left-handed individuals whereas the opposite tendency appeared for the left hand.

Gemelli (21, 22) found that the movements of subjects who were requested to trace complicated figures on a flat surface were composed of a series of discrete movements. The frequencies of these step-wise discontinuities ran parallel to the square root of the angular velocity. Practice resulted in a decrease in errors and the acquisition of more characteristic form by the movement. As the movement became more precise, the number of discontinuities became progressively greater.

Montpellier (39) reports the existence of a conflict between least effort and the tendency toward exact reproduction of the model in a tracing problem.

Graphic records made by Stetson (52) of beating movements with a baton showed a rounded point at the end of the up stroke as compared with a sharp point at the end of the down stroke.

Hsiao (31) studied the ability of children to trace between two parallel lines of four different kinds. The results indicated that curved movements were easier than horizontal or vertical movements, and regular movements were easier than irregular ones.

Repetitive reactions.—The literature on repetitive reactions appears to be moderately voluminous although many of the studies are not directly pertinent to the design of equipment. Some studies involving the task of tapping were available. Findings of experiments in this area are listed below along with a few other available items that pertain to the design of equipment.

Wide individual differences have been reported in tapping rate along with relatively constant individual performances (10, 38, 64). The maximum ranged from 8 to 13 per second after practice, according to Von Kries as reported by Bryan (10).

Rate of tapping does not appear to be a function of the distance covered within the range of 1 to 40 millimeters (10). Bryan (10) reports the highest tapping rate for an excursion of 20 millimeters.

Bryan (10) reports that the highest rate is for free tapping, next greatest for tapping with wrist and elbow, and slowest for the middle and outer joints of the forefinger. Whipple (64) reports the highest rate when the movement is performed by the elbow joint. Tinker and Goodenough (57) report that in children and adults the index- and middle-finger tapping scores are approximately the same, the little finger is slower than the other fingers, and bimanual tapping is slower than unimanual. Tapping with any finger correlates well with tapping of any other finger.

Rate of tapping with all joints increases directly with age from 6 to 16 years in a sample of 789 school children (10).

When subjects are required to accelerate at their own pace they do it in steps, with Weber's law holding for the relation between successive steps (34).

When subjects are allowed to tap at their preferred rate this ranges from 1.5 to 5.0 taps per second (38).

When a different rythm is introduced and the subject is then required to return to his preferred rate, it is found that the latter is changed in the direction of the introduced rate whether faster or slower (38, 46).

Small correspondence is found between the tapping scores of children obtained with hand and arm unrestrained and with hand and arm restrained (57).

Beder (6) reports that interpolated activity of the same or opposite arm and hand reduces speed of tapping, changes the pattern, and reduces the excursion.

Hick (28) reports that friction, as might be expected, reduces the speed of repetitive winding movements.

Pacaud-Korngold (42) found that in turning cranks in the horizontal and vertical planes with various combinations of directions for the two hands operating simultaneously, coordinated action of both hands was usually slower than movements of a single hand. Horizontal movements of one hand were not slowed appreciably by vertical movements of the other hand, but vertical movement of one hand was retarded by horizontal action of the other hand.

In studying rhythmical movements, Stetson (52) required subjects to make up and down beating movements with a baton at the fastest possible rate. The velocity of the down beat was about two or three times as great as that of the back stroke. The velocity of the beat stroke was dependent on the length of the stroke, but not on the tempo of the rhythm. The duration of the beat stroke was strikingly uniform and was independent of either the tempo or the length of the stroke. Introduction of an obstacle against which the limb struck at the end of the stroke did not appear to alter the character of the stroke.

Serial reactions.—Typical samples of studies of serial reactions are available in the literature pertaining to industrial time-and-motion investigations such as those described by Barnes (2) and Farmer (18). Most of these studies have been concerned with the analysis of performance in a particular position with a view to increasing output. Barnes and his associates have investigated such items as simultaneous symmetrical and asymmetrical hand motions, but unfortunately the original reports were not available to the authors at the time of writing. From the literature available it appears that systematic researches have not been conducted on the effect of the variables listed in an earlier section of this report on serial-type motor activity.

Continuous reactions.—The literature on the effect of a number of factors upon continuous reactions is summarized below.

In a study of the effect of asymmetrically placed aircraft controls, Honeyman and Yallop (30) found that in a three-dimensional compensatory-pursuit task, offsetting the control column to the left produced a tendency to pull it to the right and vice versa. When the rudder was displaced to the left of the central position there was a tendency to give right rudder. They conclude that if an unusual posture is required, movements of control become less well adapted. Fatigue or stress may produce even more marked deterioration.

Vince (60) has studied the relation between the direction of movement of the control handle and the direction of movement of the display pointer. In general, errors in the "unexpected" direction were statistically greater than those made in the "expected" direction. The difference between the two directions was found to be related, however, to the time interval between reactions. When only one reaction was required every 4 seconds, the numbers of errors in the two directions were about equal. At intervals of less than 1 second, however, the errors with the "unexpected" connection rose to more than twice the number with the "expected" connection. When the direction of control was reversed in the middle of a series of trials, some of the subjects did not notice the change. The subjects made fewer subsequent errors than those who did notice the change.

Vince (61) reports that in tracking a varying course with linear and nonlinear control levers, no consistent difference was found between errors under the two conditions. Also the learning curves did not appear to be affected by the change from linear to nonlinear and vice versa. However, when linear and nonlinear levers were used in correcting large misalinements of a display pointer, the difference in favor of the linear lever was statistically significant. In general, a sudden change in direction produced marked errors with the nonlinear lever. Practice tended to make performance with a nonlinear lever almost as good as that with linear.

Craik (13) found that in positional laying of constant-speed courses, lack of clarity of the display introduced errors of unusually long period. A slight acceleration increased the magnitude of the errors greatly as compared with a constant speed course. Craik lists the following factors as of importance in affecting accuracy of movement in a tracking-type situation: limb employed, mean position of the limb and its point of support, amplitude of movement required, force required, and direction of movement. No clear-cut data bearing on these points were presented.

Craik and Vince (14), in studies of psychological and physiological aspects of gun-control mechanisms, attempted to analyze the nature of the errors obtained on constant-speed courses. In the early stages of learning, the errors resembled damped oscillations. As learning progressed, the oscillations about the correct course became approximately constant in amplitude and period; and in the final stages of learning, the subject anticipated more, moved his control more smoothly, and only occasionally relapsed into the oscillatory type of reactions. On variable-speed courses, the oscillations were about twice as rapid as on constant-speed courses, the difference being highly significant statistically. Laying accuracy was not improved by spring-centering of the twist-grip type controls, nor by varying the amplification of the control in a ratio of 1.9:1.

Interrelations Among Motor Reactions

Several investigations have been concerned with the relationships among a number of motor abilities. The available reports may be summarized as follows.

Harrison and Dorcus (25) required 50 subjects to perform a number of different movements at their own chosen rates, with the purpose of the experiment being camouflaged. Results of the timing of these performances indicate that (1) there is no unitary speed trait which is characteristic of various spontaneous movements or motor adjustments of an individual, and (2) individuals tend to perform at a fairly constant rate from one time to another.

Pacaud (41) concludes that a basal relationship exists among an individual's several reaction times for arm movements, whether in the circular-horizontal or circular-vertical plane and whether isolated or variously coordinated. A particular reaction time, however, is in reality representative only of those movements of the same nature. It is misleading to speak of an individual's general reaction time, nor is it satisfactory to accept the reaction time of one group of movements as a general motor characteristic of the individual.

Seashore and associates (49, 50) have carried out extensive correlational and factorial analyses of individual differences in sensory motor coordinations. Evidence was found for the existence of certain group factors or areas within which tests were at least moderately related. Examination of the nature of these groupings of tests showed that the boundaries of the groups usually cut across those of (1) specific musculature and (2) specific sense fields. Thus, the various skilled actions of a given musculature were not found to be significantly correlated, but similar skilled actions of different musculatures were closely correlated. Similarly, the various skilled actions involving a given sense field were not ordinarily correlated, but similar skilled muscular reactions involving different sense fields were usually at least moderately correlated. In general, the factors correspond to qualitative similarities in the pattern of action and not to basic biological constants of sense field of musculatures.

BIBLIOGRAPHY

1. Abel, T. M. A comparison of tactual-kinesthetic and visual perceptions of extent among adults, children, and subnormals. *Amer. J. Psychol.*, 1936, 48, 269-296.[a]
2. Barnes, R. M. *Motion and time study.* New York: Wiley & Sons, 1940.
3. Bartlett, F. C. Instrument controls and display—efficient human manipulation. Flying Personnel Research Committee *Report No. 565*, Great Britain, December 1943.

[a] Item was seen only in abstract form.

4. Becker, J. Über taktilmotorische Figurwahrnehmung. Versuche mit 9 bis 15 Jährigen Schulkindern. (On the tactual-motor perception of children.) *Psychol. Forsch.*, 1934, **18**, 102–157.[*]

5. Beeler, De E. Maximum rates of control motion obtained from ground tests. National Advisory Committee for Aeronautics, *Restricted Bulletin No. L4E31*, Washington, May 1944.

6. Boder, D. P. The influence of concomitant activity and fatigue upon certain forms of reciprocal hand movement and its fundamental components. *Comp. Psychol. Monogr.*, 1935, 11, No. 4.[*]

7. Boring, E. G. *Sensation and perception in the history of experimental psychology.* New York: Appleton-Century, 1942.

8. Bousfield, W. A. The influence of fatigue on tremor. *J. exp. Psychol.*, 1932, 15, 104–107.[*]

9. Breternitz, W. Ueber das Tourtualsche Täuschung. (The Tourtual illusion.) *Z. Sinnesphysiol.*, 1934, 65, 146–165.[*]

10. Bryan, W. L. On the development of voluntary motor ability. *Amer. J. Psychol.*, 1892, 5, 123–209.

11. Calabresi, R. La precezione tattilo-cinetica delle grandezze a delle forme negli adulti e nei fanciulli. (The tactile and kinetic perception of size and shape in adults and in normal and abnormal children.) *Riv. di sci. appl. all'educ. fis.*, 1931, 21–37.[*]

12. Craik, K. J. W., with the assistance of Vince, M. *Psychological and physiological aspects of control mechanisms with special reference to tank gunnery.* Part I. Military Personnel Research Committee, Medical Research Council, Great Britain, B. P. C. 43/254, August 1943.

13. Craik, K. J. W. with assistance of Vince, M. *Psychological and physiological aspects of control mechanisms.* Part II. With special reference to ground-tank and A. A. tank gunnery. Military Personnel Research Committee, Medical Research Council, Great Britain, B. P. C. 44/322, March 1944.

14. Craik, K. J. W. and Vince, M. *Psychological and physiological aspects of gun-control mechanisms.* Part III. Effects of "stiffness" and of spring-centering of hydraulic velocity controls. Military Personnel Research Committee, Medical Research Council, Great Britain, B. P. C. 45/406, January 1945.

15. Davis, R. C. The pattern of muscular action in simple voluntary movement. *J. exp. Psychol.*, 1942, 31, 347–366.

16. Eaton, M. T. The effect of praise, reproof, and exercise upon muscular steadiness. *J. exp. Educ.*, 1933, 2, 44–59.[*]

17. Edwards, A. S. The finger tremometer. *Amer. J. Psychol.*, 1946, 59, 273–283.

18. Farmer, E. *Time and motion study.* Reports of the Industrial Fatigue Research Board, No. 14, London, 1923.

19. Fitts, P. M. *German air-force research on the design of sights and controls for most effective use by the operator.* Mimeographed report, Psychology Branch, Aero Medical Laboratory, August 1945.

20. Fossler, H. R. Range and distribution of tremor frequencies. *J. gen. Psychol.*, 1931, 5, 410–414.

21. Gemelli, A. *Ricerche sperimentali sulla natura e sulla diagnosi dell 'abilità manuale.* (Experimental researches on the nature and determination of manual skill.) Milan: Society editrice "Vita e pensiero," 1920.[*]

22. Gemelli, A. *Ricerche sperimentali sulla forma nei movimenti volontari.* (Experimental studies on form in voluntary movements.) *Atti pont. accad. sci. nuov. Lincei, Roma*, 1920.[*]

[*] Item was seen only in abstract form.

23. Glanville, A. D., and Kreezer, G. The maximum amplitude and velocity of joint movements in normal male human adults. *Hum. Biol.*, 1937, 9, 197–211.
24. Gough, M. N., and Beard, A. P. Limitations of the pilot in applying forces to airplane controls. *Technical note No. 550*, National Advisory Committee for Aeronautics, Washington, January 1936.
25. Harrison, R. and Dorcus, R. M. Is rate of voluntary bodily movement unitary? *J. gen. Psychol.*, 1938, 18, 31–39.*
26. Hartson, L. D. Contrasting approaches to the analysis of skilled movements. *J. gen. Psychol.*, 1939, 20, 263–293.
27. Hertel, H., Determination of the maximum control forces and attainable quickness in the operation of airplane controls. National Advisory Committee for Aeronautics, *Technical Memorandum No. 583*, Washington, January 1930.
28. Hick, W. E. Friction in manual controls with special reference to its effect on accuracy of corrective movements in conditions simulating jolting. Flying Personnel Research Committee, *Report No. 623*, Great Britain, June 1945.
29. Hick, W. E. The precision of incremental muscular forces with special reference to manual control design. Flying Personnel Research Committee, *Report No. 642*, Great Britain, August 1945.
30. Honeyman, W. M. and Yallop, J. M. Report of an investigation into the effects of asymmetry of aircraft controls. Flying Personnel Research Committee, *Report No. 621*, Great Britain, May 1945.
31. Hsiao, H. H. A study of four kinds of abilities of hand movement. *Chung Hwa educ. Rev.* (Chinese), 1933, 21, No. 35–46.*
32. Hudgins, C. V. The incidence of muscular contraction in reciprocal movements under conditions of changing loads. *J. gen. Psychol.*, 1939, 20, 327–338.
33. Hugh-Jones, P. *Some physiological aspects of tank driving controls.* Part I, The maximum push exertable on a foot pedal and how this varies with the position of the pedal relative to the seat. Military Personnel Research Committee, Medical Research Council, Great Britain, B. P. C. 44/341, May 1944.
34. Jaworski, G. and Liberson, W. Recherches sur le pouvoir de discrimination des rhythmes des mouvements volontaires. (Ability to discriminate rhythms in voluntary movements.) *Travail hum.*, 1936, 4, 163–195.*
35. Jenkins, W. O. *Research prospectus concerning the design of cockpit controls from the psychological standpoint.* Mimeographed report, Psychology Branch, Aero Medical Laboratory, February 1946.
36. Koepke, C. A. and Whitson, L. S. Power and velocity developed in manual work. *Mechanical Engineering*, 1940, 62, 383–389.
37. McNeill, H., *Motor adaptation and accuracy.* Louvain: Editions de l'Institut Superieur de Philosophie, 1934.
38. Miles, D. W. Preferred rates in rhythmic response. *J. gen. Psychol.*, 1937, 16, 427–469.
39. Montpellier, G. de. Les alterations morphologiques des mouvements rapides (Morphological alteration in rapid movements.) *Etud. Psychol.*, 1935, 4
40. Montpellier, G. de. Note sur l'acceleration dans les mouvements volontair de la main. (Note on acceleration in voluntary hand movements.) *Arch Psychol., Geneve*, 1937, 26, 181–197.*
41. Pacaud, S. Contribution a l'etude des mouvements volontaires. (Contribution to the study of voluntary movements.) *Annee psychol.*, 1942, 40, 152–170.*

* Item was seen only in abstract form.

42. Pacaud-Korngold, S. Contribution a l'etude des mouvements volontaires. (Contribution to the study of voluntary movements.) *Travail hum.*, 1940, 8, 10–43.[a]

43. Paulsen, G. B. The reliability and consistency of individual differences in motor control. *J. appl. Psychol.*, 1935, 19, 29–42.[a]

44. Poliakova, M. G. (On the influence of work on tremor.) (*The psycho-physiology of industrial work.*) Leningrad: 1936.[a]

45. Raffel, G. Visual and kinaesthetic judgments of length. *Amer. J. Psychol.*, 1936, 48, 331–334.

46. Rethlingshafer, D. Measurement of a motor set. *J. exp. Psychol.*, 1943, 32, 75–81.[a]

47. Riper, C. V. and Bryngelson, B. Speed and accuracy of clockwise and counterclockwise movements. *J. Psychol.*, 1936, 1, 247–253.[a]

48. Robinson, E. S. and Richardson-Robinson, F. A simple series of abilities. *Amer. J. Psychol.*, 1929, 41, 33–53

49. Seashore, R. H. An experimental and theoretical analysis of fine motor skills. *Amer. J. Psychol.*, 1940, 53, 86–98.

50. Seashore, R. H., Buxton, C. E., and McCollom, I. N. Multiple factorial analysis of fine motor skills. *Amer. J. Psychol.*, 1940, 53, 251–259.

51. Sperry, R. W. Action current study in movement coordination. *J. gen. Psychol.*, 1939, 20, 295–313.

52. Stetson, R. H. A motor theory of rhythm and discrete succession, I. *Psychol. Rev.*, 1905, 12, 250–270.

53. Stetson, R. H. A motor theory of rhythm and discrete succession, II. *Psychol. Rev.*, 1905, 12, 293–350.

54. Stetson, R. H. and Bouman, H. D. The action current as a measure of muscle contraction. *Science*, 1933, 77, 219–221.[a]

55. Stetson, R. H. and McDill, J. A. Mechanism of the different types of movement. *Psychol. Monogr.*, 1923, 32, 18–40.

56. Tiffin, J. and Westhafer, F. L. The relation between reaction time and temporal location of the stimulus on the tremor cycle. *J. exp. Psychol.*, 1940, 27, 318–324.

57. Tinker, M. A. and Goodenough, F. L. A comparative study of finger tapping in children and adults. *Child Development*, 1930, 1, 152–159.[a]

58. Travis, L. E. and Hunter, T. A. Tremor frequency, *J. gen. Psychol.*, 1931, 5, 255–260.

59. Troland, L. T. The principles of psychophysiology. Vol. II. *Sensation.* New York: Van Nostrand, 1930, Section 72, 336–347.

60. Vince, M. A. Direction of movement of machine controls. Flying Personnel Research Committee, *Report No. 637*, Great Britain, August 1945.

61. Vince, M. A. The psychological effect of a nonlinear relation between control and display. *SR1 (Servo) Report No.* 2, Great Britain, February 1946.

62. Weber, C. O. Space and time in kinaesthetic fields of force. *Amer. J. Psychol.*, 1927, 38, 597–606.

63. Weber, C. O. and Dallenbach, K. M. The properties of space in kinaesthetic fields of force. *Amer. J. Psychol.*, 1929, 41, 95–105.

64. Whipple, G. M. *Manual of mental and physical tests.* Part I: Simpler processes. Baltimore: Warwick and York, 1924.

65. Woodworth, R. S. *Experimental psychology.* New York: Henry Holt, 1938

66. Young, I. C. A study of tremor in normal subjects. *J. exp. Psychol.*, 1933 16, 614–636.[a]

---

[a] Item was seen only in abstract form.

BLANK PAGE

CHAPTER FOUR

The Relative Effectiveness of Presenting Numerical Data by the Use of Tables and Graphs[1]

LAUNOR F. CARTER

INTRODUCTION

The flight engineer, bombardier, and navigator must frequently refer to different mathematical functions in performing their flight duties. Since these *numerical functions* are usually presented in either graphic or tabular form, it is desirable to reduce the degree of error found in their use by determining the best means of presenting these relationships. The Flight Data Branch of the Aircraft Projects Section requested the Aero Medical Laboratory to study this problem and recommend the best methods for presenting data for use by air-crew members.

There are a number of different factors, such as type size, grouping of data, selection of units for the axes of graphs, and relationship of lines to white space, which should be investigated, but the fundamental purpose of this study was to compare the relative efficiency with which data presented in tables and graphs can be used. It is desirable to be able to recommend that given types of data for certain uses be presented by resort to tabular or graphic techniques.

EXPERIMENTAL PROCEDURE

To present comparable data in graphic and tabular form, a table and a graph for each of the following equations was prepared. The form of the material presented was as follows:

Table I and graph I, $y=1.2x$

Table II and graph II, $y=\frac{x^2}{100}$

[1] This chapter is based upon research findings reported in Headquarters AMC, Engineering Division Memorandum Reports Nos. TSEAA 694-1 and TSEAA-694-1C.

Table III and graph III, $y = cx$; $c = 1.2, 1.4, 1.6, 1.8$, and 2.0

Table IV and graph IV, $y = \frac{x^2}{c}$; $c = 60, 70, 80, 90$, and 100

The tables were prepared with five-unit increments in the argument. Each of the graphs was plotted on standard graph paper (Frederick Post Company, Number 317–S, 20 x 20 lines to the inch) with one unit

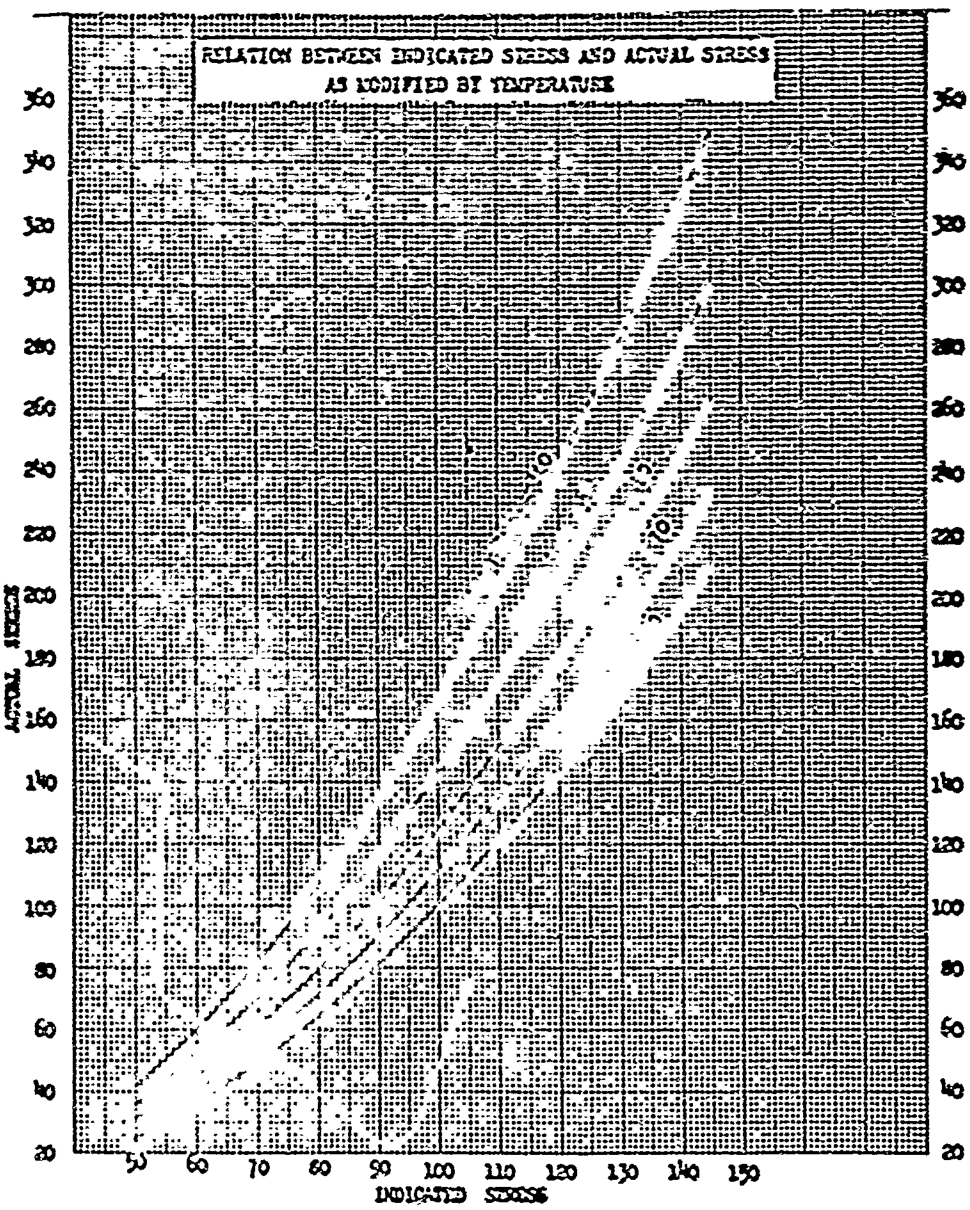


FIGURE 4.1.—Sample showing the type of graph used in the study.

on the abscissa equaling one unit in the argument and one unit of ordinate for figure I and one-half unit for all other figures equaling one unit in result. Table 4 and figure 4 of the experimental material are presented to illustrate the tables and graphs used.

Table 41.—Sample showing the type of table used in the study

[Relation between indicated stress and actual stress as modified by temperature]

Indicated stress	Actual stress				
	Temperature				
	60	70	80	90	100
50	25	28	31	36	41
55	30	34	38	43	50
60	36	40	45	51	60
65	[illegible]	47	53	60	70
70	49	54	61	70	82
75	56	62	70	80	94
80	64	71	80	91	[illegible]
85	72	80	90	103	120
90	81	90	101	115	[illegible]
95	90	100	112	128	139
100	100	111	125	143	[illegible]
105	109	122	135	157	[illegible]
110	121	134	151	173	[illegible]
115	132	147	165	189	[illegible]
120	144	160	180	206	[illegible]
125	156	174	195	225	[illegible]
130	169	188	211	241	[illegible]
135	182	203	228	258	[illegible]
140	196	219	245	276	[illegible]
145	[illegible]	234	263	300	[illegible]

The master copies of the tables were typed. The graphs were drawn with india ink. These master copies were then photographed and 25 glossy prints of each table and graph were made. Answer sheets were drawn up requiring the subject to write down the result to the arguments given in each item. For the simple tables and graphs the subjects were required to give answers both where the arguments were tabulated and where interpolation was required. For the tables and graphs presenting families of data, the subjects were required to work with arguments for tabulated values, and with arguments requiring both single interpolation and double interpolation. Identical problems were presented for both the tables and graphs but the order of problems within a series was randomized so that the subjects would not recognize that they were working the same problems. Each subject worked problems using each table and each graph. The material was presented in the following order: table I, graph I, graph II, table II, table III, graph III, graph IV, table IV. Time limits were established so that none of the subjects were able to complete all the problems presented. The time limits used for each part were as follows:

	Tabulated values	Single interpolation	Double interpolation
Table I and graph I	1 m. 30 s.	3 m.	
Table II and graph II	1 m. 30 s.	3 m.	
Table III and graph III	1 m. 45 s.	4 m.	4 m.
Table IV and graph IV	1 m. 45 s.	4 m.	4 m.

Before the material was administered the method of entering each table and graph was explained. The subjects were also given a short review of the techniques of single and double interpolation.

With this material, it was possible to investigate whether tabular or graphical presentation is superior for linear and curvilinear data, for single sets of data and for families of data, and which is superior when the subjects are required to enter the table with tabulated values and when they are required to interpolate by using nontabulated arguments.

RESULTS

The results of administering this material to 27 subjects (26 pilots and 1 navigator) were analyzed from two points of view: speed and accuracy. The measure of speed used was simply the number of problems completed in a given time, but several measures of accuracy were used. Table 4.2 shows the results of the administration when analyzed for number of problems completed. Figure 4.2 shows the average number of problems completed per minute when using the tables and graphs.

From table 4.2 and figure 4.2 it is quite evident that, as far as speed of use is concerned, the table is superior to the graph whenever the arguments used in entering the table are tabulated, but that when the arguments are nontabulated, the graphical presentation is superior. In all cases the difference between the performance on the table and on the graph is so great that the statistical probability is less than one in a thousand that the direction of these differences would be changed if the experiment were repeated on a similar group of men.

TABLE 4.2.—*Mean, standard deviation, and significance of difference between tables and graphs in terms of the number of problems completed*

[N=27]

	Mean number completed		Standard deviation		CR[1]
	Table	Graph	Table	Graph	
TABULATED VALUES					
Table I and graph I	29.41	13.04	5.18	2.06	17.16
Table II and graph II	32.82	12.52	4.41	2.58	23.13
Table III and graph III	20.01	11.67	2.43	2.66	14.44
Table IV and graph IV	20.52	9.44	3.28	2.11	21.78
SINGLE INTERPOLATION					
Table I and graph I	13.67	22.89	5.42	4.04	8.12
Table II and graph II	10.52	20.19	3.50	3.09	15.66
Table III and graph III	12.45	17.30	3.31	5.60	6.41
Table IV and graph IV	10.56	17.41	2.65	4.71	8.93
DOUBLE INTERPOLATION					
Table III and graph III	2.74	14.22	1.14	5.27	11.23
Table IV and graph IV	3.15	15.07	1.60	5.29	11.07

[1] These critical ratios were computed by the technique described on pp. 165 ff. of Peters, C. C. and Van Voorhis, W. R., *Statistical procedures and their mathematical bases.* New York: McGraw-Hill Book Co. 1940.

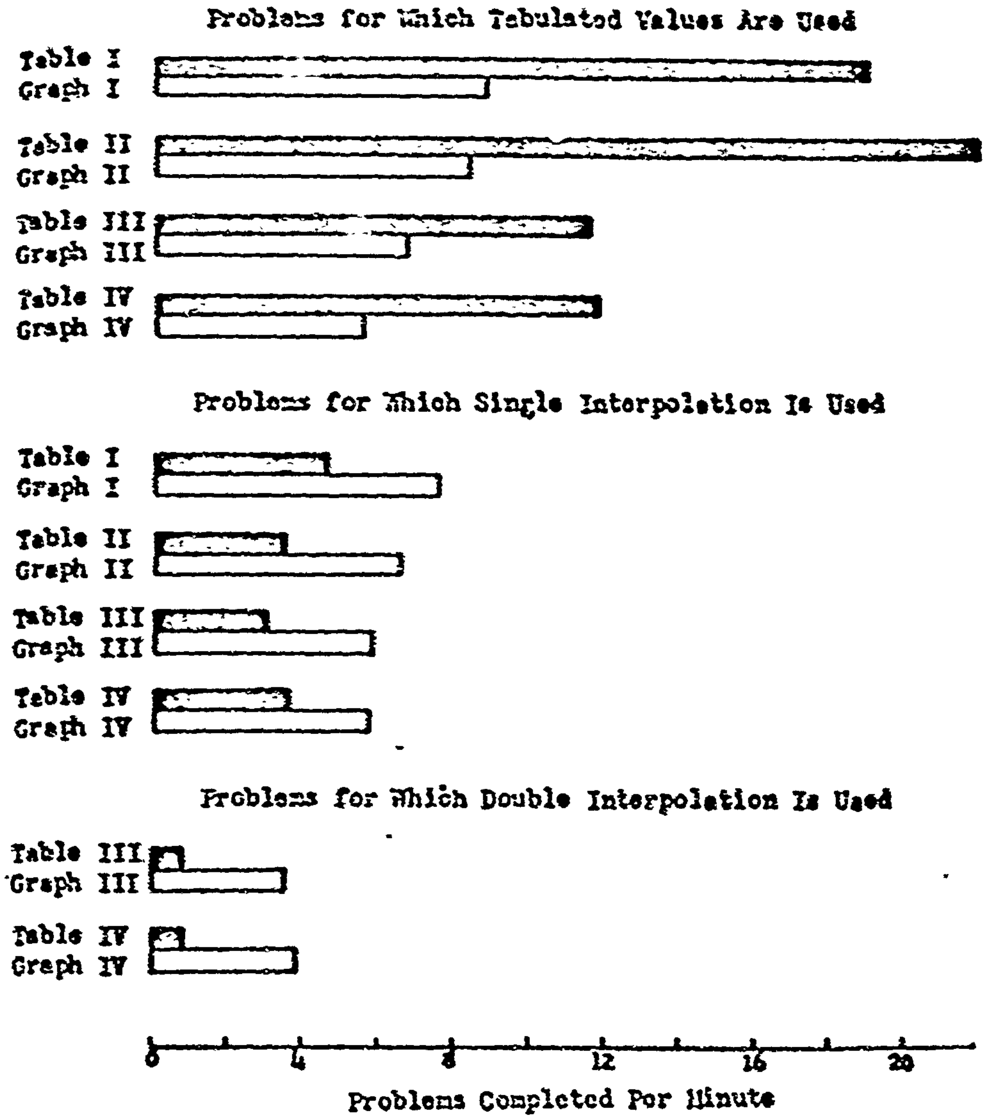


FIGURE 4.2. Relative speed in obtaining information from tables and from graphs.

The criterion for accuracy is not as clear cut as is that for speed. A number of small errors may not be as serious as one or two large errors. Table 4.3 shows the number of errors made, the average magnitude of these errors, and the average magnitude of the errors in terms of the total number of problems attempted. In all cases an error is any answer which differs from the correct answer by one or more units.

It was not thought wise to compute more elaborate statistics for the error data since the distributions were very skewed, being almost J shaped.

A second measure of accuracy is the percentage of problems attempted for which the answer was in error. Table 4.4 shows the per-

TABLE 4.3.—*Number of errors and their average magnitude*

	Total number of errors		Average value of each error		Average value of each error per problem completed	
	Table	Graph	Table	Graph	Table	Graph
TABULATED VALUES						
Table I and graph I	2	37	7.00	2.00	0.02	0.21
Table II and graph II	4	52	2.75	6.85	.02	1.05
Table III and graph III	12	65	8.42	5.94	.19	1.23
Table IV and graph IV	23	121	12.52	4.26	.53	2.02
SINGLE INTERPOLATION						
Table I and graph I	81	153	2.46	2.02	.54	.68
Table II and graph II	72	233	6.59	3.60	1.31	1.54
Table III and graph III	130	192	5.78	4.56	2.23	1.86
Table IV and graph IV	93	336	4.52	4.09	1.47	2.94
DOUBLE INTERPOLATION						
Table III and graph III	46	267	12.02	2.79	7.47	1.94
Table IV and graph IV	61	360	6.11	2.98	4.38	2.64

TABLE 4.4.—*Percentage error in using tables and graphs*

	Number of errors divided by total completed x 100		Number of errors greater than one divided by total completed x 100	
	Table	Graph	Table	Graph
TABULATED VALUES				
Table I and graph I	0.3	10.5	0.3	2.0
Table II and graph II	.5	15.4	.2	3.5
Table III and graph III	2.3	20.6	2.0	7.3
Table IV and graph IV	4.2	47.5	3.4	12.5
SINGLE INTERPOLATION				
Table I and graph I	21.9	29.6	7.3	4.5
Table II and graph II	27.5	42.7	12.3	11.7
Table III and graph III	39.6	41.1	20.5	15.2
Table IV and graph IV	32.6	71.9	18.9	29.5
DOUBLE INTERPOLATION				
Table III and graph III	62.2	69.5	49.5	24.0
Table IV and graph IV	71.8	88.4	52.9	48.9

centage error in terms of the number of problems attempted and also the percentage error where the error was greater than one unit. This latter statistic is given since in graphs II, III, and IV the smallest division on the ordinate of the graph represented two units and it is not reasonable to expect the answer to be correct within less than one unit.

The data in table 4.3 and table 4.4 support the same conclusions regarding the accuracy with which the material is used when presented in graphic and tabular form. It is apparent that when the tables are entered with tabulated values we find the table to be superior to the graph in both speed and accuracy. On the other hand, whenever the table cannot be entered directly and some type of interpolation must be used, the data gathered from the graph is just as accurate and perhaps more accurate than that gathered from the table. With the graph a large number of errors are made but they are usually very small errors.

It seems apparent that when relatively accurate answers are required and it is not possible to tabulate all the values that will be used, the material should be presented in graphic form. On the other hand, if speed is essential and some accuracy can be sacrificed, it seems better to have the material presented in tabular form and not to require any interpolation by allowing all entries to be made with the nearest tabulated argument.

FURTHER RESEARCH

The results reported above suggest that further research should be done to determine the best composition of graphs and the best length for tables. As was demonstrated, a table can be used more rapidly and accurately than a corresponding graph when the arguments used in entering the table are tabulated. However, when the arguments used are such as to require interpolation in the table, its use is slower than is the use of the graph. This suggests that it may be possible that a table in which all of the unit values are tabulated can be used more rapidly than a graph since there will be no need for interpolation with such a table. Of course such a table is more complicated to construct and may have ten or twenty times more pages than the corresponding graph. If such a table is too cumbersome, then a table in which every other point is tabulated might be superior since most of the arguments would be tabulated and those which were not would require only the simplest interpolation: that is, interpolation exactly halfway between two tabulated values.

It is also possible that the composition of graphs may be improved. Perhaps graphs in which the coordinate lines are spaced further apart would be read more rapidly and just as accurately as those in which the coordinate lines are closer together. It also seems possible that graphs should be entered on the y-axis rather than on the x-axis as is traditionally done. All of our reading habits involve reading from left to right and from the top of the page down, but the traditional method of entering graphs requires reading from the bottom up and often from right to left.

To investigate these different possibilities, a second experiment was designed. In an attempt to determine the optimal number of tabulated points in a table three different tables have been constructed: one in which every fifth point in the major argument and every tenth point in the minor argument is given, another where every second point is given for both arguments, and a third with every point given. Three graphs have been constructed in which the distance between the co-ordinate lines has been varied; one graph is drawn on 20 x 20 lines-to-the-inch graph paper, another on 8 x 8 graph paper, and the third on 4 x 4 graph paper. Finally one graph has been constructed on 20 x 20 graph paper in which the independent variable is plotted on the *y*-axis. At the time of writing, this material is being administered through the cooperation of Dr. Clarke Crannell at Miami University.

SUMMARY

1. A comparison was made of the relative speed and accuracy of individuals in using graphs and tables. Identical mathematical functions were presented in both graphic and tabular form and experimental subjects solved the same set of problems with each presentation.

2. It is concluded that the table is superior to the graph in both speed and accuracy of use whenever the arguments used in entering the table are tabulated, but that whenever the arguments are nontabulated a graphic presentation is superior to a tabular one in speed of use and that it is just as accurate.

CHAPTER FIVE

Psychological Factors Involved in the Design of Air Navigation Plotters[1]

JULIEN M. CHRISTENSEN

INTRODUCTION

One important part of the aerial navigator's job involves the measuring and plotting of various vectors and lines of position. In order to facilitate these measurements and plottings, the navigator employs an aircraft plotter. The plotter is an instrument combining a special protractor and straight edge. It is used to measure given angles and to plot lines along angles of specified numbers of degrees. Previous investigations have shown that an appreciable percentage of the errors made by navigators can be attributed to their plotting instruments. In a recent unpublished study by Robert T. Joseph of this laboratory, rated AAF navigators were required to measure and to plot 20 single-leg courses. The incidence of error (i. e., an answer differing by more than 1° from the correct answer) in this simple task was about 5 percent. In another test carried out in 1945 by the Psychological Research Project (Navigator) of the Air Training Command the common types of dead reckoning and air plot problems were administered to rated AAF navigators. The authors reported the incidence of plotter error to be as high as 31 percent.[2]

PURPOSE

The present experiment was designed so that variations in the basic components of air navigation plotters could be studied. The results are to be used as a basis for the design of a plotter which incorporates those features which contribute to error reduction and facility in the handling of the instrument.

[1] This chapter is based upon research findings reported in Headquarters AMC Engineering Division Memorandum Report No. TSEAA-694-1D.

[2] Research Bulletin 45-36, Psychological Section, Office of the Surgeon, AAFTC.

DESCRIPTION AND DEFINITION OF PLOTTER ERRORS

Plotter errors may be classified into precisely defined groups. Approximately 90 percent of the plotter errors fall into the four groups termed reciprocal, tolerance, reversal, and reference.

A *reciprocal error* is defined as that type in which the navigator reads the value of a measured or plotted line exactly 180° in error. It is apparent from examination of the plotters that one reason for this is the fact that each position on the protractor is shared by two scales whose values differ by 180°.

A *tolerance error* is defined as that type in which the navigator reads a few degrees in error either because of carelessness or inability to read the scale with the required accuracy. A tolerance of 1° on either side of the correct answer was allowed. Thus, if the correct answer was 103°, answers 102° and 104° were also considered correct.

A *reversal error* is defined as that type in which the navigator reads in the wrong direction from a numbered graduation mark. Thus, if the navigator meant to plot a value of 93°, he made a reversal error if he read in the wrong direction from the 90° mark and plotted an erroneous course of 87°.

A *reference error* is defined as that type in which the navigator measures or plots a course using East or West instead of North as a reference of zero. Such an answer is always 90° or 270° in error.

Miscellaneous errors were found to include the following types:

1. Confusion between the figures 3 and 5. The inverted flat-topped three (Ɛ) and the inverted five (ς) are quite similar in appearance.
2. Transposition errors. For example, plotting 193° for 139°.
3. Misreading a number. For example, reading 226° for 266°.
4. Inverting the plotter and using it wrong face up.

EXPERIMENTAL PROCEDURE

In order to determine the causes of plotter errors and in order to be able to recommend methods of eliminating these errors logical variations were made in the design of the basic elements of plotters. Tests were carried out to determine the relative effectiveness of these variations. Four basic comparisons were examined in this experiment. In order to examine these four comparisons it was necessary to design and construct five experimental plotters. Each experimental plotter was designed for use in the study of one basic comparison; therefore, none of the experimental models should be considered a plotter recommended for general use. These were used in conjunction with a model which, except for the scales on the base, is a copy of the plotter used at present in the Army Air Forces.

Design of the Experimental Plotter Models.

The first comparison, involving the models shown in figure 5.1, was designed in order to determine whether it would be better to have one or two straight edges on a plotter. It is practically impossible to measure a course within one plotter's width of the top of the chart with the single-edged plotter unless it is inverted. When the single-edged plotter is inverted it then becomes necessary to read all scales upside down and in a direction opposite to that used when the plotter is in the normal position. When using a two-edged plotter the operator merely has to slide the plotter down and use one of the parallel lines or the top edge. It is never necessary to use the two-edged plotter topside down.

The second comparison, involving the models shown in figure 5.2, was designed in order to determine whether it would be better to have the protractor scale read from left to right or from right to left. The scales on the present AAF plotter read from right to left. This is contrary to normal habit. The left-right reading habit is well established. Practically all reading matter, tables, rulers, etc., are read in such manner.

The third comparison, involving the models shown in figure 5.3, was designed in order to determine whether it would be better to number the graduations on the proctractor scales every 10° or every 5°. On model II the 5° graduation marks as well as the 10° graduation marks were numbered. It was thought that numbering more of the graduation marks would aid in the prevention of reversal of errors.

The fourth comparison, involving the models shown in figure 5.4, was designed in order to determine whether it would be better to have two protractor elements or one protractor element on a plotter. One element was set 180° in azimuth from the other so that the figures of one would always appear right side up. If the navigator wished, he could obtain independent readings on the two protractors and use one reading as an accuracy check on the other.

In order to facilitate drafting and increase the accuracy of the reduced printed model the designs for the six plotters were drafted on a linear ratio of 2:1. The models were drafted by Paul Galloway of the Aero Medical Drafting Unit and printed by the AMC print shop on 0.080-inch plexiglas. The Machine Shop Branch trimmed them and drilled the reference holes.

Design of the Experimental Test Materials.

The test materials designed for use with these plotters consisted of 26 measuring items and 26 plotting items. Twenty-six items were printed on a Mercator chart. The remaining 26 items were printed on a type of chart similar to the Mercator except that the meridians were approximately as far apart as the meridians on a sectional

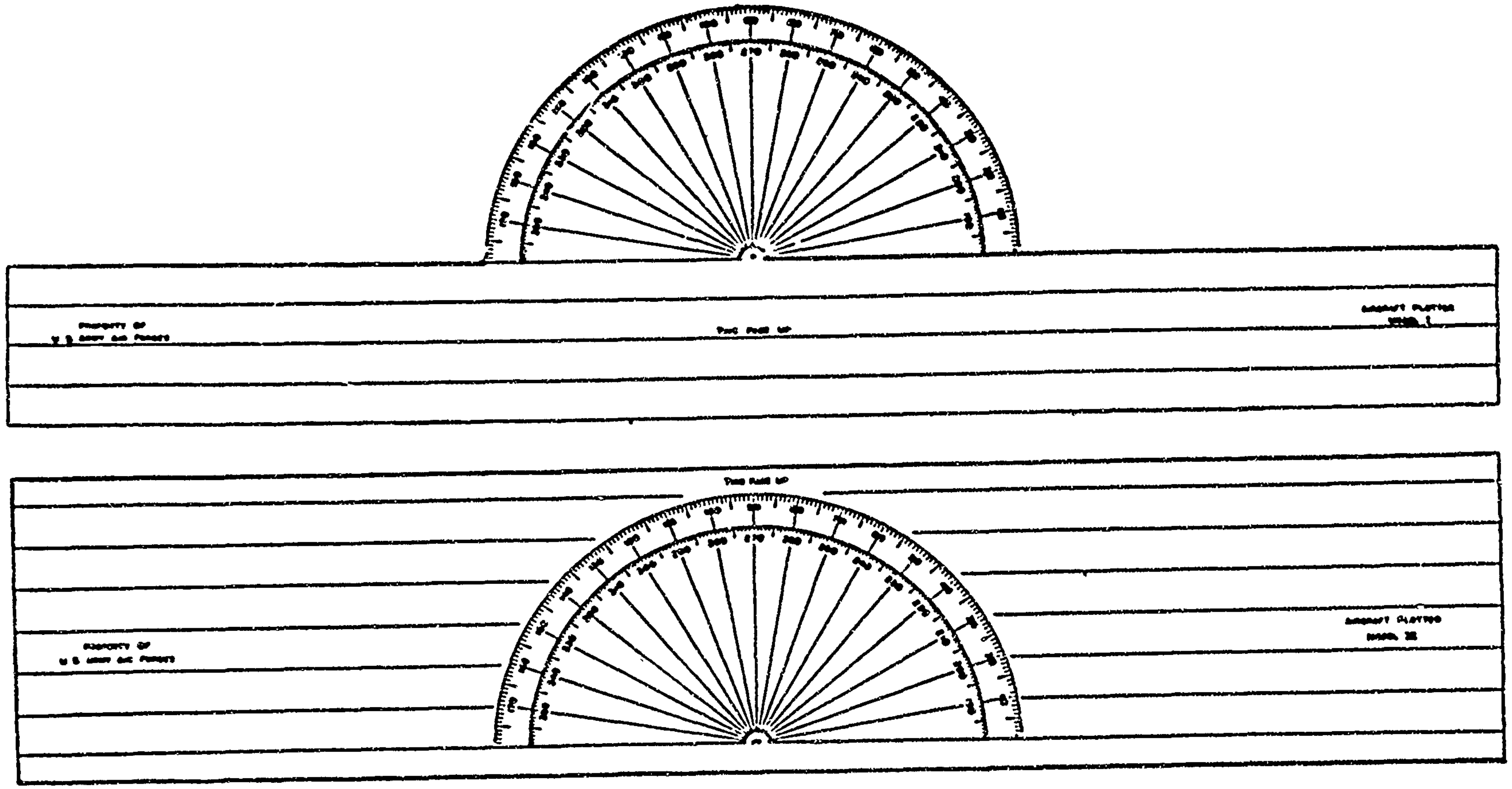


FIGURE 5.1.—Plotter models I and VI, used in comparing a single-edged plotter and a double-edged plotter.

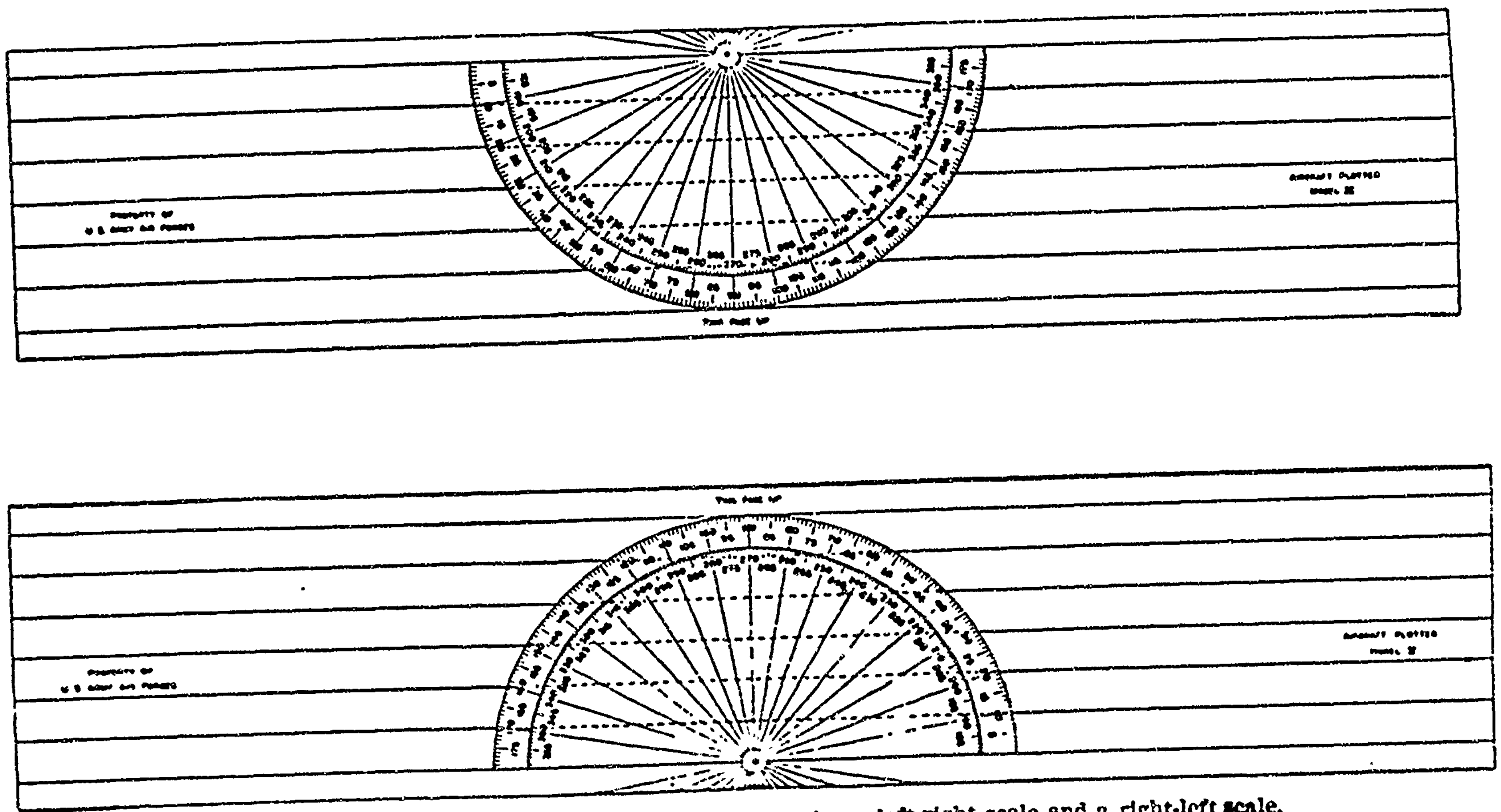


FIGURE 5.2.—Plotter models III and V, used in comparing a left-right scale and a right-left scale.

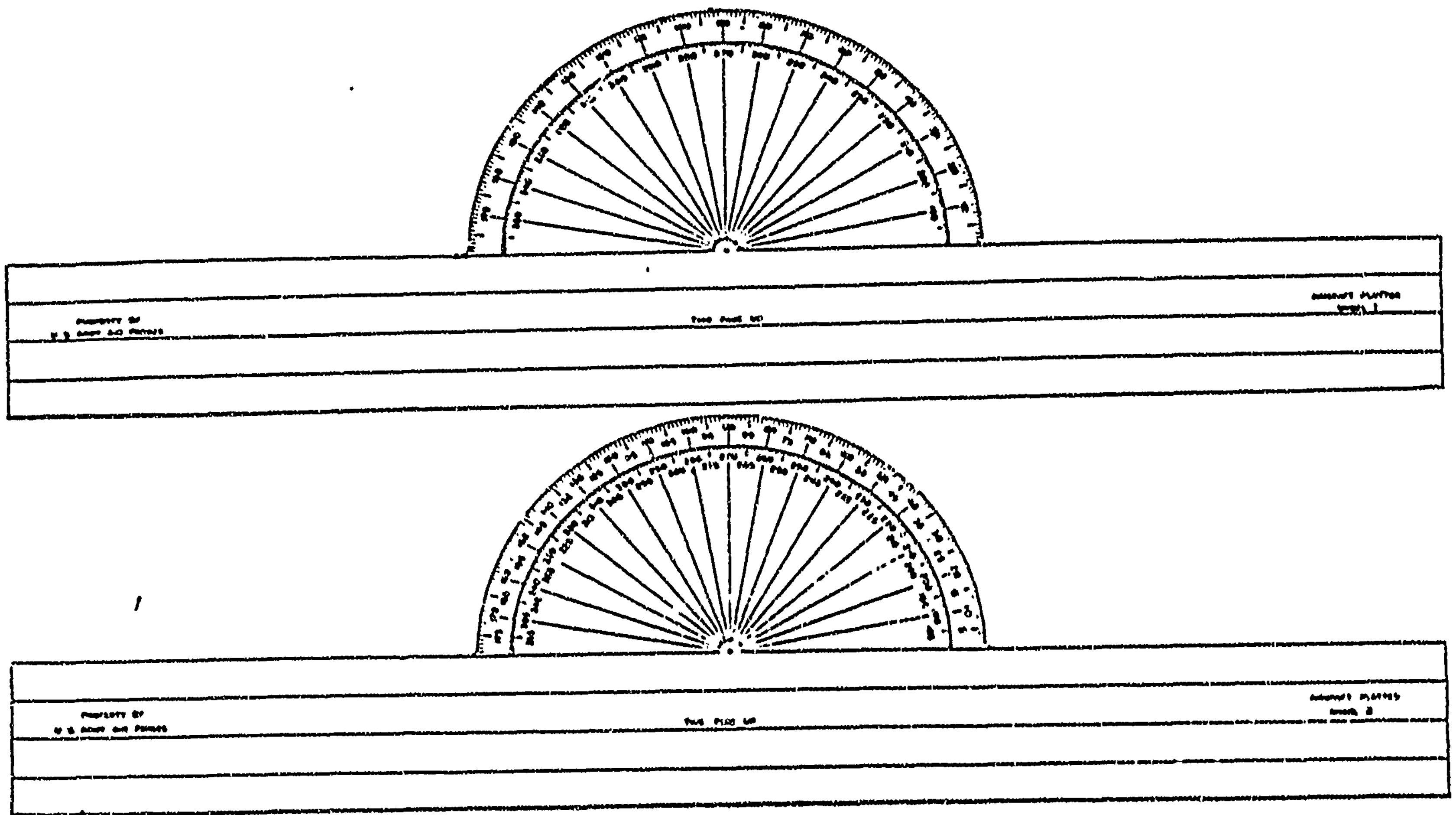


FIGURE 5.3.—Plotter models I and II, used in comparing a 10° scale and a 5° scale.

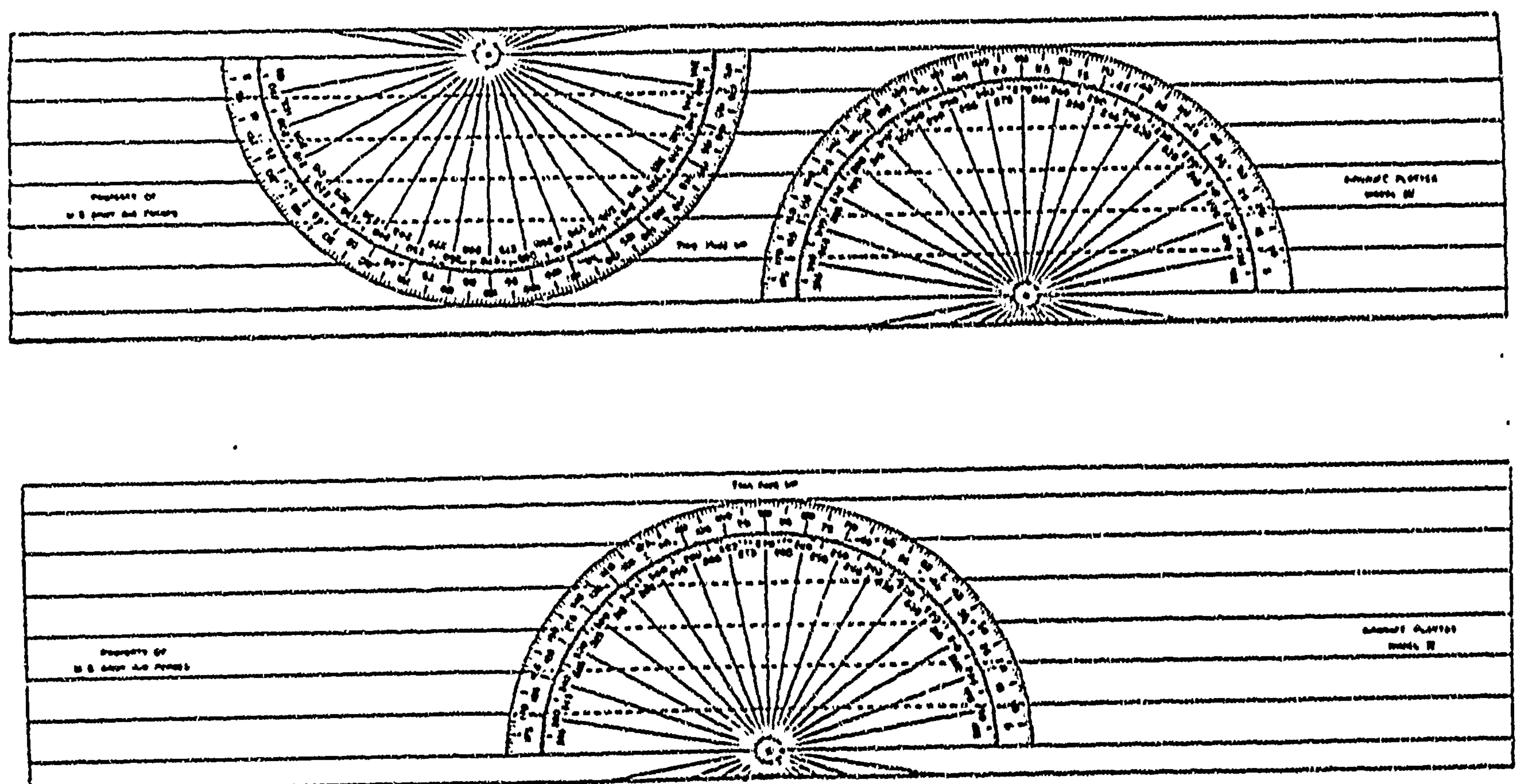


FIGURE 5.4.—Plotter models IV and V, used in comparing a double protractor element and a single protractor element.

aeronautical chart. This type chart was used in preference to a Lambert conformal chart, as it simplified teaching the high-school students to use the charts. Six similar alternate charts were constructed. Sample charts are shown in figures 5.5 and 5.6.

The first six items were used as warm-up problems, so the graded test consisted of 46 items. In order that an adequate sample of both measuring and plotting could be obtained, the subjects were required to measure and to plot in alternate order. Six alternate forms of the test were constructed. An idea of the reliability of the tests may be obtained from the fact that the Pearson product-moment correlation using different plotter models on alternate forms was of the order of 0.75. The relibility of each test is necessarily as high or higher.

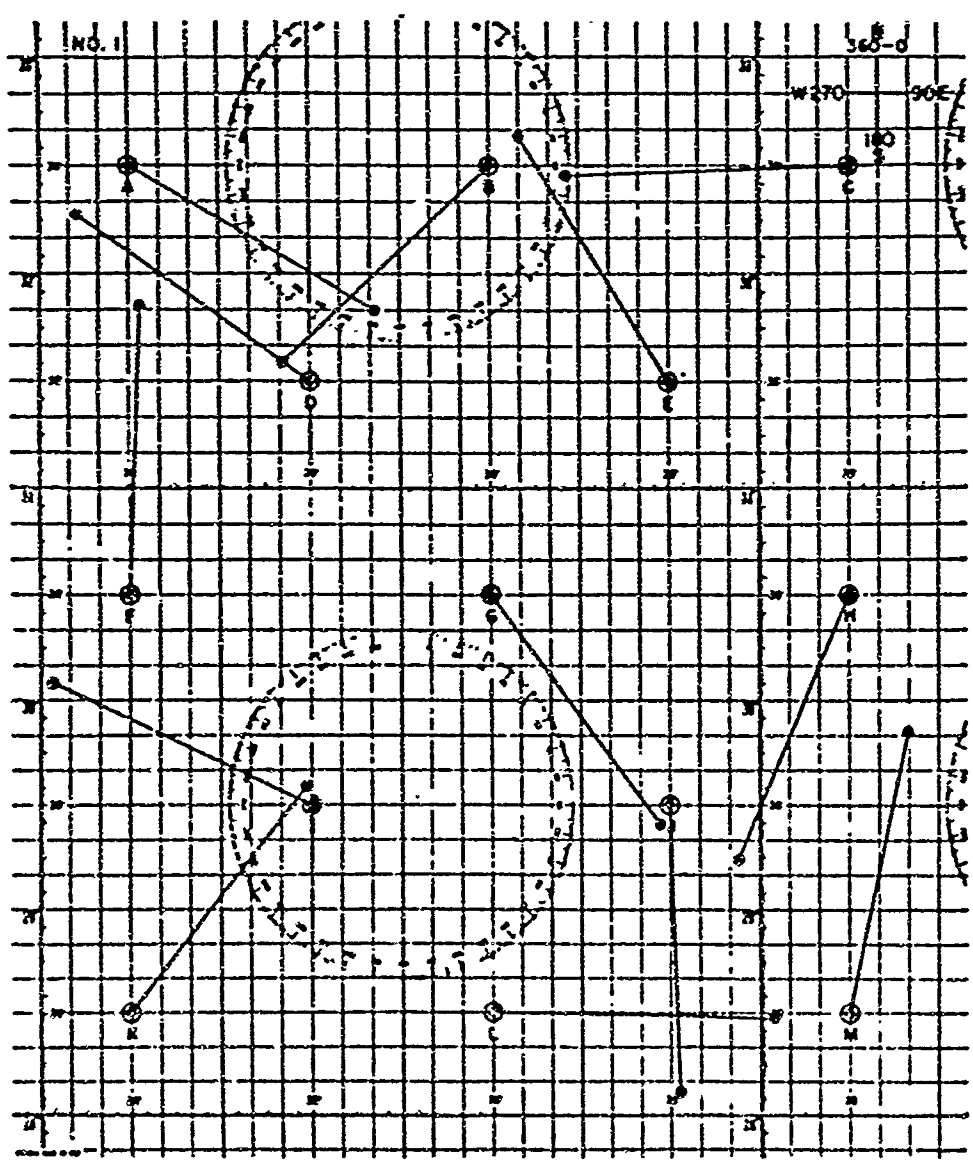


FIGURE 5.5.—Test chart No. 1.
The actual chart was 15¼" by 12⅝".

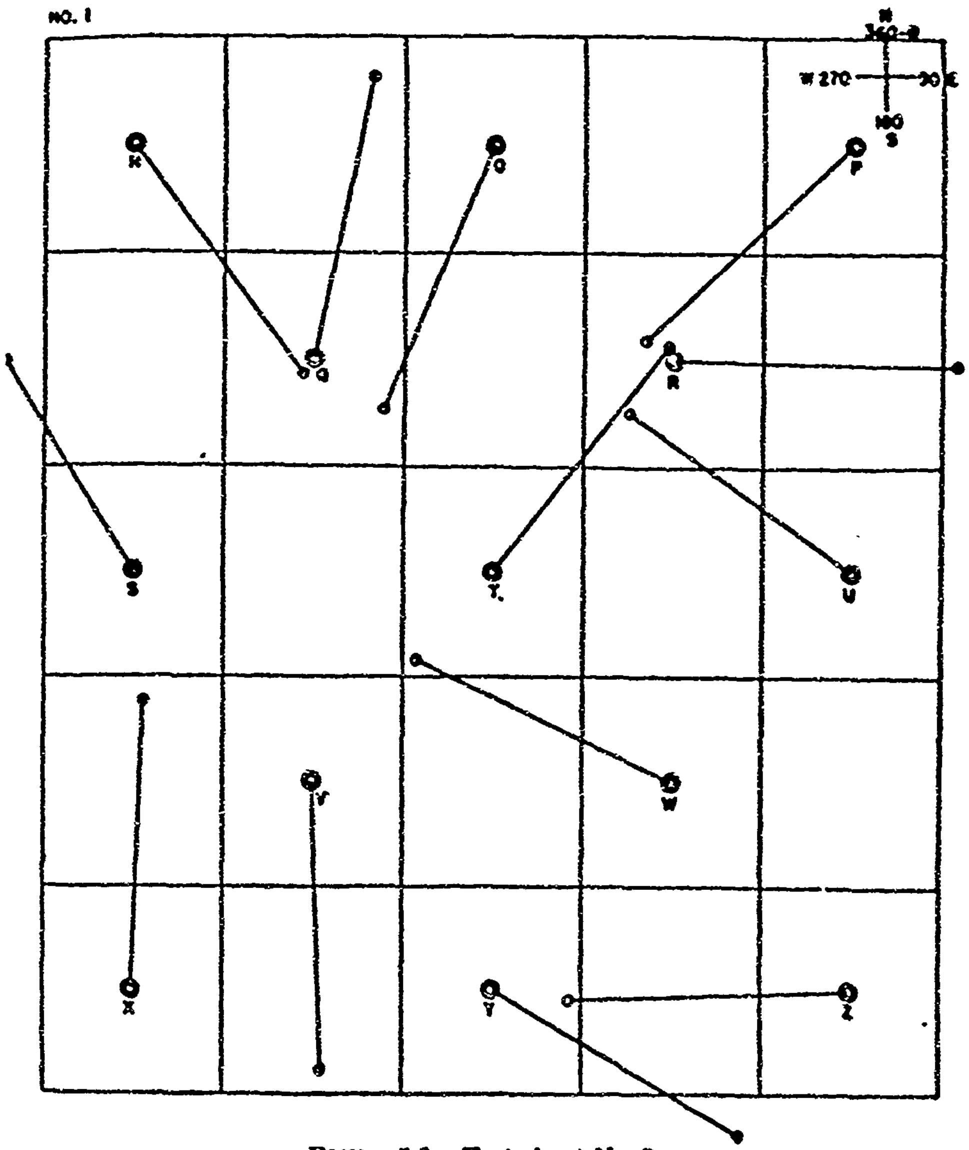


FIGURE 5.6.—Test chart No. 2.
The actual chart was 15¼″ by 12¾″.

Administration of the Tests.

One hour on each of three successive school days was spent with each of four classes at each of four Dayton, Ohio, high schools.[3] The first day was devoted entirely to teaching each class the use of the charts and the appropriate pair of plotters. Each of the subsequent days was devoted to a warm-up period followed by a test on one of the two experimental plotters used with that particular group.[4] Subjects were encouraged not to omit any problems. Each subject worked with both plotters of a designated pair, thus serving as his own control. The

[3] The author is indebted to H. L. Boda, Assistant Superintendent of Schools for Dayton, Ohio, for making possible the testing program in the Dayton High Schools.

[4] The author is indebted to James Parish, Melvin Warrick, and George Abrams for assistance in administering the tests at the four high schools.

plotters were taught and the tests were administered in an order planned so that no advantage would accrue to any of the experimental plotters. Using plotters I and II as an example, plotter I was taught first and tested first at the first school; plotter I was taught second and tested first at the second school; plotter I was taught second and tested second at the third school; plotter I was taught first and tested second at the fourth school. Compensation for differences in learning ability among the plane geometry, advanced algebra, and trigonometry classes was made by matching insofar as possible an equal number of each type of student taking a particular plotter first with an equal number of each type of student taking the alternate plotter first.

A time limit of 25 minutes was used at the first school. As 56 percent finished the test the second day, the time limit was reduced to 20 minutes for the three remaining schools. To adjust for this, the first-day means for the first school for "number of items attempted" were multiplied by 0.80. The second-day means for the first school for number of items attempted were estimated by using the means of a similar class which took the same test in another school and used the 20-minute time limit. This estimate was accurate enough as only a cursory analysis of speed was deemed necessary. It is important to note, however, that the longer time limit at the first school would not significantly affect the percent of error scores, since the students who finished the test did not go back and check their work.

ANALYSIS OF RESULTS

The types and numbers of errors are summarized in table 5.1. Inspection of this table discloses that 55 percent of all errors were errors in measurement and that 45 percent of all errors were errors in plotting. Further inspection discloses that reciprocal errors occurred most frequently. Tolerance errors were the next most common type and reversal errors were third. These three types constituted 84

TABLE 5.1.—*Total measuring and plotting errors by type*[1]

Type of error	Measuring		Plotting		Measuring and plotting	
	Number	Percent	Number	Percent	Number	Percent
Reciprocal	631	49.1	124	11.7	765	32.2
Tolerance	164	12.8	538	50.8	702	30.0
Reversal	299	23.3	212	20.0	511	21.8
Reference	148	11.5	47	4.4	195	8.3
Miscellaneous	42	3.3	139	13.1	181	7.7
Total	1,284	100.0	1,060	100.0	[2] 2,344	100.0

[1] For a definition of these types of errors see "Description and Definition of Plotter Errors."
[2] "Combination" errors are included under the two titles applicable in each case. Thus, if an individual made both a reciprocal and reversal error on one problem, the error was tabulated once under reciprocal and once under reversal. There were 113 "combination" errors.

percent of the total number of errors. Thus, variations that reduce these three types of errors should appreciably reduce the total number of plotter errors. High-school advanced-mathematics students made approximately twice as many errors as did rated navigators, but their errors were in the same general proportions of reciprocal, reversal, and tolerance, as shown in table 5.2. These three types of errors constituted 84 percent of the total for both populations.

TABLE 5.2.—*Break-down of plotter errors by types: Comparison of a group of AAF navigators and a group of high school advanced mathematics students*

Type of error	100 AAF navigators (percent of total errors)	348 mathematics students (percent of total errors)
Reciprocal	22	32
Tolerance	43	30
Reversal	19	22
Reference	(1)	8
Miscellaneous	16	8
Total	100	100

[1] This report makes no reference to reference errors. It is not known whether reference errors were never made by rated navigators or whether they were included with the miscellaneous errors. However, the incidence of this type of error is low and has little bearing on the comparison.

Further inspection of table 5.1 will show how the various errors were divided between problems involving measurement of a given vector and problems involving plotting a vector of a specified number of degrees. It can be seen that 631 out of a total of 755, or 84 percent, of all reciprocal errors were made on measurement problems; only 16 percent were made on plotting problems. Measurement would seem to be a simpler operation than plotting. The subject merely placed the plotter on the course, made necessary adjustments, and read the answer. The high incidence of this type of error might be explained by the fact that this process soon becomes highly automatic for an experienced navigator and so, without thinking, the subject apparently often read the most accessible figure and thus made a reciprocal error. Analysis shows that the high-school students made reciprocal errors by erroneous reference to the outer scale more frequently than by erroneous reference to the inner scale by a ratio of 1.21 to 1.00. This proved to be even more strikingly true in the earlier study of navigators carried out by Joseph.[5] He found a ratio of 2.6 to 1.0, using a sample of 100 experienced navigators.

Analysis of Four Basic Factors of Plotter Design

The criterion established for the determination of the relative effectiveness of various basic plotter designs was accuracy. However, the gain in accuracy must not be at the expense of such other important

[5] From an unpublished report by Robert T. Joseph, Aero Medical Laboratory, Wright Field, Dayton, Ohio.

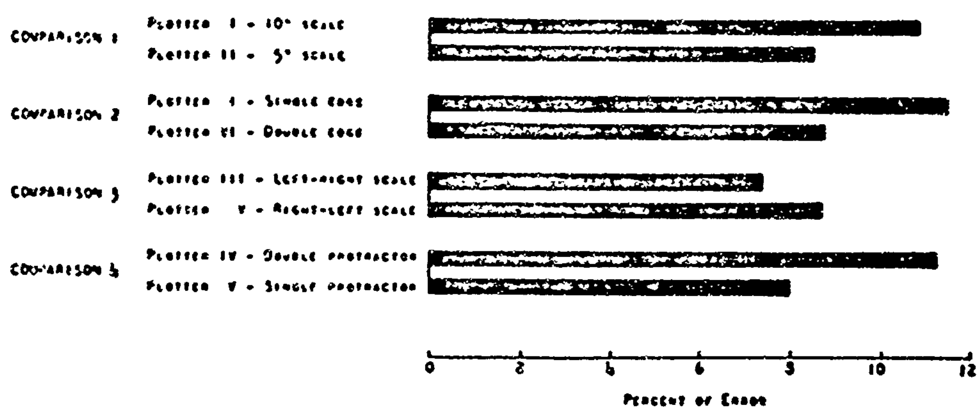


FIGURE 5.7.—Effectiveness of different plotter variations in reducing errors.

factors as speed, size, weight, adaptability to crowded working conditions, and durability.

A "percent of error" score was computed for each subject by dividing the total number of errors committed by the number of problems completed. Values of *t* were computed for the four basic comparisons by using the method of differences between paired scores. The results shown in figure 5.7 indicate that if accuracy were the only criterion the ideal plotter should have a double edge and a single protractor. The question of left-right scale versus right-left scale was not so clear. The p value was .30. However, the mean for the left-right scale was lower, and, as a choice must be made, it would seem preferable to use the left-right scale. The same is true to a lesser extent for the 5° vs. 10° scales, as the p value was only .18. More experimental evidence is being collected on the left-right vs. right-left reading habits and their application to dial and scale design.

TABLE 5.3.—*Distribution statistics of plotter errors and* t *values for four basic plotter comparisons*

Comparison	N	Percent of errors[1]	SD of percent of errors	Means of the difference scores[2]	Range of the difference scores	SD of the difference scores	SE_{Diff}	t	p
Plotter I (10° scale)	84	10.53	10.68	1.54	−27.2 to +35.3	9.90	1.09	1.41	0.18
Plotter II (5° scale)	84	8.99	9.45						
Plotter I (single edge)	97	11.29	8.85	2.56	−30.5 to +37.0	8.60	.87	2.93	.01
Plotter VI (double edge)	97	8.74	8.77						
Plotter III (left-right scale)	67	7.83	11.38	1.19	−35.4 to +33.8	9.27	1.14	1.04	.30
Plotter V (right-left scale)	67	9.02	10.47						
Plotter IV (double protractor)	99	11.15	9.88	3.00	−35.0 to +30.5	9.01	.91	3.30	.005
Plotter V (single protractor)	99	8.15	8.04						

[1] The higher the mean, the greater the percent of error; therefore, the plotter with the higher mean is the worse with respect to accuracy. The error score here is synonymous with the percent of problems wrong. Combination errors (two errors on one problem) are not counted twice.

[2] The difference score of one subject was 4.7 sigma from the mean of the difference scores. His score was omitted.

The distributions of the percent of error scores were appreciably skewed as can be seen in table 5.3. However, the distributions of the difference scores were quite normal. In the interpretation of the mean and sigma values of the difference scores the range should be noted, as a large proportion of the difference scores were negative. It was these difference scores that were used in the calculation of the *t* values.

TABLE 5.4.—*Mean number of problems completed by plotters for four basic plotter comparisons*

Comparison	N	Mean number completed	Difference
Plotter I—10 degree scale	85	34.51	1.04
Plotter II—5 degree scale		33.47	
Plotter I—single edge	97	35.70	
Plotter VI—double edge		36.19	.49
Plotter III—left-right scale	67	34.68	
Plotter V—right-left scale		34.77	.09
Plotter IV—double protractor	99	33.48	
Plotter V—single protractor		36.57	3.09

Table 5.4 presents the pertinent data of the speed analysis. Eight Pearsonian coefficients of correlation were computed between percent of error and number of problems attempted. All were negative and ranged from −.11 to −.46. The median coefficient was −.30. This finding is compatible with the psychological principle applicable to tasks of this nature; namely, that the most accurate performers are, in general, also the most rapid. Perhaps the most significant single conclusion that can be drawn from the data in tables 5.3, 5.4, and 5.5 and figure 5.7 is the fact that error scores can be reduced by making changes in basic plotter design without reducing the speed with which the subjects can handle the instrument. In fact, working speed is generally increased (table 5.4). It seems obvious, however, that the major emphasis should be placed on error reduction and not on speed, as, unless the speed differences are very great, the time the navigator spends with a plotter is relatively unimportant when compared with the importance of obtaining an accurate answer.

Other Findings of Importance

In grading the papers it was apparent that a flat-topped three (3) and a five (5) were confused when inverted. It is believed that a figure 3 with a round top (3) would not be as easily confused with a conventional 5. Such a finding should not be viewed as a contradiction of Berger's findings.[6] His recommended number designs imply that the figures will be read in normal fashion. However, the navigator often finds it necessary to read figures upside down on several

[6] Berger, Curt, "Stroke-width, form, and horizontal spacing of numerals as determinants of the threshold of recognition," *Journal of Applied Psychology*, 28, 1944, 208–231.

of his instruments. When it is likely that readings are to be taken with figures upside down, a flat-topped 3 is not recommended.

It is common knowledge to AAF navigators that the most difficult vectors to measure and to plot are those which are nearly parallel to a meridian. This was also true for the experimental group. The incidence of error on this type of problem was 21.2 percent as compared with 8.9 percent for all other types of problems. Further evidence as to the difficulty experienced with these near-meridian vectors can be obtained by examination of the problems omitted. Only 130 of the 24,823 problems read were omitted, but of these, 62, or 47.7 percent, were problems within 10° of being parallel to a meridian. Yet this type of problem comprised only 16.2 percent of the total problems read.

A RECOMMENDED PLOTTER

As a result of the analysis of the data of this experiment it is believed that an improved plotter should combine the features described in the following paragraphs. A model has been drafted in order to show how the combination of these features would look in an actual plotter. This model is shown in Figure 5.8.

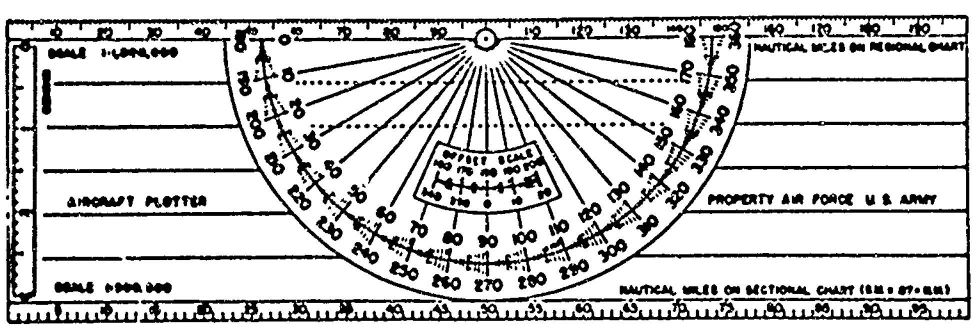


FIGURE 5.8.—A plotter design combining desirable variations of plotter elements.

This plotter has two straight edges instead of one. It has the protractor element set down into the base proper. This feature is highly desirable according to the accuracy criterion and slightly favored from the standpoint of speed. The two-edged plotter never needs to be inverted when measuring or plotting a course near the top of a chart; the user needs merely to slide it down and use the other edge if he finds that he cannot get a reading with the bottom edge of the plotter placed on the vector in question. A double-edged plotter is also far more durable than a single-edged plotter as it has no points of stress where the protractor element joins the base of the plotter. Setting the protractor element into the base also permits the employment of a larger protractor without increasing the over-all width of the plotter. This makes it possible to print the graduation marks farther apart.

This feature should enable the navigator not only to read the scales more rapidly, but also to read them with greater precision. This feature should materially reduce tolerance errors.

The left-right scale of the recommended plotter should materially reduce reversal errors. The left-right reading habit is conventional. Practically all reading materials (e. g., this line of print) are read from left to right. Most scales, such as those on common rulers, increase from left to right. In a stressful or emergency situation and moments of inattention it may be expected that an individual will revert to the most deeply ingrained habit, which is to read from left to right. With the plotter now in use this results in a reversal error.

A "staircase" scale has been employed in the design of the recommended plotter. It is believed that this type of scale will also aid in the reduction of reversal errors. It will be noticed by referring to the recommended model that the larger unit figures are matched with the longer graduation marks. A navigator should learn promptly that if a vector lies on or near a relatively long graduation mark, his answer must be in the 6–9 range in the units place, and, correspondingly, if a vector lies on or near a relatively short graduation mark, the correct digit must be in the 1–4 range in the units place. There is no reversal problem at the numbered increments. Admittedly this does not completely solve the problem of the 5-degree reversal (e. g., reading 95 for 85), but it is hoped that the "staircase" effect will cause the user to realize as he departs from a numbered graduation that the reading must be smaller if he goes down the stairs, and larger if he goes up the stairs. • Unique markings enable easy recognition of the 5-degree graduation marks. A study has been initiated in an effort to determine the effects of the staircase scale on reduction of reversal errors. The experiment is being designed so that the results will have application to a number of problems.

It is believed that a 10-degree scale should be used on the protractor element. A plotter with a 5-degree scale was used more accurately but not more rapidly than a plotter with a 10-degree scale. As can be seen in table 5.5 this increased accuracy was brought about chiefly by a reduction in reversal errors. However, it can be seen by examination of the models with the graduations numbered every 5 degrees that it was necessary to make the figures quite small. For purposes of the experiment, figures of this same size were used on all the experimental plotters. However, these small figures would be extremely difficult to read in vibrating aircraft and under conditions of poor illumination, and for this reason it seems preferable to use larger figures and place them at only the 10-degree graduation marks. It is a debatable point and if the size of the plotter were unrestricted, a 5-degree scale would be preferable. It is felt, however, that the purpose of the

5-degree scale (i. e., reduction of reversal errors) is served as well or better by employment of the staircase scale described above.

TABLE 5.5.—*Error break-down by plotter types*

[Measuring and plotting errors combined]

Plotter model number	N	Total problems completed	Total errors		Reversal		Reciprocal		Reference		Tolerance		Miscellaneous	
			Number	Percent of number completed	Number	Percent of number completed	Number	Percent of number completed	Number	Percent of number completed	Number	Percent of number completed	Number	Percent of number completed
I-10° scale	85	3,982	323	10.8	91	3.1	115	3.9	28	0.9	67	2.2	22	0.7
II-5° scale	....	3,055	260	8.5	56	1.8	82	2.7	32	1.0	68	2.2	22	.7
I-single edge	97	3,544	408	11.5	95	2.7	124	3.5	36	1.0	113	3.2	40	1.1
VI-double edge	...	3,598	315	8.8	63	1.8	102	2.8	27	.8	105	2.9	18	.5
III-L-R scale	67	2,349	173	7.4	33	1.4	61	2.6	21	.9	48	2.0	10	.4
V-R-L scale	...	2,448	212	8.7	43	1.8	74	3.0	23	.9	56	2.3	16	.7
IV-double protractor	99	3,241	363	11.2	77	2.4	108	3.3	16	.5	131	4.0	31	1.0
V-single protractor	....	3,606	290	8.0	53	1.5	89	2.5	12	.3	114	3.2	22	.6
Totals	348	24,823	2,344	9.4	511	2.1	755	3.0	195	0.8	702	2.8	181	.7

The elimination or reduction of the reciprocal error is a more complex problem. It is believed that most of the reciprocal errors are orientation errors and not plotter errors. A competent navigator has a general idea of the direction of a line of position or a vector before he measures or plots it. Thus, if he is going in a northerly direction, he knows that his course cannot possibly be in the vicinity of 180°. Or, if he is heading north and drifting to the right, he knows that the wind vector cannot indicate a wind from 90°. However, the fact that plotter design does play some part in reciprocal errors is evidenced by the finding that navigators and naive subjects both prefer the outer scale to the inner scale. The recommended plotter combines the two graduated scales. It is hoped that this will force the user to make an intelligent choice between the two sides of the scale, each of which is equidistant from the graduations, when he takes a reading. The choice between two sides of a single scale is a specific choice which is made after the mechanical operations of manipulating the plotter and determining the point of intersection between the vector and the scale have been completed.

Results of the experiment indicate that a single protractor plotter is superior to a double protractor plotter. The double protractor plotter appreciably slowed the subjects (see table 5.4) and did not increase their accuracy (see figure 5.7).

The model of the revised plotter has a partial scale near the center of the protractor. This "90° offset scale," as it is termed, is offered as a solution to the problem of measuring or plotting lines or vectors nearly parallel to a meridian. When this scale is used, the parallels of latitude instead of the meridians of longitude are used as zero reference lines. In order to avoid confusion between the offset scale and the regular scales, the offset scale extends only 20° on either side of the zero point. The regular scales can be conveniently used when the vector is greater than 20° from a meridian. It might be argued that this offset scale would be useful only on a Mercator chart, as on that chart the parallels and meridians are straight lines intersecting at angles of 90°. The answer to this is that an estimated 90–95 percent of all the navigator's measuring and plotting is accomplished on a Mercator or grid, and even on such charts as the regional and sectional, the error resulting from use of this offset scale will be found to be negligible.

It is planned to adhere one small semispherical knob 1¼ inches from each end of the plotter. These two knobs (which are not shown in figure 5.8) will make it impossible for the navigator to use the plotter wrong face up. The knobs also make the plotter easier to pick up if it is lying wrong face down, as a tap on one end will bring the other end up in the air, enabling the user to grasp it. This is especially convenient if the navigator is wearing gloves.

The suggested plotter has one scale along each edge for measuring distance in nautical miles. One scale is graduated in the ratio of 1 to 5,000,000 for use on sectional charts; the other is graduated in the ratio of 1 to 1,000,000 for use on regional charts. The factor for converting statute miles to nautical miles is also printed on the plotter. It is believed that these two auxiliary scales are the only ones that are necessary and that additional scales only foster confusion and impair the transparency of the plotter.

The plotter should be of 0.035-inch plexiglass construction. All printing should be on the bottom side so that the possibility of parallax errors in reading the scales will be eliminated. The bottom side of the plotter should be covered with a thin acetate veneer to prevent the printing from wearing away.

It is planned to compare directly the plotter recommended in this study with the plotter employed at present in the AAF. The results of such an experiment will show how much is to be gained by combining all the best features of the experimental plotters and will indicate whether the new plotter can be used with sufficiently greater accuracy then the present plotter to warrant its adoption throughout the AAF.

SUMMARY

1. The present experiment was designed in order that variations in four basic components of air navigation plotters could be studied. Six

models were constructed and tried on 348 high-school advanced-mathematics students.

2. It was found that a double-edge plotter is superior to a single-edge plotter, a left-right scale is probably superior to a right-left scale, a 5° scale is superior to a 10° scale, and one protractor element is better than two.

3. A proposed new air-navigation plotter was designed, using the findings of this study as a working basis.

CHAPTER SIX

Design of Clock Dials for Greatest Speed and Accuracy of Reading in Military (2400-Hour) Time System[1]

WALTER F. GRETHER

INTRODUCTION

People commonly experience difficulty in using the 2400-hour time system which has become standard in military practice. When military time is read from a 12-hour dial, it is necessary to add 1200 to all readings after 12:00 a. m. The mental arithmetic thus required introduces an opportunity for error and also some delay in obtaining the desired figure. On the other hand, 24-hour dials designed to give readings directly in military time are, at first glance, quite confusing to persons who have spent their entire life reading time from 12-hour clocks. In the 24-hour dial, only one of the hourly positions can appear in its conventional location. In addition, an interval of 1 hour on the hour scale corresponds to 2½ instead of 5 minutes on the minute scale. One of the major purposes of this experiment was to find out whether the 12-hour or 24-hour dial can be read more easily when readings are required in military time. A further purpose was to evaluate a number of the possible factors in the design of either type of dial which might influence speed and accuracy with which readings are obtained.[2]

EXPERIMENTAL MATERIALS AND PROCEDURE[3]

With the assistance of personnel in the Instrument and Navigation Branch of the Equipment Laboratory, 11 different designs of clock

[1] This chapter is based upon research findings reported in Headquarters AMC, Engineering Division, Memorandum Report Number TSEAA-694-8.

[2] This problem was suggested by the Instrument and Navigation Branch of the Equipment Laboratory, Engineering Division, Wright Field.

[3] Dr. John T. Cowles assisted in the preparation of experimental materials for this study. The necessary photographic work was done by the Aero Medical Laboratory Publications Unit.

dials were prepared. A sample of each of these designs is presented in figure 6.1. The first five clocks, types A through E, are variations of the 12-hour clock. The remaining 6 are variations of the 24-hour clock. The 11 clock dial designs used in the experiment were selected in order to make possible a comparison of the following variables:

1. A 12-hour vs. a 24-hour dial.
2. Use of numerals vs. no numerals on the minute scale.
3. Use of 1-minute vs. 5-minute graduations on the minute scale.
4. Use of numerals at all hourly positions vs. replacement of some numerals with mere reference marks.
5. Addition of a 13- to 24-hour scale to the 12-hour dial vs. no such scale.
6. Placement of the 24-hour position at the top vs. the bottom of a 24-hour dial.
7. Placement of the 60-minute position at the top vs. the bottom of a 24-hour dial.

Mock-ups of the 11 clocks were prepared with movable hands and then photographed with the hands in different positions to make up the actual items of a printed test. This test was made up in two parts. In part I there were 10 reproductions of each clock face. The different dial designs were intermingled in a predetermined irregular sequence so that the subject was required to change from one dial to another as he worked on the successive items of the test. A time limit was used for the entire part, and thus no speed data could be obtained for any individual dial design. Part II of the test was prepared with 10 reproductions of each dial presented successively, thus making possible the use of a time limit for each of the 11 designs presented. In part II, therefore, both accuracy and speed data could be obtained.

In selecting the time settings to be used in the actual test, an attempt was made to control the average difficulty of the items for all clocks, by equating such factors as number of a. m. and p. m. readings, number of readings at 5-minute positions, average magnitude of minute readings, and the number of hour readings at major positions (i.e, 3, 9, 12, 15, etc.). In determining the sequence in which the test items appeared in part I of the booklet, precautions were taken to insure that there was no grouping of a particular clock near the beginning or end of the test. The actual test items used in part II of the test employed time settings different from those used in part I.

This test was administered to 62 rated military personnel at Wright Field and to 100 advanced-mathematics students in a Dayton high school. All subjects took part I of the test prior to part II. In taking part II of the test, however, approximately one-half of the subjects began at the front of the test booklet. The remaining one-half of each group took part II of the test in reverse order. That is,

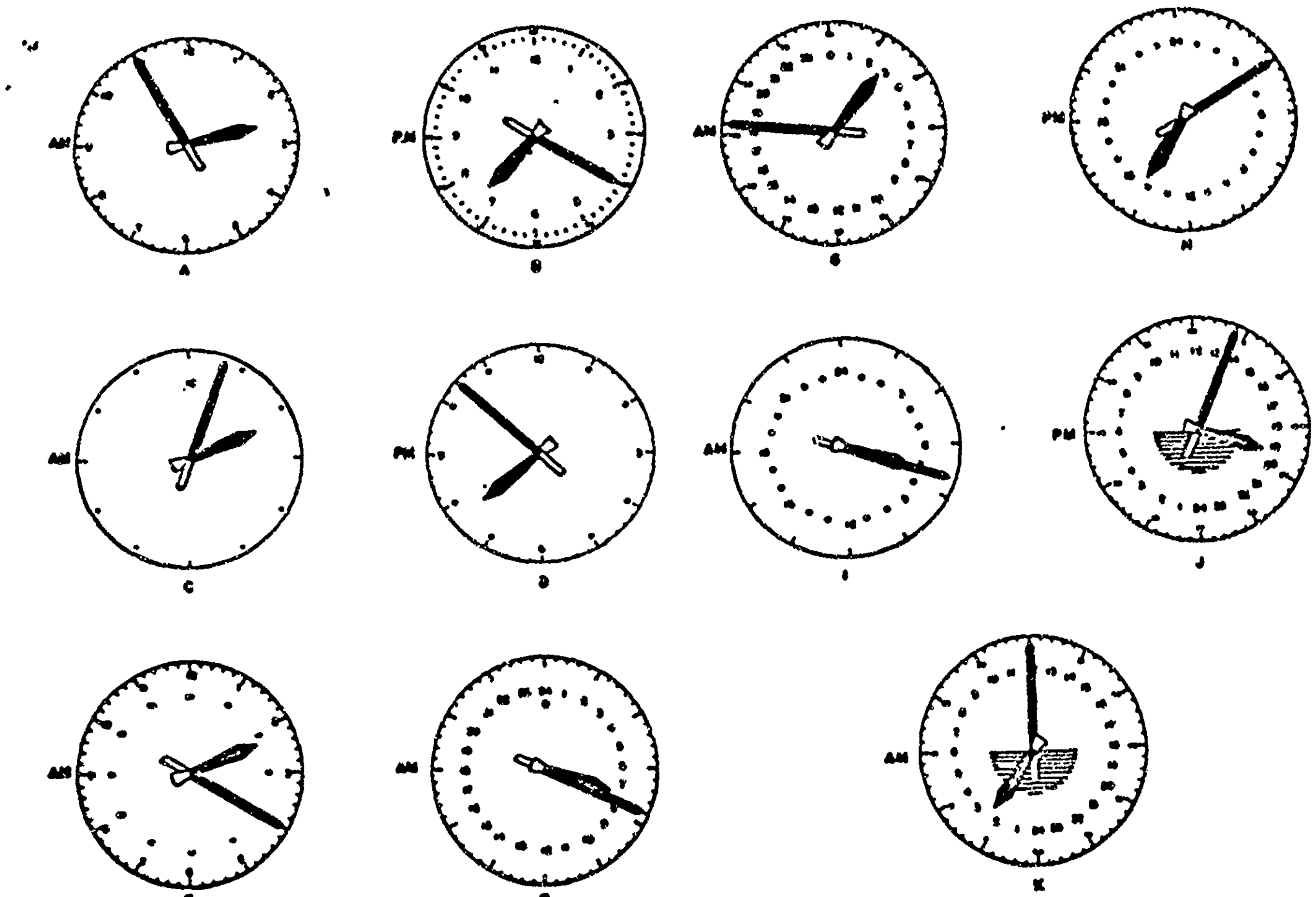


FIGURE 6.1.—Experimental designs used in the study of clock dials for reading in the military (2400-hour) time system.

they first completed the 10 items for the last dial design in the booklet in the order in which they appeared, then those for the second last dial design in the booklet, etc. For the rated military personnel, a time limit of 15 minutes was used on part I of the test and a time limit of 45 seconds on each section of part II of the test. For the high-school students, a 19-minute time limit was used for part I and a 1-minute time limit for each section of part II.

RESULTS

A summary of the major results of this experiment is presented in table 6.1, which shows the percent errors (of one or more minutes) on each clock, for both parts of the test, and for both groups of subjects. The table also shows the time per clock reading in seconds for part II of the test, for both groups of subjects. At the bottom of the table are shown the estimated differences required for significance at the 1-percent level. Wherever the differences in the results for any two clocks are equal to or greater than the differences at the bottom of the table, they can be assumed to be genuine differences and not the result of chance factors.

TABLE 6.1.—*Percent errors and time per reading for eleven experimental clock dials*

Clock type	Rated military personnel $N=62$			High-school students $N=100$		
	Part I	Part II		Part I	Part II	
	Percent errors	Percent errors	Seconds per reading	Percent errors	Percent errors	Seconds per reading
A	7.2	6.4	5.28	13.2	11.8	7.52
B	5.6	7.1	5.39	8.6	13.8	7.88
C	19.0	19.1	5.61	27.4	23.5	8.55
D	8.7	13.3	5.69	13.0	20.9	8.24
E	7.3	14.5	5.34	13.0	23.9	8.10
F	7.4	8.0	4.93	13.3	15.6	6.90
G	4.2	6.8	4.79	6.1	8.4	6.56
H	10.8	17.3	5.40	15.3	22.8	7.95
I	12.8	17.7	5.45	19.6	29.5	7.79
J	7.7	3.6	5.02	14.7	4.9	6.86
K	42.8	14.7	5.64	35.9	14.2	7.82

Significance of Differences

When the average error score for 2 clocks being compared is	*Percent* 5	*Percent* 10	*Percent* 20
The results can be considered significant (1-percent level of confidence) if the differences between clocks are equal to or greater than the following:			
For rated personnel	3.5	4.7	6.2
For high-school students	3.3	4.0	5.4
The results for time per clock reading can be considered significant (1-percent level of confidence) if the differences between clocks are equal to or greater than the following:	*Second*		
For rated personnel	0.20		
For high-school students	.31		

The results presented in table 6.1 are also presented in the form of bar diagrams in figures 6.2, 6.3, and 6.4. It will be noted in table 6.1 and figures 6.2, 6.3, and 6.4 that the data for high-school students and rated personnel present substantially the same over-all picture. In general, also, the differences which appeared among the dials in part

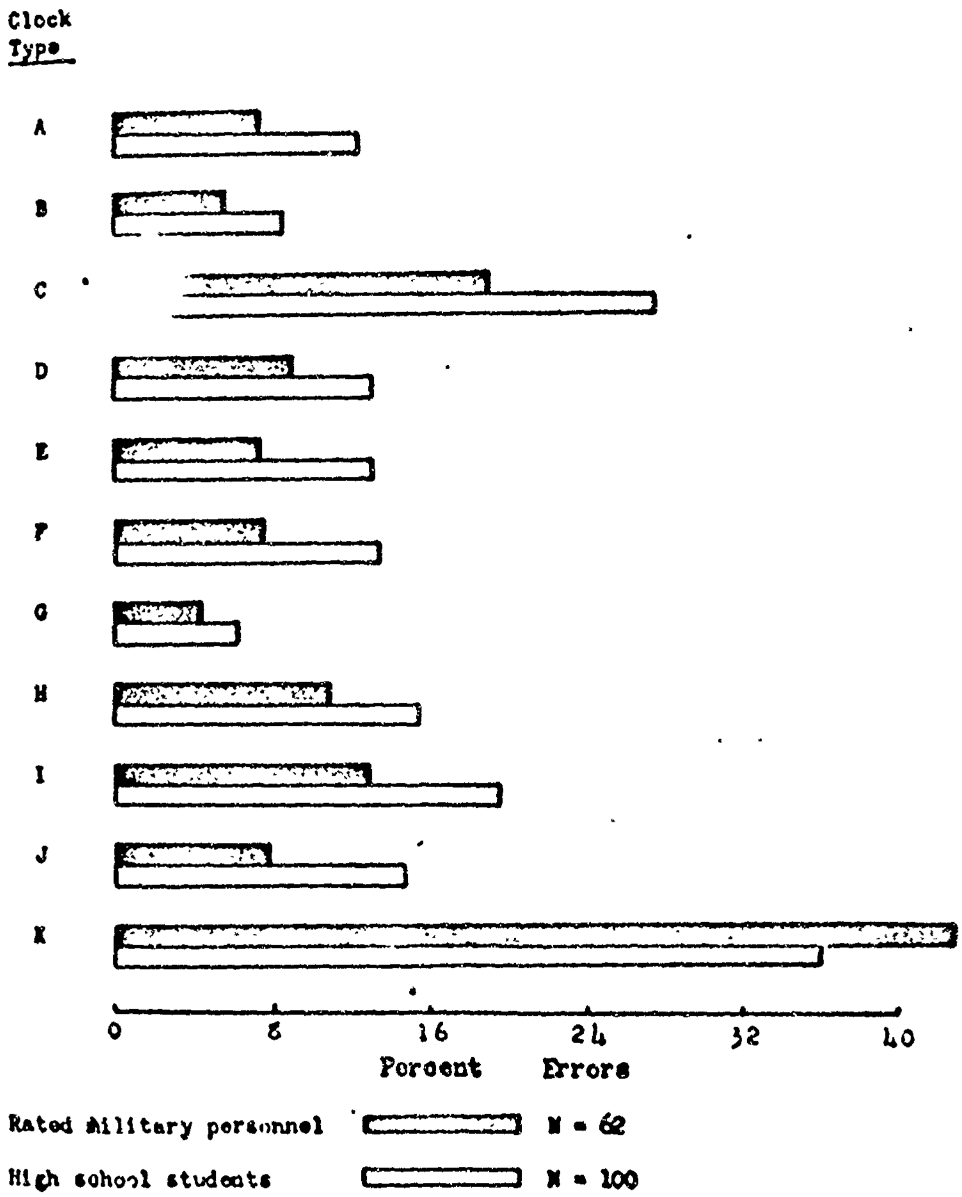


FIGURE 6.2.—Percent errors in military time readings on 11 experimental clock dials (part I).

I of the test reappear in part II. Thus, although many of the differences between dials on one part of the test and for one group of subjects are not significant, the fact that the differences are in the same direction in part II and for the other group of subjects greatly increases the likelihood that the differences are significant.

In general, accuracy of readings was somewhat lower in part II of the experiment, probably because the timing of individual sections motivated the subjects to work at greater speed.

In table 6.2, an analysis is presented of the various types of error made on the different clock dials. Most of the errors were 1 minute,

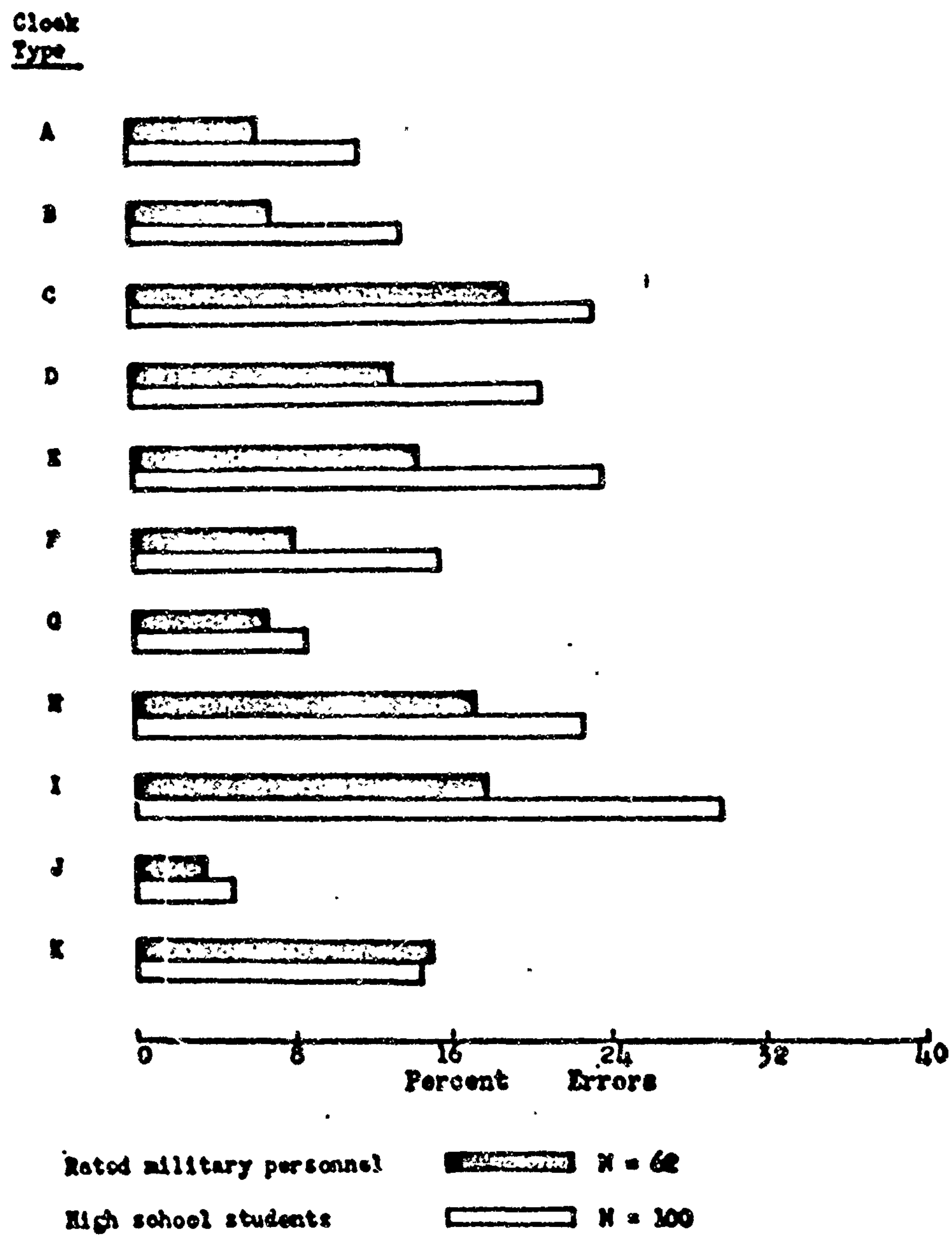


FIGURE 6.3.—Percent errors in military time readings on 11 experimental clock dials (part II).

5 minutes, 1 hour, or 12 hours in magnitude. The frequency of each of these types of error is presented for each dial for a random sample of military personnel and another random sample of high-school students.

Twelve-hour vs. Twenty-four-hour Dial

Comparison of the first five with the last six clocks in table 6.1 shows that there is no major advantage in favor of either the 12- or 24-hour

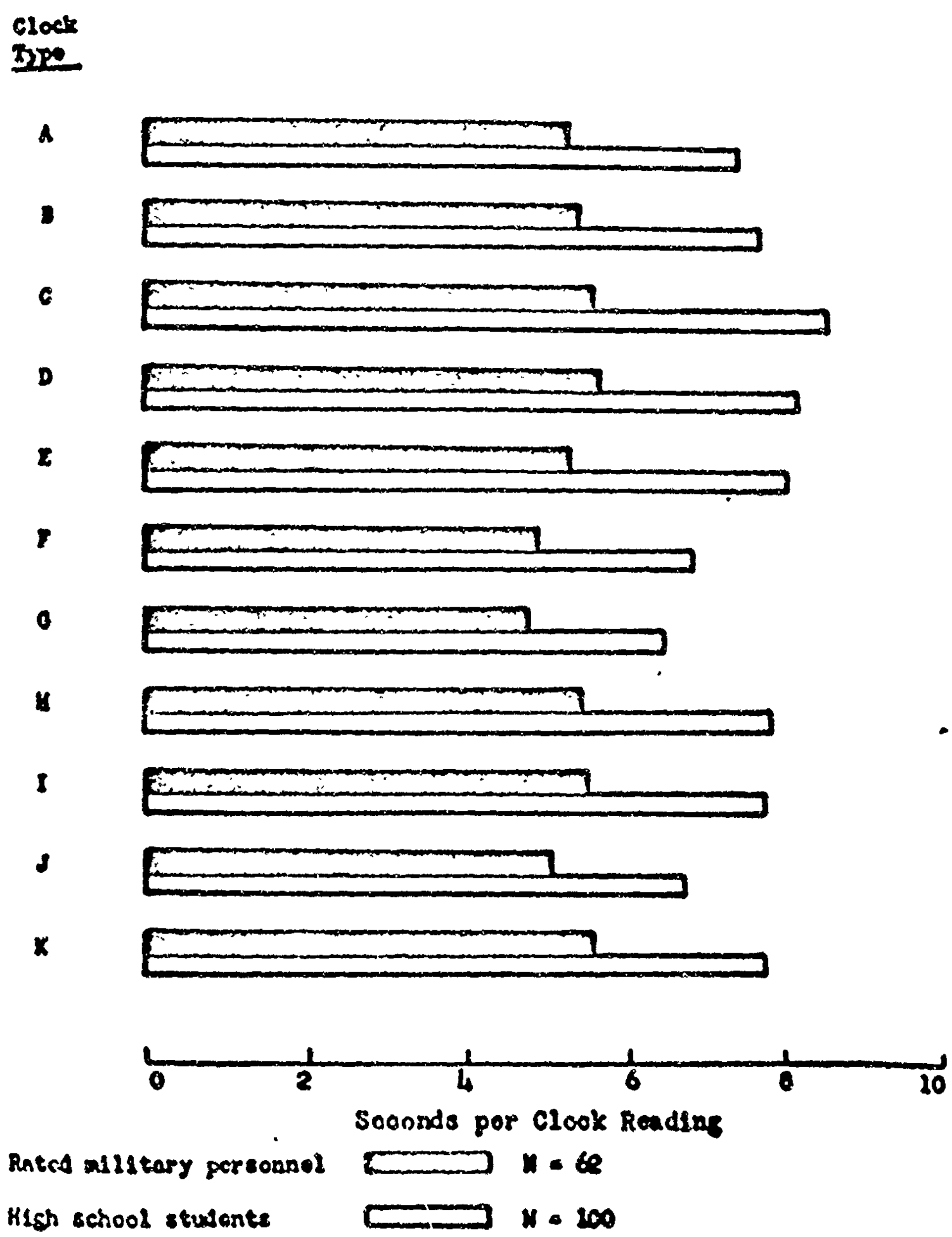


FIGURE 6.4.—Speed of military time readings on 11 experimental clock dials (part II).

dial, although 24-hour dials, types G and J, are superior to the two best 12-hour dials, types A and B. This is particularly true for speed of reading. The 24-hour clock showed somewhat more 1-hour errors, probably because of the smaller spacing of the hour numerals.

TABLE 6.2.—*Frequency of several types of error in reading 11 experimental clock dials (part II of test)*

Clock type	Rated military personnel [N=20] Size of error [1]				High-school students [N=20] Size of error [1]			
	1 minute	5 minutes	1 hour	12 hours	1 minute	5 minutes	1 hour	12 hours
A....	3	5	3	0	2	7	4	0
B....	0	3	1	4	4	1	5	3
C....	10	8	8	4	9	11	13	1
D....	3	9	8	1	5	7	5	3
E....	4	11	8	6	2	11	8	8
F....	2	4	5	0	3	7	12	0
G....	1	4	7	0	1	4	14	0
H....	10	6	6	1	7	6	12	0
I....	16	7	9	0	11	6	10	0
J....	2	1	1	1	1	0	5	0
K....	5	0	3	2	4	0	5	0

[1] Entries in table are total numbers of errors of size indicated, regardless of direction, for 20 randomly selected subjects.

Numerals vs. No Numerals on Minute Scale

The comparison of clocks A and B does not reveal any significant advantage to placing numerals on the minute scale of a 12-hour dial. In the case of the 24-hour dial, however, as indicated by comparison of clocks F and G, there does appear to be a definite advantage in favor of numerals on the minute scale. Dials without numerals on the minute scale show a somewhat higher proportion of 5-minute errors in table 6.2.

One-minute vs. 5-minute Graduations on the Minute Scale

Comparison of clocks A and C, and F and I indicates a significant difference in favor of placing graduations at 1-minute intervals when readings are required to an accuracy of 1 minute. Clocks C and I show a high proportion of 1-minute errors.

Numerals at All Hourly Positions vs. Replacement of Some Numerals With Mere Reference Marks

Comparison of clocks A and D, and clocks F and H indicates a loss in accuracy when numbers are omitted at some of the hourly divisions.

Addition of a 13- to 24-hour Scale on a 12-hour Dial

Clock E, with the 13- to 24-hour scale added was inferior to clock A without such a scale.

Placement of 24-hour Position at the Top vs. the Bottom of a 24-hour Dial

Clock G, with the 24-hour position at the top, was best in part I of the test, whereas J, with this position at the bottom, was best in part II of the test. This would suggest that in a situation where an individual can become accustomed to reading a particular clock, as in part II of the test, there is some advantage to placing the 24-hour position at the bottom of the dial.

Placement of the 60-minute Position at the Top vs. the Bottom of the 24-hour Dial

The results for clock K show quite clearly that the unconventional location of the 60-minute position at the bottom of the dial causes a high percentage of errors and should therefore be avoided.

SUMMARY AND CONCLUSIONS

The purpose of this investigation was to evaluate a number of the possible factors in the design of clock dials which affect the speed and accuracy of readings in miltary (2400-hour) time. Five experimental variations of the 12-hour dial and 6 variations of the 24-hour dial were presented as items in a printed test. This test was designed for obtaining data on both speed and accuracy of readings of the 11 different types of clock dials. This test was administered to 62 rated military personnel and to 100 advanced high-school students. From the results of this investigation the following conclusions are drawn:

1. The best 24-hour dial can be read more quickly and accurately in military time than the best 12-hour dial.

2. Numerals on the minute scale (as contrasted with lack of such numerals) increases clock-reading speed and accuracy, particularly for the 24-hour dial.

3. When readings are required to an accuracy of 1 minute, dials with 1-minute intervals on the minute scale can be read more quickly and accurately than dials with 5-minute intervals.

4. The replacement of some of the hourly numerals with mere dots or reference marks reduces the speed and accuracy of clock reading.

5. The addition of a 13- to 24-hour scale to a 12-hour dial for afternoon readings does not increase the ease of reading a 12-hour dial in military time.

6. The placement of the 24-hour position at the bottom of the hour scale appears to be slightly superior to its placement at the top of the hour scale. This dial arrangement leaves the noon (12-hour) position in its conventional location, and gives the hour hand a logical relation to the rotation of the sun about the earth.

7. Placement of the 60-minute position at the bottom of the minute scale on a 24-hour dial (as contrasted with the conventional location at the top) greatly decreases the accuracy of clock reading.

8. The results suggest that the best dial for readings in military time would be a 24-hour dial with the midnight (24-hour) position at the bottom of the hour scale, with the 60-minute position at the top of the minute scale, with 1-minute graduations on the minute scale, with numerals at 5-minute graduations on the minute scale, with numerals at all positions on the hour scale, and with shading of the lower half of the dial to represent nighttime.

BLANK PAGE

CHAPTER SEVEN

Speed and Accuracy of Dial Reading as a Function of Dial Diameter and Angular Separation of Scale Divisions[1]

WALTER F. GRETHER and A. C. WILLIAMS, Jr.

INTRODUCTION

The pilot of a modern airplane is presented with a large variety of instruments, all of which must be compressed into the relatively small area of the instrument panel. In order to make the best use of the available space, it is important that the size of various dials be in proportion to the accuracy required in their reading. There are, however, no known available data which specify the dial characteristics necessary to obtain a certain degree of reading accuracy. It was the purpose of the present experiment to determine the manner in which the speed and accuracy of dial reading vary as a function of the diameter of circular dials. Since the accuracy of dial reading must at the same time be a function of the spacing between graduations, this factor also was included as a variable to be investigated. In order to make the data applicable to the varied conditions encountered in flight, measurements were made under both simulated day and simulated night conditions.

The purpose in carrying out this experiment was to obtain data which would be useful in defining the characteristics of instrument dials to obtain specified degrees of reading accuracy. In other words, know-

[1] This experiment was carried out at the University of Illinois by Dr. A. C. Williams, Jr., under a "dollar-a-year" contract with the Air Matériel Command, Wright Field. All costs except for the experimental dials and several other components of the apparatus were borne by the University of Illinois. The initial arrangements for this project were made through Dr. John T. Cowles, who later left the university. Dr. Walter F. Grether proposed the study and worked out much of the experimental design. The need for data of the type obtained in this study was expressed by members of the Instrument and Navigation Branch of the Equipment Laboratory, Engineering Division, Wright Field. This chapter is based upon research findings reported in Headquarters, AMC Engineering Division Memorandum Report No. TSEAA-694-1E.

ing the numerical range to be covered, the graduation intervals which most conveniently divide this range, and the reading accuracy required in the use of this instrument, it should be possible from the data of this study to specify the smallest dial size which can be read to the required accuracy with a known statistical probability.

APPARATUS

For the purposes of this experiment a series of 16 simulated instrument dials was prepared, of which a sample is illustrated in figure 7.1.

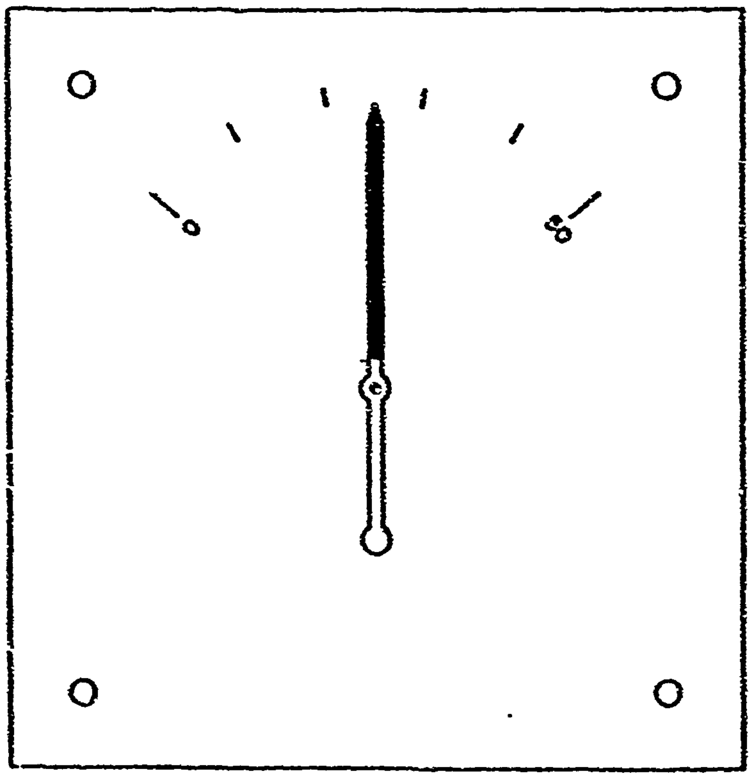


FIGURE 7.1.—Sample Experimental Dial. Actual dials were reverse of above color relationship (actually pale yellow on black). Above sample is of 4-inch diameter and 20° angular separation of scale divisions.

The four following sizes of dials were used: 1, 1⅞, 2¾, and 4 inches in diameter.[2] Each size of dial was produced with four different graduation intervals, defined in terms of the angular separation between scale marks as follows: 5, 10, 20, and 40°. Except for the variations in diameter and graduation intervals, all dials were identical. The width and length of the graduations and the size of the numerals were constant, and all dials covered a range of from 0 to 50 units, with graduation marks at the 0, 10, 20, 30, 40, and 50 positions. The width of the pointers also was constant although the length obviously varied with

[2] 1⅞-inch and 2¾-inch dials were chosen because AAF instruments have been standardized in these two sizes.

dial diameter. These experimental dials were engraved on brass plates. The plates were then painted a dull black and the engraved markings filled with orange fluorescing paint (pale yellow in daylight), as used on the latest type of AAF instruments.

The experimental dials were presented singly in a panel-opening 30 inches from and perpendicular to the subject's eyes. Daylight conditions were simulated with a fluorescent-type daylight lamp which provided an illumination of 45 foot-candles at the panel opening. For simulation of night conditions the subjects' room was completely darkened and the dial illuminated with a standard C-5 ultraviolet aircraft instrument-panel light operating at maximum intensity. No means was available for obtaining a quantitative measurement of the brightness of the scale markings under ultraviolet illumination. Covering the opening in which the experimental dials were presented was a mechanical shutter operated by the experimenter.

On the experimenter's side of the test panel was a carriage on which four of the dials could be mounted side by side. This carriage rode upon two horizontal tracks parallel to the screen. To present any one of the dials, therefore, the experimenter moved the carriage so that the desired dial would appear in the panel opening. At the experimenter's side of the carriage were four master setting dials 5 inches in diameter. On each of these dials was a pointer connected to the same shaft as the pointer on the dial to be read by the subject. On the experimenter's dials were closely spaced graduations which made possible accurate settings in tenths of the space between graduations on the subject's dials.

Also provided at the experimenter's station was a lever for manual operation of the shutter used to expose the dial to the subject. This lever was used also to operate an electric timer through a suitable switch. Thus, the timer indicated the time during which the shutter remained open. Since the experimenter closed the shutter as soon as the dial reading had been completed, the reading on the clock gave a crude measure of the reaction time on each test trial. Several other methods of measuring reaction time were tried but found to be unsatisfactory.

TEST PROCEDURE AND SUBJECTS

Eighty male college students were used as subjects in this experiment. Only men with 20-20 binocular vision (corrected or uncorrected) were accepted for the experiment. The subjects were seated in a chair in front of the screen with their eyes 30 inches from the panel opening and with the line of sight perpendicular to the panel opening in order to eliminate parallax. The subjects were divided into groups of 20, each group being tested on a set of 4 dials. The 4 dials included one of each diameter and one of each graduation interval. Each subject was given a total of 80 trials, equally divided among the 4 dials in a random sequence. Of each group of 20 subjects, 10 were tested under simulated daylight conditions and the remaining 10 under simulated night conditions.

A variety of dial settings was chosen so as to represent all portions of the dial from 0 to 50. The actual numbers to be read were the same for all dials although the order of presentation was randomized. The subjects were instructed to read the dials as quickly and accurately as possible to the nearest whole number. As can be seen in figure 7.1, the reading to the nearest whole number required estimation to the nearest one-tenth of the distance between graduations.

On each trial the experimenter set the pointer of the dial to be presented, then opened the shutter and waited for the subject's verbal response, following which the shutter was closed and the subject's reading and the clock score recorded.

RESULTS

For each dial four frequency distributions were made, each distribution having an N of 200 readings. These were (1) distribution of errors under day conditions, (2) distribution of errors under night conditions, (3) distribution of time of readings under day conditions, and (4) distribution of time of readings under night conditions. The step interval of the error distributions was an error of one digit which, in each case, was equal to one-tenth of the space between graduations. Error distributions were made without regard to sign, all errors being considered positive. Both the error and time distributions were found to be considerably skewed. This was particularly true of the error distributions where, in many cases, the modal error was zero. Because of the highly skewed nature of the distributions, the statistical treatment in this report has been limited to medians and 75th percentile points. Means were computed from the data and found to show substantially the same picture as the medians.

In table 7.1 is shown the median dial-reading error for simulated daylight conditions. The median is presented first in tenths of the graduation interval, which was actually the unit in which the subjects were required to read the dials. Also given in the table is the same error converted into degrees. In table 7.2, corresponding data are shown for simulated night conditions. The data from these two tables have been combined and are shown in graphic form in figure 7.2. A still different picture of the results is given in table 7.3 which gives the 75th percentile error in reading of the same dials.

Table 7.1.—*Median dial-reading error for simulated daylight conditions*

[N for each median is 200 readings]

Dial diameter	Graduation interval			
	5°	10°	20°	40°
	Median error in tenths of interval			
1 inch	2.18	1.78	1.43	1.32
1¾ inches	2.00	1.12	1.24	.78
2¾ inches	1.54	1.20	.94	.81
4 inches	1.49	.97	.88	.98
	Median error in degrees			
1 inch	1.09	1.78	2.86	5.28
1¾ inches	1.00	1.12	2.48	3.12
2¾ inches	.77	1.30	1.88	3.24
4 inches	.75	.97	1.76	3.93

Table 7.2.—*Median dial-reading error for simulated night conditions with ultraviolet lighting*

[N for each median is 200 readings]

Dial diameter	Graduation interval			
	5°	10°	20°	40°
	Median error in tenths of interval			
1 inch	3.10	1.88	1.49	1.21
1¾ inches	1.94	1.42	1.03	.81
2¾ inches	1.57	.93	.91	.81
4 inches	1.39	.90	.79	.91
	Median error in degrees			
1 inch	1.55	1.88	2.98	4.84
1¾ inches	.97	1.42	2.06	3.24
2¾ inches	.78	.93	1.82	3.24
4 inches	.70	.90	1.58	3.64

Table 7.3.—*75th percentile error in dial-reading for simulated day and night conditions combined*

[N for each 75th percentile error is 400 readings]

Dial diameter	Graduation interval			
	5°	10°	20°	40°
	Error tenths of interval			
1 inch	4.55	2.85	2.12	1.76
1¾ inches	3.18	1.85	1.71	1.39
2¾ inches	2.12	1.65	1.52	1.41
4 inches	2.00	1.52	1.43	1.80
	Error in degrees			
1 inch	2.28	2.85	4.24	7.04
1¾ inches	1.59	1.85	3.42	5.56
2¾ inches	1.06	1.65	3.04	5.64
4 inches	1.00	1.52	2.86	6.00

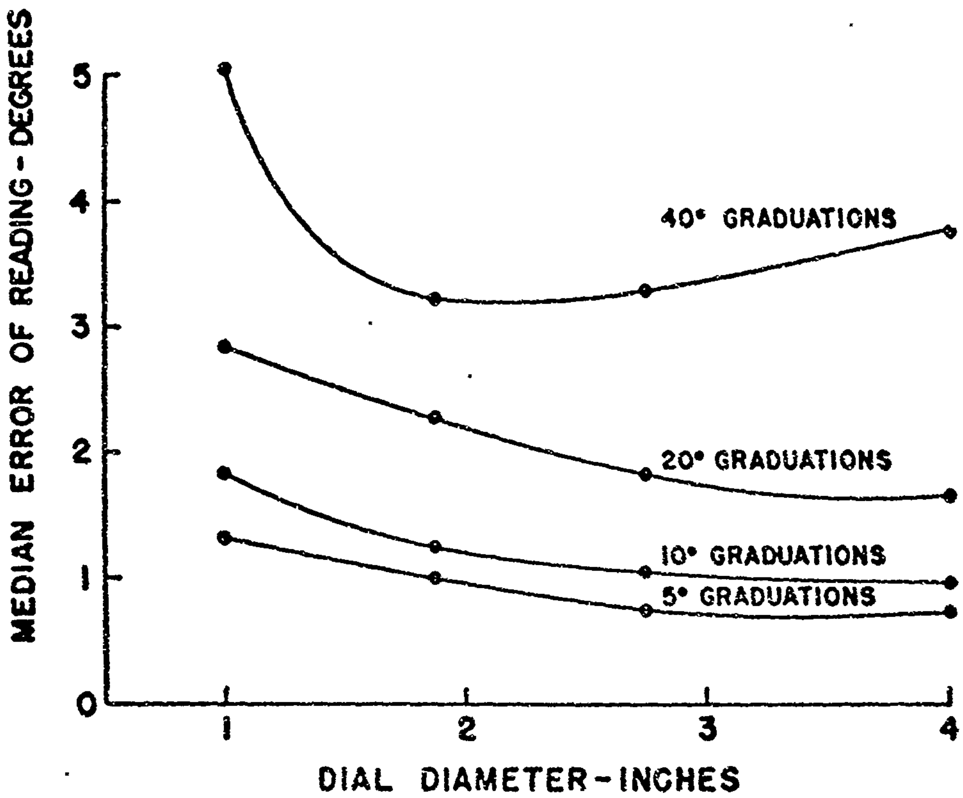


FIGURE 7.2.—Median dial-reading error in degrees as a function of dial diameter for day and night conditions combined.

In table 7.4 the median dial-reading time is shown separately for both day and night conditions.

TABLE 7.4.—*Median dial-reading time for simulated day and simulated night conditions*

[*N* for each median is 200 readings]

Dial diameter	Graduation interval			
	5°	10°	20°	40°
	Seconds per reading for day conditions			
1 inch	1.98	1.78	1.91	1.84
1⅞ inches	1.73	1.80	1.86	1.76
2¾ inches	1.83	1.85	1.73	1.77
4 inches	1.87	1.75	1.68	1.90
	Seconds per reading for night conditions			
1 inch	2.48	2.26	2.10	2.13
1⅞ inches	2.18	2.14	2.02	2.10
2¾ inches	2.35	2.05	2.03	2.06
4 inches	2.00	2.02	2.08	2.23

A still further analysis of the data is given in figure 7.3, which shows the median dial-reading error for each of the 16 dials plotted against the graduation interval in inches. For this purpose the length of

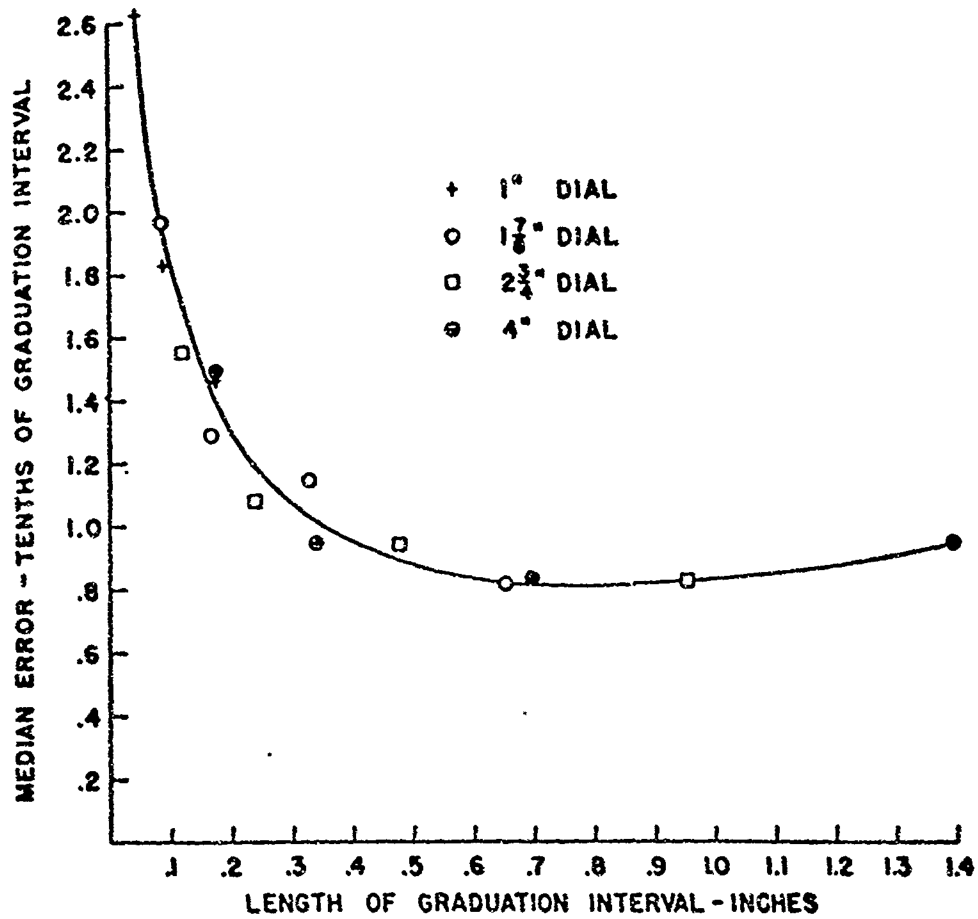


FIGURE 7.3.—Median dial-reading error in tenths of graduation interval plotted as a function of length of graduation interval for day and night conditions combined.

the graduation intervals is measured along the arc of the circle joining the outer edges of adjacent graduation marks. In this graph the two variables of dial diameter and angular separation between scale divisions have, in a sense, been reduced to the single variable of the length of graduation interval.

DISCUSSION OF RESULTS

It is apparent from comparison of table 7.1 and table 7.2 that there is no consistent difference in accuracy of dial reading for day and night conditions. Only the 1-inch dial, with 5° graduation intervals showed any real difference in accuracy under the two conditions. It was because of the general lack of difference that the two sets of results were combined in figure 7.2 and table 7.3. This lack of difference between day and night conditions came as somewhat of a surprise, since it had been anticipated that the low visual acuity in dim light would be reflected in lower accuracy of dial reading. This did not prove to be the case except for the 1-inch, 5° dial on which the graduations were most closely spaced. It must be pointed out, however, that the ultraviolet light source was adjusted to maximum intensity, and that the results might have been quite different had a lower level of intensity been used.

The error results as summarized in figure 7.2 show quite clearly that the accuracy of dial readings is a function of both dial diameter and angular spacing of scale divisions. It is interesting to note, however, that in the case of the 40° dials there is a reversal in the curve and that the error in degrees for the 4-inch dial is larger than for the 1⅞-inch and 2¾-inch dials. This suggests that a dial may be too large as well as too small for greatest reading accuracy. It is quite probable that the same phenomenon would have appeared for some of the other scale intervals had the experiment been carried out to sufficiently large dial diameters. From the data of figure 7.2 it should be possible to estimate with reasonable accuracy the median dial-reading error to be expected from any specific design of dial. It should be remembered, however, that these results are for more or less ideal conditions with little or no parallax, no vibration, and good lighting.

As can be seen in figure 7.3 the two variables, dial diameter and angular spacing of scale divisions, can be replaced by a single variable, length of graduation interval. When plotted in this manner the data for the 16 dials fit a single curve reasonably well. It would appear from figure 7.3 that the accuracy of dial reading, relative to the space between graduations, increases with the size of the graduation interval up to intervals of approximately 0.7 inch. As the graduation intervals are increased beyond this length there appears to be a slight reduction in accuracy.

In studying the time of reading for the different dials as summarized in table 7.4, it is readily apparent that the median time of reading is somewhat higher for night conditions. On the other hand, there appears to be no consistent or systematic relationship between the time of reading and either dial diameter or angular spacing of the divisions. This finding comes as somewhat of a surprise until we recall that the reading time was a relatively crude measure and included the time required for verbalization as well as for perception, plus the experimenter's reaction time in closing the shutter. It is possible that actual relationships exist which were obscured by the crudeness of the time measure. The difference in reading time for day versus night conditions could very well have resulted from the delay between opening of the shutter and the fluorescing of the scale markings. Some time was required for the fluorescent markings to reach maximum brightness after being illuminated by ultraviolet light.

SUMMARY AND CONCLUSIONS

A study was made of the speed and accuracy of reading dials ranging from 1 to 4 inches in diameter and from 5° to 40° in angular separation of scale markings. Reading of the dials was required in units equal to one-tenth of the space between scale divisions. Thus, the problem of reading was primarily one of estimating the position of the pointer between scale graduations. Measurements were made for both simulated daylight and simulated night conditions with ultraviolet illumination. College students at the University of Illinois were used as subjects. The results of this experiment may be summarized as follows:

1. Except in the case of the smallest dial and the smallest graduation intervals, accuracy of reading was not significantly different under the day and night conditions compared, but the ultraviolet illumination used in the night conditions was of relatively high intensity.

2. Accuracy of dial reading increased with increase in dial diameter. For 40° graduation intervals, however, the accuracy of dial readings tended to decrease as dial diameter exceeded 2 inches.

3. Except for very large intervals the accuracy with which the position of a pointer was estimated relative to the space between graduations increased with increase in size of the graduation intervals. The accuracy of dial readings in degrees, however, decreased as the size of the graduation intervals increased.

4. The data on accuracy of readings for all 16 dials were found to fit reasonably well along a single curve when accuracy was plotted as a function of length of graduation interval. Accuracy was found to increase up to graduation intervals of approximately 0.7 inch, and appeared to decline for longer intervals.

5. The speed of dial reading was not systematically related to either size of the dial or the graduation intervals. Although this finding may be a consequence of the crudity of the time measure, a difference was found in reading time for day as compared with night conditions which was probably an artifact resulting from the delay in fluorescence of the dial markings under ultraviolet light.

BLANK PAGE

CHAPTER EIGHT

An Experimental Evaluation of the Interpretability of Various Types of Aircraft Attitude Indicators[1]

ROGER BROWN LOUCKS

INTRODUCTION

The adaptability of the human organism is such that an individual with normal aptitude can learn to interpret almost any type of flight indicator. There are important differences, however, in the ease with which various types of aircraft attitude indicators can be read. Ground-trainer instructors report that beginning students find the conventional artificial horizon or attitude indicator difficult to interpret. Unless extremely vigilant, the beginner tends to move his control stick so as to increase the amount of roll whenever he trys to return to level flight. It has also been reported that experienced pilots with inadequate instrument training sometimes become so seriously disoriented in a cloud bank that they slip into a spin. It would appear, therefore, that there is a certain degree of inherent ambiguity in the conventional type of attitude indicator. The question thus arises as to whether or not some other type of design might not present the aircraft's attitude more clearly.

The relative effectiveness of two instrument designs can be determined by measuring the rapidity with which two representative

[1] The initial report of the investigation to be described in this chapter was circulated 22 June 1945 as Project No. 341, 27th AAF Base Unit, AAF School of Aviation Medicine, Randolph Field, Tex. The experimental work was conducted in the Department of Psychology, AAF School of Aviation Medicine, of which Col. Arthur W. Melton was chief. Recognition is due Colonel Melton for his active cooperation and assistance in this project, and to Lt. Col. Paul M. Fitts, of the Office of the Air Surgeon, for his aid in obtaining the ground trainer and the flight instruments which made the study possible. Special assistance was rendered by Lt. Earl Green, Department of Statistics, AAF School of Aviation Medicine, in the analysis of the experimental data reported in this study. Cpl. James A. Capper, who tested a majority of the subjects in the project, contributed many valuable suggestions for improving the standardization of the testing procedure. S/Sgt. Alvin M. Williams was largely responsible for instituting various maintenance procedures which served to keep the performance of the ground trainer at a stable level of performance. Assistance in the testing of subjects and the processing of the data was rendered by T/Sgt. Austin J. Jernigan, Cpl. Charles L. Phillips, and Pfc. Morton Levin.

groups of novices master the use of the respective indicators. Considered merely from the standpoint of efficiency in pilot training, the instrument which can be correctly read with the minimum of training has an important and practical advantage over other designs. It is also significant that any skill which is acquired with relatively little effort must, for that very reason, be compatible with the subject's habitual and inborn reaction tendencies. The indicator which the novice can interpret most easily provides the minimal possibility of error when used by the more experienced pilot who has been trained on that same instrument. In contrast, the indicator which is difficult to master usually causes confusion because it tends to elicit perceptual-motor responses which conflict with those required for correct interpretation. The ambiguous instrument requires the subject to inhibit or block his immediate or direct response tendency and forces him to substitute an "unnatural" or sophisticated reaction. Inasmuch as an individual's perceptual habits are usually built up over a period of years and, in many instances, are based upon inborn reaction tendencies, an instrument which requires a subject to break up such well-established response systems is to be avoided wherever possible. Not only is a needlessly long period of training required to master such types of instruments, but the possibility never can be completely eliminated that under conditions of stress or fatigue the pilot may revert to his original and more "natural" way of interpreting the instrument. Under the emotional stress of an emergency situation there is always a certain probability that a pilot will misinterpret an indicator if he has had to disrupt or block his "natural" responses in learning to use the instrument. Any indicator which is ambiguous to the novice is thus a potential source of error to even the experienced pilot, particularly in those situations which produce excessive emotional stress or fatigue. The instrument which the novice has difficulty in interpreting is thus not only inefficient from the standpoint of training but a potential source of aircraft accidents.

In attempting to assess the relative effectiveness of a particular aircraft instrument design, the performance of novices should be weighted very heavily. In the experienced pilot, the veneer of training has so thickly covered up the initial, direct, and "natural" responses to the conventional attitude indicator that it is almost impossible to recall the difficulties experienced when starting instrument training. Rated pilots can frequently give valuable suggestions as to particular situations or maneuvers which should be taken into account when evaluating a particular type of instrument. But experienced airmen generally hold widely differing views as to the effectiveness of a particular instrument design or panel arrangement. It is evident that each pilot's opinion arises out of his particular training and experience.

The most effective type of arrangement for one man may cause a serious degree of confusion in another pilot who has had a different type of training. Inasmuch as the subjective impressions of consultants tend to be so very diverse, the only sound basis for a comparative evaluation of two instrument designs is the objective performance of groups of representative individuals whose training and experience are known. As has been pointed out, the speed and accuracy which the novice displays in mastering a particular instrument reflects the degree to which that skill is compatible with his "natural" perceptual and motor habits. Since this would not be true in the case of the experienced instrument flier, because of the specialized training he has received in learning to use specific indicators, the data obtained from the performance of novices is of unique importance.

Although the fundamental data with regard to the interpretability of particular instruments must be obtained from novices, there are certain practical considerations which enter into the over-all evaluation of a specific design. Whenever a new instrument is introduced, there are large numbers of pilots, trained on conventional instruments, who must learn to use the new design. It is essential, therefore, that experienced instrument fliers be tested in order to determine the relative degree of confusion caused by a new design. Where there is the possibility of a choice between two new designs, each of which is superior to the conventional instrument as judged by the performance of novices, the one which causes the experienced pilot the lesser degree of habit interference is obviously the better of the two. In certain instances, an innovation can be introduced without causing the experienced pilot any difficulty because the change is so radical that he reacts to the modified design as if it were a unique instrument and shows no tendency to reinstate the habits established during previous instrument training.

To insure a thorough and comprehensive evaluation of an instrument requires that it be tested in a variety of situations. A particular type of indicator might be easily interpreted so long as the pilot is flying solely by instruments, and using the interior of his cockpit as a fixed point of reference. When the pilot is flying partly by instruments and partly contact, the fixed point of reference may be the true horizon. It is conceivable that an instrument which proves satisfactory in the first situation might lead to confusion in the second. For example, many pilots when flying blind see the cockpit as stationary. As a consequence, the horizon bar and pointer of the conventional attitude indicator are seen to move with regard to the fixed instrument panel. In the second situation where the pilot is flying contact part of the time and using instruments part of the time, it is possible for him to perceive the true horizon as remaining fixed and the instrument panel as rotating about the "fixed" horizon bar and pointer of the attitude indicator. A particular instrument should therefore be checked in each of the above situations. Furthermore, since visual perceptions of relative motion are influenced to a certain degree by accelerations, which affect the semicircular canals and kinesthetic sense organs, data obtained with static or moving ground-trainers should be checked by measurements taken during actual flight. Various pilots have reported a feeling of vertigo in cloud formations when using certain types of attitude indicators and it is obvious that ground-trainer work with such types of instruments should be supplemented by tests in the air.

It is of interest to review some of the assumptions which influenced the specifications set up for the original design of the conventional attitude indicator. The Daniel Guggenheim Fund for the Promotion of Aeronautics, Inc. financed an investigation of problems of fog flying. A research project was organized under the direction of Lieutenant James H. Doolittle, who was stationed at Mitchel Field, and was made available for this work through the courtesy of the United States Army Air Corps. In the first report of this work it is stated that (3, p. 14), "One of Lieutenant Doolittle's projects was the simplification of the instrument board and the procurement of a flight indicator which would be simple and more direct in its indication and require less translation and mental effort on the part of the pilot. . . . The next step toward the simplification of blind flying was the procurement of a single instrument which would replace a number of other instruments giving less direct indications. The pilot, when the visibility is good, depends almost entirely on the attitude of his airplane as seen against the horizon line for determining proper conditions of flight. It was believed that this instrument should take the form of an artificial-horizon line mounted on the instrument board and arranged in such a way that the pilot by looking at it would receive the same information in maneuvering as from the horizon itself. Before initiating the development of such an instrument, two German artificial-horizon instruments, the Anschutz and the Gyrorector, were studied by Lieutenant Doolittle and were not found to be completely satisfactory. The instruments were not only heavy and bulky but the gyros showed a tendency to tumble and the indications were not sufficiently direct. The problem was taken up with the Pioneer Instrument Co. and the Sperry Gyroscope Co. and resulted in the construction by the Sperry Co. of an instrument which gives a direct indication of attitude and appears from its first preliminary tests to be very satisfactory. This instrument is now being redesigned and will soon be placed in production." The initial account of this work was followed shortly afterward by a report on the instruments used in the Full Flight Laboratory (4). In the description of the special flight instruments used in the demonstration flight of September 24, 1929, it is stated that (4, p. 42), "The ideal attitude-indicating instrument for airplane use is one that shows a miniature three-dimensional airplane integral with the airplane itself in relation to a miniature three-dimensional space, the bottom plane of which remains parallel always with the true horizon. Although theoretically ideal, the manufacture of such an instrument presents great mechanical complications. The next best instrument for this purpose is one that gives a two-dimensional representation in which a picture of an airplane maneuvers in respect to a line parallel to and representing the horizon. The poorest scheme of all is the use of an indicating hand, the direct or indirect indications of which must be interpreted by the pilot before the necessary corrections can be made.

"While not quite ideal, the Sperry Artificial Horizon . . . is a great improvement over other existing attitude-indicating instruments . . . The face of the instrument consists of a smooth background, blue on the upper half to represent the sky, shading to a dark gray on the lower half to represent the ground. Horizontally across the middle of this field is a straight bar, both ends of which extend beyond the mask which surrounds the instrument face. This bar simulates the horizon. In front of the bar and the field is a small tail view of an airplane held at the center of the instrument face as part of the mask which tilts with the plane. If the plane climbs or noses down, the horizon bar respectively falls or rises in just the same manner as the actual horizon appears to fall or rise as the pilot looks over the nose. As the plane banks the bar remains horizontal whereas the dial and airplane silhouette tilt with the plane." (In contrast with present-day instruments, this original indicator contained no pointer or 30° marks.) "The horizon bar is thus free to rise, fall, and tilt, and can assume any possible combination of these positions. Further description indicates that (4, p. 47), " . . . the blue sky always remains above the horizon bar no matter how much the instrument or ship is tilted. The horizon bar is carried by a link from the rear part of the gyro casing through a pivot on the gimbal ring and thence out in front of the field as a light horizontal bar. The reason for this reversing linkage is that this corrects the horizon motion observed by the pilot when he climbs or dives. It is the correction of the behavior of this bar to the normal appearance of the real horizon for any position or maneuver that makes flying by this instrument easy for a new pilot and less fatiguing for a pilot who has to fly blind for long periods of time."

The preceding quotations reveal a number of assumptions, some implied, others explicit, which would appear to have influenced the design of the original Sperry attitude indicator. It is not fruitful to discuss the probable validity of such assumptions on the basis of *a priori* or theoretical considerations. The fundamental question is, rather, what do objective and quantitative data actually indicate with regard to the effectiveness of the conventional attitude indicator? And, as noted, at least three criteria must be considered in evaluating such data: First, what type of attitude indicator is most easily and rapidly mastered by the novice? Secondly, of those designs mastered most easily by novices, which particular indicators cause the experienced pilot the minimal amount of confusion? Finally, what instruments meeting the above criteria prove most generally satisfactory when tested in a comprehensive series of different flight situations?

The effectiveness of a particular instrument design must ultimately be verified under actual flight conditions. Because of individual differences in instrument-flying aptitude it may require a minimum of 60 subjects to determine whether there is a reliable and statistically significant difference between the relative effectiveness of two indicator designs. Once the superiority of a particular instrument has been established, it may require a long series of comparative measurements to isolate the basic elements in the design which cause the difference. In consequence, it may be necessary to test several hundred subjects in order to make an analysis that is reasonably adequate. It is apparent, therefore, that the cost of personnel and equipment would be prohibitive were such a study to be conducted exclusively in planes. The mere factor of the time required to process several hundred subjects in test flights makes such a project relatively impractical. It should be noted, moreover, that under actual flight conditions there are so many variables which influence pilot performance that it would be very difficult to determine the proportion of the total achievement which could be attributed to the use of a particular attitude indicator. Fluctuations in air conditions tend to introduce so much variability in performance that the effect of any one variable, such as the use of a particular instrument, may be largely masked. Finally, the problem of objectively quantifying performance during actual flight is, in itself, an extremely difficult task.

In view of the various considerations just presented, it is necessary to conduct the greater part of an instrument evaluation study by making comparisons of various indicators in the ground-trainer situation. This makes it possible to exclude many of the factors which tend to make the true flight situation so variable. The experimenter thus gains much more adequate control of the test situation and can standardize the factors which influence performance. This approach does not obviate the need to verify various instrument comparisons during actual flight. There are numerous air maneuvers which cannot be simulated by the ground trainer.

The present evaluation of the relative effectiveness of four different types of attitude indicators is based upon the objectively scored Link trainer performance of aviation students. A type C-3 instrument ground trainer was employed in this project. The standard instrument panel was replaced by a black masonite board which contained two artificial horizon indicators that were to be compared with each other. These instruments were located near the top of the panel on either side of the center line (see fig. 8.1A). A central vane or shutter made it possible to cover either instrument so that only one was exposed during a particular test period.

The subject's task was to maintain the trainer in level and horizontal flight under conditions of simulated rough air. He was informed that the fuselage was fastened so that it could not turn, even though the pedals could still be manipulated. After an initial period of instruction and demonstration, the hood was lowered and the subject was required to fly the trainer by referring to a specific attitude indicator. Following this initial test period, which was 8 minutes in length, the subject was given a 2-minute rest period. He was then required to fly the Link for a second 8-minute test period. In most of the comparisons to be reported, the even-numbered subjects were tested initially with one instrument and then flew the Link during the second test period by using the other indicator. Odd-numbered subjects were tested on these same instruments in the reverse order. This was done in order to control any major shift in the sensitivity of the trainer over a period of time. To control the effect of position, half of a series of subjects was run with the conventional Link trainer artificial horizon on the left, and the other half of the group was run with this type of indicator at the right.

The standard cam system for providing rough air was removed from its position just back of the pilot's seat, and was connected to a synchronous motor which changed its speed from one r. p. m. to four r. p. m. (see fig. 8.1C). Each 8-minute test session necessitated a continuous performance on the part of the subject, but the scoring system was automatically interrupted for 30 seconds during each successive 2-minute period, for the recording of clock scores.

Independent but similar scoring systems were used to measure the deviations from horizontal and level flight in roll and in pitch (see fig. 8.1 B, D, and F). The No. 1 roll clock records the total time that the fuselage deviates from level flight more than .83° to the left or to the right. The subject can thus roll the fuselage through an angular distance of 1.67° without causing the No. 1 roll clock to score. The No. 2 roll clock starts to score whenever the fuselage rolls more than 1.83° to the right or left from level flight (total permissible arc of

3.67°). Roll clock No. 3 starts to record when the trainer fuselage is more than 2.5° to the left or right of level flight (total permissible arc of 5.0°). Roll clock No. 4 records the amount of time that the subject keeps his stick to the right of center when the fuselage is rolled more than .33° to the left (see fig. 8.1 D and F). The No. 4 roll clock thus registers a qualitatively different score as compared with the first three roll clocks. A similar set of four clocks are set to record deviations from horizontal flight in the pitch direction. These pitch-deviation clocks function in the same way as the roll clocks and are set so as to start recording at approximately similar angular deviations. These clock scores thus represent relative inferiority of performance, or error scores. The higher the score for any clock, the poorer the performance. Clocks 1, 2, and 3, represent increasingly serious categories of error, whereas clock No. 4 registers a qualitatively distinct type of error, namely, failure on the part of the student to keep his stick in the position which would tend to correct the error in the attitude of the trainer.

It should be noted that the conventional AAF attitude indicator used in the Link instrument-trainer consists of a pendulum device rather than the gyroscopic mechanism used in aircraft (see fig. 8.1E). All instruments used in this study contained the special sensitizing modification described in AAF T. O. No. 25-5-12. The sensitivity of the artificial horizon instruments employed can be stated in terms of clock scores. The No. 1 roll clock starts to record when the moving pointer of the indicator deviates more than 5° to the right or left. The No. 2 roll clock starts to record when the moving pointer deviates more than 10° to the right or left. The No. 3 roll clock starts to record when the pointer deviates more than 15° to the right or left. The No. 4 clock, which gives a qualitatively distinct type of score, cannot be equated in these same terms inasmuch as the position of the stick determines, in part, when the timer starts to record. If, however, the moving pointer of the attitude indicator deviates more than 3° to a particular side, and the stick is not moved to the opposite side of its central position, the No. 4 roll clock will start to record.

The clock-score data to be presented in this study are derived solely from the roll clocks. Inasmuch as the pitch of the trainer is registered in a conventional way and with comparable sensitivity by all of the horizon indicators employed in this study, measurements of performance for these deviations are not of any particular importance. That is to say, the differences between instruments which are being compared at this time are variations in the manner by which the roll or bank of the fuselage is represented on the respective indicators.

A further reason for making comparative evaluations of various attitude indicators on the basis of roll scores rather than pitch scores is based upon the fact that the subject's weight under certain con-

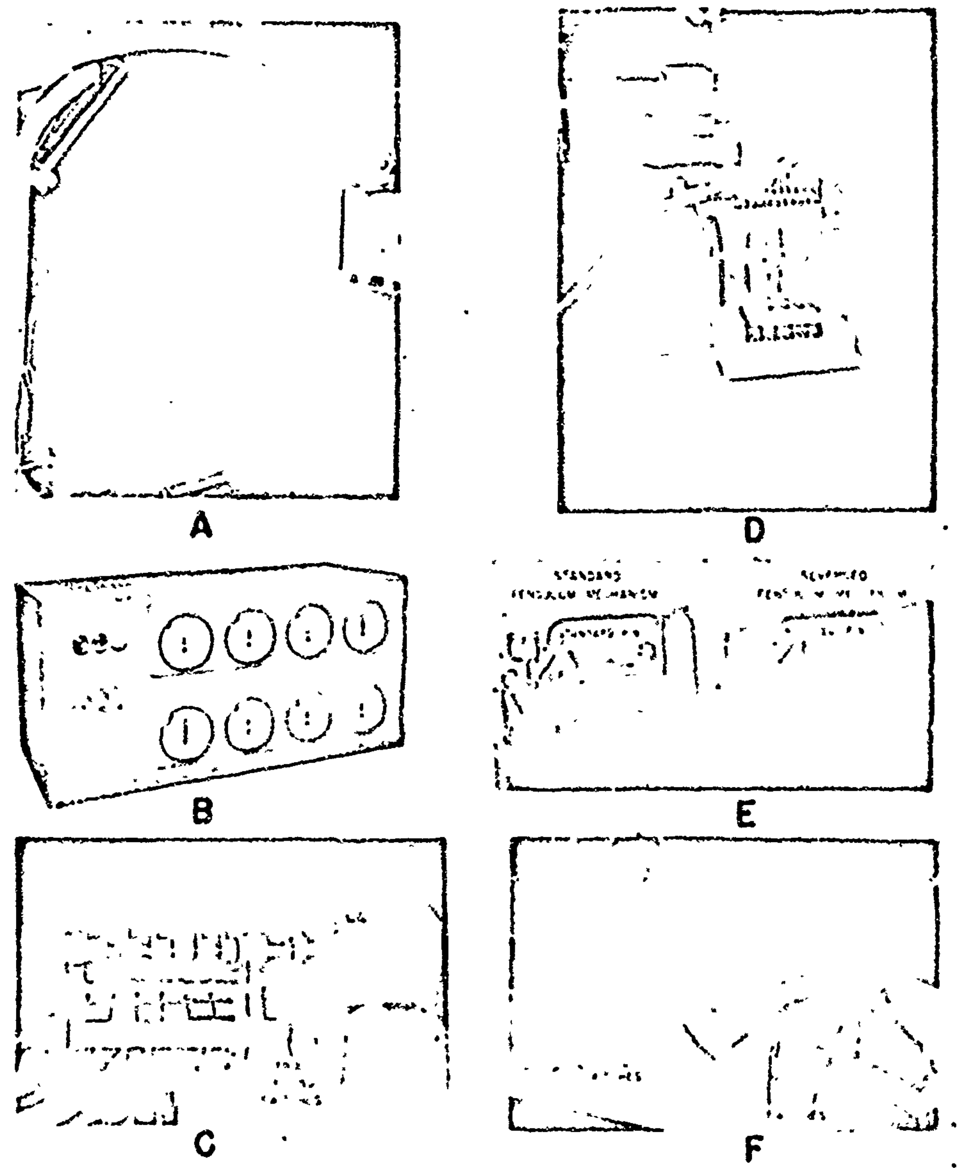


FIGURE 8.1.—Apparatus used in the experimental comparisons of flight indicators. A. Panel used in comparing artificial horizon indicators. B. Control unit for instrument trainer. C. Rough air cam mechanism. D. Assembly for scoring deviations from level flight. E. Trainer horizon pendulum mechanism for standard and reversed rotation indication. F. Assembly for scoring stick position with relation to trainer attitude.

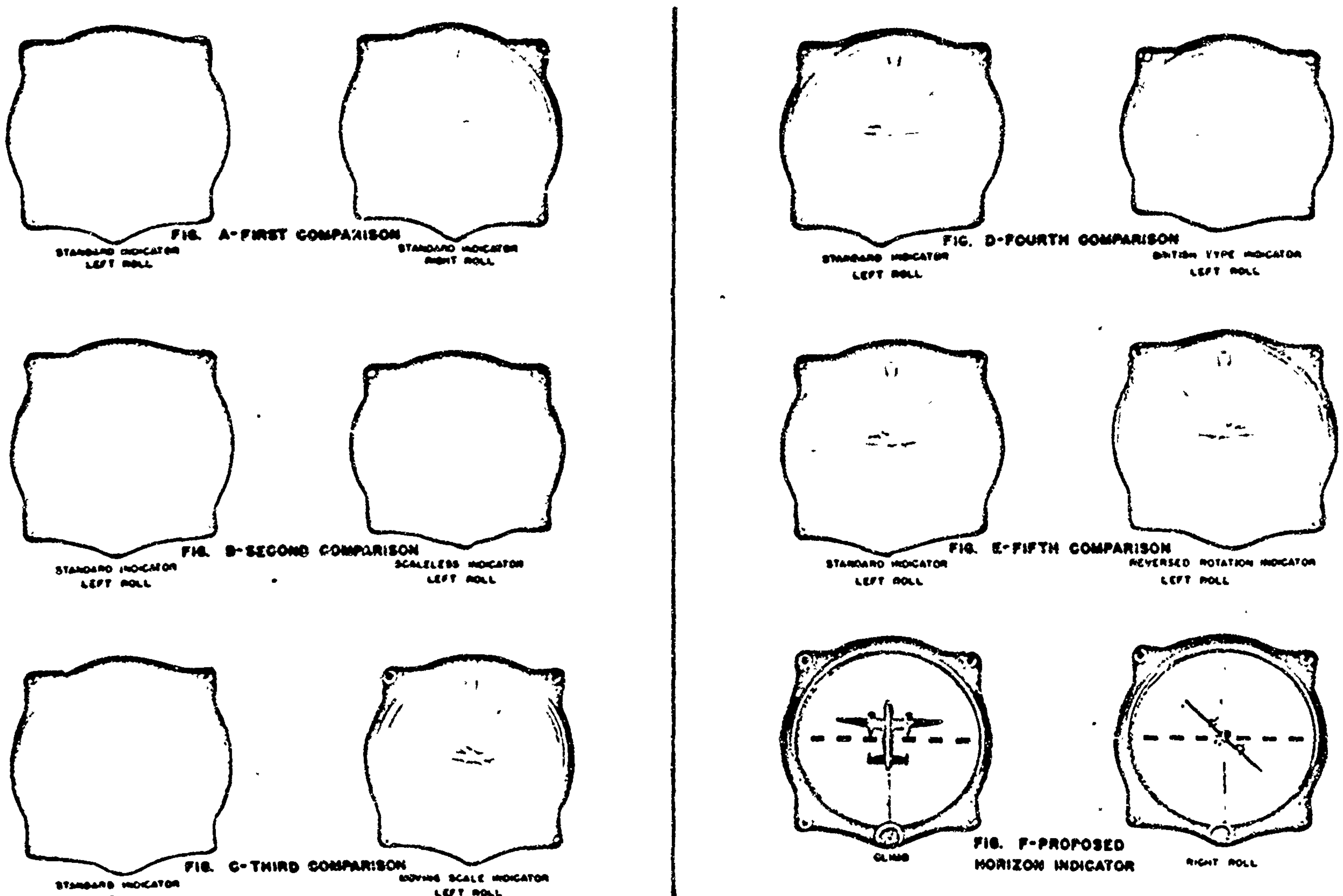


FIGURE 8.2.—Artificial horizon indicators used in the five comparisons, and the proposed horizon indicator.

ditions, tends to influence the magnitude of the pitch scores. When the C-3 trainer was first set up in the test situation there was a noticeable sluggishness of reaction if a man weighing 190 pounds or more tried to shift the trainer fuselage from an attitude of maximal climb to horizontal flight. In consequence, all of the main bellows which govern the attitude of the trainer fuselage were replaced. This appeared to reduce the sluggishness of reaction in the pitch dimension. The relationship between performance and body weight following renewal of the trainer bellows was computed for a group of 169 subjects who used the standard attitude indicator during an 8-minute test period. The correlation coefficients are as follows: No. 1 pitch clock, .11; No. 2 pitch clock, .10; No. 3 pitch clock, −.02; No. 4 pitch clock, —.17; No. 1 roll clock, .02; No. 2 roll clock, .04; No. 3 roll clock, 0.06; No. 4 roll call clock, 0.00. It would appear from these values that with the installation of the new bellows, weight had no appreciable effect on the roll scores and only a minimal influence on the pitch scores.

EXPERIMENTAL RESULTS

To determine the consistency with which performance is measured by the scoring system, odd-even reliability coefficients were computed for a group of 169 subjects who used the standard AAF attitude indicator. The reliabilities, based upon the sums of the two odd and the two even trials in the first 8-minute test period, are as follows: No. 1 roll clock, $r=.83$; No. 2 roll clock, $r=.87$; No. 3 roll clock, $r=.87$; No. 4 roll clock, $r=.82$. These coefficients, uncorrected for attenuation, show that the present test situation evokes performance that has a consistency of a high order. The intercorrelations between roll clock No. 4 and roll clocks Nos. 1, 2, and 3 are, respectively, .62, .63, and .61. The relatively moderate correlations between scores on the first three clocks and the fourth clock, taken in conjunction with the high odd-even reliabilities, tend to verify the hypothesis that this last score (i. e., for roll clock No. 4) is, to some extent, representative of a different aspect of performance than that registered by the first three clocks. The fact that the first three clock scores correlate highly with each other is ambiguous owing to the fact that when No. 2 clock scores, No. 1 must continue to score, and when No. 3 scores, both No. 1 and No. 2 must score.

Comparison 1: Two Similar Link Horizon Indicators

This comparison was designed as an experimental control. Its objective was to determine what differences in performance scores can arise which are due merely to variations in sensitivity of two similar instruments. The first half of the total population of 102 subjects used in this comparison was tested in such an order that the odd-numbered subjects used the No. 1 standard indicator in the right position for the first 8-minute test period and used the No. 2 standard indicator in the left position for the second 8-minute period. Before running the second half of the population, the instruments were reversed in position so that the No. 1 indicator was on the left and the No. 2 indicator was on the right. The second half of the population then took the test in such an order that the odd-numbered subjects were tested first with the No. 2 standard, and during the second 8-minute period used the second of the two instruments.

In analyzing the data, scores for the first 8-minute test period for indicator No. 1 are compared with indicator No. 2 scores for this same initial test period, the position of the instrument being disregarded. Similarly, scores for the second 8-minute test period obtained with the No. 1 indicator are compared with the second test period scores on the No. 2 indicator, irrespective of the position of the instrument. An inspection of table 8.1 shows that none of the four clock scores yield a significant difference as indicated by the associated t values.

TABLE 8.1—*Comparison of two standard AAF-type Link artificial horizon indicators in order to determine apparatus differences* [1]

[Comparison 1]

FIRST 8-MINUTE TEST PERIOD

Experimental group	Instrument	Number of subjects	Statistic	Clock— 1	2	3	4
A	No. 1 AAF standard	51	M	382.6	244.2	157.6	147.6
B	No. 2 AAF standard	51	M	369.0	229.3	150.8	141.9
			Diff	13.6	14.9	6.8	5.7
			t	.5	.7	.3	.5

SECOND 8-MINUTE TEST PERIOD

Experimental group	Instrument	Number of subjects	Statistic	Clock— 1	2	3	4
B	No. 1 AAF standard	51	M	312.5	152.1	73.5	124.3
A	No. 2 AAF standard	51	M	294.6	151.1	82.3	114.5
			Diff	17.9	1.0	−8.8	9.8
			t	1.1	.1	.6	1.0

SUBJECT PREFERENCE SUBSEQUENT TO LAST TEST PERIOD

No. 1 AAF standard instrument given first			No. 2 AAF standard instrument given first		
Prefer 1	No preference	Prefer 2	Prefer 1	No preference	Prefer 2
Percent 21	*Percent* 10	*Percent* 69	*Percent* 69	*Percent* 14	*Percent* 17

[1] The higher the mean clock score, the poorer the performance. A t-value of 2.32 is significant at the 2-percent level; and a t-value of 2.58 or higher is significant at the 1-percent level.

This means that as far as Link-trainer performance is concerned, the two indicators do not show a difference that has statistical significance. Any difference in sensitivity that may exist between the two instruments is unimportant. It is of interest that subjects tend to prefer the instrument they have used most recently, even though the two are identical.

Comparison 2: Standard Link Artificial Horizon vs. Scaleless Indicator

In the conventional horizon indicator a right roll produces a shift in the moving parts of the mechanism so that as the pilot views the indicator the pointer has rotated to the left of the central index mark on the fixed scale at the periphery of the instrument and the horizon has tilted to the left. A right roll of the plane thus produces a counterclockwise rotation of the moving parts of the indicator. There are various possible sources of ambiguity in this relationship between the true attitude of the plane and the correct movement of the stick as indicated by the setting of the attitude indicator. Not all of these factors will influence the responses of each subject, but they constitute a potential source of error in any group of naive subjects, and they may occasionally cause errors in experienced pilots under conditions which necessitate a distribution of attention.

When the pilot is flying exclusively by instruments, his frame of reference tends to be imposed by his immediate surroundings, his cockpit. His line of regard does not rotate with respect to the instrument panel, and the case of the attitude indicator remains in a fixed position with regard to his body. Instrument-flying instruction is designed to make the pilot disregard his body sensations because of their misleading cues. As a consequence, the unchanged relationship between the instrument panel and the pilot's line of vision tends to reinforce the traditional relationship in which the main index of the attitude indicator is at the top and represents "up." When the plane rolls to the right, the pointer and horizon bar rotate counterclockwise, as noted above, and because moving objects tend to attract attention, the novice's first and most direct perception is that the *movement* registered is a left tilt. The pointer is to the left of the fixed index mark. By tradition, that part of an instrument which moves is designated to convey the intended change in relationships. In the case of the artificial horizon, however, he must learn to regard the pointer as fixed and the scale as having moved, a relationship which is clearly antagonistic to his immediate perception of the fixed relationship of the scale to himself. Many novices report that it is difficult to remember to bring the center of the fixed scale up to the moving pointer. Others try to disregard the pointer and focus their attention on the miniature airplane which maintains the correct static relationship to the horizon bar. When the miniature plane is tilted to the right in respect to the horizon bar, a left movement of the stick or counterclockwise rotation of the wheel corresponds to the movement the aircraft itself must execute to attain level flight.

Because many experienced pilots report that they always observe the position of the miniature plane and disregard the moving pointer and because compensatory adjustments of the stick or control wheel conflict with the direction in which the pointer moves, a conventional indicator was modified by removing the pointer and obscuring the fixed scale at the periphery of the instrument (see fig. 8.2B). This modified instrument is designated as the "Scaleless Horizon Indicator." Comparison 2 thus involves an evaluation of Link performance scores when subjects use the conventional attitude indicator and when they use the "Scaleless Horizon Indicator." The experimental routine is similar to that used in comparison 1, and the differences between instruments, shown in table 8.2, are taken irrespective of position. An inspection of the mean scores for the 115 subjects used in this comparison reveals the fact that in their initial 8-minute test session, the standard instrument yields poorer performance than the scaleless instrument. The variation between subjects is so great, however, that the difference between means is not sufficiently important to overweight the individual differences in skill between subjects of the present groups. The more skilled individual can largely overcome any handicap that might be inherent in the conventional instrument under these test conditions. In a test situation requiring a greater distribution of attention he might occasionally be misled by the standard instrument whereas the modified instrument might have less tendency to confuse him.

TABLE 8.2.—*Standard Link artificial horizon vs. scaleless horizon indicator*

[Comparison 2]

FIRST 8-MINUTE TEST PERIOD

Experimental group	Instrument	Number of subjects	Statistic	Clock— 1	2	3	4
A	AAF standard	58	M	365.5	226.6	143.2	108.4
B	Scaleless	57	M	344.0	197.2	116.1	97.3
			Diff	21.5	29.4	27.1	11.1
			t	1.3	1.5	1.5	1.2

SECOND 8-MINUTE TEST PERIOD

Experimental group	Instrument	Number of subjects	Statistic	Clock— 1	2	3	4
B	AAF standard	57	M	297.6	148.4	71.0	76.9
A	Scaleless	58	M	278.3	133.9	65.3	82.0
			Diff	19.3	14.5	5.7	−5.1
			t	1.2	.9	0.5	0.6

SUBJECT PREFERENCE SUBSEQUENT TO LAST TEST PERIOD

No. 1 AAF standard instrument given first			Scaleless horizon indicator given first		
Prefer 1	No preference	Prefer scaleless	Prefer 1	No preference	Prefer scaleless
Percent 16	*Percent* 0	*Percent* 84	*Percent* 74	*Percent* 0	*Percent* 26

Comparison 3: Standard Link Artificial Horizon vs. Moving Scale Indicator

It has been noted that novices frequently report that it takes constant vigilance to remember to move the stick so as to bring the central index point of the fixed scale at the periphery of an instrument to coincide with the moving pointer, instead of moving the stick in the direction the pointer should travel to attain coincidence with the center of the fixed scale. To correct this particular difficulty the No. 2 artificial-horizon mechanism used in the first comparison was modified a second time so that the scale was attached to the inner shield or face which carries the rotating pointer (see fig. 8.2C). The original fixed scale at the periphery was obscured except for the large central index mark. Sixty subjects were used in this comparison. The experimental routine was that employed in comparison 1. Differences shown in table 8.3 are for the two instruments irrespective of position. Again it is evident that while the average error scores tend to be lower for the moving scale instrument than for the standard indicator, the variation in performance from subject to subject is such as to mask any statistically significant difference that is due to the instruments in and of themselves. It must be remembered that the subjects were attending to one instrument, and it is conceivable that the stress of attending to a number of indicators might increase differences in

TABLE 8.3.—*Standard Link artificial horizon vs. moving scale indicator*

[Comparison 3]

FIRST 8-MINUTE TEST PERIOD

Experimental group	Instrument	Number of subjects	Statistic	Clock— 1	2	3	4
A	AAF standard	30	M	374.7	212.5	154.5	107.9
B	Moving scale	30	M	348.6	202.0	122.4	95.3
			Diff	26.1	40.5	32.1	12.6
			t	1.1	1.4	1.2	1.0

SECOND 8-MINUTE TEST PERIOD

Experimental group	Instrument	Number of subjects	Statistic	Clock— 1	2	3	4
B	AAF standard	30	M	289.0	143.3	72.2	72.8
A	Moving scale	30	M	276.9	129.7	60.4	77.5
			Diff	12.1	13.6	11.8	−4.7
			t	.5	.6	.8	.4

SUBJECT PREFERENCE SUBSEQUENT TO LAST TEST PERIOD

No. 1 AAF standard instrument given first			Moving scale indicator given first		
Prefer 1	No preference	Prefer moving scale	Prefer 1	No preference	Prefer moving scale
Percent 7	*Percent* 10	*Percent* 83	*Percent* 73	*Percent* 7	*Percent* 20

performances attributable to the two horizon indicators. It should also be noted that while the movement of the stick required to bring aircraft to level flight is in the same direction as the movement of the fixed central index, were it to be conceived of as having to be rotated to the central point of the scale, the actual motion which is perceived under these circumstances is in conflict with the required rotation of the stick or wheel.

Comparison 4: Standard Link Artificial Horizon vs. British Type of Artificial Horizon

It has been suggested that one model of the British type of artificial horizon is easier to use because it has, inherently, less ambiguity in its action than the AAF instrument. To test this hypothesis, a conventional Link artificial horizon was modified so as to function as the British type of attitude indicator (see fig. 8.2D). The scale was moved from the fixed rim at the top of the instrument to the lower edge of the instrument. The miniature plane was then supported from the upper rim of the instrument as in the British type of indicator. As the pilot views the instrument, a right roll tilts the horizon bar to the left and the pointer moves around its fixed scale to the right. This instrument has the distinction that the correct movement of the stick, considered merely as to the right or the left is in the same direction as the movement of the pointer in relation to the central index of the fixed scale. Various individuals report, however, that this instrument causes a certain degree of confusion because the moving elements of the instruments are perceived as rotating in a counterclockwise direction when the plane is actually rolling to the right in a clockwise direction. A corrective counterclockwise rotation of the stick about its point of support or a counterclockwise rotation of the control wheel are in the same direction as the preceding movement of the elements in the indicator. That is to say, the motion of the indicator must be followed by the same type of movement in the control rather than by a compensatory adjustment in the opposite direction which experience in everyday life tends to establish as the correct, and therefore, the most immediate or natural reaction. In other words, the correct response is not the most immediate response to the angular rotation of the moving elements in the instrument. It must be made to an inferred position of the plane which is interpreted from the indicator.

As in the previous comparisons, the differences indicated in table 8.4 tend to favor the modified instrument but, except for clock No. 4, are not statistically significant. Individual differences are so great that the effect of the variations in the indicators, *per se*, tends to be masked. The difference for clock No. 4, which favors the British type of instrument, is significant at the 2-percent level.

TABLE 8.4.—*Standard Link artificial horizon vs. British type of artificial horizon*

[Comparison 4]

FIRST 8-MINUTE TEST PERIOD

Experimental group	Instrument	Number of subjects	Statistic	Clock— 1	2	3	4
A	AAF standard	32	M	346.9	200.4	117.1	80.0
B	British type	31	M	307.9	156.8	84.0	62.9
			Diff	39.0	43.6	33.1	17.1
			t	1.7	1.6	1.3	[1] 2.4

SECOND 8-MINUTE TEST PERIOD

Experimental group	Instrument	Number of subjects	Statistic	Clock— 1	2	3	4
B	AAF standard	31	M	263.1	151.7	75.9	72.6
A	British type	32	M	299.5	114.6	54.4	53.7
			Diff	−36.4	37.1	21.5	18.9
			t	1.6	1.6	1.1	[1] 2.4

SUBJECT PREFERENCE SUBSEQUENT TO LAST TEST PERIOD

No. 1 AAF standard instrument given first			British type indicator given first		
Prefer 1	No preference	Prefer British type	Prefer 1	No preference	Prefer British type
Percent 19	*Percent* 6	*Percent* 75	*Percent* 48	*Percent* 0	*Percent* 52

[1] A *t* value of 2.32 is significant at the 2-percent level; and a *t* value of 2.58 or higher is significant at the 1-percent level.

Comparison 5: Standard Link Artificial Horizon vs. Reversed Rotation Type of Artificial Horizon

Because in none of the instruments used in the previous comparisons do the moving elements rotate in a clockwise direction when the plane rolls to the right or exhibit a counterclockwise rotation when the plane rolls to the left, a standard Link artificial horizon was modified so as to provide this relationship (see fig. 8.2E). This was achieved merely by lengthening the pin which engages with the bank pendulum (see fig. 8.1E). Table 8.5 shows that the average clock error score made with the reversed rotation type of artificial horizon is lower than that made with the standard indicator and that the differences are statistically significant at the 1-percent level. Tables 8.1 through 8.4 show that the individuals who served as subjects in the first four comparisons generally expressed a preference for the particular instrument which they used in the most recent test period. In this fifth comparison, those subjects who were tested on both the standard and reversed rotation indicators showed a decided preference for the modified instrument regardless of whether or not that was the instrument most recently used. Thus both in the performance which it evokes and in expression of preference on the part of the subjects, the reversed rotation artificial horizon proves to be clearly superior.

TABLE 8.5.—*Standard Link artificial horizon vs. reversed rotation type of artificial horizon*

[Comparison 5]

FIRST 8-MINUTE TEST PERIOD

Experimental group	Instrument	Number of subjects	Statistic	Clock 1	Clock 2	Clock 3	Clock 4
A	AAF standard	68	M	391.2	246.7	161.1	90.1
B	Reversed rotation	69	M	332.1	183.7	102.2	67.7
			Diff	49.1	63.0	58.9	22.4
			t	[1] 3.2	[1] 3.4	[1] 3.4	[1] 3.0

SECOND 8-MINUTE TEST PERIOD

Experimental group	Instrument	Number of subjects	Statistic	Clock 1	Clock 2	Clock 3	Clock 4
B	AAF standard	33	M	354.9	206.0	121.5	70.8
A	Reversed rotation	32	M	285.1	127.7	60.4	50.8
			Diff	69.8	78.3	61.1	20.0
			t	[1] 3.3	[1] 3.9	[1] 3.5	[1] 3.1

SUBJECT PREFERENCE SUBSEQUENT TO LAST TEST PERIOD

No. 1 AAF Standard instrument given first			Reversed rotation instrument given first		
Prefer 1	No preference	Prefer reverse rotation	Prefer 1	No preference	Prefer reverse rotation
Percent 9	*Percent* 3	*Percent* 87	*Percent* 6	*Percent* 18	*Percent* 76

[1] A t value of 2.58 or higher is significant at the 1-percent level.

Seventy-two subjects of the 137 cadets in the fifth comparison continued to use the same horizon indicator during the second test period that they had used in the initial period. In the case of these 72 subjects it was found that the highly significant differences between the standard horizon indicator and the reversed rotation instrument hold only for the first 8 minutes of practice and not for the second test period. This should not be taken to mean that the difference in performance yielded by these instruments is merely a temporary phenomenon. It should be emphasized that the subject does not have a full panel of instruments to which he must attend, but only the relatively simple task of keeping the trainer in level and horizontal flight by attending to a single flight indicator. Under these circumstances, it is possible for the subjects who use the standard, and inferior, instrument to attain a level of performance in two practice periods that is comparable to that achieved by the group using the superior or reversed rotation instrument. Given two 8-minute practice periods in succession, both groups achieve a relatively high degree of proficiency. The group using the reversed rotation instrument soon reaches a leveling off in its performance, and the group with the inferior instrument overtakes them. It will be shown in a later section that when the remaining 65 subjects in this fifth comparison who were tested on both instruments shifted from one indicator in the first period to a second indicator in the next period, the effect of habit interference revealed a difference that carried over into the second 8-minute period. It should be emphasized that there is always the possibility that under conditions of stress a pilot may make an error because of the ambiguity of the conventional instrument which could have been prevented had the indicator possessed optimum interpretability.

In view of the striking differences in performance which have been obtained with the reversed rotation and standard instruments, the question arises as to the similarity of the two groups using the respective instruments. Table 8.6 shows the characteristics of the subjects when they are segregated on the basis of the particular instrument used during the first 8-minute test period. These data show conclusively that the two groups are comparable in hours of pilot training (primary ships only), pilot stanine, and in previous experience with an experimental test that involves 8 minutes of contact flying in a Link ground trainer.

TABLE 8.6.—*Equivalence of groups used in the fifth comparison involving the standard AAF instrument and the reversed rotation instrument*

	AAF standard used first				Reversed rotation instrument used first			No other Link test
	Hours pilot training	Pilot stanine	Had Link contact flying test	No other Link test	Hours pilot training	Pilot stanine	Had Link contact flying test	
Means	8.82	6.65			6.47	6.21		
SD	12.00	1.67			7.83	1.87		
Number subjects	68	[1] 62	15	53	69	[1] 63	12	57
Percent			22	78			17	83

[1] Stanines not available for 6 subjects.

Can the differences in performance obtained with the two instruments in this fifth comparison be accounted for in terms of the relative sensitivities of the two instruments? A calibration of the two indicators demonstrates that the reversed rotation horizon is of somewhat greater sensitivity than the standard. Table 8.7 shows that when the

TABLE 8.7.—*Relative sensitivity of standard AAF Link horizon indicator and reversed rotation instrument as utilized in comparison 5*

Number of clock starting to record	1	2	3	4
AAF standard indicator				
Clockwise rotation in degrees	5	10	15	3
Counterclockwise rotation in degrees	5	10	15	3
Reversed rotation indicator				
Clockwise rotation in degrees	9	14	20	4
Counterclockwise rotation in degrees	7	13	18	4

fuselage has rolled to the point where the first clock starts to score, the standard horizon registers a displacement of 5°. At this point the modified indicator gives a reading of 7° to 9°, depending on the direction of roll. The fact that the pointer of the reversed rotation instrument has moved further than the pointer of the standard indicator might, conceivably, give the subject an earlier and more emphatic warning that he has drifted off level flight.

To establish the significance of variations in sensitivity, the reversed rotation indicator was desensitized by lengthening the pin attached to the gimbal ring assembly which engages with the bank pendulum. The length of the standard pin as measured in the vertical dimension is $\frac{5}{16}$ inch. Its net moment arm is $\frac{3}{16}$ inch as measured down from the rear bearing of the gimbal ring. The first model of the reversed rotation instrument, used in the fifth comparison, has a pin approximately $\frac{21}{32}$ inch long with a net moment arm of approximately $\frac{5}{32}$ inch as measured above the rear bearing of the gimbal ring. The desensitized instrument has a length of $\frac{11}{16}$ inch with a net moment arm of approximately $\frac{3}{16}$ inch as measured above the rear bearing of the gimbal ring. The calibration of the desensitized version of the reversed rotation indicator is presented in table 8.8. It is clear from these data that the pointer of the desensitized instrument moves a shorter distance than the pointer of the standard instrument when the fuselage banks to one side or the other.

TABLE 8.8.—*Relative sensitivity of standard AAF horizon indicator and desensitized reversed rotation instrument as utilized in comparison 6*

Number of clock starting to record	1	2	3	4
AAF standard indicator				
Clockwise rotation in degrees	5	10	15	3
Counterclockwise rotation in degrees	5	10	15	3
Desensitized reversed rotation indicator				
Clockwise rotation in degrees	4	8	12	2
Counterclockwise rotation in degrees	4	7	12	2

Seventy-eight additional subjects were run for 8 minutes in comparison 6 to determine the direction of the difference in performance when the desensitized reversed rotation instrument is compared with the standard. Table 8.9 gives the data from this last group. All of the clock scores for the desensitized reversed rotation instrument are superior to the scores for the standard. All differences are significant at the 1-percent level.

The differences appear to be greater than in the original group V comparison involving a more sensitive reversed rotation instrument. The data of the original 137 subjects and the data from this supplementary group of 78 subjects, making a total of 215, are in accord in establishing the superiority of the reversed rotation instrument. The

TABLE 8.9.—*Standard Link artificial horizon vs. desensitized reversed rotation type of artificial horizon*

(Comparison 6)

Initial 8-minute test period

[No previous instrument experience]

Instrument	Number of subjects	Statistic	Clock 1	Clock 2	Clock 3	Clock 4
AAF standard	40	M	411.4	273.2	197.7	97.3
Desensitized reversed rotation	38	M	312.4	156.4	89.8	52.7
		Diff	99.0	116.8	107.9	44.6
		t	[1] 5.1	[1] 4.8	[1] 4.3	[1] 5.5

[1] A *t*-value of 2.58 or higher is significant at the 1-percent level.

experimental test of variations in sensitivity has shown that the obtained differences cannot be explained as due to this factor. It is concluded, therefore, that the superiority of the reversed rotation instrument must be due to its mode of functioning.

DISCUSSION

Significant Factor in the Superiority of Horizon Indicator with Reversed Rotation

It should be noted that the superiority of the reversed rotation type of artificial horizon has been exhibited in spite of the fact that when the aircraft assumes a right-roll attitude, the indicator registers this maneuver by showing the miniature airplane with its left wing dipped below the horizon bar. Although this static relationship must ultimately be interpreted as representing a left-roll attitude, the dynamic relationship of the moving elements is such as to suggest a right roll. In consequence, it would appear that the direction of rotation of the moving elements in the instrument comprises the factor which the novice reacts to most immediately—a factor which the more experienced pilot has learned to disregard. It also seems reasonable to believe that if the correct static pattern were presented along with the appropriate dynamic relationship of the moving elements, e. g., when the horizon remains fixed and the miniature plane rotates, the resulting instrument might be superior to the reversed rotation horizon used in the last comparison of the series. Data on such an instrument have been reported by a British investigator.

Since circulation of the initial report on the present study, two papers have been made available to the writer, describing a study conducted at the request of the instrument department of the Royal Aircraft Establishment, Farnborough (1, 2). The British investigation differs from the present study in several respects. Its objective was to compare a single experimental indicator with a conventional instrument. Two groups of 20 cadets were employed as subjects.

The cadets, who had had no instrument instruction, were required to fly a standard Link trainer on a straight and level course while the machine was periodically deflected by its standard rough-air mechanism. Although the reports contain no reference to a modification of the turning mechanism, there is a statement to the effect that ". . . the only indication before the subjects was one or other of the two attitude indicators which were being tested" (2). It is not clear, therefore, whether the British subjects were required to maintain a heading or whether they were free to devote all of their attention to the attitude indicator. Performance was scored by means of two mechanical integrators connected directly to the fuselage of the trainer. The scores, in arbitrary units, are proportional to the angular deviation in attitude multiplied by time. In addition to the two mechanical integrators for measuring roll and pitch deviations, there was a graphic recording system for tracing the actual deflections. Each subject was required to fly the trainer for a total interval of 12 minutes, and performance was measured in successive 2-minute samples. The data in table 8.10 represent the performance of two separate groups of cadets, each of which was tested on a single instrument.

The type of attitude indicator compared with the conventional instrument consisted of a two-dimensional model aircraft which moved in relation to two fixed horizon bars at the edges of its wings. The author of the reports was of the opinion that with such an instrument, "the pilot identifies himself with the model instead of imagining himself to be a fixed point in space with a moving real horizon—a conception which does not fit the facts and which may be difficult to grasp" (1). As demonstrated by the figures of table 8.10, both the graphic records and the integrator scores show that the modified design is significantly superior to the conventional design. It was also found that the preference of a group of cadets tested on both instruments was clearly in favor of the new design. The general

TABLE 8.10.[1]—*Comparison of indicators in terms of graphic records and integrator scores*

Recording method		Pitch		Roll		Recording method		Pitch		Roll	
		Indicator		Indicator				Indicator		Indicator	
		Old	New	Old	New			Old	New	Old	New
Graphic......	M.......	27.2	15.3	121.8	95.8	Integrated....	M.......	22.3	20.0	29.8	25.1
	SD......	34.6	17.1	52.9	50.2		SD......	7.31	4.9	9.54	6.33
	Diff.....	11.9		26.0			Diff.....	2.3		4.7	
	SE_{Diff}...	4.31		8.16			SE_{Diff}...	1.0		1.31	

[1] The mean values of all the observations for new and old indicators and for both methods of recording. The difference between old and new indicators is unlikely to be due to chance since in every case the difference between the mean is more than twice the standard error. From Flying Personnel Research Committee *Report No. 611* (a) by R. C. Browne.

trend of the findings is thus in keeping with the results of the present study in the sense that an indicator which contains a moving element rotating in the same direction as the aircraft rolls was found to be clearly superior to the conventional instrument.

In the present study, a series of four different modifications in attitude design were studied in an effort to analyze and isolate the basic elements which make one instrument superior to another. Including the control measurements, a total of six comparisons involving 555 subjects provided the experimental data upon which the present discussion has been based.

Further Research Suggested by the Present Study

To round out the present investigation it would be highly desirable to study three additional attitude-indicator models. The exact number of comparisons required would depend upon the trend of the experimental data. Model No. 1 of this proposed series is shown in figure 8.2F. It should have the following characteristics: (1) moving element, consisting of a three-dimensional miniature aircraft model enclosed within a transparent sphere; (2) fixed horizon reference line and fixed scale at periphery of stationary transparent sphere surrounding the miniature aircraft; (3) direction of rotation of moving element to be varied in successive comparisons; (4) panel indicator activated by autosyn mechanisms interposed between trainer fuselage and base.

Model No. 2 of the proposed series would be characterized by the following features: (1) moving element of indicator similar to that in the new Sperry Attitude Gyro, i. e., a sphere; (2) fixed horizon line, fixed scale at periphery, sphere marked as AAF models; (3) direction of rotation to be varied in successive comparisons; (4) autosyn activating mechanisms.

Model No. 3 of the proposed series would include the following characteristics: (1) moving element of indicator consisting of transparent hemisphere with meridians embossed on inner surface, and so placed that it surrounds the anterior half of a miniature plane which is fixed in position and heads away from the pilot; (2) reference horizon line on the hemisphere moving up when the parent ship dives and down during a climb; (3) direction of rotation to be varied in successive comparisons; (4) autosyn activating mechanisms.

It was not found possible to test these three models in the time allotted to this investigation. The definitive analysis of the basic relationships insuring optimum interpretability must therefore await further study.

Basic Requirements for Attitude Indicators in the Light of Present Experimental Data

From the various considerations reviewed in the preceding discussion it would appear that the most easily interpreted horizon indicator for blind flying is one in which the moving elements directly register a shift in attitude by a displacement in the same direction. It tends to follow from this basic requirement that the horizon line must remain fixed in relation to the pilot's instrument panel. It would seem that the best results could be achieved by the construction of an instrument which contains a three-dimensional minature aircraft that banks or pitches about a horizon which is fixed in relation to the instrument panel and not with regard to the true earth's horizon. Thus when the attitude of the plane is that of a right roll, the miniature aircraft rotates clockwise; when the plane rolls to the left, the miniature model rotates in a counterclockwise direction. During a climb, the nose of the miniature model rises and the tail sinks (see fig. 8.2 F).

It should be emphasized that all of the data of the present investigation as well as the findings of the British study were obtained from subjects who were flying blind and unable to see outside the trainer. It is essential that in future work these findings be verified in a situation which involves a combination of blind and contact flying. As an initial check, it would be desirable to test subjects in a situation somewhat like that used by the Canadians in developing a screening test for student pilots. The Canadians employed a special Link trainer with biplane wings. This was flown contact in relation to a cyclorama painted to simulate a view of the earth as seen from an elevation of several thousand feet. To adapt this test situation to the needs of the proposed study, it would be necessary to install only the proper instrument panel, rough-air mechanism, and scoring system as developed for the present study. The subjects would be required to shift from contact to instrument flying in a standardized but irregular sequence when given the appropriate signal by way of their earphones.

Finally, those instruments which proved superior in the combination contact-and-blind-flying trainer situation should be checked under actual flying conditions. Both novices and experienced pilots could be tested for specified periods in which the task was merely one of flying level and horizontal. Ultimately, those acrobatic maneuvers which cannot be simulated by the Link trainer should be studied, using experienced pilots as subjects and a check-list inventory scored by an observer.

In the situation where the task is merely one of flying level and horizontal, performance could be scored by some system which records roll and pitch deviations. One relatively simple system for obtaining such types of scores employs mercury switches which close circuits to various counters when the plane rolls or pitches through successively increasing angles. Each counter circuit is interrupted at a fixed rate. The different counter scores thus represent the time interval during which the aircraft deviated from straight and level flight by various specified amounts. Probably a more refined method for recording such data could be developed by using the pitch and roll components of an electrically operated gyro-horizon to close successive timing circuits.

Possible Habit Interference in Shifting from Old-Style to New-Type Indicators

The data in tables 8.1 to 8.5 indicate that there is a general tendency for minus signs to appear in the differences for the second 8-minute test period as subjects change from one instrument to another. This means that those subjects who used a modified indicator following practice with the conventional horizon did less well than those subjects who used the conventional horizon indicator after having had an 8-minute practice period with the modified instrument. The negative differences point to the effect of interference from habits that were established during the first 8-minute test period. They were clearly significant, statistically, only in the case of the reversed rotation horizon indicator used in the fifth comparison. One possible explanation of the predominant direction of the interference is suggested by previous studies of the learning function. The brief initial learning period in which the subjects use the conventional indicator requires so much effort and attention, and the subjects get so much practice in correcting their errors, because of the inherent difficulty of the standard instrument, that they tend to fix their responses relatively well. In contrast, the reversed rotation indicator fits in so well with past habits that less effort is required to perform the task satisfactorily. It is thus somewhat easier to break up these habits, temporarily, when the subjects transfer to the conventional instrument in the second 8-minute test period. There is some indirect substantiation for this interpretation, possibly, in the fact that the subjects express a decided preference for the modified instrument and regard it as easier.

The question as to how difficult it would be for experienced pilots to master an attitude indicator of the new type, referred to just above, cannot be answered by armchair theorizing or a casual expression of opinion. A satisfactory and dependable answer requires objective and quantitative evidence.

SUMMARY AND CONCLUSIONS

It has been pointed out that although pilots can learn to use almost any type of artificial-horizon indicator, the novice may be retarded in acquiring proficiency in blind flying, and even the experienced pilot may have difficulty during emergencies because of the inherent ambiguity in the conventional type of instrument.

Four different types of artificial-horizon indicators have been compared with the conventional instrument in terms of the objectively scored Link ground-trainer performance as the subject attempts to maintain the fuselage in level and horizontal flight under conditions of rough air. These indicators comprise the following types of instruments: (1) a scaleless horizon indicator in which the moving pointer and fixed scale of the conventional instrument have been removed; (2) a moving scale indicator in which the scale rotates but the reference index mark remains fixed; (3) a British type of horizon indicator in which the fixed scale and pointer are at the base of the instrument dial; (4) a reversed rotation instrument in which the moving pointer and the horizon bar tilt or rotate in the opposite direction from that of the standard AAF flight indicator.

All of the modified horizon indicators tend to show some slight superiority over the conventional instrument as contrasted with the differences arising when two instruments of the same design are compared. In the case of one category of measurement, the British type of horizon shows a difference in its favor which has some degree of statistical significance. The indicator in which the moving pointer and the horizon bar rotate in an opposite direction to that of the conventional AAF flight indicator is the one instrument that shows a superiority which has unquestionable statistical significance. The reverse rotation indicator is also the one instrument which is most consistently preferred by the subjects.

An analysis of the various factors which have been studied in the series of comparisons between the modified instruments and the standard suggests that the most easily interpreted attitude indicator *for blind flying* will be one in which the horizon remains fixed with regard to the instrument panel and the moving element rotates in the same direction as the plane rolls. It would appear that for blind flying the moving element of the indicator should consist of a three-dimensional miniature aircraft which banks or pitches within a horizon circle which is fixed in relation to the instrument panel rather than with regard to the true horizon. Thus, when the attitude of the plane is that of a right roll, the miniature aircraft rotates clockwise; when the plane rolls to the left, the miniature model rotates in a counterclockwise direction. During a climb, the nose of the miniature aircraft rises and the tail sinks. It should be noted, however, that before a final and definitive analysis can be made with regard to the basic factors essential for an ideal design, at least three additional types of instruments, described previously, should be studied. It must also be emphasized that additional work is required before it can be stated which type of instrument will insure minimal confusion when used in situations other than blind flying or when used by experienced pilots who have been trained on conventional instruments.

BIBLIOGRAPHY

1. Browne, R. C. Comparative trial of two attitude indicators. Preliminary Note. Flying Personnel Research Committee. *Report 611.* February 1945.

2. Browne, R. C. Comparative trial of two attitude indicators. Flying Personnel Research Committee. *Report 611 (a)*. April 1945.

3. Daniel Guggenheim Fund for the Promotion of Aeronautics, Inc. *Solving the problem of fog flying: A record of activities of the Fund's Full Flight Laboratory to date.*[1] October 9, 1929. 52 + ii.

4. Daniel Guggenheim Fund for the Promotion of Aeronautics, Inc. *Equipment in experiments to solve the problem of fog flying: A record of the instruments and experience of the Fund's Full Flight Laboratory.*[1] March 1930. 67.

[1] The Guggenheim Fund reports appear to have been privately printed.

BLANK PAGE

CHAPTER NINE

Direction of Movement in the Use of Control Knobs to Position Visual Indicators[1]

MELVIN J. WARRICK

Numerous situations exist in aviation where visual indicators are positioned by rotation of a control knob. In a simple control-indicator system there are two possible alternate relationships between the direction of motion of a control and the resultant motion of its indicator. For example, a clockwise motion of a control could move an indicator either clockwise or counterclockwise. For greatest efficiency of operation it seems desirable that the relationship between indicator and control motion be such as to conform to the response tendencies of the majority of the operators. For further discussion of this problem see chapters 2, 3, 10, and 17. The two studies reported in this chapter represent an initial investigation of this problem and were planned and carried out concurrently.

EXPERIMENT 1

Purpose

The purpose of the first experiment was to determine the direction in which operators most frequently turn control knobs to accomplish a certain direction of motion of an indicator in several arrangements of controls and indicators.

For this initial experiment a knob was selected as being a commonly used type of control. A straight row of light positions was used as the indicator, primarily because other types of displays might so resemble an aircraft instrument as to elicit a response biased by experience with that instrument. Five of the more common spatial relationships between the indicator and control were arbitrarily selected for study.

[1] This chapter is based upon research findings reported in Headquarters AMC, Engineering Division Memorandum Report No. TSEAA-694-4C. The assistance of Joseph Bakalus and Helmut Muehlhauser in constructing the apparatus and in testing the subjects greatly facilitated the experiment.

Apparatus and Procedure

An apparatus was built containing five separate panels each having an indicator (a row of lights) and a control knob, as shown in figure 9.1. Each panel presented a different relationship between the location of the indicator and the location of the control.

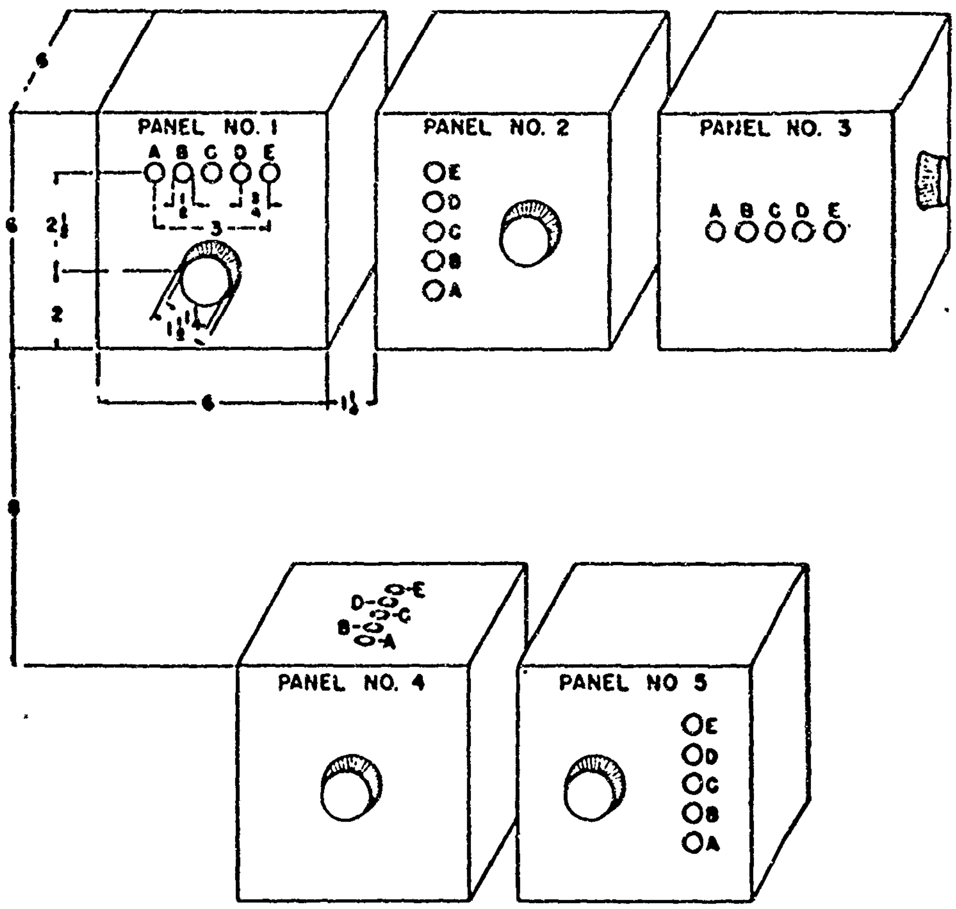


FIGURE 9.1—Control-indicator panels used in determining direction of response preferences.

It will be noted in figure 9.1 that the panels are numbered from left to right and the lights on each are numbered in sequence from left to right, bottom to top, or front to rear. An indicator position will hereafter be referred to by number and letter to designate the panel and the light position.

The equipment operated in the following manner. A light appeared on any one of the panels in a position other than C (center). The subject, using his right hand only, moved the light to center with the appropriate knob. The control for each knob was so wired that no matter which way (clockwise or counterclockwise) the subject rotated the control, the light moved toward the center. However, once the light was moved from its original position, the relationship between the direction of control motion and the resultant direction of the indicator motion was established and remained the same until the problem was solved or the light returned to its initial position. Either a clockwise or a counterclockwise response was correct in that it moved the light to center and solved the problem.

A series of 40 successive problems was presented in a randomized order of panels and light positions (except that the same panel never succeeded itself). Random time intervals not exceeding 1 second occurred between the solution of one problem and the onset of the next. Each light position (excluding the center) appeared twice during the series. Thus there were eight problems on each panel, four on each side of center. With no wrong or unrewarded responses possible, with a randomized order of presentation, and with a definite emphasis on speed of reaction, it was hoped that the subject's responses to a particular problem would be least likely to be influenced by his responses to previous problems.

Fifty subjects, all rated military pilots, were employed in this study. They were divided into 10 groups. Each group started at a different indicator position (A or E) in the series of 40 problems. They were instructed to move each light to center as rapidly as possible and were told that their score would be the amount of time it took them to complete the series. Cumulative reaction times over the series were obtained and recorded. The subjects were shown the unidentified time scores of other subjects before the test as a stimulant to faster reactions. They were not told that they could move the control either way to move the light to center.

Results

In table 9.1 are presented the percent of clockwise and of counterclockwise responses to (1) each problem; (2) problems on the same side of center combined; and (3) all the problems of each panel combined.

There are two possible directions in which the knob can be rotated to move the indicator in the desired direction. If neither direction of response is preferred over the other, the subject would be as likely to respond with a clockwise as with a counterclockwise movement of the control, unless other factors such as difficulty in making certain movements are influencing the responses. Thus, the responses to particular problems would be split 50 percent clockwise and 50 percent counterclockwise within the limits of chance variation. With 50 subjects the probability is only one in a hundred that the number of clockwise (or counterclockwise) responses would exceed 66 percent. Thus if there are more than 66 percent clockwise responses or more than 66 percent counterclockwise responses, it is assumed that there is a preferred response in that direction in the particular situation.

TABLE 9.1.—*Percent of clockwise (C) and counterclockwise (CC) responses to the five indicator panels*

[N=50][1]

Panel	Indicator position	Individual positions: Percent C	Individual positions: Percent CC	Positions combined: Percent C	Positions combined: Percent CC	Entire panel: Percent C	Entire panel: Percent CC
1	A	[2] 95	5	[2] 94	6	56	44
	B	[2] 92	8				
	D	19	[2] 81	19	[2] 81		
	E	19	[2] 81				
2	A	[2] 85	15	[2] 88	12	56	44
	B	[2] 90	10				
	D	23	[2] 77	24	[2] 76		
	E	26	[2] 74				
3	A	57	43	56	44	57	43
	B	54	46				
	D	57	43	58	42		
	E	60	40				
4	A	[2] 78	22	[2] 78	22	[2] 69	31
	B	[2] 77	23				
	D	61	39	60	40		
	E	60	40				
5	A	32	[2] 68	30	[2] 70	57	43
	B	29	[2] 71				
	D	[2] 82	18	[2] 83	17		
	E	[2] 84	16				
				All panels		59	41

[1] Had an N equal to the number of responses rather than the number of subjects been used an "individual-position" percent figure would be significant at the 1-percent level if it exceeded 61.6 percent ($N=100$); a "positions-combined" figure significant if it exceeded 58.2 percent ($N=200$); an "entire-panel" figure significant if it exceeded 56.0 percent ($N=400$); and the total significant if it exceeded 52.6 percent ($N=2{,}000$). Thus in terms of a sample of responses, not individuals, the significance of the results indicated in the table is greatly enhanced.

[2] Significant at or above the 1-percent level of confidence for $N=50$.

The obtained response tendencies might be due either to the variability between individuals or the variability within individuals. For example, the percent of clockwise responses to positions A and B combined, of panel 3 may be due to 56 percent of the subjects consistently responding clockwise and 44 percent consistently responding counterclockwise whenever one of these lights came on. Or the percent of clockwise responses may be due to each subject responding sometimes clockwise, sometimes counterclockwise. Or it may be some combination of these two possibilities. (The response frequencies to positions A and B and to positions D and E of each panel (table 9.1) are so consistently similar that they may be considered as like problems.)

Since each light on each panel appeared twice, the subjects made four responses to each like pair of positions. In table 9.2 is presented the number of subjects responding four, three, two, one, and no times clockwise (no, one, two, three, and four times counterclockwise) to each pair of lights on either side of center for each panel.

TABLE 9.2.—*Number of subjects responding with four, three, two, one, and zero clockwise (C) responses to problems on each side of center of each panel*

[N=50]

Panel	Indicator positions combined	Number of subjects making 4 C responses	Number of subjects making 3 C responses	Number of subjects making 2 C responses	Number of subjects making 1 C responses	Number of subjects making 0 C responses
1	A and B [1]	42	6	1	0	1
	D and E [1]	5	2	4	4	35
2	A and B [1]	36	7	4	2	1
	D and E [1]	1	6	8	11	24
3	A and B	16	11	3	8	12
	D and E	20	6	6	7	11
4	A and B [1]	22	18	7	0	3
	D and E	18	7	12	4	9
5	A and B [1]	8	2	7	10	23
	D and E [1]	34	7	3	3	3

[1] Indicates that the number of subjects responding predominately in the direction indicated in the table cannot reasonably be explained by chance. Note that the majority of subjects tended to respond clockwise to combined positions A and B, and D and E of panel 4.

Examination of table 9.2 indicates that there is considerable variability within individuals in responding to like problems. Thus, the response tendences presented in table 9.1 are not entirely a result of combining individuals with consistent but opposite response tendencies. The indicator-control arrangement of panels 3 and 4 which in table 9.1 yielded the most ambiguous response tendencies are also the ones in table 9.2 which show the greatest variability within individuals and between individuals. On panel 3, for example, it will be noted that 16 subjects always responded clockwise to positions A and B, that 12 subjects always responded counterclockwise to these lights, and that 22 subjects responded with at least one control movement in each of the two possible directions.

Two other major questions can be raised: (1) is the preponderance of clockwise responses an artifact elicited by the design of the apparatus itself, and (2) are the obtained responses to a particular problem biased by responses to a previous problem? These questions cannot be answered without further experimentation. However, when only the first response of each of the 50 subjects was considered, it was found that 54 percent of the responses were clockwise, which does not differ significantly from the frequency (59 percent) when all 40 responses of each subject were considered. This would indicate that there was probably no significant tendency for the number of clockwise responses to increase during the series as a result of experience on the apparatus.

Examination of the data presented in tables 9.1 and 9.2 indicates the following conclusions: (1) In the control-indicator arrangement of panel 1 the operators usually turned the control clockwise to move the indicator to the right and counterclockwise to move the indicator to the left. (2) In the control-indicator arrangement of panel 2 the operators usually turned the control clockwise to move the indicator up and counterclockwise to move the indicator down. (3) In the control-indicator arrangement of panel 3 the operators showed no uniform response preference; the control was turned clockwise slightly more than 50 percent of the time regardless of the direction in which it was desired that the indicator move. (4) In the control-indicator arrangement of panel 4 the operators most frequently turned the control clockwise regardless of which direction they wanted the indicator to move. (5) In the control-indicator arrangement of panel 5 the operators usually turned the control clockwise to move the indicator down and counterclockwise to move the indicator up. On all panels operators tended to make clockwise responses more frequently than counterclockwise responses. This was particularly true of the responses to panel 4.

Conclusions of Experiment 1

In the situations where an indicator moved in the same plane as its control (panels 1, 2, and 5), the operator usually moved the control so that the part of it nearest the indicator moved in the direction in which he was attempting to move the indicator. In both situations where the control did not move in the same plane as its indicator (panels 3 and 4) the response tendencies were ambiguous. Thus, in equipment design, the indicator-control arrangements of panels 1, 2, and 5, if used in accord with the results of this experiment, are to be preferred over those of panels 3 and 4.

EXPERIMENT 2

Purpose

From preliminary investigation it seemed likely that the control-indicator arrangements of panels 3 and 4 of experiment 1 would yield ambiguous responses. (This was borne out by the results.) It also seemed likely that it would occasionally be necessary to use these arrangements.

To study these two relationships in more detail, experiment 2 was planned and conducted at the same time as experiment 1. The purpose of experiment 2 was to determine which of the alternate motion relationships is most efficient in terms of speed or minimum errors in these two specific arrangements.

Apparatus and Procedure

The apparatus used in this investigation is shown in figure 9.2. It consisted of a large box with the control knob in the center of one side and a row of 11 indicator lights across an adjacent side parallel to the axis of rotation of the knob. The same box, in different positions, was used to obtain the two indicator-control arrangements. Each light, except the center and two extreme lights appeared an equal number of times in an irregular sequence of 40 trials. The subject was required to turn the knob and move the light to center. A clock cumulated the reaction time for the 40 successive trials. If the

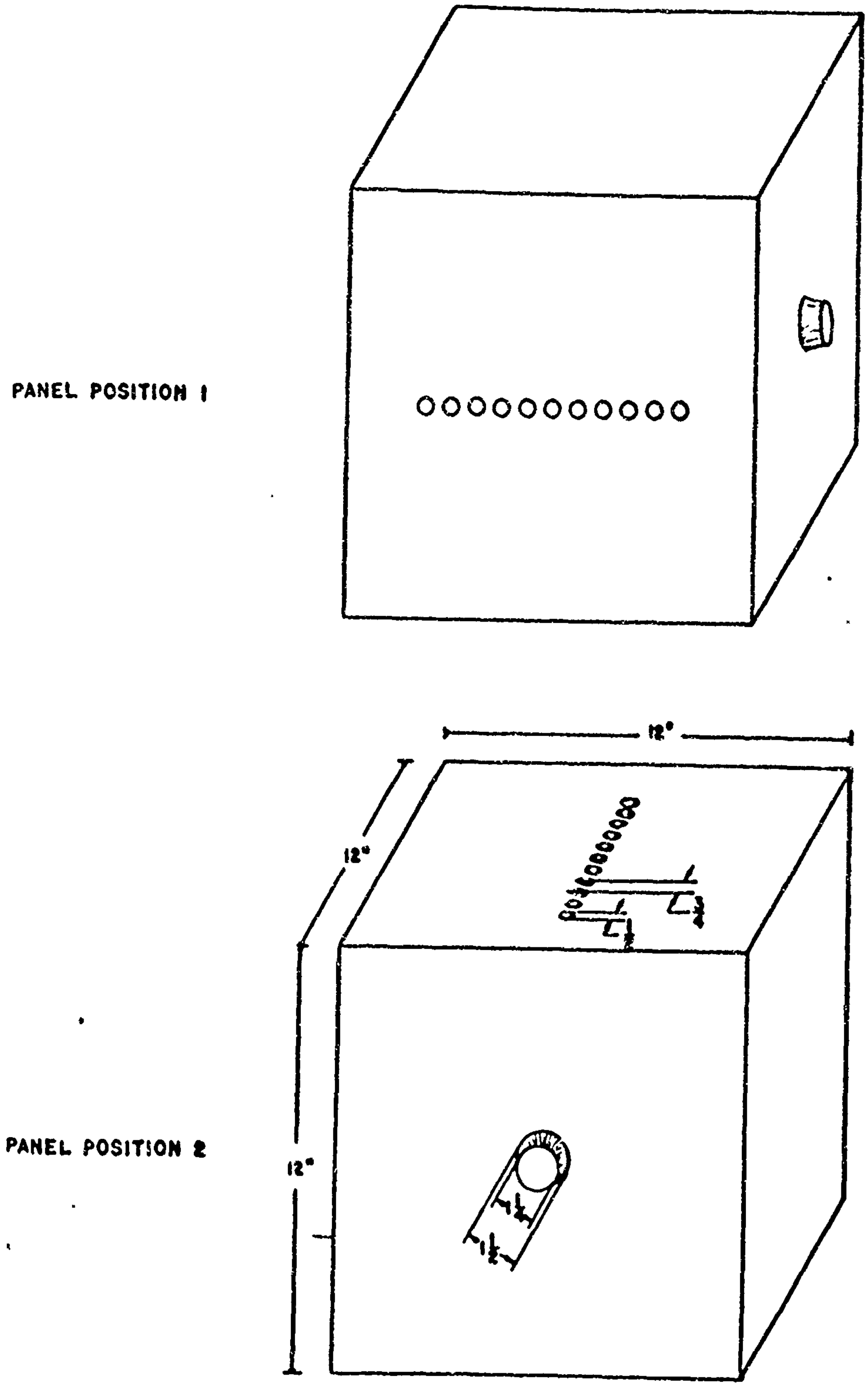


FIGURE 9.2.—Control-indicator arrangements used in experiment 2.

control was turned the wrong way, the light moved from the center and an error was accumulated on a counter. An error was recorded only if the knob was turned in the wrong direction enough to light an indicator bulb. A switch allowed the administrator to reverse the direction of the indicator movement in relation to the control movement. The test was self-paced with an interval between the solution of one problem and the onset of the next of approximately 0.6 second.

Forty-two rated military pilots were tested. Half the subjects started with one indicator-control relationship and half the subjects started with the alternate relationship. In order to test each subject using each control-indicator relationship and to counterbalance for interaction effects, the relationship between indicator and control was reversed after the first and third of the four successive 40-trial sequences. Thus, half the subjects were run in a ABBA order and half in a BAAB order on each panel position. The subjects were told that the relationship between indicator and control *might* be reversed from one series of trials to the next. They were told their time scores at the end of each series of trials but were not told that a count of their errors was being made.

Results

In table 9.3 are presented the means and the standard deviations of the time and error scores for each relationship of the two panel positions. The differences between the mean error scores and between

TABLE 9.3.—*Relative effectiveness of alternate control-indicator motion relationships in two arrangements of control and indicator*

PANEL POSITION NO. 1 ($N=42$)

Relationship	Mean	SD	t	r
ERRORS				
"A" (80 trials)[1]	7.74	3.95		
"B" (80 trials)[2]	8.62	5.55	0.88	+0.12
CUMULATIVE REACTION TIME (IN SECONDS)				
"A" (80 trials)[1]	59.82	6.72		
"B" (80 trials)[2]	62.88	8.16	[3] 2.90	+.59

PANEL POSITION NO. 2 ($N=42$)

Relationship	Mean	SD	t	r
ERRORS				
"A" (80 trials)[4]	11.17	6.26		
"B" (80 trials)[5]	8.79	5.97	[6] 2.22	+0.37
CUMULATIVE REACTION TIME (IN SECONDS)				
"A" (80 trials)[4]	60.78	6.36		
"B" (80 trials)[5]	60.06	6.90	.78	+.60

[1] "A" relationship, a clockwise motion of the knob moved the light to the left and a counterclockwise motion moved the light to the right.

[2] "B" relationship, a clockwise motion of the knob moved the light to the right and a counterclockwise motion moved the light to the left.

[3] Significant at the 1-percent level of confidence.

[4] "A" relationship, a clockwise motion of the knob moved the light away from the operator and a counterclockwise motion of the knob moved the light toward the subject.

[5] "B" relationship, a clockwise motion of the knob moved the light toward the operator and a counterclockwise motion of the knob moved the light away from the subject.

[6] Significant at the 5-percent level of confidence.

the mean time scores for each panel position are presented in terms of t.

In panel position 1 an operator worked slightly, but significantly, faster when the relationship between indicator and control was such that a clockwise motion of the control resulted in an indicator motion to the left, and a counterclockwise motion of the control resulted in an indicator motion to the right. The number of errors was also slightly but not significantly reduced by this arrangement.

In panel position 2 an operator made significantly fewer errors when the relationship between indicator and control was such that a clockwise motion of the control moved the indicator toward him and a counterclockwise motion moved the indicator away from him. The speed was also increased, but not significantly, by this arrangement. It will be noted that the differences in speed or errors on either panel, even though significant, were extremely small.

A comparison of the first 40 trials was made between the 21 subjects starting with one motion relationship and the 21 subjects starting with the alternate motion relationship of each panel. The results agreed with the results from the total group as presented in table 9.2.

Conclusions of Experiment 2

In the control-indicator arrangement of panel position 1, operators respond more rapidly, and perhaps with fewer errors, when a clockwise motion of the control results in motion of the indicator to the left, and a counterclockwise motion of the control results in motion of the indicator to the right. In the control-indicator arrangement of panel position 2, operators respond with fewer errors, and perhaps with greater speed, when a clockwise motion of the control results in motion of the indicator toward the operator and counterclockwise motion of the control results in motion of the indicator away from the operator. However, the practical difference in efficiency resulting from use of one control-indicator relationship as opposed to its alternate in these two indicator-control arrangements is so slight as to be easily outweighed by other considerations important in the design, installation, and use of indicators and controls.

GENERAL SUMMARY AND CONCLUSIONS

The purpose of these experiments was to determine the preferred relationship between the direction of motion of a control and its indicator in five arrangements of indicators and controls. These arrangements are shown in figure 9.1.

Fifty rated military pilots were required to move a light to center from various positions in a row of lights by a right-hand operated control knob. Each of the five arrangements of control and indicator were presented to subjects in a randomized order. The apparatus was built so that no matter which way the control was moved, the light would move toward center.

It was found that arrangements 1, 2, and 5, where the indicator moved in the same plane as its control, yielded consistent responses, the subjects preferring the motion relationship such that the part of the control nearest the indicator and the indicator moved in the same direction. It was found that arrangements 3 and 4 (in which the indicator did not move in the same plane as its control) did not yield consistent responses.

A second experiment was performed using only arrangements 3 and 4 (panel positions 1 and 2 of fig. 9.2) to determine which alternate motion relationship would result in the greatest speed or the fewest errors. Forty-two rated military pilots were tested, using each alternate motion relationship of each control-indicator arrangement.

It was found that the subjects worked significantly faster and with slightly fewer errors if the motion relationship of arrangement 3 (fig. 9.1) was such that a clockwise motion of the control knob moved the indicator to the left and the converse; and that the subjects made significantly fewer errors and worked slightly faster if the motion relationship of arrangement 4 (fig. 9.1) was such that a clockwise motion of the control knob moved the indicator toward the subject and the converse.

Because of the systematic consistency of response to the control-indicator arrangements illustrated by panels 1, 2, and 5 of experiment 1 these arrangements are recommended for use in that order providing, of course, that the motion relationship is such that the part of the control adjacent to the indicator and the indicator move in the same direction.

CHAPTER TEN

A Study of the Most Effective Relationships between Selected Control and Indicator Movements[1]

LAUNOR F. CARTER AND NORMAN L. MURRAY

INTRODUCTION

Grether (see ch. 2) and Brown and Jenkins (see ch. 3) have discussed the theoretical aspects of the display-movement, control-movement problem. An experimental approach to the problem has been described by Warrick (see ch. 9). A somewhat different method of attacking this problem with a more complicated display-control relationship is reported in this chapter. In the operation of equipment such as the visual bomb sight and the radar tail-gun sight, it is necessary for the operator to manipulate controls in such a manner as to cause a target and cross hairs to coincide. In the performance of this task the question often arises as to whether it is more natural to think of bringing the cross hairs to the target or the target to the cross hairs. The question is indicative of the confusion sometimes attendant in the operation of equipment of this type. It has many times been observed that naive subjects operating equipment of this type for the first time make many errors in control movement, i. e., move the controls in a direction calculated to bring a target and cross hairs into coincidence only to find that the displacement between the two has increased. It has further been observed that with practice the difficulty is overcome and a learned response is made in the correct direction. It has been hypothesized, however, that under conditions of stress the operator will revert to the old or natural set and it is desirable, therefore, to determine the most natural control movements for bringing the target and cross hairs into coincidence.

This experiment was designed to investigate the relationship between the line of movement of the spot on the face of the oscilloscope and the axis of movement of two control knobs. Since it was also

[1] This chapter is based upon research findings reported in Headquarters AMC, Engineering Division Memorandum Report No. TSEAA-694-7.

possible to vary the relationship between the direction of display movement and the direction of rotation of the control knob, this problem also has been investigated.

APPARATUS AND PROCEDURE

The apparatus[2] consisted of five distinct units: an RCA 5-inch cathode ray osilloscope, a pair of control knobs, a stepping switch, plane-of-movement switches, and counters and timing mechanisms. The assembled apparatus (without the timers and counters) is illustrated in figure 10.1. A commercial cathode ray oscilloscope was

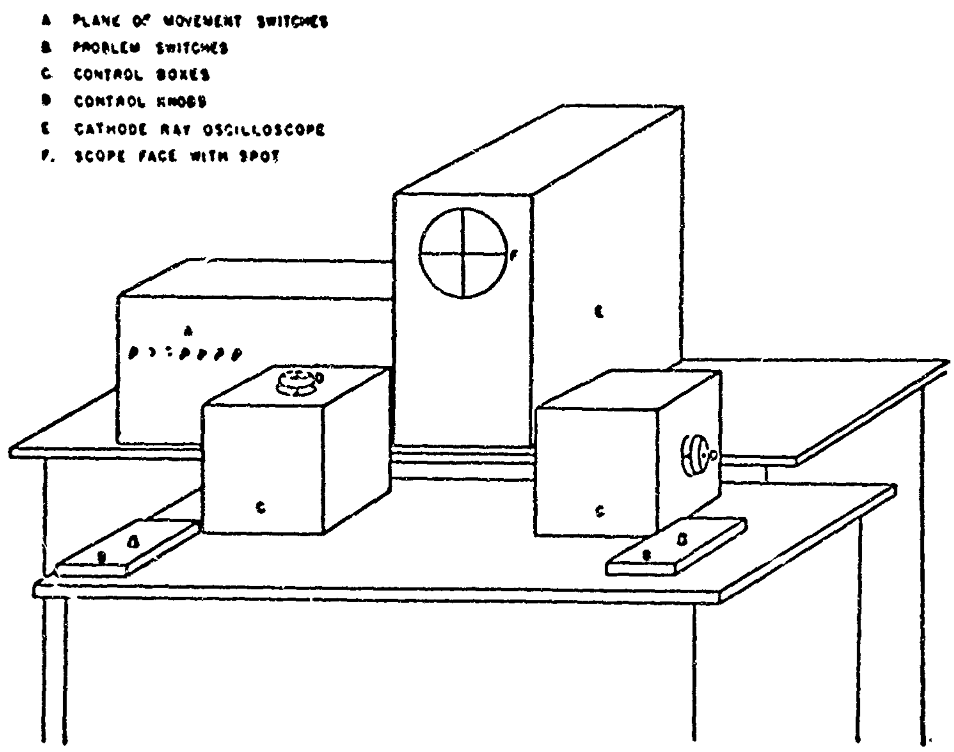


FIGURE 10.1.—Apparatus used in the study of control-indicator movement relationships.

modified so that the potential on the plates could be varied and would hold a direct-current charge. By varying the potential, the spot on the face of the scope could be displaced either right and left or up and down. Cross hairs were positioned on the face of the scope so that they intersected at the center, dividing the face into four quadrants.

The two mechanisms known as control knobs were identical and are illustrated in figure 10.2. Each consisted of a knurled knob mounted on a shaft. The knob bushing was oversized 0.005 inch allowing it to rotate independent of the shaft. Two sets of points (electric contacts) supported on a U-shaped bracket were mounted on the con-

[2] The authors wish to acknowledge the assistance of Helmut Muehlhauser both in the construction of the apparatus and in the running of the subjects.

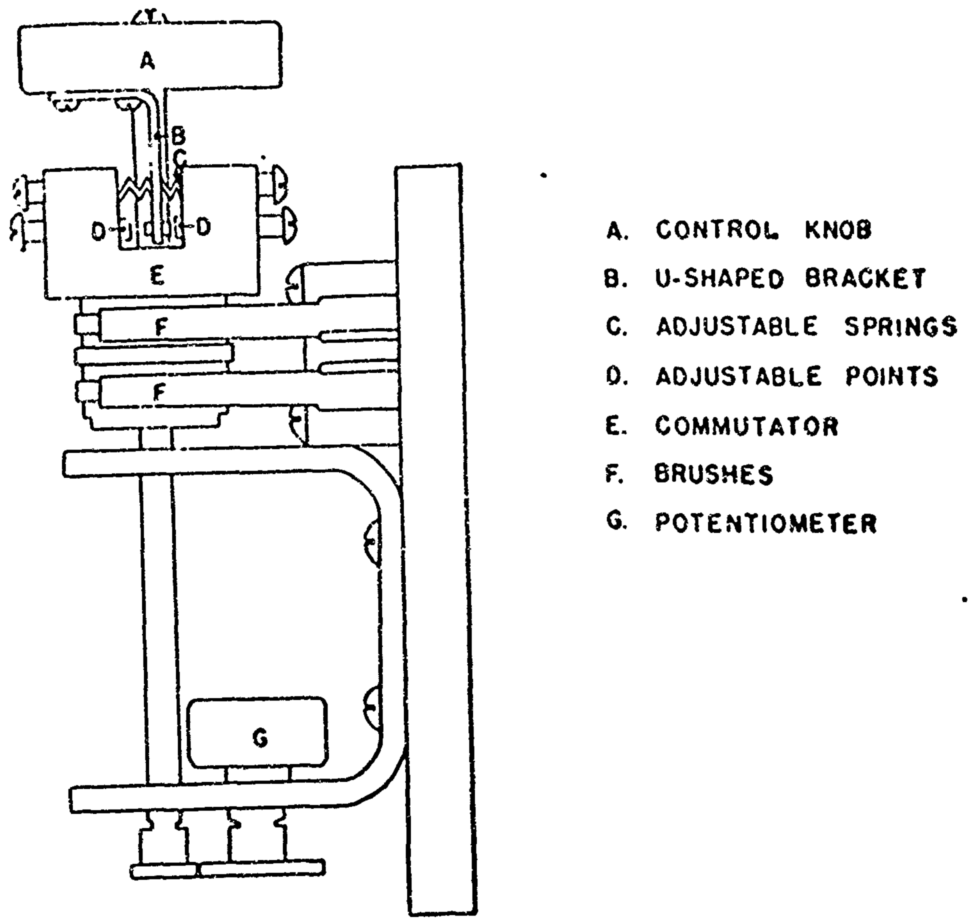


FIGURE 10.2.—Diagram of control mechanism.

trol knob in such a way that the arms extend perpendicular to the plane of rotation and parallel to the shaft. The arms of the bracket were 180° apart, one extending below the shaft and one above.

A three-ring commutator was also mounted on the shaft and fixed in position by a setscrew. The arms of the control-knob bracket meshed with a slot in the face of the commutator in such a way that clutching action was obtained along with electrical switching. The two contacts at the top of the slot in the face of the commutator were wired to one of the slip rings. Each of the contact points at the bottom of the slot were wired to separate rings. The contact points of the control knob were positioned midway between the top and bottom points of the commutator by adjustable springs. Thus, when no torque was applied to the knobs, they were self-centering, contact was broken, and no current passed. If, however, force was applied to the control knob, contact was made between one of the upper points on the face of the commutator and the opposite point on the lower face, thereby completing a circuit. The current was led off by brushes to the clocks and counters of the scoring device. It should be pointed out here that as force was applied the control knob slipped on the shaft and moved in the direction of the applied force. As the spring tension of the centering device was overcome, contact was made between diagonally opposite points in the face of the commutator, completing a circuit between separate rings of the commutator, and the current was in turn led off by brushes to the clocks and counters concerned. Although the clocks and counters began to record both time and direction of motion as soon as the spring tension was overcome, it was not until the applied force was transmitted by the points of the control knob to the opposite points of the commutator that the inertia and friction of the shaft was overcome and it began to rotate. A gear train driven by the shaft controlled the rotation of the wiper arm of a potentiometer which was wired in series with a battery and the plates of the oscilloscope. It can be seen then that there was an initial latency between the movement of the control knob and the visual displacement of the spot on the face of the scope. This time delay was desirable as it afforded the subject no clue to an incorrect response before actual recording was begun.

A 40-position stepping switch was provided which, when it stepped to a new position, varied the resistance in the circuit, thereby displacing the spot on the face of the scope and presenting a new problem to the subject. The subject was then required to center the "spot" which he did by balancing the system, i. e., by rotating the wiper of the potentiometer. The stepping switch was stepped to a new position by depressing two spring-loaded switches.

Switches were also provided for changing the polarity of the current, thus making it possible to change the direction of motion of the "spot," in relation to the control knobs.

The clock and counter mechanism consisted of three clocks and four counters which were provided to record time and direction of motion of the control knobs. One clock recorded the time that the left control knob was in motion either clockwise or counterclockwise. Another clock recorded the same information for the right control knob. A third clock recorded the total operating time, i. e., the time necessary to complete the 40 problems. One counter recorded all clockwise motions of the left control mechanism; another counter recorded all counterclockwise motions of that control mechanism. A third counter recorded all clockwise motions of the right control mechanism; and another counter recorded all counterclockwise movements of the right control mechanism.

The subjects, Army Air Forces pilots, were divided into two groups of 24 each. The first group was tested under condition C, in which the axis of movement of the control knobs was perpendicular to the plane of movement of the spot. Thus, under condition C a movement of the left control knob caused a horizontal movement of the spot and a movement of the right control knob caused a vertical movement of the spot. The second group was tested under condition N, in which the line of movement of the spot was parallel to the axis of control movement. The control boxes were fixed in position and maintained their relationship throughout the experiment. The left control box was fixed in a position such that the axis of rotation of the control knob was vertical, and the right control box was fixed in position such that the axis of rotation of the control knob was horizontal. Under condition C rotation of the left control knob produced horizontal displacement of the spot either to the right or to the left, while rotation of the right control knob produced vertical displacement of the spot either up or down. Under condition N, rotation of the left control knob produced vertical displacement of the spot either up or down, and rotation of the right control knob produced horizontal displacement of the spot either right or left. There were four distinct relationships between display movements and control movements under condition C and four distinct relationships under condition N. The four relationships under condition C were (with R and L for right and left and U and D for up and down) CRD, CRU, CLD, and CLU. Similarly under condition N the relationships were NDR, NUR, NDL, and NUL. All of the above relationships refer to clockwise movement of the knobs; of course, a counterclockwise movement of the knob resulted in the opposite movement of the spot.

The subjects were seated before the apparatus and instructed in the operation of the apparatus. First, the problem switches were pointed out and it was explained that they were to be depressed in order to begin each succeeding new problem. The operator next pointed out the spot and the cross hairs and demonstrated how they are made to coincide by the simultaneous manipulation of the control knobs. Here again it was pointed out that once the spot was centered the subject was to remove his hands from the control knobs and depress the two problem switches to present a new problem.

There were 40 problems in each series. The subjects were run through one series, given a short rest during which time the direction of control movement was changed (a new relationship presented), and then the subjects were run through another series, until all four series were completed.

A counterbalanced experimental order was used involving 48 subjects. Twenty-four of the subjects were run under condition C and 24 under condition N. Within the two conditions the order in which each of the four relationships was presented was completely counterbalanced.

RESULTS

Three different problems can be investigated with the data collected. First, the best display-movement, control-movement relationship under condition C, where the line of movement of the spot was perpendicular to the axis of control movement. Next the best relationship under condition N, where the line of movement of the spot is parallel to the axis of movement of the control knob. Finally the relative accuracy and speed of response under condition C and condition N can be determined.

Under condition C there were four control relationships, CRD, CRU, CLD, and CLU. Table 10.1 shows the means, standard deviations, correlations, and critical ratios both for the average number of starts and for the average time per control knob. Figure 10.3 gives a graphic presentation of the means for both the number of starts and the average time. The number of starts given in table 10.1 is the number of starts in either direction averaged over the four counters. The time is given in terms of the average amount of time either of the control knobs was turned, that is, the times on the two clocks averaged and expressed in minutes. An inspection of table 10.1 and figure 10.3 reveals that there are very significant differences in the number of starts required to center the spot with different control relationships.

TABLE 10.1.—*Means, standard deviations, correlations, and critical ratios for the number of starts and time scores for condition C*

[N=24]

Comparison	Number of starts			
	Mean	SD	r	CR
CRD	40.57	8.41	0.53	4.59
CRU	53.42	16.17		
CRD	40.58	8.41	.54	5.51
CLD	50.67	10.10		
CRD	40.58	8.41	.68	12.36
CLU	61.96	11.55		
CRU	53.42	16.17	.76	1.25
CLD	50.67	10.10		
CRU	53.42	16.17	.64	2.37
CLU	61.96	11.55		
CLD	50.67	10.10	.57	5.48
CLU	61.96	11.55		
Time (minutes)				
CRD	1.27	0.48	0.60	1.73
CRU	1.41	.42		
CRD	1.27	.48	.67	3.77
CLU	1.55	.39		
CRU	1.41	.42	.49	1.71
CLU	1.55	.39		

When there is correspondence between the axis of movement of the control knobs and the line of movement of the spot the relationship CRD is by far the best, that is, a relationship where a clockwise movement of the knobs causes a right and down movement of the spot.

This relationship allows the completion of the problems with the smallest number of false starts and in the shortest time. The relationship CLU is the poorest relationship, with the two other relationships, CRU and CLD, falling about halfway between the best and poorest

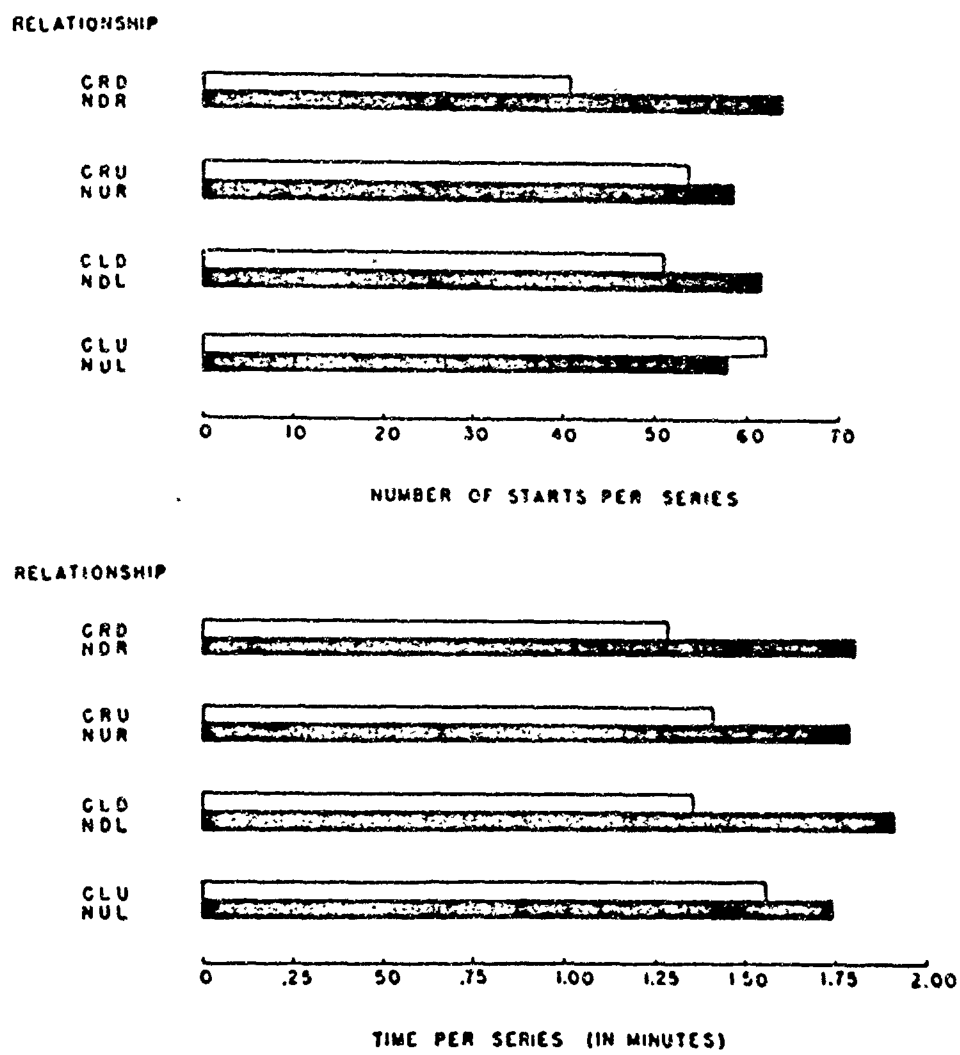


FIGURE 10.3.—Comparison of number of starts and total time scores for each condition studied.

relationship. The best display-control relationship is associated with about one-half the number of false starts associated with the poorest relationship. To determine the approximate number of false starts in any series the approximately 20 correct starts in each series must be subtracted from the total number of starts. It will be noted that differences between the times from relationship to relationship are not as significant as the differences between the number of starts. However, both the times and number of starts indicate the superiority of the CRD relationship.

Under condition N there were four control relationships similar to those in condition C. These relationships were NDR, NUR, NDL, and NUL. For example, NDR differed from CRD because a movement of the control knob mounted in the horizontal plane caused a vertical movement of the spot, and movement of the control knob mounted in a vertical plane caused a horizontal movement of the spot.

Table 10.2 shows the means, standard deviations, correlations, and critical ratios for the number of starts and for the average times for two relationships.

TABLE 10.2.—*Means, standard deviations, correlations, and critical ratios for the number of starts and time scores for condition N*

[N=24]

Comparison	Number of starts			
	Mean	SD	r	CR
NDR	63.92	28.75	0.72	1.19
NUR	58.70	12.85		
NDR	63.92	28.75	.87	.79
NDL	61.62	25.77		
NDR	63.92	28.75	.78	1.65
NUL	57.83	23.35		
NUR	58.70	12.85	.74	.78
NDL	61.62	25.77		
NUR	58.70	12.85	.70	.25
NUL	57.83	23.35		
NDL	61.62	25.77	.80	1.84
NUL	57.83	23.35		
Time (minutes)				
NDL	1.91	0.82	0.40	1.04
NUL	1.74	.55		

Table 10.2 shows that there are no significant differences between the number of starts for any of the relationships in condition N. Similarly there are no significant differences between the times, since the critical ratio given is between the two most divergent mean times.

TABLE 10.3.—*Means, standard deviations, and* t *values for the number of starts and time scores for condition C and condition N*

[N=24]

Comparison	Number of starts		
	Mean	SD	t
CRD	40.58	8.41	3.79
NDR	63.92		
CRU	53.42	16.17	1.26
NUR	58.70	12.85	
CLD	50.67	10.10	1.94
NDL	61.62	25.77	
CLU	61.96	11.65	.79
NUL	57.83	23.35	
CRD	40.58	8.41	3.40
NUL	57.83	23.35	
Time (minutes)			
CRD	1.27	0.48	3.49
NDR	1.80	.58	
CRU	1.41	.42	2.46
NUR	1.77	.57	
CLD	1.37	.41	2.89
NDL	1.91	.82	
CLU	1.65	.39	1.37
NUL	1.74	.55	
CRD	1.27	.48	3.24
NUL	1.74	.55	

While these results are based on only 24 subjects for each relationship it seems apparent that any differences found with a larger number of subjects would not be of great practical importance.

Table 10.3 shows the means, standard deviations, and critical ratios between condition C and condition N for the different relationships. The first four comparisons are between the corresponding movement relationships for the two conditions, that is, a clockwise movement of the knob caused the spot to move in the same direction in each case. The last comparison in each part of the table is between the empirically best control relationship under each condition.

Table 10.3 shows that while the number of starts or times under condition C are usually less than under condition N they are not always significantly different. However, there is no doubt that the empirically best relationship under condition C is significantly better than the best relationship under condition N. This is true for both the number of starts and the time taken. It should be noted that the poorest relationship under condition C causes more starts than the best relationship under condition N.

DISCUSSION AND CONCLUSIONS

These results demonstrate the very large change in efficiency and speed of response that can be effected by changing the relationship between the movement of a control and the resulting movement of the display. Even though the control may be associated with the proper plane of movement of the display its effectiveness may be greatly reduced by having it associated with the display by an inappropriate relationship between its movement and the movement of the display.

Vince[3] has demonstrated that in a very simple display-response problem there are considerably fewer errors in direction of response made when the control-display relationship is in the "expected" direction. This superiority of "expected" display-control relationship held only when the subjects were required to respond very rapidly; as soon as the subjects were given considerable time in which to make this response the "expected" and "unexpected" relationships gave similar results. The present results seem to be consistent with Vince's findings, since in a rather complicated display-control relationship problem it was found that the relationship between the axis of movement of the control and the line of movement of the display was important and also that the relationship between the direction of movement of the display and of the control was important in reducing errors and time. It would seem that, if the display and response required are very simple and if considerable time is allowed for responding, the display-control

[3] Vince, M. A., Direction of movement of machine controls. Flying Personnel Research Committee, *Report No. 637*, August 1945.

relationships may not be important, but if the display or response required is complex or if the operator is hurried, the display-control relationships are of extreme importance.

Data presented by Warrick (see ch. 9) have been interpreted as showing that when the axis of movement of the control is perpendicular to the line of movement of the display a consistent operator preference with respect to direction of response exists, but that when the two movements are not in the same plane there are less marked preferences. The present results seem to lend weight to such an hypothesis since the fewer starts and shorter time were taken by those operators working with the relationship where the axis of control movement was perpendicular to the line of the display. However, the results also show that the advantage of having proper axis of movement relationships may be nullified by improper line-of-movement relationships.

The results of this experiment suggest the following hypotheses:

1. In a display-control relationship it is important to insure that the axis of movement of a rotary control is perpendicular to a line parallel to the line of movement of the display.

2. The relationship between the plane of movement of the display and the axis of the control movement will be effective only if there is also a proper relationship between the direction of movement of the control and the direction of movement of the display.

3. When the controls are located between the operator and the display and the axis of the control movement is perpendicular to a line parallel to and below the line of movement of the display, a knob controlling horizontal movement of the display should be so related to the display that a clockwise movement of the control causes a movement of the display to the right, and a knob to the right of the operator controlling the vertical movement of a display should be so related to the display that a clockwise movement of the control causes a downward movement of the display.

While the hypotheses above appear to be consistent with other research, they should be verified and extended. Particularly should other axis-of-movement, direction-of-movement relationships be determined. Subjects should be run with only one control knob and one direction of movement of the spot. It would also be instructive to substitute levers for the present round control knobs. It is often stated that although a particular display-control relationship may be awkward for the naive subjects, trained operators have so learned the required responses that that particular relationship has become "natural" for them. This hypothesis should be investigated as well as the possibility that the trained operators will revert to their original "natural" response relationships under stress.

SUMMARY

The purpose of this study was to investigate the relationship between the line of movement of a spot on the face of an oscilloscope and the axes of two control knobs located between the subjects and the display. Forty-eight AAF pilots were tested on each of the variables in a counterbalanced sequence. The following conclusions and hypotheses are formulated:

1. When the axis of movement of the control knobs is perpendicular to a line parallel to and below the line of movement of the spot, the best relationships are as follows: (a) for a control knob at the operator's right, controlling up-down movement of the spot, downward movement of the display should be associated with clockwise movement of the control knob, and (b) for a control knob on the operator's left, controlling right-left movement of the spot, movement of the display to the right should be associated with clockwise movement of the control knob.

2. In this study, when the line of movement of the spot was parallel to the axis of movement of the control knobs there were no significant differences for any of the relationships.

3. The empirically best relationship when the axis of control movement is perpendicular to a line parallel to and below the line of movement of the spot was significantly better than the best relationship when the axis of control movement is parallel to the line of movement of the spot.

CHAPTER ELEVEN

Comparative Interpretability of Two Methods of Presenting Information by Radar[1]

H. Richard Van Saun

INTRODUCTION

The physical development of radar, television, and other new and proposed methods of presenting schematic or pictorial information to be utilized by human beings has brought with it numerous psychological problems of perception, orientation, and reaction. Some of these psychological problems are sufficiently unique as to require special investigation. The whole field of electronically instrumented display is in a state of flux, and any experiment, even one on a specific problem, is potentially of widespread application. The present study, which deals with a relatively specific problem of radar presentation, is also of general interest insofar as it throws light on the perceptual problems facing the radar operator.

Operational employment of radar has been considered to require equipment having different attributes for various specific uses. As a navigational aid, it has been assumed that the radar scope should present a display of as great an area of the terrain around the aircraft as possible.[2] The inclusion of azimuth or north stabilization in the equipment, causing the scope display to be oriented with north at the top has also been regarded as a desirable characteristic for navigational use, inasmuch as such an arrangement enables the navigator to transpose the radar display directly to his conventionally north-oriented maps. For other special purposes, however, such as the use of radar for bombing, blind landing, or intensive terrain study,

[1] This chapter is based upon data reported in Headquarters AMC, Engineering Division Memorandum Report No. TSEAA-694-5.

[2] The Plan Position Indicator, or PPI scope, has been primarily used for navigational purposes. This scope presents a 360° polar grid display of the terrain below and around the aircraft for distances up to 100 miles. This display resembles a rough map with land areas, built up areas, and mountains appearing as successively brighter images and water areas appearing as dark or empty spaces.

a restricted area is of interest to the operator. This area, it is presumed, should be presented as large and with as much definition as the equipment is capable of producing. Moreover, it seems desirable for bombing and other special purposes that the radar display be oriented to the heading of the aircraft [3] rather than north as is the case for navigation.

Radar installations for the two jobs, navigation and bombing, have generally been designed in accordance with the assumptions expressed in the previous paragraph. The advent of new, high speed bombardment aircraft, where design characteristics seriously limit the space available for personnel and thus require the consolidation of the duties of navigator, radar operator, and bombardier into one job to be performed by one man, has raised certain problems regarding the design of the equipment to be so used. Two such problems, broadly stated, are first, what type of radar bombing presentation makes for the most rapid and accurate identification of targets first presented on the PPI navigation scope and second, what is the effect on the speed and accuracy of identifying targets when the PPI scope is north-stabilized and the bombing scope is heading-stabilized? It is with these two problems that the present study is concerned.

In attacking the problem of the optimal sector scope to be used in conjunction with the PPI scope, prototypes of the two basic display systems, cartesian grid and polar grid [4] were compared on the basis of the speed and accuracy with which targets presented on the PPI scope were recognized on the sector scopes.

In attacking the problem of the effect on efficiency of shifting the attention from a north-stabilized PPI scope to a heading-stabilized sector scope, a comparison was made of the speed and accuracy with

[3] The radar scope in air-borne operation is conventionally located with the face or viewing screen in a vertical or near vertical plane. The picture presented on the scope is usually heading stabilized, i. e., stabilized so that the radar image of a portion of the terrain lying directly ahead of the aircraft appears first at the top of the scope and (disregarding the effect of wind) gradually passes straight down over the face of the scope and off the bottom as the aircraft flies over the particular area. By tying the radar scope into the fluxgate compass, however, it is possible to effect azimuth or north stabilization which changes the orientation of the radar display so that north is always at the top of the scope, and the radar image of the terrain passes across the face of the scope in accordance with the direction the aircraft is flying. For example, if the aircraft were flying due west the radar image would move across the scope face from left to right.

[4] There are two fundamental systems for the presentation of radar information such as is used by the bombardier and navigator. These are cartesian-grid displays and polar-grid displays. The former presents a square or rectangular picture wherein equal steps of angular direction from the aircraft are indicated by equally spaced parallel vertical lines and equal steps of distance are indicated by equally spaced parallel horizontal lines. The polar-grid display presents a circular (or sector-of-a-circle picture) wherein equal steps of angular direction from the aircraft are indicated by equally spaced lines radiating from a common focus at the center of the circle and equal steps of distance are indicated by equally spaced concentric circles whose centers are common to the center of the circle. Examples of these systems of radar display are found in this report as figures 11.1, 11.2, and 11.3 for polar grid and figure 11.4 for cartesian grid.

which simulated targets were recognized when the scopes had the same orientation and when the PPI scope was north-stabilized and the sector scope was heading-stabilized. The practical significance of the results have immediate and important application in the design and use of radar bombing scopes.

METHOD AND PROCEDURE

Inasmuch as various components of the equipment investigated in this study were hypothetical or existed only in the blueprint stage and inasmuch as a practical, proven, paper-negative technique of producing simulated radar-scope photographs had been developed and utilized in printed tests of radar-operator proficiency by the Psychological Research Project (Radar),[5] the study was designed around a series of simulated radar-scope photos.

To obtain comparable data on the two variables under investigation, 200 items, each consisting of a PPI scope photo with a terrain feature indicated as a target and a sector scope photo with the same target and a number of misleads indicated by letters, were organized into various sections and parts of the test according to the arrangement presented in table 11.1.

TABLE 11.1.—*Composition of test used in evaluating radar presentations*

Section A: All scope photos heading-stabilized.
- Part 1: 50 items, consisting of PPI scope photos followed by polar-grid sector scope photos.
- Part 2: 50 items, consisting of PPI scope photos identical to those used in part 1, followed by cartesian-grid sector scope photos.

Section B: PPI scope photos azimuth-stabilized, sector scope photos heading-stabilized.
- Part 3: 50 items, consisting of PPI scope photos followed by polar grid sector scope photos.
- Part 4: 50 items, consisting of PPI scope photos identical to those used in part 3 followed by cartesian-grid sector scope photos.

As indicated in table 11.1, section A of the test was designed to yield comparative data on the two sector scopes under the condition of heading stabilization of all scopes. Section B was designed to yield similar data under the condition of azimuth stabilization of the PPI scopes and heading stabilization of the sector scopes.

The preparation of the scope photos to fulfill the requirements set up in the test organization was accomplished as described in the following paragraphs.

Seventy-five basic paper negatives for the photos were prepared, 25 each for the PPI scope photos, polar-grid sector scope photos, and

[5] Cook, S. W., ed. *Psychological Research on Radar Observer Training*, AAF Aviation Psychology Program Research Reports, No. 12. Washington: Government Printing Office, 1947. Ch. 5.

cartesian-grid sector scope photos. The size of each type of negative was dictated by the useful area of a 5-inch oscilloscope tube face. Thus the PPI scope photo negatives were 4½ inches in diameter, the cartesian-grid sector scope photo negatives were squares measuring 3³⁄₁₆ inches on a side and the polar-grid sector scope photo negatives were 60° sectors of circles measuring 4½ inches in the greatest dimension.

The outlines of terrain features were drawn on the PPI negatives and shaded to simulate scope photo negatives. (To produce the bright areas of the photo, areas were darkened on the negative.) The sector scope negatives were prepared in the same way, reproducing a 60° sector of the PPI negative. Range marks[6] and heading markers[7] were added to the negatives by drawing suitable circles, arcs, and vertical and horizontal lines. (See figs. 11.1, 11.2, 11.3, and 11.4 for examples of heading markers and range marks.)

In preparing the contact prints from the PPI negatives, a film-transparency azimuth stencil was used to give the prints an azimuth scale, and a small triangular piece of black paper was used as an arrow to indicate targets. By varying the azimuth scale and the arrow, each of the PPI negatives was made to yield two sets of four different problem-presenting photos, to fulfill the test organization requirement set up in table 11.1.

Prior to printing the sector scope photos, the relevant PPI prints were examined to enable the exact designation of the target on the sector scope photos. The targets were designated on the sector scope negatives by letters. In the case of the polar-grid scopes, the letter was one of the series from A to J inclusive. In the case of the cartesian-grid scopes, the letter was one of the series from Q to Z inclusive. The letters used to designate targets were selected at random from within the letter series. Misleads, usually nine in number, were also lettered on the negatives, utilizing the remaining letters of each respective series. The same terrain features were indicated as misleads on both comparable sector scope photos. The order of the letters was randomized. Since all sector scope photos were to be heading-stabilized according to the test organization, four prints of each negative were made with the proper azimuth scale. Figures 11.1, 11.2, 11.3, and 11.4 illustrate the four types of simulated scope photos used in the study.

[6] Range marks may be caused to appear on the scope faces as equidistant bright circles, arcs, or lines, depending on the type of scope. They are used to measure distances of terrain features from the aircraft.

[7] A heading marker may be caused to appear on the scope as a bright line extending from the position of the aircraft to the azimuth scale It indicates the longitudinal axis of the aircraft when the scope is heading-stabilized and, in the case of PPI scopes, it indicates the direction the aircraft is flying when the scope is azimuth- or north-stabilized.

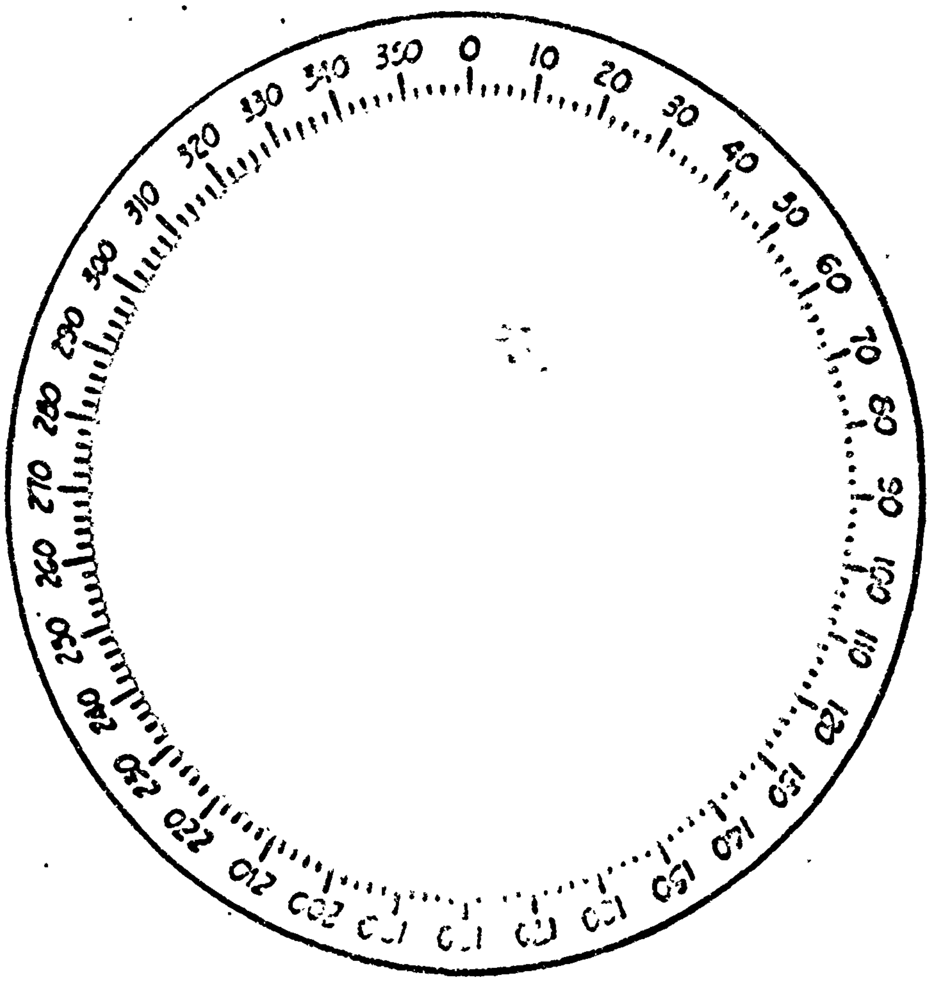


FIGURE 11.1.—Simulated PPI scope photo, heading-stabilized. This is an example of the problem-presenting photos used in the test. A target is indicated by a triangular arrow at about 354 degrees, approximately one and a half range marks from the center of the display. (The bright area in the center of the picture is caused, in actual operation, by strong echoes from the area beneath the aircraft. It can be nearly eliminated by proper set operation and function.)

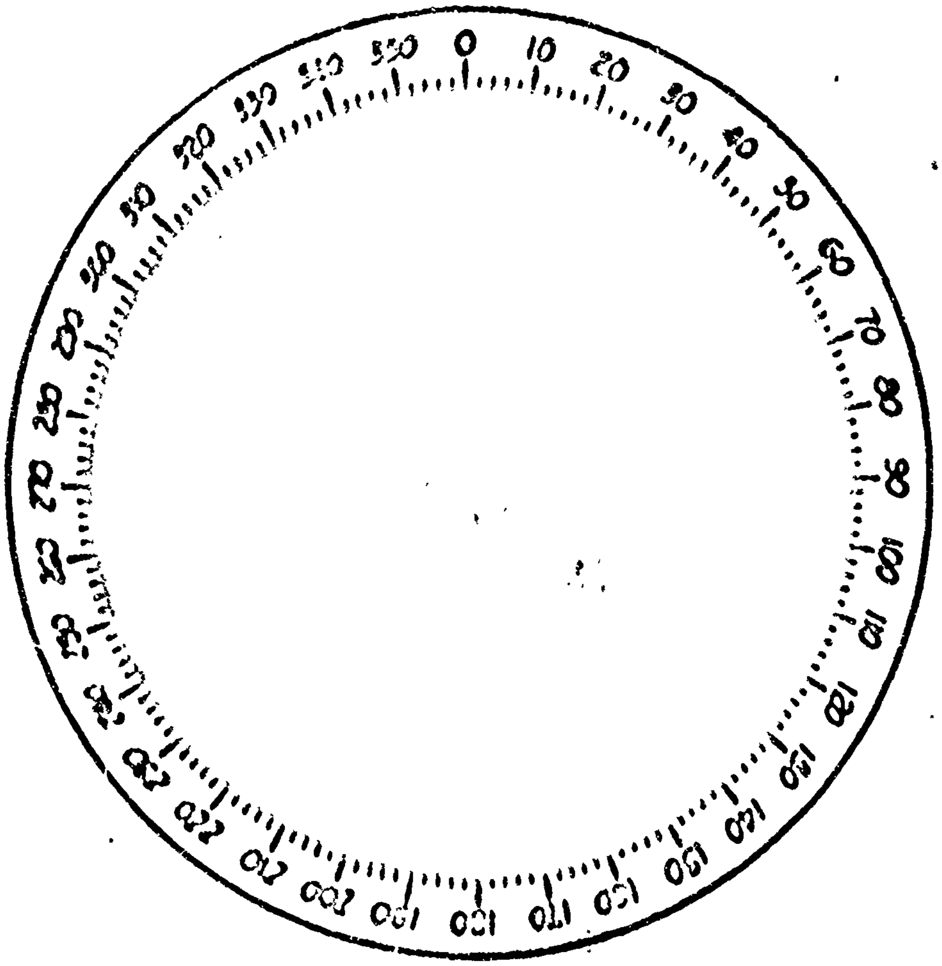


FIGURE 11.2.—Simulated PPI scope photo, azimuth-stabilized. This is also a problem-presenting photo. It presents the same terrain as figure 11.1, but indicates how the radar display would appear if the radar set were azimuth-stabilized and the aircraft heading were 108 degrees. Note that the heading marker points to this value.

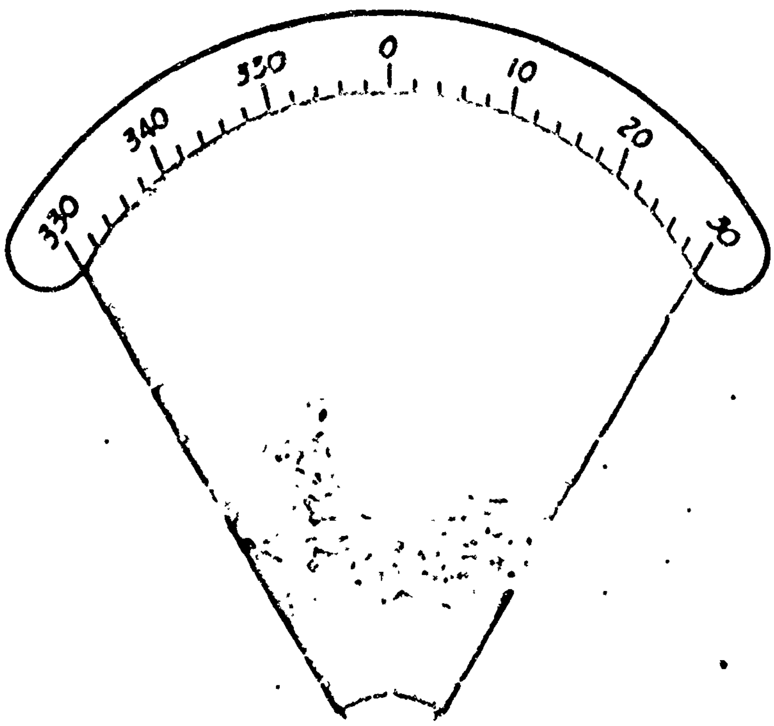


FIGURE 11.3.—Simulated polar-grid sector scope photo, heading-stabilized. This photo presents a polar-grid sector of the display presented in figures 11.1 and 11.2. The sector is 60° wide and is bi-sected by the heading marker. Note that the terrain features are enlarged. The letter "A" indicates the terrain feature corresponding to the target presented in the previous pictures. In the test, heading-stabilized sector scope pictures were used with both heading- and azimuth-stabilized PPI scope photos.

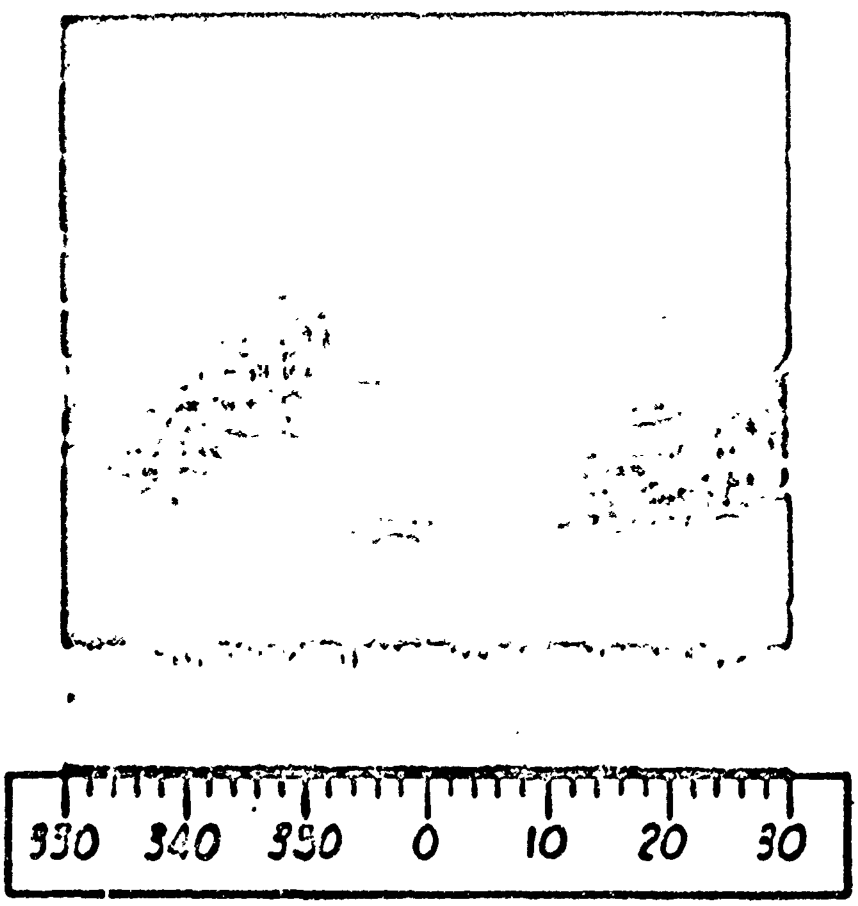


FIGURE 11.4.—Simulated cartesian-grid sector scope photo, heading-stabilized. This photo presents a cartesian-grid display of a sector of the photos in figures 11.1 and 11.2. As indicated by the azimuth scale and range marks, the area represented in this photo is the same as that in figure 11.3. The heading marker bisects the display. Note that the terrain features are enlarged and distorted horizontally. (In this type of display the azimuth scale is at the bottom; the heading of the aircraft is still toward the top of the picture, however. This discrepancy in the location of azimuth scales was dictated by the differences in actual set installations.) The letter "R" indicates the terrain feature corresponding to the target presented in figures 11.1 and 11.2. Note that the misleads on this photo correspond to those in figure 11.3.

The photos were then assembled into the proper pairs, PPI scope followed by sector scope, roughly ordered according to difficulty, and inserted in loose-leaf binders. Each pair of photos constituting an item was inserted so that the PPI, problem-presenting photo was on the right-hand page, followed by the sector photo on the left-hand page, so that it was impossible to view them simultaneously. The PPI photos in each of the four parts of the test presented the same target, item for item. Furthermore, each test part was divided into halves to enable the computation of measures of test reliability. The terrain of each item in the second half of each test part matched that of the first half, item for item, except that a different target was indicated in the items in the second half.

The test was administered in two sessions to 4 groups of 12 pilots each according to the schedule presented in table 11.2.[8]

TABLE 11.2.—*Test administration schedule*

FIRST SESSION

Group	I	II	III	IV
Test component order	Part 1 Part 2	Part 2 Part 1	Part 3 Part 4	Part 4. Part 3.

SECOND SESSION

Group	I	II	III	IV
Test component order	Part 3 Part 4	Part 4 Part 3	Part 1 Part 2	Part 2. Part 1.

Part 1: Heading-stabilized PPI scope to heading-stabilized polar-grid sector scope.
Part 2: Heading-stabilized PPI scope to heading-stabilized cartesian grid sector scope.
Part 3: Azimuth-stabilized PPI scope to heading-stabilized polar-grid sector scope.
Part 4: Azimuth-stabilized PPI scope to heading-stabilized cartesian-grid sector scope.

In general, because of administrative convenience, the two testing sessions were separated by 24 hours. A few random subjects, however, were administered the test with a shorter intervening period between the sessions, and one subject with a longer intervening period between sessions. In every case, however, there was some intervening activity, if only a short break for a cigarette. Inspection of the data for these subjects did not indicate any material deviation from the general configuration of answers.

Each part of the test was administered with a 10-minute time limit. At the end of 5 minutes the subject was started on the second half of the items even though he had not finished the items in the first half.

Prior to the first testing session, subjects, predominantly without radar experience, were given extensive standardized instruction in which a sample PPI scope photo and both types of sector scope photos were employed. Both speed and accuracy were stressed as being important in locating the targets on the sector scopes.

[8] Miss Sally Bedworth assisted in administering the tests and computing the statistics in the study. Mrs. Sue Diggles assisted in the preparation of the scope photo negatives.

The procedure followed in taking the test, and so outlined in the instructions, was for the subject to study the PPI scope photo until he had the size, shape, and location of the target well in mind and then turn to the sector scope photo and select the letter most nearly indicating the terrain feature he believed was the target. He then printed the letter on his answer sheet opposite the number of the item and proceeded immediately with the next item. He was not allowed to turn back to the PPI photo.

RESULTS

The results of administering this material to the 48 pilots were analyzed in terms of the number of targets correctly recognized on the sector scopes and in terms of the number of errors in target recognition within the time limits of each part of the test. Tables 11.3 and 11.4 present the results of the test administration. Figure 11.5 is a series of bar graphs summarizing the pertinent differences in performance with the two sector scopes.

TABLE 11.3.—*Means, standard deviations, and significance of differences between the sector scopes in terms of the average number of targets correctly recognized and the average number of errors in target recognition*

Type of sector scope	PPI and sector scopes heading-stabilized				
	Mean number of targets correctly recognized	SD	SE_{diff}	CR	r
Polar grid	25.75	8.59	1.14	[1] 2.68	.60
Cartesian grid	22.60	8.89			
	Mean number of errors				
Polar grid	11.54	5.07	0.65	[1] 5.02	.67
Cartesian grid	14.79	5.86			
	PPI scopes azimuth-stabilized: sector scopes heading-stabilized				
	Mean number of targets correctly recognized	SD	SE_{diff}	CR	r
Polar grid	17.81	8.44	0.95	[2] 2.25	.71
Cartesian grid	15.67	8.94			
	Mean number of errors				
Polar grid	13.46	6.16	0.78	0.96	.64
Cartesian grid	14.21	6.50			

[1] Significant at 1-percent level.
[2] Significant at 5-percent level.

TABLE 11.4.—*Means, standard deviations, and significance of differences between sector scopes under conditions of azimuth stabilization of PPI scopes and heading stabilization of PPI scopes*

PPI Scopes	Polar-Grid sector scope				
	Mean number of targets correctly recognized	SD	SE_{diff}	CR	r
Heading stabilized	25.75	8.59	1.27	[1] 6.25	.46
Azimuth stabilized	17.81	8.44			
	Mean number of errors				
Heading stabilized	11.54	5.07	.76	[1] 2.53	.57
Azimuth stabilized	13.46	6.16			
	Cartesian-grid sector scope				
	Mean number of targets correctly recognized	SD	SE_{diff}	CR	r
Heading stabilized	22.69	8.89	1.52	[1] 4.62	.30
Azimuth stabilized	15.67	8.94			
	Mean number of errors				
Heading stabilized	14.79	5.86	.74	[2] .78	.94
Azimuth stabilized	14.21	6.60			
	Sector scope data combined				
	Mean number of targets correctly recognized	SD	SE_{diff}	CR	r
Heading stabilized	48.44	15.61	2.79	[1] 5.36	.35
Azimuth stabilized	33.48	16.07			
	Mean number of errors				
Heading stabilized	26.33	9.97	1.37	.98	.63
Azimuth stabilized	27.67	11.47			

[1] Significant at 1-percent level.
[2] Difference is inconsistent with other results, more errors being made under conditions of the PPI scope being heading-stabilized than under conditions of the PPI scope being azimuth-stabilized. However, the proportion of errors is much less for the heading-stabilized condition.

703333—47——12

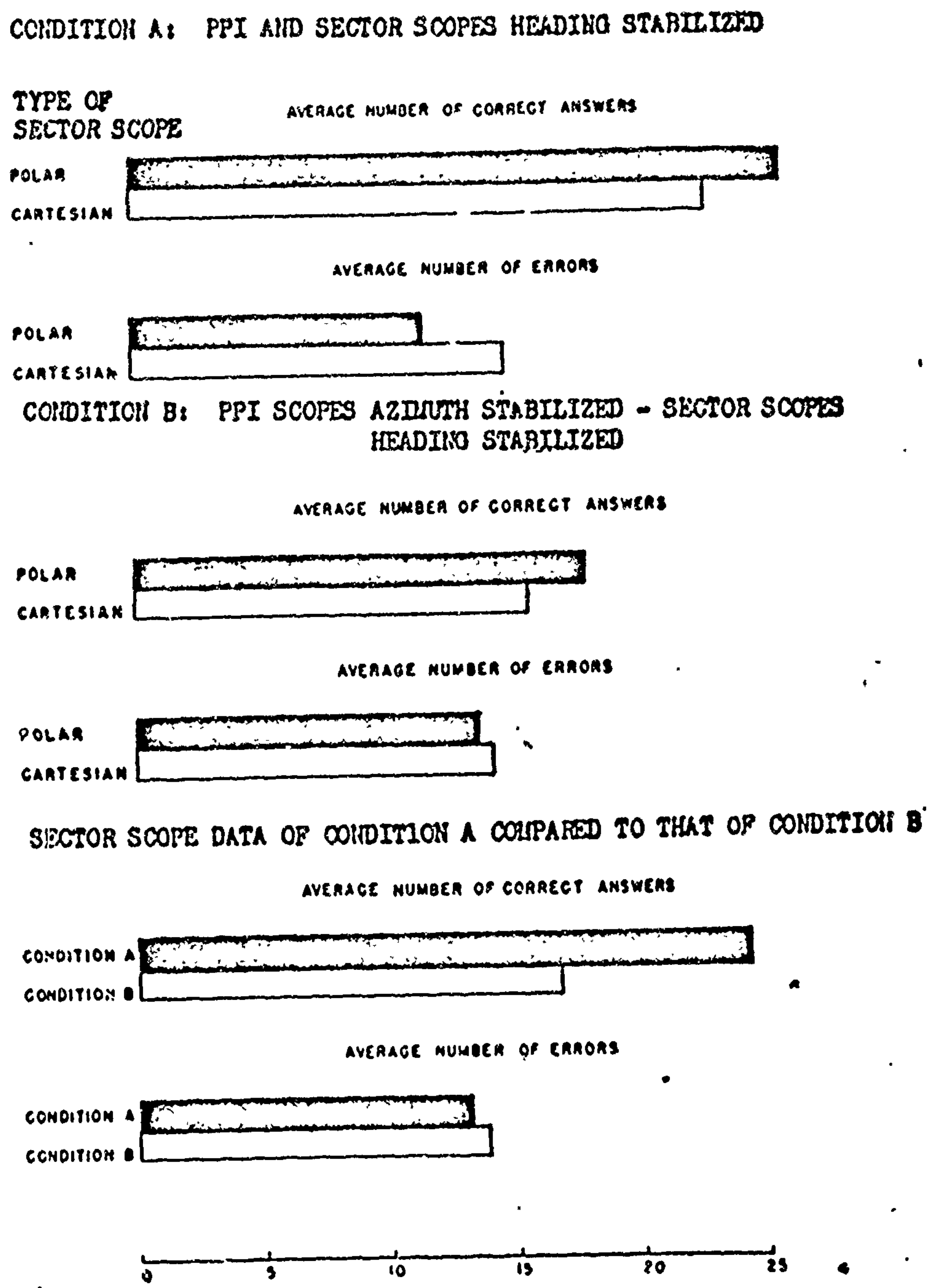


FIGURE 11.5.—Graphic presentation of the results of the sector scope comparisons.

First-half versus second-half test reliabilities were computed in terms of the number of targets correctly recognized and in terms of the number of errors in target recognition for each of the four parts of the test. These relationships are shown in table 11.5.

TABLE 11.5—*Reliabilities of the different parts of the test*

	rll (split-half) [1]
Section A: All scopes heading-stabilized:	
Part 1: Polar-grid sector scope:	
Targets correctly recognized	0.90
Errors in target recognition	.69
Part 2: Cartesian-grid sector scope:	
Targets correctly recognized	.82
Errors in target recognition	.70
Section B: PPI scopes azimuth-stabilized; sector scopes heading-stabilized:	
Part 3: Polar-grid sector scope:	
Targets correctly recognized	.87
Errors in target recognition	.77
Part 4: Cartesian-grid sector scope:	
Targets correctly recognized	.83
Errors in target recognition	.81

[1] Corrected by Spearman-Brown formula.

Learning curves were prepared in terms of the average number of targets correctly recognized and in terms of the average number of errors in target recognition. These averages were computed from all the data on the basis of the order in which the material was presented to the subjects, thus counterbalancing the relative difficulty of the different parts of the test. The curves are shown in figure 11.6.

DISCUSSION

The results of the analysis of the number of targets correctly recognized and the number of errors in target recognition presented in tables 11.3 and 11.4 definitely indicate that in terms of the number of targets correctly recognized the polar-grid sector scope is superior to the cartesian-grid sector scope, both when the PPI scope and the sector scope have the same orientation (both scopes heading-stabilized) and when the scopes are differently oriented (PPI scope azimuth-stabilized and sector scope heading-stabilized). One of the critical ratios between the means for this factor is significant at the 1-percent level and the other is significant at better than the 5-percent level.

The data concerning the number of errors in target recognition are not as clear cut as are those for the number of targets correctly recognized. Only the critical ratio between the mean number of errors when the sector scopes and PPI scopes have the same orientation is significant (1-percent level). The general pattern of critical ratios for error scores on this variable, however, are all in the same direction and the actual proportion of errors is considerably smaller. These facts tend to support the superior interpretability of the polar-grid sector scope.

It is believed that the superior performance exhibited with the polar-grid display is due to the fact that it expands the image of the terrain proportionately in all directions thus retaining the original configura-

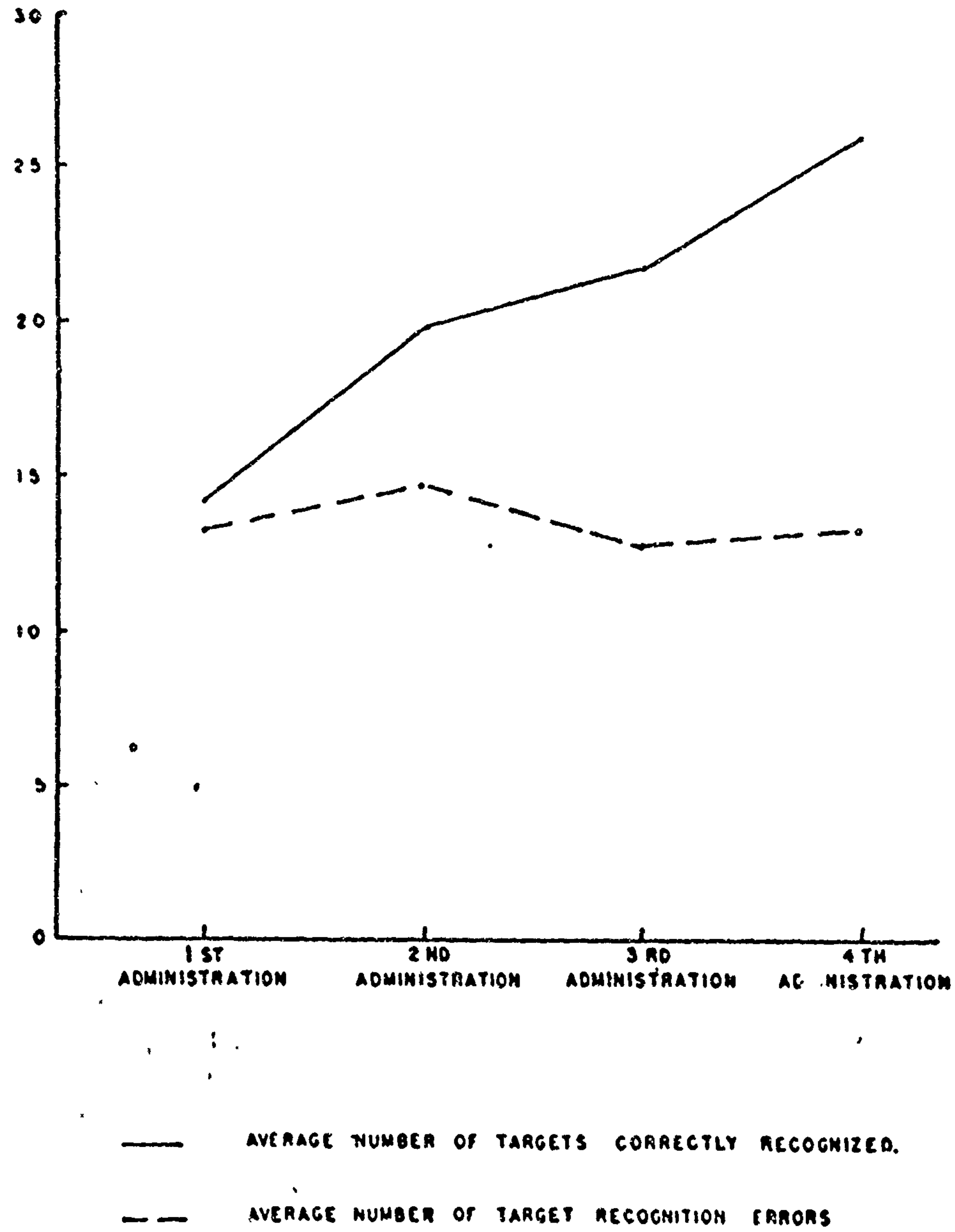


FIGURE 11.6.—Graphic presentation of the change in performance with learning.

tion, whereas the seemingly advantageous horizontal expansion of the cartesian-grid display results in a distorted image that sufficiently masks the original configuration of terrain features so as to make them difficult to recognize.

The data concerning the effect of PPI scope orientation on sector scope interpretability presents a picture similar to that discussed in the previous paragraphs. Table 11.4 shows that, in terms of the number of targets correctly recognized, the situation wherein both PPI and sector scopes have like orientation (both scopes heading-stabilized) is far superior to the situation where the scopes are differently oriented (PPI scope azimuth-stabilized and sector scope heading-stabilized).

More targets were correctly recognized in the former situation on both types of sector scopes, the critical ratios all being significant at the 1-percent level.

The error scores are again less informative, only the critical ratio between the mean number of errors on the polar-grid scope showing a significant difference (1-percent level) between the two conditions. In the other situation, concerning the cartesian-grid sector scope, fewer errors were made when the PPI and sector scopes were differently oriented. Other than this instance, however, the data tend to support the superiority of performance when both PPI scope and sector scope are heading-stabilized.

The superior performance demonstrated with displays having identical orientation is to be expected, inasmuch as the change in orientation of the sector scope, manifest when it is heading stabilized and when the PPI scope is azimuth stabilized, adds but another confusing factor to the altered figure-ground relationship inherent in the problem of shifting the attention from one scope to the other.

The reliability coefficients presented in table 11.5 indicate adequate test reliability. The learning curves presented in figure 11.6 indicate that a relatively large amount of learning took place during the test administration, the proportion of errors materially decreasing as testing progressed. The presence of learning suggests that the test might well be readministered to a group of experienced radar operators to check the results. Inasmuch as the effects of learning were counterbalanced by the order of administration [9] (see table 11.2), it is doubtful if such a readministration of the test would materially affect the results of the present study.

Possible future research, in addition to the readministration of the present test to experienced radar operators discussed in the previous paragraph, might profitably be directed toward investigating the effect that simultaneous presentation of the PPI and sector displays would have on the relative interpretability of the two types of sector scopes.

Another profitable avenue of future research might be an investigation of the effect memory-point markers would have on the relative interpretability of the two types of sector scopes. Memory-point markers, as proposed, consist of some indication such as bright cross hairs, which may be caused electronically to appear on a target in the PPI display and maintain their relative location when the scope is converted to a sector presentation, or appear simultaneously on a separate sector presentation.

[9] In combining the data for the various parts of the test it was assumed in one case that taking parts 1 and 2, in that order, was equivalent to taking the parts in the order 2 and 1. The same assumption was made in another case concerning the experience value of parts 3 and 4 as compared to that of parts 4 and 3.

Furthermore, it would be very desirable to repeat the present study and to conduct additional investigations with dynamic display, i. e., displays more closely simulating actual operation by changing in relation to the aircraft's movement over the terrain. Such investigations would provide a criterion for the methodology of using static scope photos in studies of this type and would make possible the study of dynamic factors of scope interpretation *per se.*

SUMMARY

A test of 200 items, each consisting of a simulated PPI scope photo and a simulaed sector scope photo, was administered to 48 pilots in an effort to establish (1) the relative superior interpretability of one of two types of sector scopes, a polar-grid sector scope and a cartesian-grid sector scope, and (2) the effect on sector scope interpretability of like orientation of the PPI and sector scopes compared to unlike orientation of the scopes.

The results of the administration support the following conclusions:

1. The polar-grid sector scope is superior in interpretability to the cartesian-grid sector scope in terms of the average number of targets correctly recognized on the sector scopes.

2. Both sector scopes are more easily interpreted when the PPI scope and sector scopes have the same orientation in terms of the average number of targets correctly recognized on the sector scopes.

CHAPTER TWELVE

A Psychophysical Investigation of Ability to Reproduce Pressures[1]

WILLIAM O. JENKINS

INTRODUCTION

The present series of studies are concerned with the accuracy with which pressures can be reproduced or discriminated on aircraft-type controls in a laboratory situation. The problem is of interest both for the theoretical aspect of the operation of Weber's Law in pressure discrimination, and for the design of control mechanisms which involve the pressure cue, particularly pilot controls in aircraft.

The execution of maneuvers in an aircraft requires the pilot to learn and to apply varying degrees of control pressure in various directions. Introduction of booster systems into aircraft control mechanisms reduces the absolute amount of pressure which needs to be applied. The practical question arises whether these reduced pressures result in a greater or lesser degree of variability and error in control operation by the pilot.

It is widely reported by pilots that "feel of controls" is a basic cue in flying. It appears reasonable to assume that accuracy in discriminating pressures is an important element of "feel of controls." A desirable aim in the designing of control systems is to achieve that "feel of controls" which will yield most accurate performance. From the engineering standpoint the problem becomes one of maximizing "feel of controls" according to the criterion of accuracy of human operation. From a behavioral point of view the question concerns the operation of Weber's Law over a wide range of pressure; that is, the determination of the relative and absolute accuracy of discrimination or reproduction of pressures applied by the hand or foot in different directions.

The limitations of generalizing the results of a laboratory experiment of this kind to actual conditions of flight are recognized. A number of cues which are probably used by the pilot in controlling the

[1] This chapter is based upon research findings reported in Headquarters AMC, Engineering Division Memorandum Reports Nos. TSEAA-694-3, TSEAA-694-3A, and TSEAA-694-3B.

aircraft in the latter situation are absent in the former. These cues include extent and rate of control movement, and visual and auditory stimuli. On the other hand, the laboratory situation permits the isolation and control of the variable under examination whereas at this stage these possibilities do not exist in flight. Research should be undertaken on the other variables alone and in combination with the one studied here before design principles for practical use are made available to engineers (see ch. 3 for a treatment of a number of variables basic to the motor side of equipment design).

A vast amount of data has been accumulated in psychophysical studies of human capacities for discrimination (5, 12). While considerable information has been gathered in related fields, namely weight discrimination and discrimination of pressures applied to body surfaces, there do not appear to be any studies directly concerned with the problem of pressure discrimination in the present sense and the application of Weber's Law in this area. Several studies (3, 6, 7) have been concerned with the problem of the maximum force exertable on hand and foot controls. The findings are summarized in chapter 3.

In a study related to the present one, Hick (8) investigated the precision with which small increments or decrements could be made to relatively constant basic muscular forces in a nearly isometric situation. Subjects were required to push or pull a hand control against basic forces ranging up to 5 pounds. With a basic force of 4 pounds a constant error in the direction of overshooting was found of the order of 5 to 15 percent. The overshooting was most marked when the subject was letting off force, particularly when relaxing from a steady pull. The standard deviation appeared to be independent of the basic force within plus or minus 3 percent, but an examination of Hick's data indicates that the Weber Fraction was not constant throughout the small range of standard values employed.

The studies reported below were concerned with the accuracy with which pilots and nonpilots could reproduce pressures ranging from 1 to 60 pounds on a stick-type control, a wheel-type control, and rudder-pedal-type controls.

APPARATUS AND PROCEDURE

The first study dealt with accuracy of pressure reproduction with a stick-type control. Preliminary testing indicated that a hydraulic system was inadequate for use in the study because of friction and lag and because the cue of extent of movement was not minimized. It was also found that sufficiently sensitive equipment was not immediately available to record through strain gages changes in resistance induced by application of pressure to the handle of a semirigid stick.

The method[2] found acceptable for use was an optical system in which a semirigid steel control stick, one half inch square, was mounted as shown in figure 12.1. The stick was welded to a plate which, in turn, was bolted to a reinforced stand. Pressures applied to the handle of the control stick resulted in slight movements which were transmitted over two cables, one for lateral and one for longitudinal action, to a mirror pivoted for movement in both dimensions.

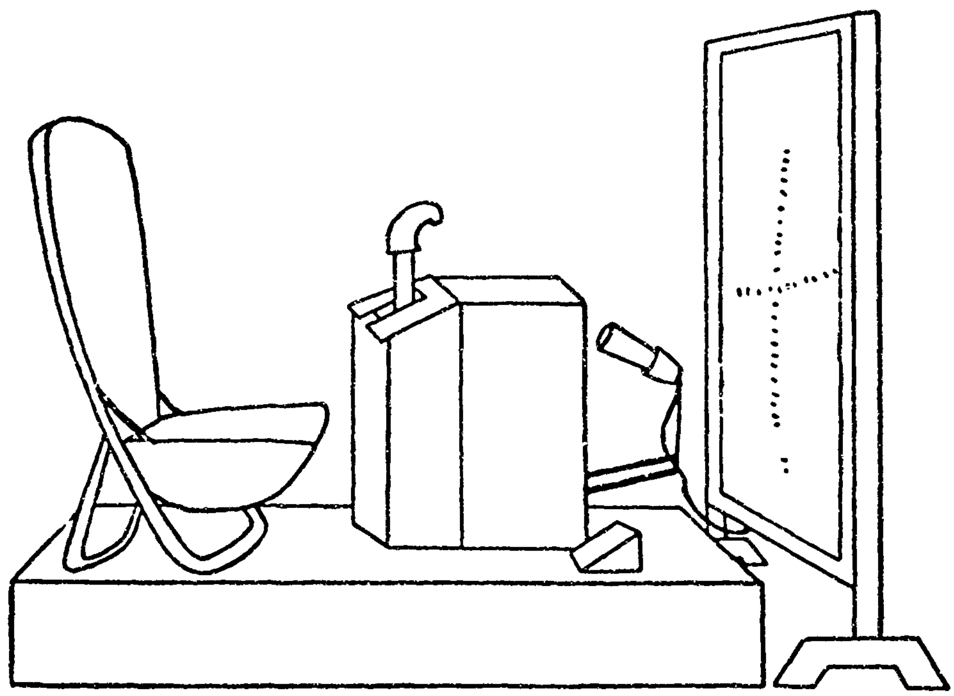


FIGURE 12.1.—Sketch of the apparatus employed in a psychophysical study of ability to reproduce pressures on a stick control.

These cables worked against two springs which were connected from the mirror to its mounting. A light beam from an adjustable galvanometer lamp was projected onto the mirror and reflected from it to a screen. This indicating system was approximately isometric in nature with ¾-inch of movement resulting at the stick handle when a pressure of 50 pounds was applied.

The screen was calibrated directly in pounds by means of weights placed in a pan which was attached to the handle of the stick by a cable passing over a pulley. The reliabilities of the calibration and of the recording procedure were found to be satisfactory.

The indicating systems employed in the studies of performance with wheel-type and rudder-pedal-type controls were essentially similar to that described above. In the first case an aircraft wheel was welded to the same semirigid shaft employed with the stick. The shaft was

[2] The aid of J. R. Brick, M. B. Allenstein, and A. Pepitone in designing and constructing this apparatus is gratefully acknowledged. The hydraulic system in the preliminary investigation was designed by J. R. Brick and the study was carried out by A. Pepitone.

welded to a plate which was rigidly fixed in a metal framework in such a manner that the shaft was parallel to the floor and the wheel could be operated with both hands for aileron (right-left) action. The shaft turned in a ball bearing pressed into a block of wood bolted to the floor of the apparatus. The torque effect about the shaft was measured.

In the study involving the rudder pedals a rigid 1-inch bar of steel was welded at right angles to a ¾-inch steel shaft which, in turn, was welded to a plate. The plate was bolted to the floor of the apparatus. Corrugated metal pedals were welded to the ends of the bar. The torque around the ¾-inch steel shaft was measured. This apparatus was less sensitive than the apparatuses employed in the other studies, moving only ½-inch for 60 pounds pressure applied to one of the pedals. The units of the calibrated scale were not as fine, being in 1-pound units as compared with ½-pound units in the other studies.

In determining the position of the seat relative to the controls and the height of the control anthropological principles of cockpit seating were followed **(10)**.

An Army Air Forces specification **(1)** sets forth the maximum and minimum desirable pressure limits in aircraft control mechanisms. In the present investigation the range of standard pressures employed went below the prescribed lower limit, and in most cases approximated the upper limit. The only major discrepancy was in the case of the rudder pedals, where the upper limit was 60 pounds as compared with the specified value of 180 pounds.

The standard pressures employed in the case of the stick were, for fore-and-aft action, 1, 5, 10, 20, and 40 pounds; for lateral action, they were 1, 5, 10, 20, and 30 pounds. Subjects had considerable difficulty attaining pressures of 40 pounds laterally, so that 30 pounds was made the upper limit for both right and left action. The corresponding values for performance with the wheel operated with both hands to the right and left were 1, 5, 10, 20, 30, and 40 pounds; with the rudder pedals worked separately with each foot, they were 5, 10, 20, 40, and 60 pounds. In the latter case it was found that merely resting the foot on the pedal caused a deflection of the light beam beyond 5 pounds, so that 1 pound was employed with only a few subjects for exploratory purposes. There were, of course, only 2 directions of control operation for the wheel and rudder pedals in these studies.

The order in which the various combinations of pressures and directions were presented to the subjects was randomized. A few adjustments were made in the orders so that sufficient information could be collected concerning the effect of a preceding trial at a high pressure on a succeeding trial at a low pressure, and vice versa.

The procedure was as follows. Each subject was strapped in the cockpit seat, which was adjusted until he reported he was comfortable.

He listened to the instructions explaining the purpose and procedure of the experiment and then put on a pair of blacked-out goggles. For the first test, with a particular pressure and direction of control operation, the subject was given practice trials until he was approximating the desired pressure. During these trials, he gradually applied pressure and was told when he had attained the specified value. The practice trials were not scored. The practice session was followed by 20 successive test trials (15 in the case of the wheel and rudder pedals) for the same pressure and direction. During these trials the subject applied pressure gradually out to the point that he thought was the desired value and then returned the control to the zero position. He was told the pressure he had attained at the end of each successive trial. The same practice and test procedure was employed for the other pressures and directions.

It will be noted that the technique is the psychophysical "Method of Average Error." One modification was introduced which consisted in not permitting the subject to make adjustments about the pressure under test as in the usual psychophysical situation, but rather to have him move the stick gradually out towards the desired value until he attained what he thought was the specified pressure, and then return the control to zero. The maximum excursion of the light spot was scored. This change was introduced in order to simulate somewhat more exactly the performance of a pilot in manipulating controls in flight.

Three independent groups of 20 Army Air Forces pilots each were employed as subjects with the stick, wheel, and rudder controls. A group of 13 nonpilots was also tested with the stick-type control with the procedure described above. The performance of two special groups was measured with the stick control. The first consisted of 11 AAF pilots tested with a learning procedure in which the practice trials were omitted in order that the course and extent of learning in this task could be determined. The second group was composed of 28 AAF pilots of the standard and learning groups who applied pressures without knowledge of results for 3 successive trials at each pressure in each of the directions after completion of the regular experiment in order to determine performance trends in the absence of this cue.

The difference limens (DL's) were computed by the standard technique for the method of average error by computing the standard deviation of each individual's performance about his own average at each point, combining individual sigmas to obtain the SD of the group, and dividing the standard deviation by the standard at each point. The constant errors (CE's) were, of course, taken as the difference between the standard and average attained pressures at each point.

RESULTS

The results from the three investigations are presented separately in the following paragraphs with a final section devoted to a comparison of performance with the three types of controls.

Stick-type Control

The data were collected separately for the four directions of control operation (front, back, right, and left). The analysis began with these results which are summarized in table 12.1 for the group of 20 pilots tested with the standard procedure. An examination of the standard deviation and DL values in table 12.1 reveals that there are no consistent differences favoring one of the directions. The only difference exceeding the 5-percent level of significance is that between the standard deviation for 1 pound to the front in comparison with the other 3 standard deviations involving 1 pound. Comparable results were found for the other 3 groups. In view of these findings the data for the different directions of control operation have been combined in the succeeding treatment.

The basic data of this study are summarized in table 12.2 and figures 12.2 and 12.3. It can be seen that the standard deviations (SD) increase directly as a function of the magnitude of the standard

TABLE 12.1.—*Standard deviation (SD), difference limen (SD/S), and constant error (CE) in pounds for 20 pilots tested with 4 different directions of stick-type control operation*

Standard pressure in pounds	Front			Back			Right			Left		
	SD	SD/S	CE	SD	SD/S	CE	SD	SD/S	CE	SD	SD/S	CE
1	0.16	0.16	0.07	0.20	0.26	0.10	0.20	0.20	0.08	0.20	0.20	0.09
5	.48	.10	.13	.56	.11	.10	.49	.10	.09	.50	.10	.11
10	.88	.09	.14	.85	.08	.17	.66	.07	−.01	.91	.09	.05
20	1.34	.07	.08	1.34	.07	.08	1.50	.08	−.02	1.60	.08	.03
30	......	......	......	......	......	......	2.01	.07	−.20	1.98	.07	.07
40	2.42	.06	−.47	2.34	.06	.04	......	......	......	......	......	......

TABLE 12.2.—*Standard deviations (SD), difference limens (SD/S), and constant errors (CE) in pounds for four directions of stick operation combined for 20 pilots and 13 nonpilots tested with a standard procedure, 11 pilots tested without practice trials (learning), and 28 of the same pilots tested without knowledge of results*

Standard pressure in pounds	20 pilots: Standard procedure			13 nonpilots: Standard procedure			11 pilots: Learning procedure			28 pilots: No knowledge of results		
	SD	SD/S	CE	SD	SD/S	CE	SD	SD/S	CE	SD	SD/S	CE
1	0.21	0.21	0.08	0.32	0.32	0.13	1.30	1.30	0.49	0.33	0.33	0.62
5	.51	.10	.11	.67	.13	.20	1.25	.25	.32	1.13	.23	1.46
10	.83	.08	.09	.99	.10	.22	1.47	.15	.22	1.53	.15	1.62
20	1.46	.07	.04	1.66	.08	−.01	2.64	.13	−.14	1.83	.09	.56
30 [1]	1.99	.07	−.07	2.18	.07	−.19	3.79	.13	−1.15	1.94	.06	−1.53
40 [2]	2.40	.06	−.22	2.61	.06	−.22	4.69	.12	−1.11	2.85	.07	−.03

[1] Right and left directions only.
[2] Fore and aft directions only.

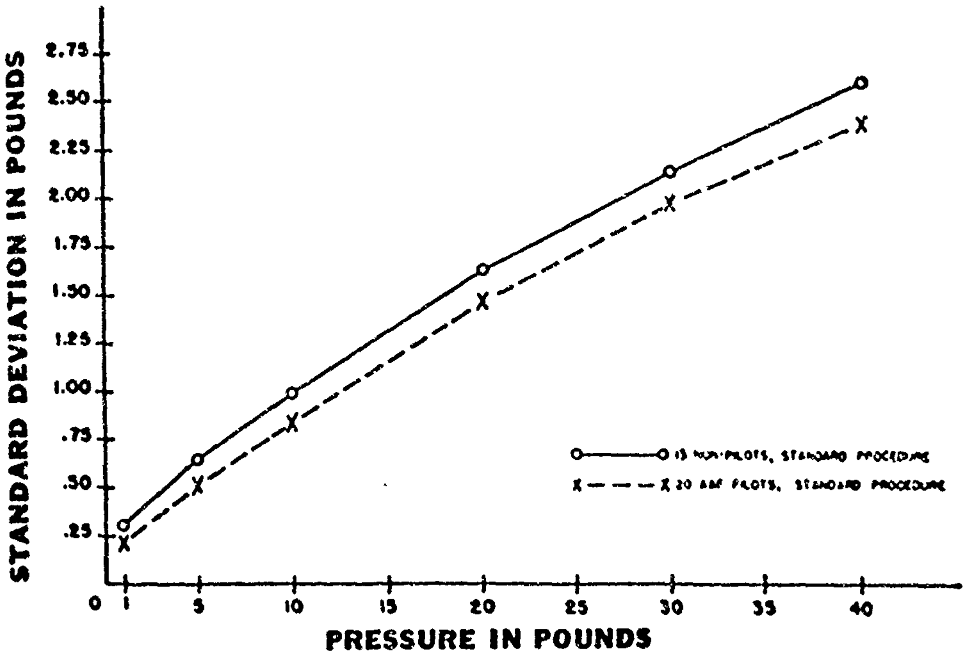


Figure 12.2.—Standard deviations plotted as a function of the standard pressures for 13 nonpilots and 20 pilots.

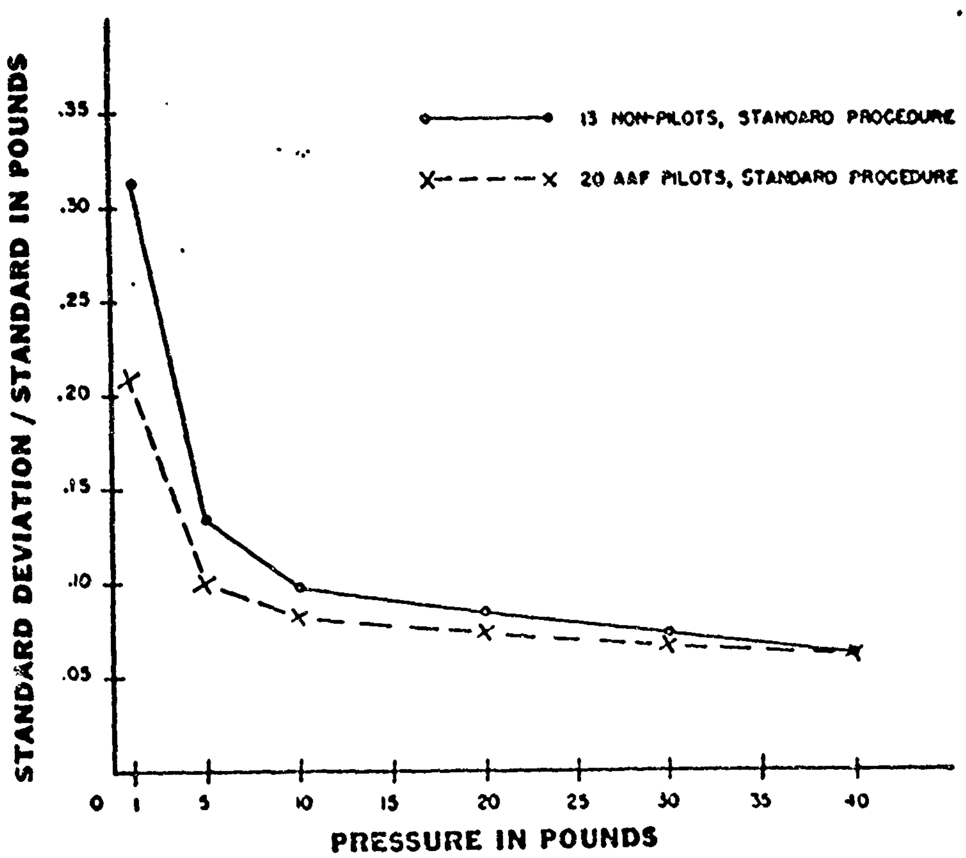


Figure 12.3.—Difference limens (standard deviation/standard) plotted as a function of the standard pressures.

pressure. While it was found that a straight line fitted these data somewhat better than a power function, it is obvious that a larger number of points are needed, particularly at the high pressures, before a final conclusion concerning the shape of the curve can be drawn (2). It is possible that the exponential function would have fitted the points better than the straight line if lower and higher pressures had been employed since, as Guilford (4, p. 79) has pointed out there is ". . . a stubborn theoretical demand that the curve shall pass through the origin." Hovland's (9) data concerning estimation of line lengths indicate a better fit for a power function as contrasted with a straight line.

It can be seen from figure 12.3 that the Weber Fraction is not a constant, but decreases markedly from 1 to 10 pounds. The difference in the DL's between 1 pound and 10 pounds is significant at the 1-percent level of confidence for all groups. The difference between the DL's for 1 and 5 pounds is significant at approximately the 10-percent level. For the pilot and nonpilot groups (of 20 and 13 subjects respectively) tested under standard conditions the DL for 5 pounds was not significantly greater than the DL for any other standard at the 5-percent level. Beyond 10 pounds Weber's Law appears to hold moderately well with the curves exhibiting approximate linearity for both of the standard groups.

It is of interest to note that individual differences were not large either in kind or magnitude. The shapes of the DL curves for individual subjects were quite similar, exhibiting the same downward trend from 1 pound to an approximate asymptote at 10 pounds for 92-percent of the curves (33 subjects of the standard groups each tested in 4 directions of control operation).

Another item apparent from table 12.2 is the fact that the pilot group tested under standard conditions exhibited a lower standard deviation, and, correspondingly a lower DL at all 6 points as contrasted with the nonpilot group. When the data for the 4 different directions were compared it was found that the pilot group yielded lower values at 19 of the 20 points, with the one exception being a tie. The chances of this situation's arising by chance are less than 1 in 100 (11).

The difference in DL's between the pilots and nonpilots can be seen to decrease from 1 pound to 40 pounds. The critical ratio of the over-all difference is 2.8, a value significant beyond the 1-percent level. The major background factors differentiating the pilots from the nonpilots appeared to be that the pilot group had undergone the psychological and physical selection procedures and had had considerable flying experience.

The performance of 11 pilots tested without practice trials in order to determine the course and extent of learning is also summarized in table 12.2. It can be seen that the DL's follow the same general function for this group as for those tested under the standard procedure, but that the absolute values are shifted upward in all cases. It might be noted that a majority of the upward shifting appears to be attributable to the early trials in a given test session where deviations from the standard were greatest. Such a trend was not found in the groups tested under the standard procedure where practice trials were given. The course of learning indicated a tendency for the 11 pilots to start above low pressures and learn to come down to them, and, conversely, start below high pressures and come up to them as learning proceeded. Small amounts of learning occurred in the middle range of pressures, i. e., 10 and 20 pounds. Learning continued for 8 to 10 trials before a plateau was achieved. These findings are in line with the data concerning constant errors which are treated below.

The results for the men tested without knowledge of results yielded values at the high pressures comparable to those obtained when the pilots were tested with it, as can be seen in table 12.2. At pressures of 1, 5, and 10 pounds, however, the SD's and DL's of this group are significantly poorer at the 5-percent level as compared with the corresponding data for the standard pilot group. The shape of the function of DL's plotted against standards is approximately the same as that for the other groups.

The data regarding the magnitude and direction of the constant errors for the four groups are summarized in tables 12.1 and 12.2. It can be seen that in most cases CE's for low pressures are positive and those for the high ones are negative. On the average the constant errors are smaller for the pilot group than for the nonpilot group and those for the group tested without knowledge of results are greatest of all followed closely by those of the group tested under the learning procedure. None of the differences in CE's between the performance of the pilot and nonpilot groups tested under standard conditions is significant at the 10-percent level of confidence. Over all, the CE's for the learning group and for those pilots tested without knowledge of results are significantly greater at the 5-percent level than the corresponding figures for the standard groups. The direction of the constant errors was not greatly different for the four planes of operation. The magnitude of the errors was slightly, but not significantly, greater for the fore-and-aft dimension than for the right-and-left plane.

For the pilot group a comparison was made of amount of flying experience with performance in the present experiment. In the group of 20 pilots, 8 men had less than 800 flying hours and 12 had flown more than 1,000 hours. There were no differences significant at the 5-percent level or better for the following measures: mean performance and standard deviation of performance for 40 pounds and for 1 pound. The differences for other measures were practically zero.

In setting up the order of presentation of the various pressures in combination with the four directions of movement for the stick control a random order was employed. In order to check for interaction effects an analysis was made of performance of a high pressure which had been preceded by a low one in contrast to a high pressure preceded by a medium or high one. A corresponding analysis was made for low pressures. The differences were in the expected directions; e.g., when a low pressure preceded another low one, the mean of the second was smaller than when the low pressure was preceded by a heavy one, but no differences were significant at the 5-percent level of confidence. It appears likely that interaction effects did not distort the present findings to any appreciable extent.

It was thought desirable to check the effect of body weight on performance in the present task, particularly for the high pressures. Correlations were computed between mean attained pressures and standard deviations of these on the one hand and body weight on the other. The resulting values were distributed about equally around a mean of zero with a range of −0.35 to 0.30. None of them was significantly different from zero at the 1-percent level of confidence.

The intercorrelations among standard deviations and constant errors were computed. Weber's Law assumes a perfect correlation of the errors of observation whereas the Fullerton-Cattell Square Root Law assumes zero correlation (5, p. 138). The obtained values for standard deviations for the 20 pilots and 13 nonpilots for the aft direction ranged from 0.15 (1 pound versus 40 pounds) to 0.53 (5 pounds versus 20 pounds) with an average of 0.30. Five of the 10 correlations were significantly different from zero at the 5-percent level, but only two coefficients were significant at the 1-percent level in this small sample. There was no systematic variation in the cofficients. These data on this small sample suggest a closer approximation to Guilford's Generalized Psychophysical Law (4) than to either of the other two possibilities.

The correlations among CE's ranged from −0.21 to 0.51 with an average of approximately 0.00. The coefficients between standard deviations and CE's became progressively smaller from 1 pound to 40 pounds. The value for 1 pound was 0.88 while that for 40 pounds was −0.09.

Wheel-type Control

The standard deviation values, the difference limens, and the constant errors are summarized in table 12.3 for performance at each level of pressure for the two directions of control operation.

TABLE 12.3.—*Standard deviations (SD), difference limens (SD/S), and constant errors (CE) in pounds for 20 pilots operating a wheel-type control laterally*

Standard pressure in pounds	Right			Left			Combined		
	SD	SD/S	CE	SD	SD/S	CE	SD	SD/S	CE
1	0.22	0.22	0.13	0.24	0.24	0.15	0.23	0.23	0.14
5	.44	.09	.24	.45	.09	.22	.44	.09	.23
10	.66	.07	.26	.69	.07	.32	.67	.07	.29
20	1.22	.06	.25	1.17	.06	.30	1.20	.06	.28
30	1.86	.06	.28	1.31	.05	−.02	1.69	.06	.13
40	2.08	.05	.19	2.00	.05	.01	2.04	.05	.10

It is apparent from table 12.3 that the standard deviations increase in a roughly linear fashion. With regard to the DL's it may be seen that these decrease in a nonlinear manner to 10 pounds and are roughly constant beyond this point through 40 pounds. The DL for 1 pound was found to be significantly greater at the 1-percent level than that for 40 pounds. None of the other DL's differ significantly at this level from any other. The DL's for the two directions of action of the wheel are not significantly different from one another at the 10-percent level.

The constant errors are about the same for the two planes of operation through 20 pounds. Above that value they are smaller for action to the left. The numbers, however, are not significantly different at any statistically acceptable level. For the two planes of action combined the CE's are all positive and show a maximum at 10 pounds.

Individual differences were not large, with a vast majority of the DL curves for individual subjects showing the same downward trend from 1 to 10 pounds followed by a levelling off beyond that value.

Rudder-pedal-type Control

The basic data derived from this experimentation are summarized in table 12.4. The standard deviation values increase with an increase in the standard pressure in an approximately linear fashion. It can be seen that Weber's Law holds approximately beyond 10 pounds. Below this value the DL's increase. For the data of both feet combined, the DL for 5 pounds is significantly higher at the 5-percent

TABLE 12.4.—*Standard deviations (SD), difference limens (SD/S), and constant errors (CE) in pounds for 20 pilots operating rudder-pedal-type controls*

Standard pressure in pounds	Right			Left			Right and left		
	SD	SD/S	CE	SD	SD/S	CE	SD	SD/S	CE
5	0.47	0.09	0.31	0.50	0.10	0.25	0.49	0.10	0.28
10	.68	.07	.19	.76	.08	.27	.72	.07	.23
20	1.04	.05	.26	1.12	.06	.42	1.08	.05	.34
40	2.12	.05	−.05	1.97	.05	.05	2.05	.05	.00
60	2.83	.05	−.36	2.51	.04	.10	2.67	.04	−.13

level than the DL's for 40 and 60 pounds. Further support for the lack of constancy of the Weber Fraction below 20 pounds may be found in the data for the 5 men tested with 1 pound. The DL was 0.28 for the right foot and 0.22 for the left foot with a combined value of 0.26. This finding supports the data at 5 pounds in suggesting that Weber's Law does not hold below 10 pounds in the present case.

Further inspection of table 12.4 reveals that the relative accuracy of the right foot was somewhat better at low pressures than that of the left foot, but it should be noted that none of the differences on the group of 20 pilots is significant at a statistically acceptable level.

With regard to CE's it can be seen that over all those for low pressures tend to be positive and those for high pressures are negative or zero. The values for the two feet separately show negative CE's for high pressures for the right foot, but a slightly positive one when the subjects were working with the left foot. The other CE's are about the same for the two feet.

COMPARISON OF PERFORMANCE WITH THE THREE CONTROLS

The curves of DL's plotted against standard pressure for the performance of the three groups tested with the stick, wheel, and rudder pedals are shown in figure 12.4. It can be seen that the general shape

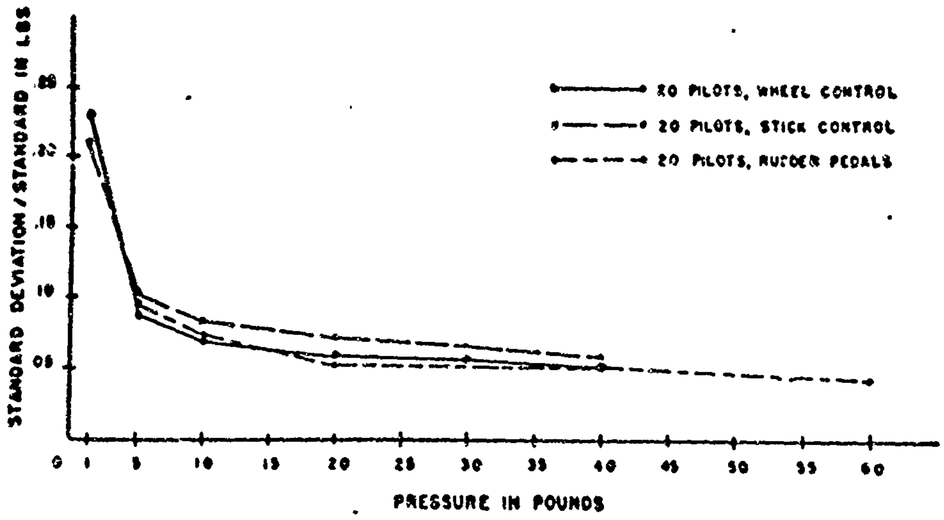


FIGURE 12.4.—Difference limens (standard deviation/standard) plotted as a function of the standard pressure for 3 groups of 20 pilots each tested with stick, wheel, and rudder pedal controls.

of the curves is the same in all three cases. Although the values decrease slightly beyond 10 pounds it appears that the Weber Fraction is roughly constant beyond this figure. Performance with the wheel as contrasted with that with the stick yielded lower DL's at five of the six points. It cannot be concluded, however, that performance with the wheel is more accurate than that with the stick since neither the differences between individual points nor the over-all difference are significant at the 5-percent level of confidence.

Performance with the rudder pedals was about the same as that with the wheel and somewhat better than that with the stick, although less accurate discrimination with the feet as compared with the hands might have been expected. None of the differences among the three sets of data was significant at the 5-percent level. The finding that performance with the rudder pedals is as good as that with the stick and wheel may be partially accounted for in terms of the fact that there was additional support in resting the heel on the floor when working with the pedals, thus permitting the application of pressure with the ankles rather than with the entire leg, the analogue of which was not available for the pilots reproducing pressures with their hands. Differences reported previously between the apparatuses employed (particularly the fineness of the calibrated scales) may be involved in this finding.

All three groups showed positive CE's for low pressures (see tables 12.2, 12.3, and 12.4). The CE's for high pressures were negative for the pilots working with the stick and rudder pedals and slightly positive for the performance of those men who operated the wheel. For the stick the CE's showed an approximately regular decreasing trend from 1 to 40 pounds; for the pedals and wheel, however, the CE's increased to a peak around the middle of the pressure range and decreased beyond this point. It should be noted that none of the differences among the three groups was significant at the 5-percent level of confidence for the number of cases employed.

SUMMARY AND CONCLUSIONS

This series of studies was concerned primarly with the accuracy with which blindfolded pilots were able to reproduce pressures on a stick-type control, a wheel-type control, and rudder-pedal-type controls. A nonpilot comparison group was tested with the stick as were two other groups to determine the course of learning and the effects of lack of knowledge of results. An approximately isometric indicating system was employed in which pressures applied to a semirigid control resulted in the transmission of slight movements through cables to a pivoted mirror. A light beam was reflected from the mirror to a screen calibrated in pounds. Five pressures (six with the wheel) in the range of 1 to 40 pounds were used with the stick, and five values from 5 to 60 pounds in the case of the rudder pedals. The technique was essentially the psychophysical "Method of Average Error."

The findings may be summarized as follows:

1. The accuracy of performance as measured by the standard deviation increased directly as a function of the magnitude of the standard pressure in the results of all three studies. The shape of the curve did not deviate significantly from linearity although too few points were involved for a final conclusion to be drawn.

2. Weber's Law did not hold throughout the range of pressures employed. The difference limens (standard deviation divided by the standard) decreased appreciably from 1 to 10 pounds for the data of the three investigations. Beyond the latter value the Weber Fraction was approximately constant at about 0.06.

3. Individual differences were relatively small with the records of most subjects exhibiting a negative growth function when the DL's were plotted against the standards. Differences in the magnitude of the DL's between individuals were not appreciable.

4. The constant errors for all three groups for low pressures were positive; for high pressures they were negative for the subjects operating the stick and rudder-pedal controls and slightly positive for those pilots who applied pressures to the wheel. Differences in CE's between groups were not significant.

5. A comparison of the performance of pilots and nonpilots working with the stick control indicated more accurate performance, i. e., lower standard deviations and correspondingly lower DL's at practically all points for the former group. The over-all difference was highly significant.

6. There were no appreciable or statistically significant differences between data derived from operation of the three controls in different directions. This finding held for both DL's and CE's.

7. There was no statistically significant evidence that performance with one control yielded more accurate performance than operation of any other control. The fact that accuracy of operation of the rudder pedals was equal to that with the stick and wheel appeared to be due, in part, to the support derived from resting the heel on the floor in this case and apparatus differences.

8. The course of learning indicated a tendency for the 11 pilots to start above low pressures and learn to come down towards the standard, and, conversely, to begin below high pressures and come up to them as learning proceeded. Small amounts of learning occurred in the middle range of pressures employed. Learning continued for 8 to 10 trials before a plateau was attained. The shape of the Weber function was similar to that obtained from the subjects tested under the standard procedure, but the magnitude of the values was greater for the learning group.

9. Twenty-eight pilots were tested, after the completion of the regular experiment, for three trials at each standard pressure without knowledge of results. The shape of the curve of DL's plotted against standards was approximately the same as for the other groups, but the values were appreciably increased, particularly at the low pressures.

10. It was concluded that amount of flying experience, body weight, and the order in which the pressures and directions were presented were not contributing appreciably to the results.

11. The intercorrelations among standard deviations and constant errors suggested that the data approximated Guilford's Generalized Psychophysical Law more closely than Weber's Law or the Fullerton-Cattell Square-Root Law.

On the basis of the findings of these studies recommendations were made to higher authority that, in control systems where pressure or force is the major cue employed by the operator in working with stick, wheel, or pedal controls operated by the arms or legs, and where relative consistency of performance is an important consideration, it appears desirable to minimize the frequency with which operators must apply pressures in the range from 0 to 5 pounds. The use of a wide range of pressures, up to a limit of 30 or 40 pounds, was also recommended in order to maximize the number of just noticeable differences for the operator. Pressures much higher than this limit were recommended as undesirable because of the likelihood of fatigue effects with frequent applications of high forces.

BIBLIOGRAPHY

1. Army Air Forces. Stability and control characteristics of airplanes. AAF Specification No. R-1815-A, April 1945.
2. Ezekiel, M. *Methods of correlation analysis.* New York: John Wiley, 1941 (2nd ed.).
3. Gough, M. N., and Beard, A. P. Limitations of the pilot in applying forces to airplane controls. *Technical Note No. 550,* National Advisory Committee for Aeronautics, Washington, January 1936.
4. Guilford, J. P. A generalized psychophysical law. *Psychol. Rev.,* 1932, 39, 73–85.
5. Guilford, J. P. *Psychometric methods.* New York: McGraw-Hill, 1936.
6. Hertel, H. Determination of the maximum control forces and attainable quickness in the operation of airplane controls. *Technical Memorandum No. 583,* National Advisory Committee for Aeronautics, Washington, September 1930.
7. Hugh-Jones, P. *Some physiological aspects of tank driving controls.* Part I. The maximum push exertable on a foot pedal and how this varies with the position of the pedal relative to the seat. Military Personnel Research Committee, Medical Research Council, Great Britain, B. P. C. 44/341, May 1944.
8. Hick, W. E. The precision of incremental muscular forces with special reference to manual control design. Flying Personnel Research Committee, *Report No. 642,* Great Britain, August 1945.

9. Hovland, C. I. A note on Guilford's generalized psychophysical law. *Psychol. Rev.*, 1938, 45, 430–434.
10. Patt, D. I., and Randall, F. E. *Principles of cockpit seating.* Memorandum Report No. 695-58C, Aero Medical Laboratory, Wright Field, Dayton, Ohio, February 1948.
11. Statistical Research Group, Princeton University. *The statistical sign test.* Memorandum No. 48.1 M, Applied Mathematics Panel, National Defense Research Committee, February 1945.
12. Woodworth, R. S. *Experimental psychology.* New York: Henry Holt, 1938.

CHAPTER THIRTEEN

The Coding of Airplane Control Knobs

JOSEPH WEITZ [1]

INTRODUCTION

How does the development of one set of habits affect the learning of a new set of habits similar to but not identical with the first? This is a problem confronting a pilot each time he changes from one type of airplane to another. In his original training, the pilot may learn that the throttle is to the right of the mixture control whereas in the next plane he finds the throttle in a quite different position. Does this affect the speed and accuracy of his performance? Actually this problem involving changes in position of controls is part of the larger problem of cockpit standardization. The present chapter, however, is concerned with the specific problem of habit interference due to a lack of standardization of controls.

In many cases, the change involved is not only one of position, but the shape and color of the control handle is also changed. That is, the throttle may have a round red knob in one airplane and in another have a square blue handle. It is possible that this too may lead to some interference with performance.

From the above considerations several questions may be asked which can be answered under experimentally controlled conditions. (1) Does changing the shape of the control handles produce more or less interference than changing the position? (2) If changing position leads to interference, does coding the shape and color of the handles differentially, according to function, have any effect on this interference? (3) If shape coding does lessen the interference caused by position changes, will it also be effective in the absence of visual cues?

It did not seem advisable to investigate these three problems initially in an airplane. Therefore, an experimental set-up was designed

[1] The research studies reported in this chapter were carried out while the author was assigned to the Department of Psychology, AAF School of Aviation Medicine, Randolph Field, Tex. The work was initiated as the result of discussions with Col. Arthur W. Melton, Chief of the Department, and Lt. Col. Paul M. Fitts in Headquarters, AAF.

from which conclusions could be drawn and applied to the cockpit situation.

EXPERIMENT 1

Procedure

The first problem (1) was investigated in the following manner. The apparatus used was a mock-up of a control column in which there were four controls. The handles were so constructed as to be easily removable and interchangeable (see fig. 13.1). These controls were used in operating the SAM Self-pacing Discrimination Reaction Time Test.[2] This test is one consisting of a panel which permits the presentation of pairs of lights, one red and one green. The red light can be above the green, to the right of the green, to the left of the green, or below the green. If the correct control is pulled backward a new setting of lights appears; if the incorrect control is pulled the lights are turned off for 2 seconds, and no response can be made during that period. The efficiency of the subject's performance in this study was measured by the number of correct responses made in four 1-minute periods. There was a 15 second rest period between each of the four trials.

The experiment was designed to determine the interference resulting (1) when the positions of the controls were changed and (2) when both the shape and the position of the controls were changed. A comparison situation was also required; this was one in which neither the shape nor the position of the controls was changed. The three ways in which the controls were varied were called conditions A, B, and C. These are shown in table 13.1.

TABLE 13.1.—*Arrangement of controls for the three conditions in experiment 1*

Condition	Position of control	Shape of control	Signal for operation
A	1	Barrel	Red right of green.
	2	Truncated cylinder	Red below green.
	3	Sphere	Red left of green.
	4	Cube	Red above green.
B	1	Truncated cylinder	Red below green.
	2	Barrel	Red right of green.
	3	Cube	Red above green.
	4	Sphere	Red left of green.
C	1	Cylinder	Red right of green.
	2	Cylinder	Red below green.
	3	Cylinder	Red left of green.
	4	Cylinder	Red above green.

It can be seen that conditions A and C had the same correct positions but differed in the shape of controls. Condition A differed from B in position of the correct control but was the same with respect to the shape of the appropriate handle. Condition C differed from condition B in both shape and position of the correct controls.

[2] This test was designed by Judson S. Brown.

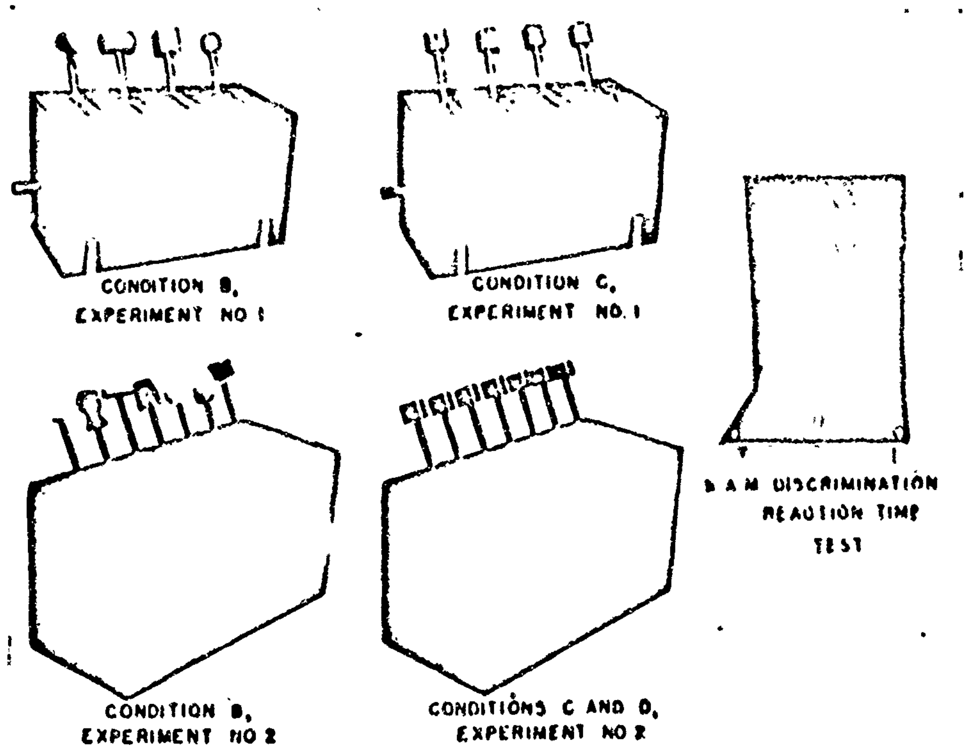


FIGURE 13.1.—Control panels and self-pacing discrimination reaction time apparatus used in experiments 1, 2, and 3.

The procedure was as follows. One group of subjects (group I) was given four trials on condition A, was allowed to rest 1.5 minutes, and then was given four trials on condition B. Here the interfering factor was the position of the appropriate control; the shapes of the handles were the same in both conditions. Group II had four trials on condition C, then after 1.5 minutes of rest was given four trials on condition B. For this group not only was the position of the correct control varied, but the shape of the appropriate handle was also changed. Group III had four trials on condition B and then, after a rest of 1.5 minutes, four more trials on condition B. This group served as the comparison group and its performance would show the advantage, if any, of maintaining the same shape and position of controls. It will be noted that condition B was administered to all three groups as their second test run. If the initial four trials had no differential effect on the subsequent four trials then it would be expected that the average score on the last four trials would be equal for all groups.

Results

The results are shown in table 13.2. In comparing the results of performance on the second four trials (B) for the three groups, it can be seen that some differences did occur. The difference between the last four trials for groups I and III gives a critical ratio of 1.50, indicating little loss in performance when shape coding was constant even though position was changed. Changing both shape and position (group II) compared with no change (group III) resulted in a statistically significant amount of habit interference, as shown by the critical ratio of 3.18.

TABLE 13.2.—*Mean number of correct responses for each condition, experiment 1*

	Group					
	I		II		III	
Change involved	Position		Shape and position		None	
Condition	A to B		C to B		B to B	
M	93.7	125.9	90.9	115.7	93.4	131.5
SD	26.5	27.3	26.7	28.5	26.1	23.8
N	45		45		45	

The same trends are obvious with respect to error scores. This can be seen in table 13.3. Here again the largest difference occurred when the number of errors on the second run of B for the group where no change occured was compared with the group having both shape and position varied. The critical ratio between these two sets of performances was statisfically significant at better than the 5-percent level of confidence (CR=2.24). It should be remembered in considering the error data that in an airplane one error may be of great practical significance.

TABLE 13.3.—*Mean number of errors for each condition in experiment 1*

	Group					
	I		II		III	
Change involved	Position		Shape and position		None	
Condition	A to B		C to B		B to B	
M	23.5	15.3	22.4	16.9	21.0	12.2
SD	13.9	14.1	12.4	11.1	12.0	9.4
N	45		45		45	

It would seem then that changing both shape and position of the controls concomitantly had the greatest detrimental effect on speed and accuracy of performance. However, these data do not show conclusively which of these variables had the greater effect.

Two other groups of subjects were run to determine the effect of varying the shape of the controls while holding position constant. Group IV was tested first on condition A and then on C; Group V was tested first on condition C and then on A. Here it will be noted that the correct position of the controls was identical in both cases, but the shape of the appropriate handle was changed. The results of these experiments are shown in table 13.4. It can be seen that there was little loss in performance on either of the second sets of trials when compared to the second run of condition B in which both shape and position were held constant. It may be assumed, then, that if the relative position of the controls is maintained constant there will be little if any habit interference regardless of changes in shape of the control handle.

TABLE 13.4.—*Mean number of correct responses for each condition, experiment 1*

	Group					
	IV		V		III	
Change involved	Shape		Shape		None	
Condition	A to C		C to A		B to B	
M	94.8	133.9	95.6	131.6	93.4	134.5
S D	28.2	36.2	26.6	31.1	26.1	28.8
N	20		20		45	

EXPERIMENT 2

Procedure

In many cases, because of engineering difficulties, it is not feasible to maintain constant the position of the controls from one plane to another. The next study (2) was designed to determine whether or not shape and color coding would reduce the interference caused by changing position.

The apparatus employed was essentially the same as that used in the first study. However, in the second study the mock-up control column had seven different controls (see fig. 13.1), of which only four were used. These controls were used again in operating the SAM Self-pacing Discrimination Reaction Time Test. Scoring was accomplished in the same manner as in the first study with the only difference being that eight trials were used in each sequence instead of four.

The following four basic test situations were used:

Condition A.—Eight trials on handles coded with respect to appropriate response.

Condition B.—Eight trials on handles coded the same as condition A but different with respect to correct position on the control column.

Condition C.—Eight trials on noncoded handles (all the same shape and color) with positions the same as those in condition A.

Condition D.—Eight trials on noncoded handles (all the same shape and color) with the correct positions the same as those in condition B. These four conditions are shown in table 13.5. It will be noted that condition A differed from condition B with respect to correct position, but was the same with regard to shape and color of the appropriate control handle. Condition C differed from condition D with respect to correct position in the same way that A differed from B, but in the case of conditions C and D there was no differentiation of the control handles in terms of shape or color.

Four experimental groups were tested as follows in order to determine the habit interference resulting from the combination of these conditions:

Group I.—Condition A followed by condition B. This involved habit interference resulting from the changed position of the controls when the handles were coded with respect to shape and color.

Group II.—Condition B followed by eight more trials of condition B. The second trial served as a comparison for the performance of group I on condition B.

Group III.—Condition C followed by condition D. This involved habit interference resulting from a change in position of controls when the handles were not coded with respect to shape and color.

Group IV.—Condition D followed by eight more trials of condition D. The second eight trials of condition D served as a control for the performance of group III on condition D.

For all groups there was a rest interval of 1.5 minutes between the two sets of eight trials.

Table 13.5.—*Arrangement of controls for the four conditions of experiment 2*

	Position	1	2	3	4	5	6	7
Condition A	Shape	a	b	c	d	e	f	g.
	Color	Blue	Gray	Black	Yellow	Green	Red	White.
	Correct when	Never correct	Red left of green	Never correct	Red below green	Never correct	Red right of green.	Red above green.
Condition B	Shape	d	a	f	c	g	b	e.
	Color	Yellow	Blue	Red	Black	White	Gray	Green.
	Correct when	Red below green	Never correct	Red right of green	Never correct	Red above green	Red left of green	Never correct.
Condition C	Shape	Cylinder	Cylinder	Cylinder	Cylinder	Cylinder	Cylinder	Cylinder.
	Color	Black	Black	Black	Black	Black	Black	Black.
	Correct when	Never correct	Red left of green	Never correct	Red below green	Never correct	Red right of green.	Red above green.
Condition D	Shape	Cylinder	Cylinder	Cylinder	Cylinder	Cylinder	Cylinder	Cylinder.
	Color	Black	Black	Black	Black	Black	Black	Black.
	Correct when	Red below green	Never correct	Red right of green.	Never correct	Red above green	Red left of green	Never correct.

Results

The effect of coding on performance when the position of the correct controls is changed can be determined by comparing the performance of group I on the second eight trials with that of group II on the second eight trials (condition B), and by comparing the second eight trials for group III with the second eight trials of group IV (condition D). Table 13.6 shows the results for the four groups.

TABLE 13.6.—*Mean number of correct responses for each condition, experiment 2*

	Group							
	I (coded)		II (coded)		III (noncoded)		IV (noncoded)	
Change involved	Position		None		Position		None	
Condition	A to B		B to B		C to D		D to D	
M	255.7	314.2	238.7	337.8	230.6	230.2	190.9	209.9
SD	66.9	48.9	53.7	46.5	58.1	58.3	68.7	63.4
N	25		25		26		25	

Since the changes in position of controls for groups I (A to B) and III (C to D) are identical it follows that if the shape and color coding used with group I have the effect of reducing habit interference, then there should be less loss in efficiency of performance when comparing the last eight trials for group I with those for group II than there is when comparing the last eight trials for group III with those for group IV. The data in table 13.6 verify this expectation. The difference between groups I and II for the last eight trials is not statistically significant (critical ratio=1.74), whereas, the difference between the mean scores on the last eight trials for groups III and IV is highly significant (critical ratio=4.05). There is little loss in performance when position is changed if the control handles are coded with respect to color and shape. There is a much larger and a statistically significant habit interference when the positions are changed and there is no color or shape coding. The difference is even more striking when the mean score on the last eight trials for group I is compared directly with the mean score on the last eight trials for group III. Here the changes in position are identical, the only difference is that in group I coding is in effect, whereas there is no coding of the handles for group III. The difference between these two sets of means is highly significant (critical ratio=5.56), showing a definite superiority in performance with the coded handles.

The results in terms of correct responses are further borne out when the error scores for the various conditions are compared. In table 13.7 are shown the means and standard deviations of the total error scores for each condition. In terms of the number of erroneous reactions made there was no loss in efficiency when the positions of the controls

TABLE 13.7.—*Mean number of errors for each condition in experiment 2*

	Group							
	I		II		III		IV	
Change involved	Position		None		Position		None	
Condition	A to B		B to B		C to D		D to D	
M	27.8	10.2	31.4	14.6	37.2	33.3	39.3	16.2
SD	25.1	6.0	17.2	11.3	23.7	24.0	25.1	15.8
N	26		25		25		23	

were changed if the handles were coded. The difference between the average error score on the last eight trials for groups I and II is not statistically significant (critical ratio=1.52) and the direction of the difference is the opposite from that expected. In the case of noncoded handles, there is a statistically significant difference (critical ratio=2.99) between the average error score for the last eight trials for groups III and IV. Here, the difference is in the expected direction, namely a greater number of errors when the position of the handles was

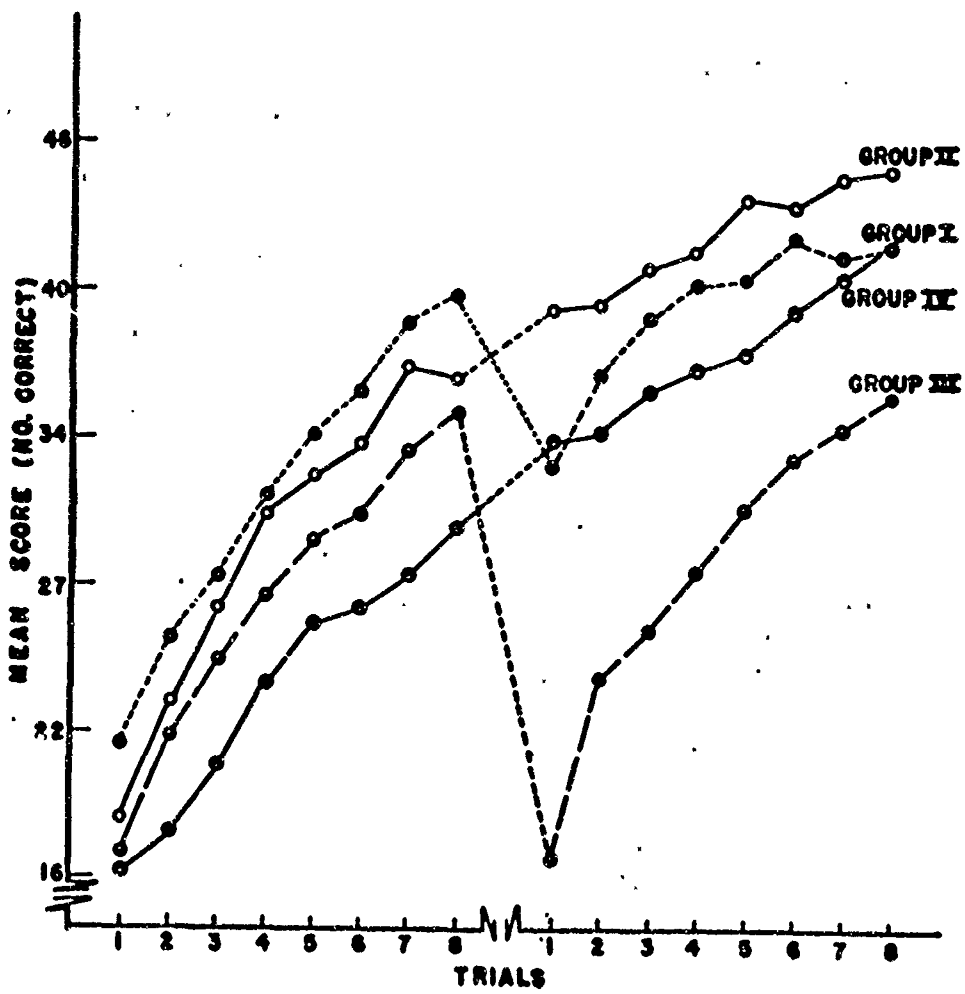


FIGURE 13.2.—Average number of correct responses per trial in experiment 2.

changed. Finally, a direct comparison of the mean error score in the last eight trials for group I and group III reveals a highly significant difference in performance in favor of coded handles (critical ratio= 4.67).

A graphic presentation of the interference caused by changing the positions of the correct controls is shown in figures 13.2 and 13.3. It is apparent from figure 13.2 that learning took place in all groups during the first eight trials of the test. That is, the average frequency

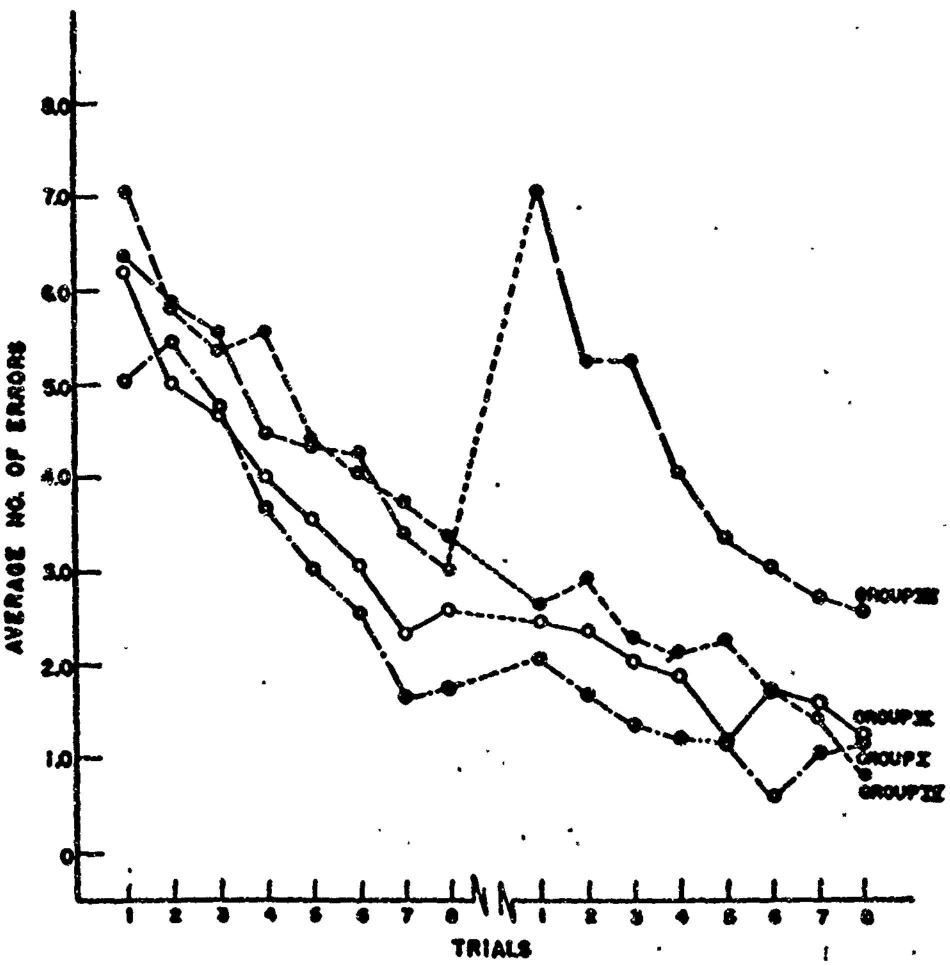


FIGURE 13.3.—Average number of errors per trial in experiment 2.

of correct responses increased from trial to trial during the first eight trials. The second group of eight trials is the critical set to which attention should be directed. It will be noted that during the second group of eight trials, points representing the average correct responses for group I consistently fell below the points representing the average correct responses for group II, thus indicating some slight amount of habit interference resulting from the change in the position of the correct controls even when the handles were coded for shape and color. Comparison of the curves for groups III and IV reveals, by contrast, the large amount of habit interference which resulted from changes in the positions of the correct controls when the handles were not coded. These curves also reveal that the maximum habit interference occurred during the first transition trials on the conditions which involved changes in the position of the correct controls, and it is further indicated that this initial habit interference was exaggerated when the handles were not coded. The difference between the scores on the first transition trial for groups I and III is highly significant (critical ratio=7.62).

These conclusions hold equally well when consideration is given to accuracy of performance as shown in figure 13.3. The number of errors decreased on the first eight trials for all groups. On the first transition trial group I showed a very slight increase in error, but it can be seen that the remainder of the curve shows the performance of this group to have been certainly as good as its comparison group (group II). On the contrary, when the handles were not coded (group III), there was a very large increase in the number of errors on the first transition trial, and the tendency to error in this group did not disappear during the following seven trials on that arrangement of control handles. That is, group III never made as few errors as its comparison group, group IV. The difference between the error scores on the first transition for groups I and III is highly significant (critical ratio=5.16).

It is evident that changing the position of a control in a series of controls having handles of similar shape had a detrimental effect upon the speed and accuracy of performance. However, this effect was minimized to a very marked degree when the control handles were coded with respect to color and shape and the individual being tested was permitted to use vision in the selection of the correct control handles.

EXPERIMENT 3

Procedure

The third question to be answered was, does shape coding lessen the interference due to position changes even though visual cues are restricted? This study (3) was designed in the same way as study 2 with one exception: in this case, the control box was placed behind a screen which concealed the controls but did not restrict the movement of the hand in manipulating the controls. The same four conditions were used and the same four groups (different subjects, of course) were used.

Results

Table 13.8 shows the results for the four groups with respect to the total number of correct responses made. Since the changes in position of the controls for groups I and III are identical it follows that if coding has a beneficial effect there should be a smaller loss of proficiency on condition B for group I when comparing it to condition B for group II than there is upon comparing condition D for group III and group IV.

TABLE 13.8.—*Mean number of correct responses for each condition, experiment 3*

	Group							
	I (coded)		II (coded)		III (noncoded)		IV (noncoded)	
Change involved	Position		None		Position		None	
Condition	A to B		B to B		C to D		D to D	
M	191.24	230.58	166.00	278.02	163.68	177.76	130.18	219.48
SD	65.78	66.21	58.13	70.55	60.69	56.01	40.53	62.52
N	50		50		50		50	

It would appear from these comparisons that there was a considerable loss in performance due to changing position and that there would seem to be no advantage shown for the coded handles. It should be pointed out, however, that the groups using the coded handles showed a somewhat superior performance throughout. When the second eight trials for group I are compared with the second eight trials for group III, it can be seen that there was a marked superiority in favor of the group using coded handles. The initial correct positions and the change in positions were identical for these two groups and yet there were 99.9 chances in 100 that group I would show a superior performance on the transition trials (critical ratio=4.31). From the data, it is obvious that the coded handles led to better performance even when no transition occurred. This is observable when the second eight trials for groups II and IV are compared. Here, again, there are 99.9 chances in 100 that there is a true difference in favor of the group using coded handles (critical ration=4.39). The beneficial effect of shape coding even though visual cues are restricted is further evidenced by comparing the error scores made by the different groups. Their error scores are shown in table 13.9. Here it can be seen that the difference between the number of errors for condition B for group I as compared to the number of errors for condition D for group III

TABLE 13.9.—*Mean number of errors for each condition in experiment 3*

	Group							
	I (coded)		II (coded)		III (noncoded)		IV (noncoded)	
Change involved	Position		None		Position		None	
Condition	A to B		B to B		C to D		D to D	
M	37.92	23.70	48.92	23.36	46.02	37.42	49.78	31.03
SD	21.56	19.53	25.25	21.96	19.72	18.45	23.38	24.94
N	50		60		60		60	

is highly reliable (critical ratio=3.61). Since the changes in position were identical for these two groups, the shape coding had the effect of significantly reducing the number of errors on the transition trials when comparing the performance on the coded and noncoded handles.

It is evident that changing the position of the appropriate controls led to marked habit interference when visual cues were eliminated. This interference effect was reduced in its initial amount after transition and was less persistent when the control handles were coded in terms of shape. This study indicates that the coding of control handles in terms of shape is a desirable procedure even though visual cues are eliminated, as in the case of flying in a darkened cockpit.

SUMMARY AND CONCLUSIONS

It may be concluded on the basis of these three studies that if changes are made in the positions of the controls after an individual has learned one set of positions, marked habit interference results. If the controls are coded with respect to color and shape, this interference is minimized even when visual cues are restricted. If both position and shape are changed, interference is maximized. Consequently, if position of the controls must be changed from one airplane to another, it would be of definite advantage to use shape and color coding of handles according to function.

REFERENCES

1. Weitz, J. Effect of the shape of handles and position of controls on speed and accuracy of performance. *AAF School of Aviation Medicine Project No. 266, Report No. 1,* June 1944.
2. Weitz, J. Effect of shape and color coding of airplane controls on speed and accuracy of performance. *AAF School of Aviation Medicine Project No. 336, Report No. 1,* December 1944.
3. Weitz, J. Effect of shape of handles and position of controls on speed and accuracy of performance when visual cues are restricted. *AAF School of Aviation Medicine Project No. 351, Report No. 1,* March 1945.

CHAPTER FOURTEEN

The Tactual Discrimination of Shapes for Coding Aircraft-Type Controls[1]

WILLIAM O. JENKINS

INTRODUCTION

The present studies were designed to determine the extent of confusion among several series of control knob shapes, including shapes now in use in aircraft and a number of experimental ones, in a situation where the subjects employed primarily tactual cues, and the use of vision, kinesthesis, size, and position was minimized or eliminated. On many occasions aircrew members, particularly pilots, must react rapidly and accurately to one of a closely bunched group of controls where positional cues are minimized and the operator is attending to certain instrument indications. Errors in this regard have produced many accidents, particularly in transition training where unfamiliar aircraft are being flown. One of the most common accidents results from confusion between flaps and landing gear controls which are in close proximity in some aircraft and are not coded with respect to any of the several possibilities.

Investigations by Weitz at the Army Air Forces School of Aviation Medicine (see ch. 13) have shown that accuracy of performance is significantly affected by position cues and by shape and color coding of aircraft-type control knobs. The present studies provide specific information, without particular samples of knob shapes, concerning which shapes yield the fewest recognition errors where size, position, vision, and mode of operation of the control are held constant or eliminated as cues.

---

[1] This chapter is based upon research findings reported in Headquarters AMC, Engineering Division, Memorandum Reports, No. TSEAA-694-4, TSEAA-694-4A, and TSEAA-694-4B.

APPARATUS AND PROCEDURE

In the first major study (study I), 25 plastic shapes, 1½ inches in the largest dimension, were each mounted on a rod which was bolted to the periphery of a turntable. The knobs and their mode of presentation are shown in figure 14.1. In the second major study (study II), 22 plastic shapes, 1¼ inches in the largest dimension, were mounted on the turntable with their shafts parallel to the ground as shown in figure 14.2. In both studies the procedure was followed of examining standard control knob shapes in the cockpit and on AAF equipment such as radar and radio sets and attempting to select shapes for study which maximized the characteristics typical of a group of related knob shapes.

In these studies the practical interest was in finding sets of knob shapes yielding a minimum number of errors for use in the cockpit and on radio and radar equipment. For this reason separate studies were made of different series of knob shapes mounted vertically and horizontally. It is quite possible that mode of mounting is not a critical factor; evidence reported below suggests it is not.

The procedure employed was as follows. Each subject was seated before the turntable and the instructions were read to him. A given knob was then presented to the blindfolded subject who felt it for 1 second. The experimenter then rotated the turntable to a predesignated point and the subject went from knob to knob, feeling each one in turn until he found and reported what he thought was the test shape. The same procedure was followed for each of the knobs, once while the subject used his bare hand and once while he wore a medium weight flying glove (A-11-A).

The conditions of bare hand and glove were counterbalanced so that half the subjects started their test under one condition and the other half under the other. In addition, the order in which the test and comparison knobs were presented was varied systematically in order to check intraserial effects.

In both investigations two types of errors were recorded. The first was the obvious case in which a subject incorrectly identified one shape as another. The second was called a hesitation error and was defined as the case in which a subject spent more than the allotted second in handling an incorrect shape, but did not identify that shape as the correct one. This type of pause might well be undesirable in the operation of controls in the cockpit.

A paired-comparisons follow-up study was conducted using the eight best shapes of the first investigation. This investigation is described in a later section.

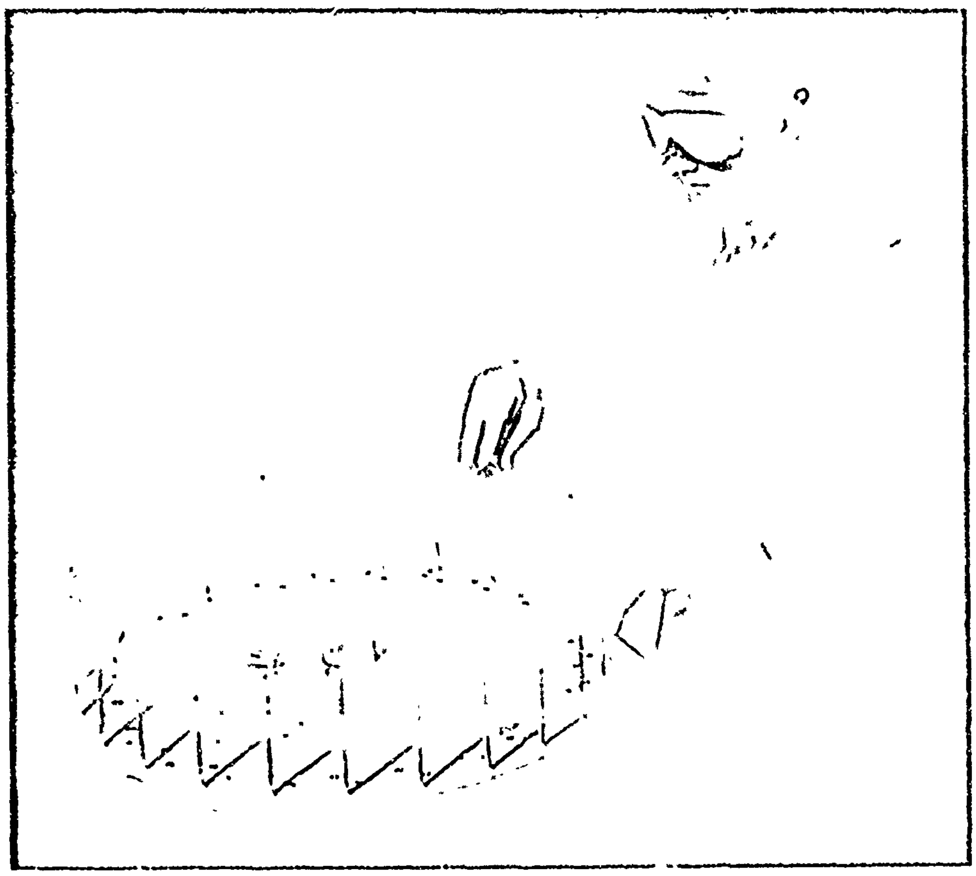


FIGURE 14.1.—Vertically mounted knobs employed in a study of shapes for use in coding aircraft-control knobs.

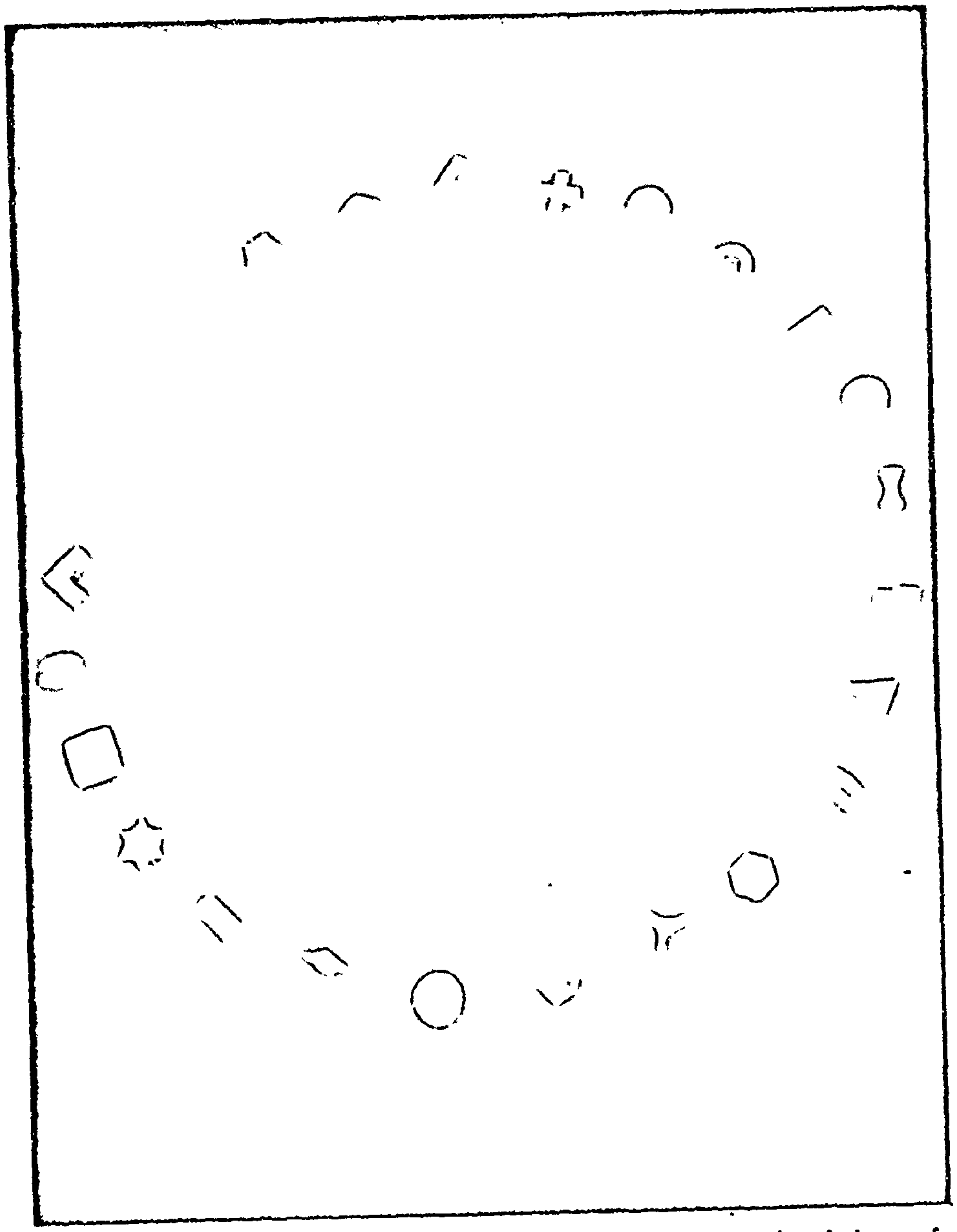


FIGURE 14.2.—Horizontally mounted knobs employed in a study of shapes for use in coding aircraft-control knobs.

The subjects in these investigations consisted of two separate samples of 40 Army Air Forces pilots who appeared to be representative of the pilot population.

RESULTS

It was found that the order in which the test and comparison shapes were presented did not affect the number and pattern of errors appreciably, so that the data for the several different conditions have been combined in the succeeding analyses. The findings have been divided into two sections: those concerned with distribution statistics, and those dealing with the pattern of errors.

Distribution Statistics

The distribution statistics for the two studies are summarized in table 14.1. It can be seen from this table that the percentage of error

TABLE 14.1.—*Distribution statistics concerning the errors made by 40 pilots in two series of knob shapes employed in investigations of shapes for coding aircraft-type controls*

Condition	Study I			Study II		
	Mean number of errors	SD	Percent error	Mean number of errors	SD	Percent error
Error, bare hand	2.9	2.2	12	3.7	2.6	17
Error, glove	4.8	2.7	19	4.9	2.8	22
Hesitation, bare hand	7.2	4.1	29	1.5	1.5	7
Hesitation, glove	9.7	6.3	40	2.1	2.1	10

ranges from 12 percent to 22 percent and the corresponding figures for hesitation-type errors vary from 7 percent to 40 percent. A satisfactory spread of scores was obtained in every case. In this connection it might be noted that practically every subject made some errors of both types.

The differences in table 14.1 between performance under conditions of bare hand and while wearing the flying glove are significant in most cases at the 5-percent level of confidence or better.

In order to obtain an estimate of the reliability of the method, rank order correlation coefficients were computed between the frequency of errors and hesitations under conditions of bare hand and while wearing the flying glove. The figure for errors was 0.70 while for hesitations it was 0.72. For errors and hesitations combined the value was 0.75. This degree of consistency between the performance under the different conditions indicates a satisfactory reliability for the testing technique for the present purposes.

One secondary analysis was performed in regard to the frequency of errors. This break-down consisted of comparing the proportion of errors for pilots in study I having less than 900 hours of flying experience with that for pilots having more than 1,100 flying hours.

While the more experienced pilots made a slightly higher proportion of errors under both conditions of bare hand and glove than did the group with fewer flying hours, the differences were well within the range of values expected on the basis of sampling fluctuations.

Pattern of Errors

It was found that the pattern of errors was comparable for errors and hesitation-type errors, for conditions of bare hand and glove, and for the different orders in which the test and comparison knobs were presented so that the data for all conditions have been combined in the following treatment.

	16	2	6	17	15	1	13	14	4	10	3	20	25	8	7	9	24	11	23	22	18	5	19	21	12
16												2			3				1	1			2		
2											1	14				1						1	1	1	
6											1						5								30
17													4			11		21				1			
15						1				18	2											11	2	23	1
1					1				2		2			23			1				1			1	1
13		1											7			9		4		1		1	1		
14	1	1	1						10								2		2		2				
4						1		3									8			2	17	3	9	4	
10		1		1	16	1	1									1				2		5	5	32	
3					1															25	2	1	1	1	6
20		29													2				6	4			1		1
25		2		1			2		1							1		11	1	1					
8	2		1			15					2										2	1		1	2
7	2	6	1									2		2		1		2				2			
9				5			14			2		1			1			16							
24			1		2				7		1										7				3
11		1		15	1		4		1		1		5	1	1	2							1		
23	3	1					2	2	2			6			1		1	2			1	1	11	1	
22		1	1		2						36	1										1	7	1	7
18				1	1	3			12		8			1		1	6			3			2	6	2
5					11	1			1	27					2			1		1	4		7	5	2
19		1	1		3		1	1	1	4	1		1				1	1	7	3	7	4		2	1
21			1	1	36				1	33								3		1	2	7	1		
12	1	1	25								12	1		7	1		1		2	3	1	1	1	1	

FIGURE 14.3.—Error pattern among 25 vertically mounted knob shapes.

The next step in the procedure was to rank each knob shape according to two criteria. The first was the total number of hesitations plus errors made by the subjects when a particular shape was presented. The second was the total number of knobs with which a given shape was confused regardless of the number of errors or hesitations in-

										1					1			1		1		
												1		4								
									1	5											1	
									1	3			1			3			2		6	
																	14		2	7		
											23	3		3								
													1						42	1		7
					1			1		4		1									1	
				1	1				5				1	1								
							11					1						2				
			2				2				1			10								
					10			1			1					4					1	1
			3							1		17										
					2			1								3	2	10	1	1	3	1
					5				1		1		1				6					8
		1				10		1	1										1	6		3
		2			1						2				13	1				4		3
					2			45									1				3	7
						11									2	1	9	3	2			2
	1			1	5			5	2						1			1	10			1
						2		5								3	2		13	1	1	

FIGURE 14.4.—Error pattern among 22 horizontally mounted knob shapes.

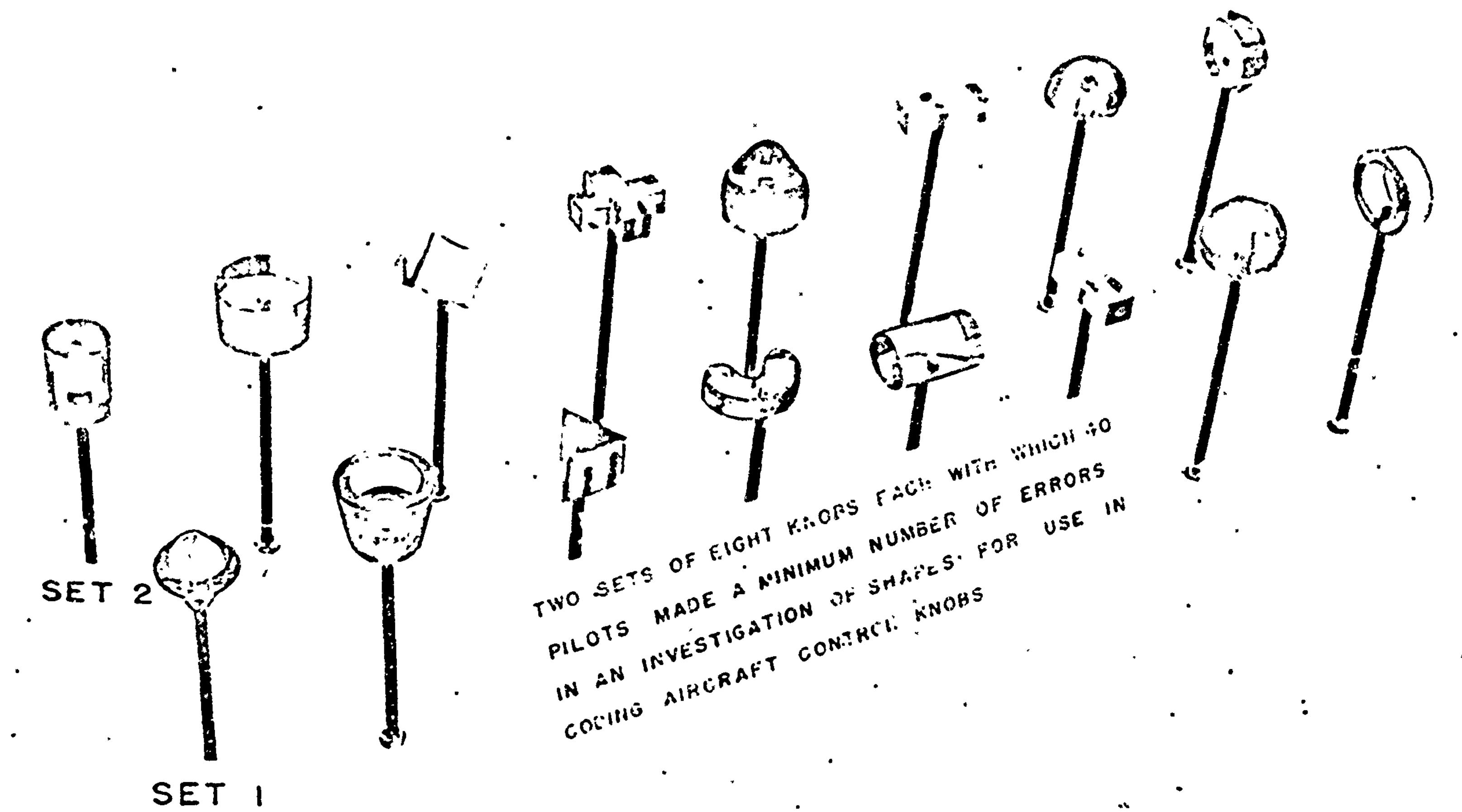


FIGURE 14.5.—Eight shapes yielding the fewest errors (set 1) and alternate shapes (set 2) in a group of 25 vertically mounted knobs.

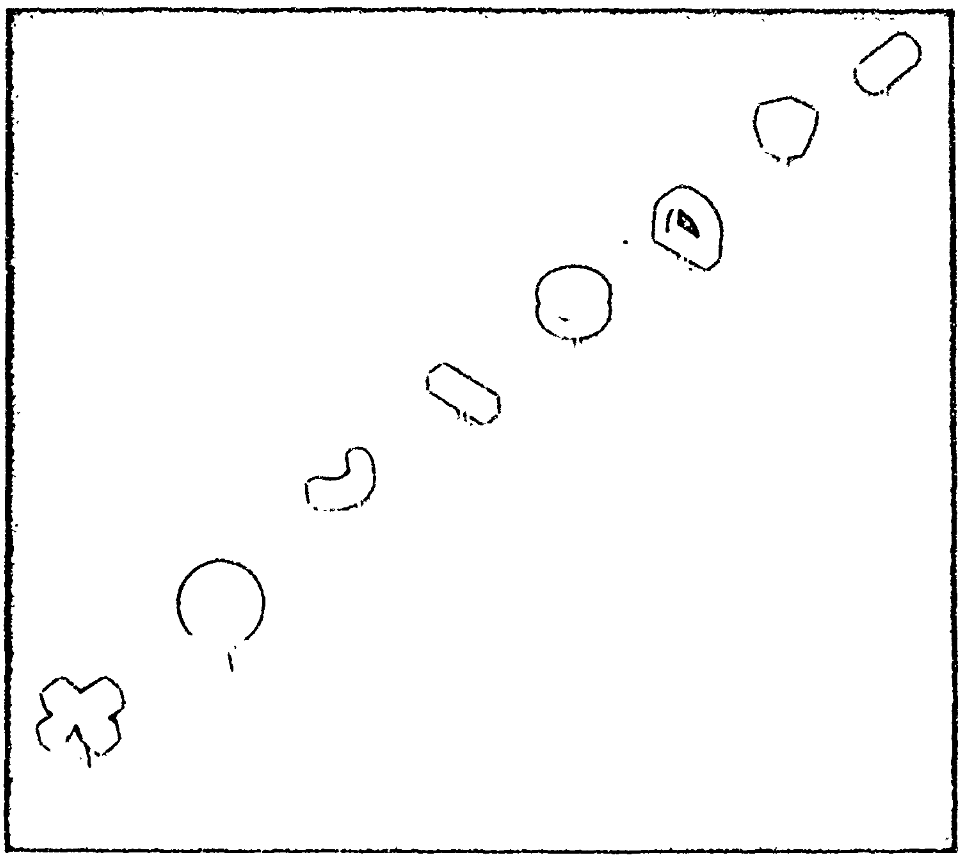


FIGURE 14.6.—Eight shapes yielding zero error in a group of 22 horizontally mounted knobs.

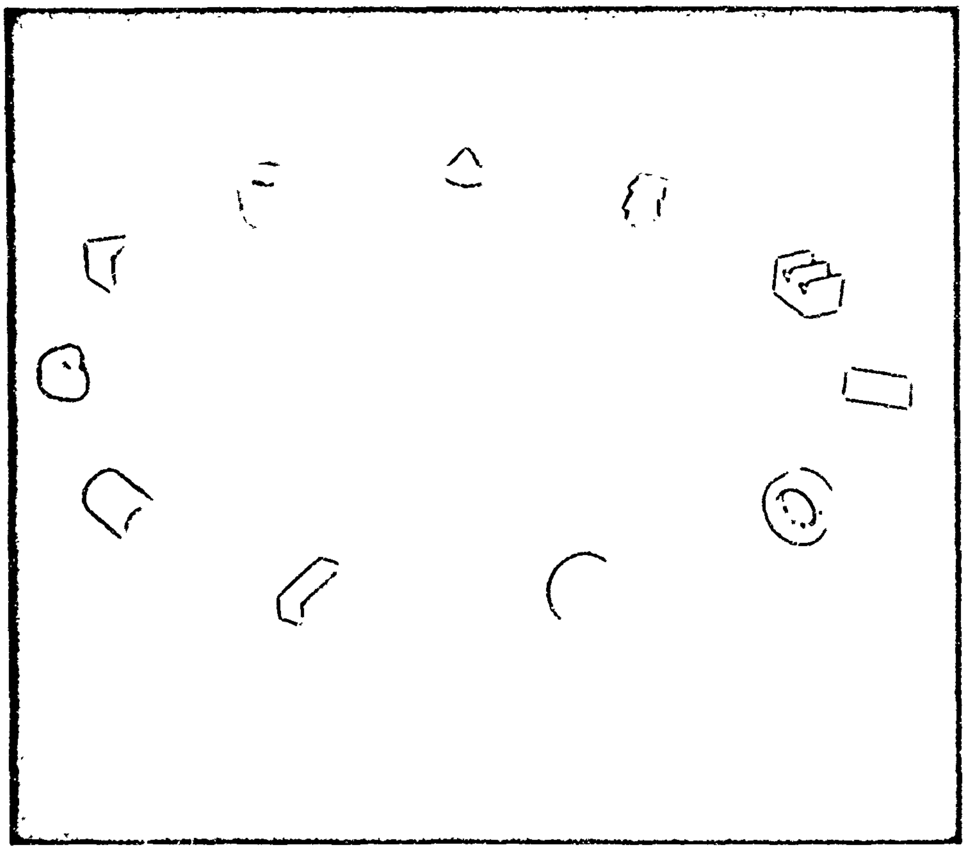


FIGURE 14.7.—Eleven vertically mounted knobs employed in a paired comparison follow-up study of shapes for use in coding aircraft-control knobs.

volved. It might be noted in passing that the rank-difference correlation between these two measures was of the order of 0.50. A two-way table was then constructed in which those knob shapes yielding the smallest number of errors and smallest number of knobs with which they were confused were placed side by side. Further refinements in grouping were accomplished by an empirical procedure designed to obtain the largest possible number of knobs with the fewest intra-group confusions. The resulting two-way tables for studies I and II are shown in figures 14.3 and 14.4. From this analysis two sets of eight knobs were derived in study I in which the number of errors were minimized and one set in study II. These groups are set off by heavy lines in figures 14.3 and 14.4, and the knobs are shown in figures 14.5 and 14.6. The best set of eight knobs in study I yielded a total of six errors or one-half of 1 percent of the total number of errors. In study II the best set of eight knobs yielded zero errors.

An examination of figures 14.3 and 14.4 reveals that in general knobs with similar shapes tend to be confused with one another but not with knobs of different shapes. There is a suggestion that the shapes tend to be grouped into families on the basis of shape, for example, shapes characterized by corners, edges, and flat surfaces (cubes, wedges, etc.), with the errors occurring within a family, but not across family lines.

Suggestive evidence that the two modes of mounting employed in these studies were not critical for the findings may be found in a comparison of the best eight shapes of each study in figures 14.5 and 14.6. It can be seen that four of the best knobs in each study are the same in shape and one is similar. One of the best shapes of study II appears in set 2 of study I. The remaining knobs were not used in both studies.

A follow-up study was made of the best eight knobs in study I along with three novel shapes employed in an "ideal" cockpit developed by the United States Navy Department. The set-up is shown in figure 14.7 with the three Navy knobs in the top right-hand corner. The method of paired comparisons was used with the usual precautions being taken of counterbalancing the order in which a given pair of knobs was presented, and varying the order of the series. Each of the 11 shapes was paired with every other shape including itself once, but no knob followed itself at any place in the series. A total of 30 AAF pilots was tested while wearing blacked-out goggles, once with the bare hand and once while wearing a medium weight flying glove (A-11-A).

In the 1,980 comparisons a total of 9 errors was made by the 30 pilots or about one-half of 1 percent. It is of interest to note that 8 of the 9 errors involved the 3 new Navy knobs which were added to the earlier set, but the N is too small for the difference to approach an acceptable level of significance.

DISCUSSION

On the basis of the present findings, recommendations have been made to the appropriate authorities that action be taken to decide which knob shapes are to be employed on which controls in the cockpit and on radio and radar equipment. The need for standardization with regard not only to shape, but also to size, position, color, and mode of operation is obvious and has also been recommended.

There is a need for further research in this area particularly on such problems as the optimal control-handle shapes for different modes of control operation, the use of mode of operation as a cue for differential reaction, and the type of color coding (e. g., brightness differentials) yielding most accurate and rapid performance under conditions of both day and night flying.

SUMMARY

These studies were undertaken as a basis for selecting sets of control knobs of different shapes for use in the cockpit and on radio and radar equipment which could be recognized immediately and accurately by touch in a situation where the use of vision, size, position, and mode of operation was held constant or eliminated as a cue.

Two sets of 25 vertically-mounted and 22 horizontally mounted knob shapes, including some knobs now in use in aircraft and a number of experimental ones, were presented to 2 separate groups of 40 blindfolded pilots who compared each shape with every other one in the series. One test was made with the bare hand and another while wearing a medium-weight flying glove. The findings may be summarized as follows:

1. A percentage of errors ranging from 12 to 40 percent was found. Practically every pilot made several errors.
2. A significantly greater number of errors was made while wearing the flying glove as contrasted with the condition of bare hand.
3. The correlation between frequency of error with the bare hand and that while wearing the flying glove was 0.75, indicating a satisfactory reliability for the technique.
4. Pilots with larger numbers of flying hours made slightly, but not significantly, more errors than those with fewer hours.
5. Two sets of eight knobs were found in the first study which yielded a minimum number of intra-set errors and there were no confusions among the best eight shapes of the second study.

6. A paired comparisons study of the best eight knobs of the first study along with three additional knobs yielded a very small number of errors.

Recommendations were made to the appropriate authorities that the shapes yielding the fewest errors in these studies be employed on aircraft equipment, and that the use of these knobs be standardized on equipment of a given kind.

CHAPTER FIFTEEN

A Study of Location Discrimination Ability[1]

PAUL M. FITTS

INTRODUCTION

How accurately can an individual, without looking where he is reaching, move his hand to a precisely designated location? The answer to this question is important for the design and location of machine controls that must be reached for and grasped without the assistance of direct vision. Many situations arise in flying where it is desirable for individuals to use controls in this manner, such as when the pilot is landing and does not want to take his eyes away from the ground long enough to look down into the cockpit in order to locate a control, or when the radar operator is watching the scope face and does not want to look away in order to adjust the gain control. Yet errors in control operation are so frequent that most aircrew members consider it inadvisable to operate a control without first making a visual check to determine its identity.

It has been suggested that the most efficient design and arrangement of controls for operation without the aid of vision would involve coding of control knobs with respect to shape and color, differentiating of controls with regard to their mode of operation, and optimal location of controls around the body. It is also proposed that these characteristics should then be standardized in all aircraft. In the two preceding chapters of this volume have been reported investigations of the problem of shape coding. The present chapter contains a report of investigations of the accuracy of aircrew personnel in reaching to different positions about them when in the seated position.

The purpose of the investigation was to determine the magnitude of errors made in positioning the hand to a number of locations of equal distance but varying directions from the body. The study also included an investigation of the direction of positioning errors and the

[1] This chapter is based upon research findings reported in Headquarters AMC, Engineering Division, Memorandum Report No. TSEAA-694-4D.

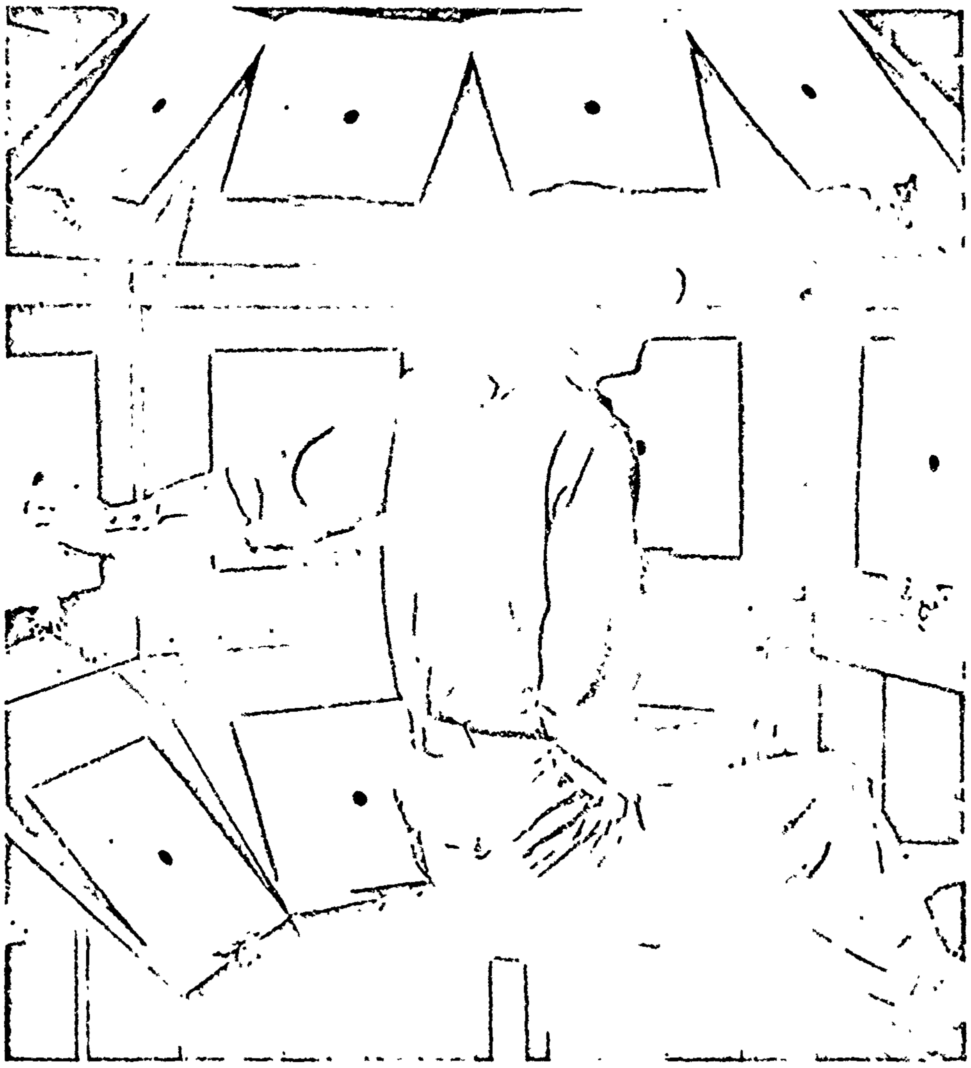


FIGURE 15.1.—View of left side of target stand used in Location Discrimination Study. For this photograph the right-hand side of the stand was removed to allow a view of the subject and the targets.

constancy of these error patterns. The study was planned in relation to a practical equipment design problem, but it is felt that several basic questions of human motor abilities are involved.

APPARATUS

The apparatus employed in the present study consisted of a wooden framework on which were mounted 20 paper targets arranged symmetrically around the midpoint of the shoulders of a seated subject. A photograph of the apparatus, with one side removed to give a view of the subject, is shown in figure 15.1.

Each target was placed at a distance of 30 inches from the reference point and was mounted perpendicular to a line connecting it to the reference point. Seven targets were located level with the reference point, one directly in front of the subject and three on either side. The side targets were 45°, 90°, and 135°, respectively, away from the one directly in front. These seven targets are referred to as "center" (C) targets. Seven targets were placed 45° higher than the level of the reference point and directly above each of the center targets. Six other targets were placed 45° lower than the reference point and directly below the six center targets on either side. The target space directly in front of and below the reference point was required for the subject's legs. The space directly behind the subject was not used, nor was the space directly above his head or directly beneath him.

A 30-inch distance from the mid-point of the shoulders to the targets was chosen because anthropological data on AAF personnel [2] show that 30 inches from the mid-point of the shoulders is the maximum distance at which controls can be placed if it is desired that more than 95 percent of the population be able to reach them easily.

The paper targets were 14 inches square. The target pattern consisted of a series of six concentric circles. The radius of the inner circle or bull's-eye was one-half inch and the radius of each succeeding circle increased by one additional inch. Each target was divided into quadrants by two diagonal lines which passed through the center at angles 45° from the vertical. Since it was the purpose of the investigation to study the ability of individuals to reach to various locations while looking elsewhere, and not while in complete darkness, a constant visual reference was provided. Two small red lights located forward of, level with, and equidistant from the subject's eyes, and 45° apart, were used for this purpose. A special pair of goggles was provided which permitted subjects to see only the red lights. Apertures in these goggles made it necessary for subjects to hold their heads in a forward-looking position in order to keep the reference lights in

[2] Randall, F. E., et al. *Human body size in military aircraft and personal equipment.* AAF Technical Report No. 5501, AAF Air Matériel Command, Wright Field, June 1946.

view. This insured a relatively constant head position during the test trials.

Each subject was provided with a marker, about the size of a toggle switch handle, which had a sharp point that permitted him to mark the tarkets by pressing against the paper.

An adjustable stool and a footrest were provided. A stool without a back was employed since the pilot or other aircrew member often leans forward in his seat while operating his equipment and in this case obtains no position cues from the back of his seat.

PROCEDURE

Test Administration

Subjects were seated in the apparatus and the height of the stool was adjusted until the mid-point of each individual's shoulders was at the reference point.[3] The purpose of the study and the experimental procedures were explained. Subjects were told that different targets would be designated in turn and that their task would be to reach with one smooth motion to the position where they thought the target was located and to mark the first point of contact. Left-hand targets were marked with the left hand; forward and right-hand targets with the right hand. For purposes of this report each vertical row of three targets is numbered from one to seven beginning with the extreme left-hand target. The middle target in each vertical row of three is called the "center" (C) target, the upper one is called the "up" (U) target, and the lower one the "down" (D) target.

Each subject was given a series of 60 practice trials, three trials to each of the 20 targets. In the practice sequence two successive trials were given to each target so that subjects could observe their errors on the first trials and attempt to make corrections on the second. The final 20 practice trials, one to each target, were given in a randomized order. Subjects were permitted to remove their goggles after each trial and look where they had marked. These practice trials were given immediately before the test proper.

During the actual test each subject was given 100 trials, followed by a 5-minute rest period and then another 100 trials. The sequence of trials was randomized, but no target ever followed itself. Each target received a total of 10 trials. Subjects were permitted to remove their goggles after each trial and secure immediate knowledge of results. A signal light which flashed every 7 seconds was used by the examiner in standardizing the rate at which targets were marked. The examiner sat directly behind the subject. If the subject missed the target entirely, the examiner recorded the miss and the direction

[3] George Korinek and Peter S. McKewen ran most of the subjects in this experiment and assisted in developing the apparatus and analyzing the data. Mrs. Kathryn D. Young and Shirley Connell also assisted in the study.

of the error. If the subject attempted to mark a wrong target, the examiner stopped him and the trial was repeated.

Ten of the subjects used in the major investigation were run for 200 additional trials on each of 2 successive days in order to obtain data on changes in accuracy scores and changes in error patterns with additional practice. The standard procedure was followed during this period and knowledge of results was given after each trial. The order in which targets were marked was in a different randomized order for each successive day.

The group of 50 subjects used in the major study was unselected as regards hand preferences. On a short handedness questionnaire, 1 man was markedly left-handed and 5 men showed a mixed preference.

Scoring Procedures

Targets were scored for over-all accuracy by the following procedure. A mark in the center or bull's-eye of the target was scored zero and the marks in subsequent circles were scored one, two, three, four, and five, respectively. Any mark which fell outside of the outermost circle was scored six. Since all marks that were farther than five inches from the center circle were arbitarily recorded as six, the scoring procedure did not give full weight to misses that were quite far from the center. The average accuracy score, nevertheless, is a close approximation of the average distance in inches of the marks from the exact center of each target. A small score indicates accurate performance.

Targets were also scored in terms of the direction of each mark from the center. Since targets were laid off into quadrants, a top (No. 1), a right (No. 2), a bottom (No. 3), and a left (No. 4), it was possible to record the total number of marks in each sector. This total for each quadrant was broken down into the number of marks within and outside of the outermost target ring. The latter figure gives the number of attempts in which the response was more than 5½ inches away from the exact center of the target.

RESULTS

Analysis of Accuracy Data

The average accuracy score for all 50 subjects on each of the 20 targets was 3.27 or an average error of somewhat more than 3 inches. A total of 15.6 percent of the 10,000 trials made by all subjects resulted in marks outside of the outermost circle. In other words, 15.6 percent of all trials were in error by more than 5½ inches. The accuracy scores, however, varied greatly from target to target. In order to study systematically the influence of target position upon accuracy of location discrimination, three comparisons were made: (1) a comparison of average accuracy in reaching for all targets on the right, and for all targets on the left; (2) a comparison of accuracy in reaching for targets directly forward, and for targets to either side of the forward position; and (3) a comparison of average accuracy in reaching for targets above, even with, and below the level of the shoulder. For the right-left comparison an average score for all right-side targets was compared with the average score for all left-side targets. For the front-back comparison, all targets which were at equal angular distances from the front including both right- and left-side targets were grouped. Similarly for the up-down comparison all targets at the same angular elevation where grouped (except that targets 4U and 4C were omitted, since there was no 4D target). In making each of these comparisons a new distribution of combined scores was made for each group of targets in question, and means, standard deviations, and intercorrelations for these distributions were computed.

In table 15.1 are summarized t[illegible] basic findings for each individual target. The mean accuracy scores, and the percent of marks that occurred in each of the four sectors of each target are given. In table 15.2 are given the mean accuracy scores, standard deviations, and correlations of scores for various combinations of targets. In figure 15.2 the same data are represented graphically. In this figure a diagram of the different targets has been made in which the diameter of each target is proportional to the mean accuracy score. The four small circles in the quadrants of each target have a relative diameter proportional to the total number of marks in each sector.

TABLE 15.1.—*Mean accuracy scores, standard deviations, proportion of marks in each target sector, and proportion of major errors in each target sector for the 20 different target areas*

[N=50]

Target	Mean accuracy score	SD	Proportion of total marks in each sector				Proportion of total marks that were major errors in each sector			
			1	2	3	4	1	2	3	4
1U	4.04	0.78	5.2	11.4	45.7	37.7	0.4	2.8	12.8	10.2
1C	3.72	.80	37.3	29.5	12.0	21.2	11.4	9.0	2.2	4.5
1D	4.08	.99	60.8	22.9	4.2	12.1	21.4	5.4	.2	2.1
2U	3.25	.79	21.2	17.2	27.8	33.8	2.7	1.9	5.6	3.9
2C	3.42	.79	33.5	22.5	17.8	26.2	5.3	4.2	2.8	5.5
2D	3.02	.60	22.5	24.6	25.1	27.8	1.4	2.2	1.8	3.6
3U	3.49	.99	7.4	11.0	59.1	22.5	0.0	.8	14.9	3.4
3C	3.13	1.09	22.9	12.5	25.9	38.7	2.2	1.2	3.9	9.1
3D	3.03	.97	30.9	11.4	9.2	42.5	4.2	.6	0.0	5.8
4U	2.46	.79	15.8	22.3	53.5	8.4	0.0	0.0	3.8	0.0
4C	2.14	.90	40.7	24.9	20.8	13.6	1.6	.4	.8	.2
5U	3.21	1.05	7.6	30.3	51.1	11.0	.4	6.6	6.6	.6
5C	3.07	.88	19.2	38.2	27.1	15.5	1.2	6.9	3.9	1.4
5D	2.76	.77	34.6	37.0	13.5	14.9	4.2	4.4	.6	1.2
6U	3.21	.77	14.6	32.9	35.7	16.8	1.0	5.4	6.0	1.0
6C	3.18	.71	21.5	29.1	21.7	27.7	2.2	4.6	2.6	3.6
6D	2.91	.70	13.3	21.5	39.8	25.4	.6	1.8	1.6	2.4
7U	4.18	.94	5.8	36.2	48.1	9.9	0.0	11.3	18.5	.8
7C	3.61	.80	27.6	23.6	16.8	32.0	4.0	3.6	1.4	8.8
7D	3.30	.94	49.2	20.8	6.4	23.6	9.7	2.8	0.0	3.1

Left-right comparison.—From an inspection of table 15.2 it will be noted that the average accuracy score on the nine right-hand targets

TABLE 15.2.—*Means, standard deviations, correlations, and t values for combined scores on selected groups of targets*

[N=50]

Target group	M	SD	r	Diff.	t
All left-hand targets	3.46	.47	0.81	0.19	4.4
All right-hand targets	3.27	.50			
All upper targets	3.56	.50	.58	.21	2.7
All center targets	3.35	.50			
All center targets	3.35	.50	.42	.17	2.0
All lower targets	3.18	.50			
Rows 1 and 7 combined	3.81	.50	.48	.66	9.6
Rows 2 and 6 combined	3.15	.44			
Rows 2 and 6 combined	3.15	.44	.61	.04	.5
Rows 3 and 5 combined	3.11	.70			

was better than the average score on the nine left-hand targets. The difference is significant at the 1-percent level of confidence.

The right-left comparisons are based on a total of 90 trials to each side for each subject. A correlation of 0.81 between scores for the two sides indicates a satisfactory level of reliability of the test procedure, since the reliability of the total test would be considerably higher.

Front-back comparison.—From an inspection of table 15.1 it will be noted that the two targets located directly in front of the subject (targets 4U and 4C) were marked with considerably greater accuracy than any other targets. The average score for these two targets was better, at the 1-percent level of confidence, than the score on any other individual target or group of targets. Referring to table 15.2 for a comparison of targets at different distances to either side, it will be noted that the average score on targets located 45° away from center (rows 3 and 4) did not differ appreciably from the average score on targets located 90° away from center (rows 2 and 5). The small difference that was found could easily have occurred by chance. Accuracy in marking the rows of targets located 135° to the rear (rows 1 and 7) was considerably poorer than for targets nearer the front. The difference is highly significant statistically and is of practical importance. It can be concluded from these findings that individuals can reach most accurately to positions directly in front of them, that they can reach with somewhat less accuracy to positions as much as 90° to the side, and with far less accuracy to positions more than 90° to the side.

Up-down comparison.—It will be noted in table 15.2 that targets located at 45° below the level of the shoulders were reached for with somewhat greater accuracy than horizontal targets, and that targets located 45° above the horizontal were located least accurately. The difference between up and center targets is significant at the 1-percent level of confidence and the difference between the center and down targets at the 3-percent level. Of the two targets directly in front of the individual the center one was located more accurately than the upper one.

Analysis of Error Patterns

Data on the distribution of marks in each quadrant of each target are given in table 15.1 and are shown graphically in figure 15.2. If marks had been distributed equally about the center of each target, 25 percent of the marks would have fallen in each quadrant. Applying the Chi Square Test to this hypothesis, it was found that in only 1 target (target 2D) out of the 20 could the total distribution of all marks be accounted for by chance variation from an unbiased pattern.

Several systematic characteristics of the error patterns can be noted. On all except three of the 11 targets marked with the right hand, including the two center targets, individuals tended to reach too far to the rear; that is, errors were made toward the back more frequently than toward the front. On seven of the nine targets marked with the left hand, errors also were made more frequently toward the rear. It also was found that on all seven upper targets the majority of marks were too low. In the case of four of the six targets located downward the most frequent error was to reach too high. However, in the case of the two targets located downward and directly on either side (targets 2D and 6D) the predominant tendency was to reach too low.

An inspection of the error pattern data given in table 15.1 and represented graphically in figure 15.2 shows a high degree of mirroring with regard to the direction of the error patterns for corresponding right- and left-hand targets. In eight cases out of nine, when the error pattern on a left-hand target was predominately toward the front or toward the back, the error pattern in the corresponding right-hand target was found to be predominately in the same direction. To a lesser extent corresponding up and down targets gave mirrored error patterns.

Analysis of Learning Data

In the learning study it was found that the 10 subjects tended to improve in over-all accuracy from day to day. The average accuracy score of the group on all 20 targets was 3.18 for the first day, 2.85 for the second day, and 2.73 for the third day. The difference between first and third day was highly significant. Improvement was shown in locating practically all targets, the differences in scores for individual targets between the first and third day being highly significant.

The error-pattern data for the three successive days were analyzed as follows. The two sectors of each target that had been marked most frequently by the 50 subjects in the first study were determined. The total number of marks which the learning group made in these two

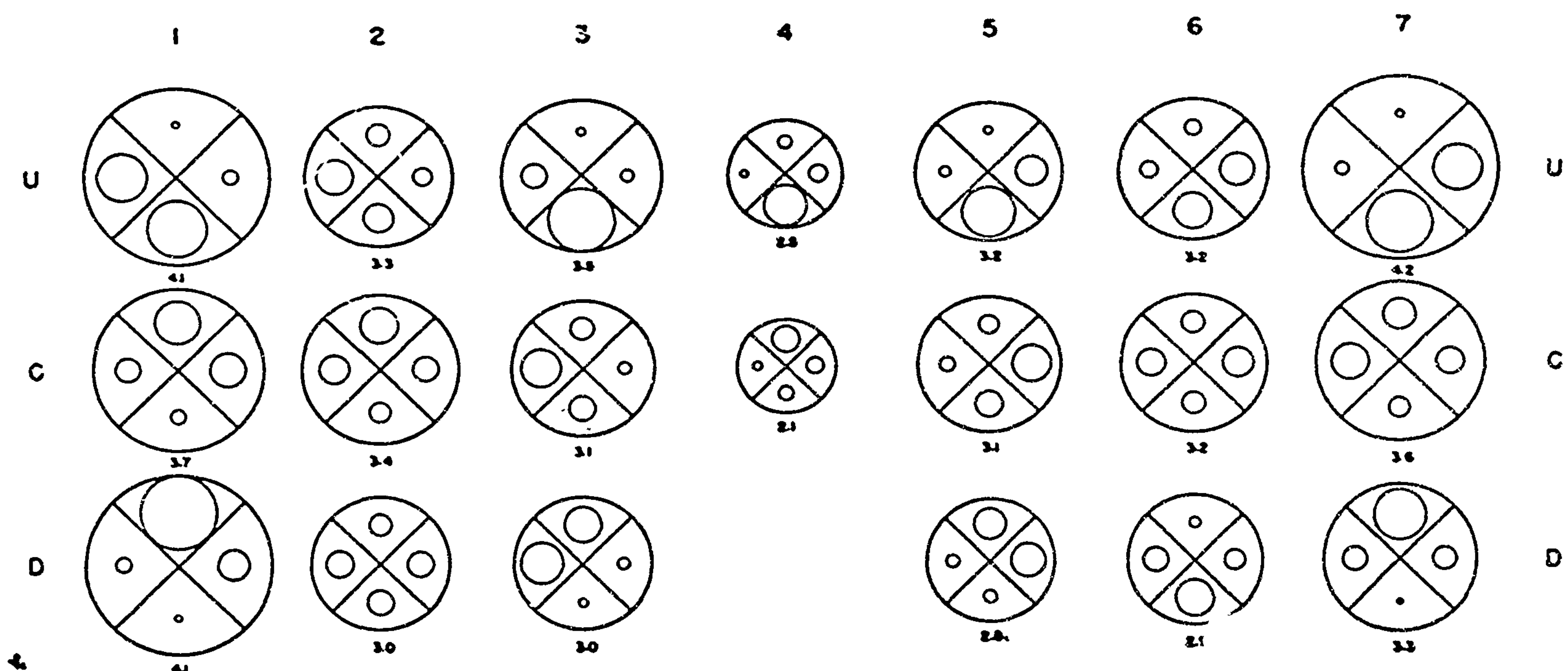


FIGURE 15.2.—Relative accuracy scores for different areas, and error patterns around each target. The size of each large circle is proportional to the size of the accuracy score for that area, a small circle representing better accuracy. The relative size of the four small circles within each large one is proportional to the number of marks made in each quadrant of each target.

sectors of each target was determined for each day, including the first day when they were a part of the larger group. These data were combined for all 20 targets. It was found that on the first day the learning group placed 66.6 percent of all marks in the two sectors of each target marked most frequently by the total group of which they were a part. On the second day the learning group placed 66.4 percent of all marks in these same quadrants and on the third day, 64.1 percent. Further analysis showed that for 8 of the 20 targets the original group placed more than 70 percent of all marks in the 2 most frequently marked sectors of each target. The percent of all marks made by the learning group on these same 8 targets was 75.9 percent for the first day, 76.6 for the second, and 73.3 for the third. It should be remembered in connection with these findings that subjects were given knowledge of results after each trial and had opportunity to correct their errors. It should be pointed out also that any errors of measurement (unreliability of testing procedures) on the first day would have resulted in a lower incidence of marks (closer to the expected 50 percent) on subsequent days for the two sectors marked most frequently on the first day. From the learning data it can be concluded, therefore, that error patterns were highly stable. The directional characteristics of the patterns did not change appreciably with additional practice accompanied by knowledge of results, even though the overall scatter of marks became less.

DISCUSSION OF RESULTS

The findings with regard to the areas to which typical individuals can reach most accurately could possibly be explained by reference to the muscle groups employed and the relative ease of reaching to the different locations. An alternative explanation could be made by reference to previous habits and more frequent experience in reaching forward and downward than in reaching backward and upward. The slight superiority of the right hand over the left hand is in agreement with findings from many other studies.

The data on accuracy of discrimination may have been influenced by the fact that the visual reference lights were located in a forward position. It would be worth while to simulate the condition existing when an individual has his head turned and is looking to one side, and to study location-discrimination ability for various areas while visual reference is other than directly forward. It would be of interest, also, to study the effect of no lights and of only one light.

Two different hypotheses can be suggested to account for the error-pattern data. One hypothesis is that the error patterns are primarily a function of the proprioceptive and tactual cues available and the muscle groups employed in reaching to different locations. It is possible, on the other hand, that direction of errors for any particular target is a function of the overall pattern or gestalt of targets employed. The finding that most individuals make errors toward the center of each vertical row of three targets suggests, for example, that there may be a tendency to make errors toward the center of a group of associated controls.

These questions cannot be settled by reference to the present data, and must await the results of further studies. It is proposed to study the effect on location discrimination ability of different points of origin for the reaching movement and to study the accuracy of location discrimination in three dimensions for a number of points closer to the body than the maximum extent of reach. The hypothesis is suggested, in connection with the latter study, that individuals will be able to position their hands more accurately in cases where it is possible to use a cue of extent of movement or to judge the position of the hand in relation to the effort involved in making movements of different amounts. It should be remembered that in the present investigation distance of movement was constant, since the surface of the target was perpendicular to the direction of movement of the hand, and therefore that extent of movement was not employed as a cue.

With regard to the application of the findings of this study to the location of control equipment, several conclusions are made. For greatest accuracy in locating a control without the aid of direct vision the control preferably should be located directly in front of the individual, and if this is not possible, then no farther to the rear than directly at the side. Controls preferably should be located below the level of the shoulders and if this is not possible then certainly not much above shoulder height. With regard to the distance by which controls should be separated, if it is desired to make it very unlikely that an operator will grasp a control by mistake, it appears that roughly a 6-inch separation is necessary for controls that are located at arms' reach and in one of the preferred areas, and approximately twice that separation for controls located to the rear of the subject or at the side above the level of his shoulders. Smaller distances could be used if operators are to be highly trained.

SUMMARY

1. A study was made of the accuracy of 50 pilot subjects in reaching to 20 targets mounted in different areas around them and at distances of 30 inches from the mid-point of the shoulders. Visual reference was provided, but vision of the target or the body was excluded.

2. When the data were analyzed in terms of average accuracy of location discrimination, it was found that accuracy was best for for-

ward areas. For areas on either side, accuracy was best for areas lower than the level of the shoulders and for areas near the front.

3. Consistent error patterns were found for each target area. Individuals tended to reach too low for targets located above the level of their shoulders, too far to the rear for targets located on either side, and too high for targets below shoulder level.

CHAPTER SIXTEEN

Principles of Control Arrangement for Sequential Operation[1]

NORMAN L. MURRAY

INTRODUCTION

In navigating and bombing operations radar operators use complicated control equipment involving levers and knobs which are manipulated in response to complex visual stimuli. Frequently, in an attempt to meet space and weight requirements, very unsystematic control arrangements are evolved. It is hypothesized that in the operation of complex equipment, arrangement of controls and indicators according to functional patterns would be best from the standpoint of learning and ease of operation. Especially should this be true where speed of operation is important, as in the operation of "Ferret" equipment where the operator is called upon to make a whole series of adjustments in a matter of seconds. The reactions in this type of task belong in the categories of what Brown and Jenkins in chapter 3 call adjustive and connective serial reactions.

It is important to determine how best to combine separate movements into a series or pattern to produce the greatest efficiency in terms of speed, accuracy, and ease of learning. The purpose of this experiment was to determine the effects of two variables on the basic problem. These were regular and irregular spacing of knobs and clockwise, counterclockwise, and alternate (clockwise-counterclockwise) rotation of the knobs. The first study involved connective reactions, and the second adjustive reactions.

APPARATUS AND PROCEDURE

The School of Aviation Medicine Finger Dexterity Test (CM116A) was employed as it seemed well suited to a preliminary investigation of

[1] This chapter is based upon research findings reported in Headquarters AMC, Engineering Division, Memorandum Report No. TSEAA-691-7A. The assistance of Judson S. Brown in the design of the experiment and that of William O. Jenkins in analyzing the data and preparing the report is gratefully acknowledged.

this problem. This test consists of a board in which 48 square holes are cut; round-headed pegs with square bases are inserted in the holes. The basic task is to lift the pegs from the holes, rotate them 180°, and reinsert them in the holes.

Eight problems or tasks were set up. The first 6 problems dealt with the direction of rotation of 24 pegs arranged in a linear pattern on the first and second rows of the board. The last 2 problems involved the manipulation of 24 pegs arranged regularly and irregularly over the entire board. In problem 1 the subjects were required to work from left to right across the board rotating the pegs through 180° in a clockwise direction (CW) and then back along the second row from right to left. Problem 2 was identical with the preceding one except that the rotation was counterclockwise (CCW). In problem 3 the subjects were required to work from left to right across the board and back from right to left along the second row. The odd pegs were rotated clockwise and the even pegs counterclockwise through 180° (A). In problem 1–a the subjects were required to work from right to left across the board rotating the pegs through 180° in a clockwise direction and then back along the second row from left to right. Problem 2–a was identical to the preceding one except that the rotation was counterclockwise. In problem 3–a the subjects were required to work from right to left across the board and back from left to right along the second row. The odd pegs were rotated clockwise and the even pegs counterclockwise through 180°. In problem 4 the subjects manipulated 24 pegs arranged regularly on the 4 rows of the board with every other peg removed. The subjects worked from left to right along the first row, from right to left along the second row, then left to right along the third row and back across the fourth row from right to left. Problem 5 was identical with the preceding one except that the pegs were arranged in irregular order. In the latter two problems there were six pegs per row and the distance traveled was the same in both cases. The three pegboard arrangements are illustrated in figure 16.1.

It will be noted that paired problems 1 and 1–a, 2 and 2–a, and 3 and 3–a, respectively, are essentially the same. They differ in that the subjects were required to begin at the upper left and work to the right in 1, 2, and 3, and to begin at the upper right and work to the left in 1–a, 2–a, and 3–a. Comparisons are made in the results section between the l-r and r-l conditions.

The subjects, Army Air Forces pilots, were divided into 2 groups of 24 each. The first group was tested with problems 1, 2, 3, 4, and 5. The second group was tested with problems 1–a, 2–a, 3–a, 4, and 5. The subjects were seated at a table and the peg board was placed in such a manner that the center of the board in its long dimension was alined with the midline of the body. The subjects were given instruc-

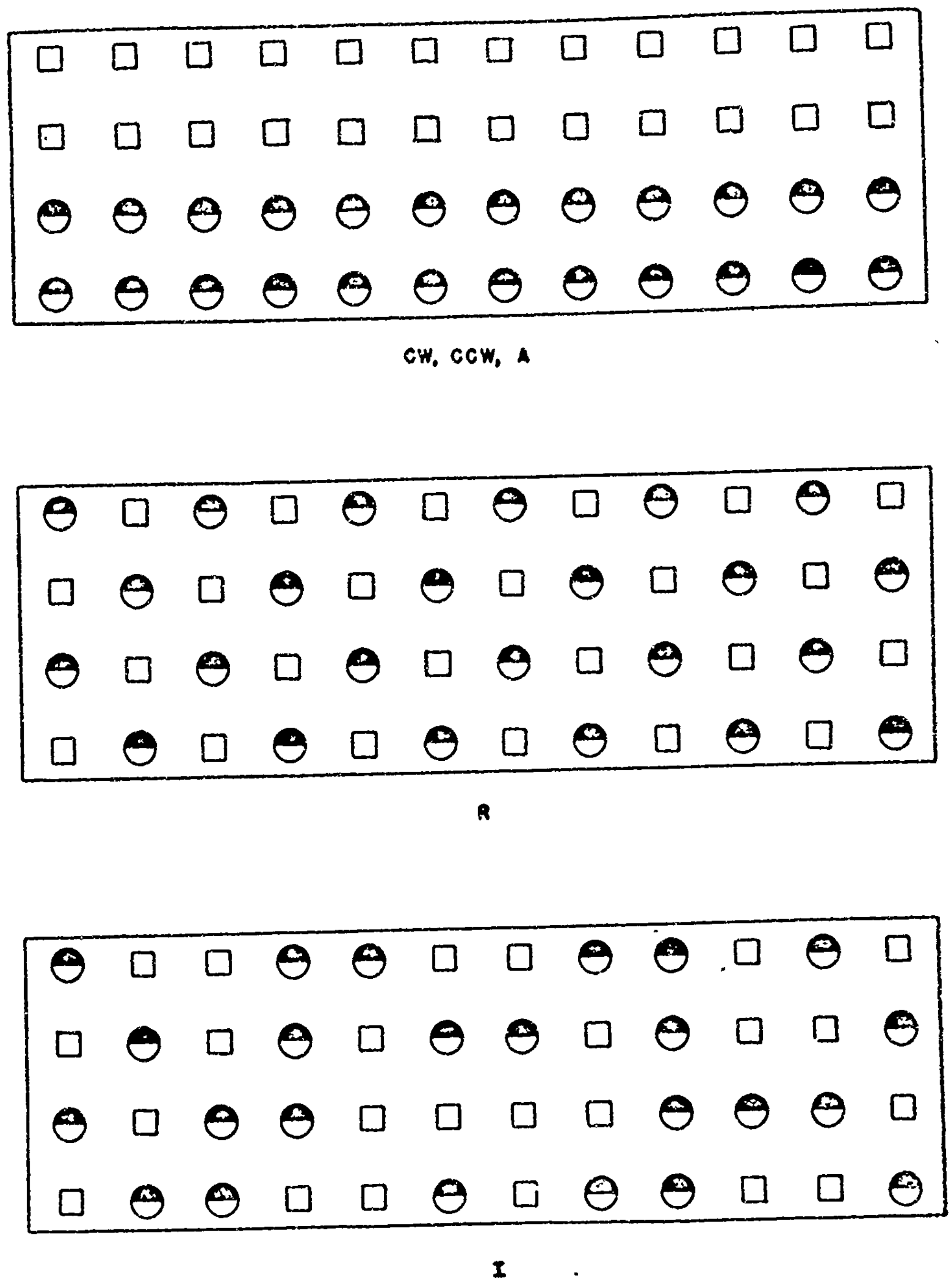


FIGURE 16.1.—Pegboard arrangements for problems CW, CCW, A, R, and I.

tions which defined the specific task at the beginning of each new problem and were told the number of times it was to be performed. Four trials were given on each problem and each trial was timed with a stopwatch. The time scores from the four trials were averaged and the individual mean scores were used in computing the group means, standard deviations, correlation coefficients, and critical ratios.

A counterbalanced experimental order was used, involving 48 subjects. As has been pointed out, 24 of the subjects were tested on problems 1, 2, and 3, in which the subjects worked from left to right, while the second group of 24 were run on problems 1–a, 2–a, and 3–a, in which the subjects worked from right to left, and the data from the two groups were combined. All subjects were run on problems 4 and 5 which are referred to hereafter as R (regular) and I (irregular). Half of the subjects started with four trials on I (I-R group) and the other half started with four trials on R (R-I group); in each case the first four trials were followed by four of the other condition.

RESULTS

Direction of Peg Rotation

In the first six problems, clockwise, counterclockwise, and alternate (clockwise-counterclockwise) rotation of the pegs was investigated. Since in the first six problems one-half of the group, i.e., 24 subjects, began at the left and worked to the right, and the other 24 subjects began at the right and worked to the left, it seemed desirable before combining the data to compare the left starters with the right starters. An examination of the mean time scores reveals that there is no significant difference in mean scores for the groups that began from left to right as compared with the ones beginning from right to left. The critical ratios were 0.4, 0.0, and 1.7, respectively, for the two orders of starting on the CW, CCW, and A problems. It seems appropriate, therefore, to combine the data for these two conditions.

It can be seen from table 16.1 that CW is slightly better than CCW but not significantly so, the critical ratio being 0.87. It is interesting to note here that some subjects remarked that CCW seemed to be an easier task than CW, although on no single occasion did the time scores

TABLE 16.1.—*Summary of the data for clockwise (CW), counterclockwise (CCW), and clockwise-counterclockwise (A) rotation of pegs in control arrangement experiment*

[N=48]

Test condition	M	SD
CW	26.35	2.57
CCW	26.63	2.47
A	27.56	2.89
	r	
	CCW	A
CW	0.61	0.56
CCW	....	.59
	CR	
	CCW	A
CW	0.87	3.26
CCW	....	2.44

support the observation. CW and CCW are both statistically significantly better than A, the data yielding a critical ratio of 3.26 for the former difference, and one of 2.44 for the latter.

Regular Versus Irregular Spacing

In the study of the two variables R and I appreciable learning effects were encountered. From table 16.2 and figure 16.2 it can be seen

TABLE 16.2.—*Means and standard deviations of each trial for regular (R) and irregular (I) spacing of pegs for the group starting with R (R–I group), that starting with I (I–R group), and the two combined*

Group	Condition	Trial 1 M	Trial 1 SD	Trial 2 M	Trial 2 SD	Trial 3 M	Trial 3 SD	Trial 4 M	Trial 4 SD
R-I	R	31.54	3.10	29.30	3.50	28.40	2.91	28.74	3.22
	I	28.29	2.96	27.93	3.13	27.67	3.06	27.58	3.29
I-R	R	28.27	3.30	29.05	4.07	28.52	2.95	28.04	3.75
	I	33.50	5.58	30.31	3.84	29.81	3.52	29.15	2.97
Combined	R	30.41	3.39	29.18	3.79	28.46	2.93	28.39	3.51
	I	30.90	5.17	29.12	3.70	29.74	3.47	28.37	3.23

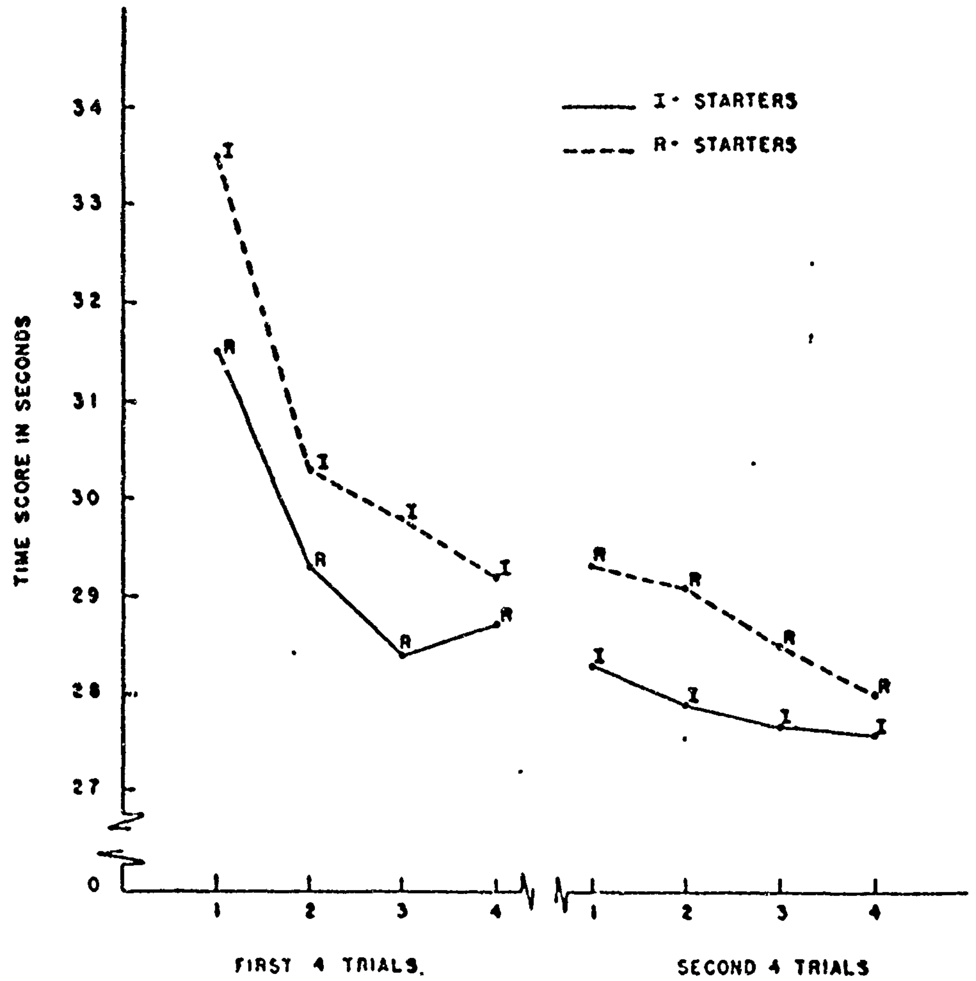


FIGURE 16.2.—Mean time scores from trials 1 to 4 for group I-R and group R-I.

that a considerable amount of learning occurs from trial 1 to trial 4 in the first set of four trials, but there is no appreciable learning for the second set of four trials. The degree of learning in the first set of four trials is statisically significant at the 1-percent level for both the R-I and I-R groups from trial 1 to trial 4.

For the R–I group, i. e., subjects who began with four trials of regular spacing followed by four of irregular, performance on the I condition was significantly better than on R, and for the I-R group, R was significantly better than I. The means, standard deviations, correlations, and critical ratios for these comparisons are given in table 16.3.

TABLE 16.3.—*Means and standard deviations for all four trials averaged for regular (R) and irregular (I) spacing of pegs for the R–I group, the I–R group, and the two combined, and correlations and critical ratios between conditions*

Group	Condition	N	Mean	SD	Difference	r	CR
R-I	R	24	29.53	2.70	1.64	0.83	5.03
	I	24	27.89	2.77			
I-R	R	24	28.73	3.12	1.99	.03	4.00
	I	24	30.72	3.65			
Combined	R	48	29.13	2.98	0.17	.56	.38
	I	48	29.30	3.52			

From these data and from figure 16.2 it can be seen that the subjects do better on the second four trials whether they start on variable R or I. Group R–I did slightly, but not significantly, better than group I–R on both R and I conditions. When data for both groups are combined, performance under R and I conditions is not significantly different (CR=0.38).

It is noticeable that the variability between individuals for condition I on the first four trials is considerably greater than that for R. The difference in variability for trial one is significant at the 2-percent level of confidence. Over all, the critical ratio of the difference in variability between R and I on the first four trials is 1.7.

DISCUSSION AND CONCLUSIONS

The results demonstrate that adjustive serial reactions, either clockwise or counterclockwise, are superior to reactions requiring alternation of these movements. It is concluded from this finding that a group of controls that must be adjusted rapidly in sequential order should, if possible, all operate in a similar direction.

Clockwise rotation in the present experiment was not significantly better than counterclockwise rotation. It is suggested, however, that this variable might be studied further when all connective movements are made in the same directions, i. e., always to the right or always to the left. There was no significant difference between the scores of those who began at the left and worked to the right and those who began at the right and worked to the left. The procedure in this experiment, it should be recalled, was for one-half the subjects to work from left to right followed by right to left, and the other half to work from right to left followed by left to right, and each subgroup worked an equal number of rows in each direction. As has been suggested, it would be interesting to determine the results if the subjects were required to work across both rows of the peg board in one direction only.

The lack of any significant differences in speed of performance with the regular and irregular arrangements of pegs could be explained in several ways. It could be that there is a lack of any real difference between regular and irregular spacing, although the curves in figure 16.2 seem to indicate that learning effects are distorting the results and tend to obscure the experimental variables. The test was carried out under direct visual control of the subjects and it may be that under such conditions the factor of regular versus irregular spacing is unimportant. It would be of interest to repeat the study in the absence of vision. Also it is possible that in the particular peg-turning test employed, the time required for connective movements is small in relation to the time required to make adjustive movements, and hence the test was not sufficiently sensitive for studying factors affecting the speed of connective movements.

SUMMARY

1. In this study two variables pertaining to serial reactions were investigated. The variables were regular versus irregular spacing of pegs and clockwise versus counterclockwise versus alternate clockwise-counterclockwise rotation of pegs. Forty-eight Army Air Forces pilots were tested on each of the variables in a counterbalanced sequence.

2. The results indicate that (a) in this study clockwise rotation of pegs is not significantly better than counterclockwise rotation; (b) clockwise or counterclockwise rotation of pegs is significantly superior to alternation of these movements; and (c) there was no significant difference in performance between regular and irregular spacing of pegs.

CHAPTER SEVENTEEN

Efficiency of Several Types of Control Movements in the Performance of a Simple Compensatory Pursuit Task[1]

WALTER F. GRETHER

INTRODUCTION

In many types of control operations performed by human beings, movements of the limbs are made more or less continuously in response to visual indicators. The flying of an airplane is one such type of control operation. Relatively little is known concerning the factors which influence the efficiency with which such movements can be made. Many variations are possible in the particular muscle groups being employed, depending on the limbs or portions of limbs being used, and the direction of movement. Many factors in the controls themselves also may be varied, such as friction, backlash, damping, spring centering, mass, extent of movement, arc of rotation, and direction of movement relative to the indicator.

A number of research studies on this general problem are known to have been carried out during the war. Hick (4) working in Great Britain investigated the advantages of friction versus no friction in crank-type controls under conditions simulating jolting and no jolting. This investigator concluded that friction increased efficiency of operation under jolting conditions. For simple to-and-fro movements friction was unfavorable under no-jolting conditions. Vince (7), also working in Great Britain, showed that in general a nonlinear relation between a control and a display is undesirable. In another study Vince (6) showed that fewer errors are made when a control moves in the expected direction in relation to the movement of the indicator.

[1] This chapter is based upon research findings reported in Headquarters AMC, Engineering Division, Memorandum Report Number TSEAA-694-9.

A number of German workers under the direction of Henschke (2) studied a considerable number of variables which might be expected to influence control efficiency. Among the conclusions reached by this German group were the following: (1) control is less efficient with the feet and legs than with the arms and hands; (2) control with the entire arm and shoulder including the wrist and hand is more efficient than use of the fingers only; (3) control is best when the joints are at a moderate degree of flexion; (4) friction, mass, and backlash are all undesirable in controls; and (5) a single control grasped by both hands and moved in two or three dimensions can be controlled with greater precision than can the necessary number of separate controls having unidirectional movement (3). These German studies were, however, carried out with small numbers of subjects and apparently were not given adequate statistical treatment to establish significance of the differences. For this reason the German conclusions cannot be accepted as final.

The present series of studies was in part stimulated by the German and British work cited above and is merely the beginning of a comprehensive program of investigation. The purpose of these studies was the measurement of efficiency of several types of control movements in the performance of a simple compensatory pursuit task. This type of task had been used in an experimental aircrew selection test developed at the AAF School of Aviation Medicine (1), and had been shown to have high reliability and a relatively flat learning curve. Both of these were considered significant advantages for the present investigations. Another reason why a very simple control task was chosen in preference to a more complex one was that it was considered advisable to begin with single-dimensional movement and reserve multidimensional movement for later studies.

APPARATUS

The apparatus consisted basically of a Single Dimension Pursuit Test (supplied by the AAF School of Aviation Medicine) which was modified to suit the needs of these particular experiments. The subject was seated in a simplified cockpit shown in figure 17.1 which provided a pair of rudder pedals (from a P-47 aircraft) and an interchangeable wheel and stick (from a Link trainer) as controls. The instrument panel contained a single autosyn indicator with two pointers. On the black dial of the indicator were two white reference marks or targets, one at the top and one at the left side. The task of the subject was to keep one of the pointers centered on its corresponding reference mark by use of one of five possible control movements. These movements were as follows: (1) reciprocating fore-and-aft movement of the right and left rudder pedals, with resting of heels on floor per-

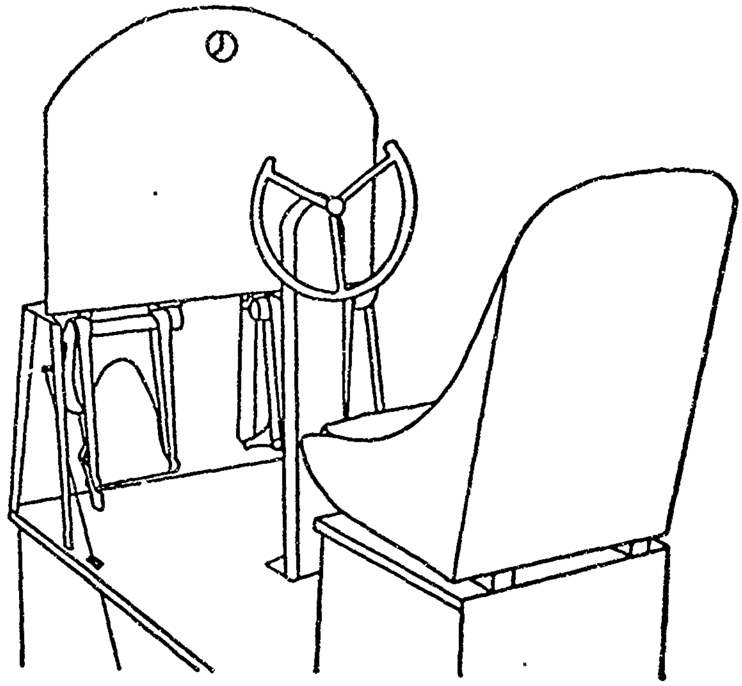


FIGURE 17.1—Compensatory pursuit apparatus used in the study of efficiency of several types of control movements.

mitted; (2) lateral (aileron) movement of stick with preferred hand; (3) fore-and-aft (elevator) movement of stick with preferred hand; (4) rotary (aileron) movement of wheel with both hands; and (5) fore-and-aft (elevator) movement of wheel with both hands. With all rudder and aileron movements the upper reference mark was used. The pointer moved to the right when the right rudder was pushed, when the stick was moved to the right, or when the wheel was rotated clockwise, and vice versa. Thus the upper pointer was operated in much the same manner as the rate-of-turn indicator in an airplane. The mark at the left side of the indicator was used only for fore-and-aft movements of the wheel and stick, and the pointer moved downward in response to forward movement of the control, thus simulating the rate-of-climb indicator. These seemed to be the most natural of the available relationships between indicator and control movements, particularly for pilots.

Random movements on the indicator were produced by the cam follower of the SAM Single Dimension Pursuit Test, which was linked to an autosyn transmitter. The cam in this test was mounted directly on the shaft of a synchronous motor which was supported on a small wheeled carriage riding on parallel tracks. Compensation for the movements of the cam follower, and hence the autosyn indicator, could be made by appropriate movements of the motor and cam. To accomplish this the carriage bearing the motor was linked to a drive shaft by a rack and pinion gear. Three continuous cables from the rudder, elevator, and aileron controls were led around pulleys on this shaft. Selection of the control to be linked to the motor carriage was by insertion of a pin through the hub of the appropriate pulley into the drive shaft. The ratio of control to indicator movement could be varied by exchange of gears on the drive shaft.

The measure of performance on this test was the time during which the pointer on the indicator was actually held on the white reference mark. To obtain this measure a scoring contact, moving with the cam follower, rode over a stationary inlaid contact and activated an electronic relay. This relay in turn energized the clutch of a Standard Electric time clock. A width of scoring contact was chosen which would allow the average subject to score during approximately 50 percent of the testing time. The width of the white reference mark on the indicator was then chosen so that the scoring would occur whenever the pointer overlapped any part of the white area.

One of the two pointers on the indicator was always used with the reference mark at the top of the dial, and the other pointer with the mark at the left side. Selection of the pointer to be used was by means of a rotary selector switch. A third position on the switch reversed the direction of movement of the upper pointer.

Timing of test trials was by means of a microswitch operated once during each revolution (one RPM) of the motion-controlling cam. This switch, connected in parallel with a "start" switch on the experimenter's panel, controlled the current to the cam and clock motors. The test trial was begun by closing the "start" switch, and was ended whenever both this switch and the cam microswitch were open. Thus the test could be stopped accurately at any multiple of 1 minute.

EXPERIMENT 1.—COMPARISON OF RUDDER, STICK, AND WHEEL CONTROLS WITH UNEQUAL EXTENT OF MOVEMENT

Purpose

The purpose of this experiment was to compare the efficiency of operators in performance of a pursuit task using the following five control movements: rudder, elevator (stick), elevator (wheel), aileron (stick), aileron (wheel).

Procedure

Twenty-four male subjects were used in this experiment. Each subject reported for 2 experimental sessions, usually 1 in the forenoon and 1 in the afternoon. Half of the subjects used rudder and stick controls (3 test conditions) in the morning and rudder and wheel controls (3 test conditions) in the afternoon. For the remaining half the sequence was reversed. Each of the experimental sessions was begun with 1 minute of practice on each of the 3 movements (rudder, elevator, aileron) to be performed during the test session. The actual test consisted of six 5-minute test periods with 3 test conditions presented in counterbalanced order. The 6 possible counterbalanced orders of 3 conditions were equally distributed among the 24 subjects. For each of the 1-minute practice periods and the 5-minute test periods the time score was recorded on a suitable data blank.

The maximum distance of the control movement required to keep the indicator centered was approximately as follows for each of the five controls: rudder, 4 inches; elevator (stick), 8 inches; elevator (wheel), 8 inches; aileron (stick), 8 inches; aileron (wheel), 11 inches. These distances have roughly the same ratio to one another as do the maximum distances of control movement in aircraft. There is considerable variability among aircraft in the ratios of these distances.

Results

The results of experiment 1 are summarized in table 17.1, which gives the average time scores for the 24 subjects under the 5 test conditions. In the table are also given the differences between the average scores for the 5 conditions and corresponding critical ratios. Those differences and critical ratios which are significant at the 1-percent or higher level of confidence are marked by an asterisk.

TABLE 17.1.—*Comparison of control efficiency on a simple compensatory pursuit task using rudder, stick, and wheel controls with unequal extent of movement* [1]

[Experiment 1. N=24]

	Type of Movement				
	Rudder	Elevator (stick)	Elevator (wheel)	Aileron (stick)	Aileron (wheel)
PERCENT TIME ON TARGET					
Mean	51.68	61.12	60.97	55.93	54.74
SD	8.00	12.72	11.62	10.41	9.08
DIFFERENCES					
Rudder	....	*−9.44	*−9.29	*−4.25	*−3.06
Elevator (stick)	....	....	+ .15	*+5.19	*+6.38
Elevator (wheel)	....	....	....	*+5.04	*+6.23
Aileron (stick)	....	....	....	....	+1.19
CR					
Rudder	....	*5.63	*6.78	*3.19	*2.67
Elevator (stick)	....	....	.12	*4.56	*5.31
Elevator (wheel)	....	....	....	*4.20	*6.22
Aileron (stick)	....	....	....	....	.91

[1] All differences and critical ratios marked by an asterisk (*) are significant at the 1-percent or greater level of confidence. Critical ratios were computed from the standard deviations of the distributions of individual difference scores rather than from the distributions of raw scores. Thus each subject served as his own control in the statistical analysis.

In determining the significance of the differences in table 17.1 the differences between scores made under the five test conditions were computed for each individual subject. Critical ratios were then computed from the means and standard errors of the distributions of these differences (for exact method see Peters & Van Voorhis (5), pp. 165–

167). By this statistical procedure each subject served as his own control.

Discussion

It is apparent from table 17.1 that control with the rudder (feet and legs) was significantly less efficient than with the stick and wheel (hands and arms). Also, the aileron (lateral and rotary) direction of control was less efficient than elevator (fore and aft) control for both the stick and wheel. Although performance was slightly better with the stick than with the wheel, for both directions of control, the differences were small and could have easily occurred by chance.

These findings are, however, subject to a number of qualifications. It will be recalled that the extent of movement was not equalized for all of the controls and this fact may have affected the results. This possibility was checked in experiment 2 in which the extent of movement was equalized on the rudder and both dimensions of stick movement. Other possible influences which could not be precisely equalized were slack, friction, and inertia in the control systems. Evaluation of these factors is needed in future studies.

Experience during the war with similar tests used for selection and classification demonstrated that the major cause of differences in scores (aside from the individuals themselves) was variation in the effective width of the scoring area. In this and the following experiments the same scoring contact was used for all controls, thus eliminating this as a possible cause of differences in performance. Both friction and inertia (mass) were so low in this experiment that it is doubtful that they could have been significant influences. Measurements of slack showed this to be virtually zero for the rudder, ½ inch for elevator control (both wheel and stick), ¾ inch for aileron movements with the stick, and 1½ inches for the aileron movements with the wheel. The poorer performance on the rudder could therefore hardly have been caused by slack (unless slack is desirable). The somewhat poorer performance for aileron as compared with elevator movements with the wheel might have been caused wholly or in part by differences in slack.

In this first experiment many of the subjects complained of eye strain resulting from the 5-minute test period. For this reason, shorter test periods were used in the subsequent experiments.

EXPERIMENT 2.—COMPARISON OF RUDDER AND STICK CONTROLS WITH EQUAL EXTENT OF MOVEMENT

Purpose

The purpose of this experiment was to compare the efficiency of control on the pursuit task using the following types of control movements with equal extent of movement: rudder, elevator (stick), and aileron (stick).

Procedure

Thirty-six rated pilots served as subjects in this and the following three experiments. Each of these men reported for a ½ hour test session on each of four successive days. This experiment was administered on the second of these days, following the test reported as experiment 3. The ½-hour session consisted of six 2-minute test periods with test conditions presented in counter-balanced order as in experiment 1. The three test conditions used were (1) rudder, (2) elevator (stick), and (3) aileron (stick). The maximum extent of movement required on each control to keep the pointer centered was 4 inches.

Results

In table 17.2 is presented a summary of the results for experiment 2. The average scores for the several test conditions, the differences between the scores, the critical ratios, and the significance of the differences are presented as in table 17.1.

TABLE 17.2.—*Comparison of control efficiency on a simple compensatory pursuit task using rudder and stick controls with equal extent of movement*[1]

[Experiment 2. N=36]

	Type of movement		
	Rudder	Elevator (stick)	Aileron (stick)
PERCENT TIME ON TARGET			
Mean	55.78	67.93	59.83
SD	5.19	6.15	7.05
DIFFERENCES			
Rudder		*−12.15	*−4.05
Elevator (stick)			*+8.10
CR			
Rudder		*17.17	*2.89
Elevator (stick)			*11.57

[1] All differences and critical ratios marked by an asterisk (*) are significant at the 1-percent or greater level of confidence. Critical ratios were computed from the standard deviations of the distributions of individual difference scores, rather than from the distributions of raw scores. Thus each subject served as his own control in the statistical analysis.

Discussion

The results of experiment 2 agree closely with those of experiment 1 in showing that rudder control was least efficient, and that aileron control was somewhat less efficient than elevator control, even when the extent of movement was equalized.

EXPERIMENT 3.—EFFECT OF A REVERSAL IN DIRECTION OF INDICATOR MOVEMENT UPON EFFICIENCY OF CONTROL

Purpose

The purpose of this experiment was to determine the effect upon control efficiency in performance of a pursuit task of a reversal in the direction of indicator in relation to control movement.

Procedure

Thirty-six rated pilots served as subjects. These were the same men used in experiments 2, 4, and 5. This experiment was run during the first of the 4½-hour test sessions. Each subject was given eight 1-minute test periods during each test session. Half of the subjects were tested on the rudder during the first four test trials and on the aileron during the last four test trials. For the remaining half of the subjects this sequence was reversed. Tests using the rudder were made under two conditions which were counterbalanced in ABBA order for the first half of the subjects and BABA order for the remaining half of the subjects. The same counterbalancing procedure was used for the tests on the aileron control.

The two test conditions used for the rudder were (1) rudder direct, in which the upper indicator moved to the right when the right rudder pedal was pushed and to the left when the left rudder pedal was pushed, and (2) rudder reversed, exactly the reverse of the above-mentioned condition. The two test conditions used for the aileron were (1) aileron direct, in which the upper pointer moved to the left when the stick was moved to the left and the pointer moved to the right when the stick was moved to the right, and (2) aileron reversed, which was exactly the reverse of the above-mentioned condition. The maximum movement required on the rudder to keep the pointer centered was 4 inches, and the movement required on the aileron was 8 inches. The stick rather than wheel was used for aileron control. Time scores were recorded on a suitably prepared data sheet at the completion of each 1-minute test period.

Results

In table 17.3 are presented time scores, the differences in scores for direct and reverse conditions, and the critical ratios for this experiment.

TABLE 17.3.—*Comparison of control efficiency on a simple compensatory pursuit task using rudder and stick controls with direct and reversed movement on the indicator*[1]

[Experiment 3. N=36]

	Type of movement			
	Rudder direct	Rudder reversed	Aileron direct	Aileron reversed
PERCENT TIME ON TARGET				
Mean	49.65	44.25	54.25	52.35
SD	6.12	8.86	5.41	6.29
DIFFERENCES				
Direct—Reversed	*+5.40		+1.90	
CR	*4.09		1.48	

[1] All differences and critical ratios marked by an asterisk (*) are significant at the 1-percent or greater level of confidence. Critical ratios were computed from the standard deviations of the distributions of individual difference scores, rather than from the distributions of raw scores. Thus each subject served as his own control in the statistical analysis.

Discussion

It will be noted in table 17.3 that the difference for direct and reverse control on the rudder was significant. For the aileron, however, although the average score was somewhat higher for the aileron direct condition, the difference was not statistically significant. The lack of significance of the difference for the aileron control may very well be attributable to the previous experience of the subjects in use of the artificial horizon. On this indicator the upper marker moves to the right when the aileron is displaced to the left and vice versa (see ch. 8). Thus rated pilots had all had considerable experience with the aileron reverse condition during actual flight.

EXPERIMENT 4.—EFFICIENCY OF RUDDER CONTROL AS A FUNCTION OF ANGLE OF KNEE FLEXION

Purpose

The purpose of this experiment was to compare the efficiency of rudder control in performance of a pursuit task under three conditions of angular flexion at the knee.

Procedure

The 36 rated pilots used as subjects served on this experiment during the third of the four ½-hour sessions during which they were tested on this apparatus. The actual test consisted of six 2-minute test periods, with three test conditions presented in counterbalanced order. The 6 possible counterbalanced orders of the 3 conditions were equally distributed among the 36 subjects.

The three test conditions were maximum extension, medium extension, and minimum extension. The degree of extension was defined in terms of the angle formed by the knee joint as measured by a protractor which could be placed upon the ventral surfaces of the thigh and lower leg. The angles formed at the knee when the rudder pedals were equalized were for the three conditions of 135°, 120°, and 105°. This appeared to be the greatest angular range which could be used without having the subject's reach exceeded during maximum control movements and without causing excessive discomfort. These changes in leg angle were accomplished by shifting the position of the subject's seat. The time-score data were recorded on a suitable data sheet as were also the subjects' statements as to the relative comfort of the three positions. The maximum movement of the rudder required to keep the upper pointer stationary was 4 inches.

Results

The average time scores, the differences, and the critical ratios for experiment 4 are shown in table 17.4.

The subjects' statements as to the relative comfort of the three leg angles are summarized table 17.5.

TABLE 17.4.—*Comparison of rudder control efficiency in the performance of a simple compensatory pursuit task for three angles of leg flexion*[1]

[Experiment 4. N=36]

	Angle of leg flexion		
	Maximum extension (135°)	Medium extension (120°)	Minimum extension (105°)
PERCENT TIME ON TARGET			
Mean	60.53	60.35	59.48
SD	8.26	8.22	7.02
DIFFERENCES			
Maximum extension		+.20	+1.05
Medium extension			+.85
CR			
Maximum extension		.32	1.56
Medium extension			1.67

[1] Critical ratios were computed from the standard deviations of the distributions of individual difference scores, rather than from the distributions of raw scores. Thus each subject served as his own control in the statistical analysis.

TABLE 17.5.—*Relative comfort of three angles of leg flexion during performance of a simple compensatory pursuit task by use of rudder control*

[Experiment 4. N=36]

	Relative comfort		
	Greatest comfort	Intermediate comfort	Least comfort
FREQUENCY OF JUDGMENTS FOR THE FOLLOWING CONDITIONS			
Maximum extension (135°)	8	13	15
Medium extension (120°)	25	9	2
Minimum extension (105°)	3	14	19

Discussion

It will be noted in table 17.4 that the efficiency of rudder control was approximately equal under the three conditions of leg extension. All of the differences are very small and lack statistical significance. It will be noted in table 17.5 that the medium angle of leg flexion was judged to be most comfortable and that the maximum and minimum angles of leg flexion used in this experiment were judged to be approximately equal in comfort.

EXPERIMENT 5.—EFFICIENCY OF ELEVATOR CONTROL AS A FUNCTION OF ANGLE OF ARM FLEXION

Purpose

The purpose of this experiment was to compare the efficiency of elevator control in performance of a pursuit task using a wheel under three conditions of angular flexion at the elbow.

Procedure

The 36 rated pilots used as subjects served in this experiment on the last of the four ½-hour test periods. The actual test consisted of six 2-minute test periods with 3 test conditions presented in counterbalanced order. The 6 possible counterbalanced orders of the 3 conditions were equally distributed among the 36 subjects.

The three test conditions were maximum extension (135°), medium extension (120°), and minimum extension (105°). This appeared to be the greatest angular range which could be used without having the subject's reach exceeded or the wheel strike the seat during maximum control movements. The angle formed by the elbow was measured in the same manner as the angle of the knee had been measured in experiment 4, except that the measuring instrument was placed on the dorsal surface of the upper and lower arm. As in experiment 4 the time scores and the subjects' reports of the relative comfort of the three positions were recorded. The variations in angle formed by the elbow were accomplished by changes in the position of the subjects' seat. The maximum extent of wheel movement required to keep the upper pointer centered was 4 inches.

Results

The time scores, differences, and critical ratios for this experiment are presented in table 17.6. The subjects' statements as to the relative comfort of the three positions are summarized in table 17.7.

TABLE 17.6.—*Comparison of elevator control efficiency in the performance of a simple compensatory pursuit task for three angles of arm flexion*

[Experiment 5. N=36]

	Angle of arm flexion		
	Maximum extension (135°)	Medium extension (120°)	Minimum extension (105°)
PERCENT TIME ON TARGET			
Mean	65.65	66.38	65.95
SD	7.34	8.08	7.90
DIFFERENCES			
Maximum extension		−.73	−.30
Medium extension			+.43
CR			
Maximum extension		1.38	.60
Medium extension			.77

TABLE 17.7.—*Relative comfort of three angles of arm flexion during performance of a simple compensatory pursuit task by use of elevator (wheel) control*

[Experiment 5. N=36]

	Relative comfort		
	Greatest comfort	Intermediate comfort	Least comfort
FREQUENCY OF JUDGMENTS FOR THE FOLLOWING CONDITIONS			
Maximum extension (135°)	1	5	30
Medium extension (120°)	19	16	1
Minimum extension (105°)	16	15	5

Discussion

It will be noted in table 17.6 that the greatest time scores were obtained under the medium extension condition. As in experiment 4, however, the differences are exceedingly small and are not statistically significant. It will be noted in table 17.7 that the medium and minimum extension conditions of elbow flexion used in this experiment were considered about equally comfortable, whereas there is almost universal agreement that the condition of maximum extension was least comfortable. The most reasonable conclusion from this experiment would seem to be that there are no significant differences in efficiency of control under the three degrees of arm flexion used in this experiment, but that the condition of maximum flexion should be avoided because of the discomfort to the operator.

GENERAL SUMMARY OF FINDINGS

1. Control with the hands and arms was found to be more efficient than that with the feet and legs. Of the two directions of hand and arm control used, the fore-and-aft (elevator) movements were found to be more efficient than lateral or rotary (aileron) movements (experiments 1 and 2).

2. Efficiency of control was approximately equal for stick and wheel controls for both elevator and aileron directions of movement (experiment 1). The stick was operated with one hand and the wheel with two.

3. The use of a direct relationship between control and indicator movement on the rudder (i. e. pointer moving to the right when right rudder is pushed) resulted in significantly better performance than a reversal of this relationship. A reversal of the relationship on the aileron control caused only a slight (and not statistically significant) reduction in control efficiency (experiment 3).

4. Three median angles of leg flexion, 135°, 120°, and 105°, resulted in approximately equal efficiency of control with the rudder, although the position of medium extension (120°) was judged by the subjects to be most comfortable (experiment 4).

5. Three median angles of arm flexion, 135°, 120°, and 105°, resulted in approximately equal efficiency of elevator control with the wheel. The condition of maximum extension (135°) was judged to be least comfortable, and the other two conditions were judged to be about equally comfortable (experiment 5).

REFERENCES

1. Brown, J. S. The S. A. M. Single Dimension Pursuit Test, Models A, B, C, D, and E (CMS01-1, B-2, B-3, B-4, and B-5). *AAF School of Aviation Medicine Research Report, Project No. 42, Report No. 1.* June 1944.

2. Fitts, P. M. *German Air Force research on the design of sights and controls for most effective use by the operator.* Mimeographed report, Psychology Branch, Aero Medical Laboratory, Aug. 1945.
3. Henschke, U., and Karlson, P. *A new control system for fighter aircraft.* Translation from German, by Air Documents Division, Intelligence (T-2), Headquarters, Air Matériel Command, Translation No. 564, released May 1946.
4. Hick, W. E. Friction in manual controls with special reference to its effect on accuracy of corrective movements in conditions simulating jolting. Flying Personnel Research Committee, *Report No. 623*, Great Britain, June 1945.
5. Peters, C. C., and Van Voorhis, W. R. *Statistical procedures and their mathematical bases.* New York: McGraw-Hill Book Company, Inc., 1940.
6. Vince, M. A. Direction of movement of machine controls. Flying Personnel Research Committee, *Report No. 637*, Great Britain, Aug. 1945.
7. Vince, M. A. The psychological effect of a nonlinear relation between control and display. *SRI (Servo) Report No. 2*, Great Britain, Feb. 1946.

BLANK PAGE

CHAPTER EIGHTEEN

An Experimental Comparison of the Accuracy of Sighting and Triggering with Three Types of Gun-Sight Handgrip Controls[1]

A. P. Johnson and J. L. Milton

INTRODUCTION

The investigations reported in this chapter were undertaken as part of a comprehensive research program on the design of gun-sight controls for most efficient use by the operator. The studies were concerned with the comparison of several types of controls for use when the primary task of the operator is to move the sight itself in such a manner as to keep the reticle on a moving target.

The standard handwheel controls employed on the B-29 Pedestal Sight during the war were subjected to considerable criticism by the men who used them. In operating this sight the gunner's right hand must operate two wheels on the same shaft, a serrated one controlling ranging (framing) and a smooth one which aids the left hand in elevation tracking. Frequently, in order to track the target properly in elevation and in order to range correctly at the same time, the serrated wheel must be rotated in a direction opposite to that in which the smooth wheel is moved. This often causes one movement to interfere seriously with the other. A further difficulty arises because of the small radius of the handwheels which are on the axis of rotation of the sight. This lack of adequate leverage is asserted to result in fatigue and irregular elevation tracking. If the right hand is used only to operate the ranging wheel, then the left hand must carry the major portion of the task of tracking in azimuth and in elevation. As the wheels are rotated for elevation tracking, the thumbs of the right and left hands must be moved over flexible rubber covers in order to press

[1] This chapter is based upon research findings reported in Headquarters AMC, Engineering Division, Memorandum Reports No. TSEAA-694-2 and TSEAA-694-2A.

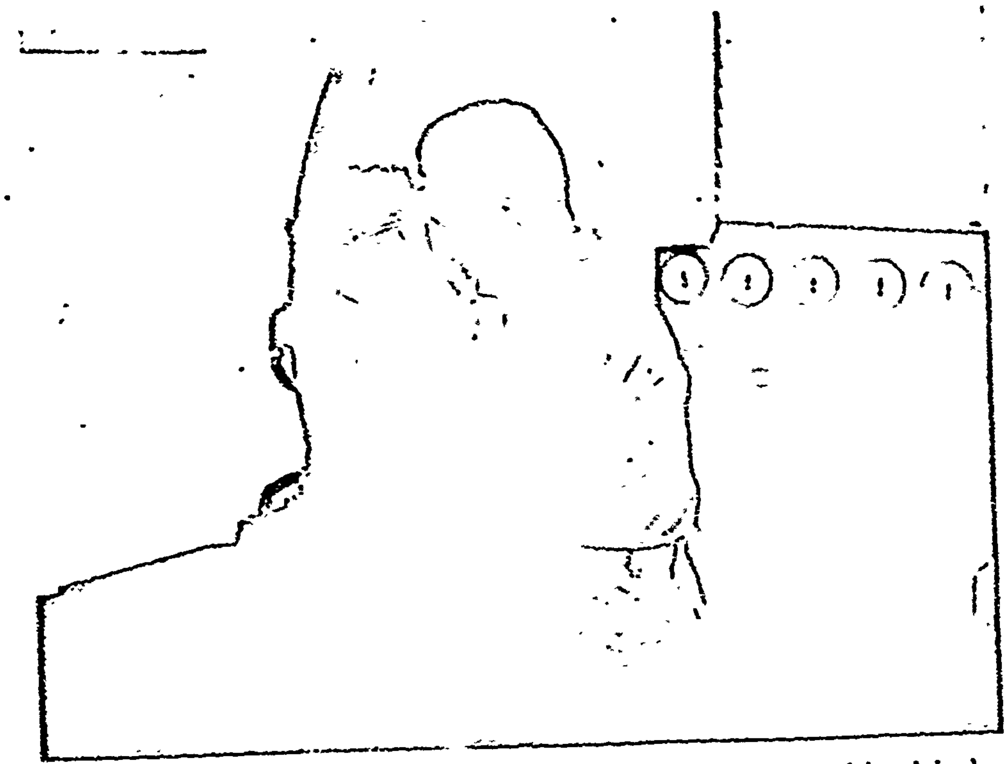
the triggers. At extreme positions t... thumbs are frequently below or above a satisfactory triggering po...tion. Unless the thumbs are raised well above the rubber trigger cov...s, friction against the rubber causes the thumbs to drag from one ...levation position to another. Furthermore, the triggers both require a great deal of pressure to operate. Persons using these controls complain that their hands become fatigued if they operate the sights for more than very short periods of time.

In 1945 Project AC-94 of the Applied Psychology Panel, N. D. R. C., published reports (2, 4) describing the design of two alternative sets of hand controls for the B-29 pedestal sight. One report (4) gave the results of an experimental study of the efficiency of sighting performance with a set of modified controls. Another report (2) described a still different design for hand controls but gave no experimental data on the performance of gunners in using this revised design. On the basis of these reports the Armament Laboratory, Engineering Division, Air Matériel Command, constructed two sets of hand controls. The Aero Medical Laboratory was requested to repeat the initial N. D. R. C. study with the first set of hand grips and also to study the second design. Studies were undertaken, therefore, to determine the relative superiority, if any, of the two new designs over the old type B-29 hand controls, and to determine which of the proposed new designs was superior. These studies were believed worth while for several reasons. It was considered desirable to repeat the validation study reported by Project AC-94 and to add triggering which was not included in the previous study. Since the present research utilizes an apparatus that is different from the one employed by Project AC-94, the repetition of the earlier study is of further interest from the standpoint of comparing research findings with two different methods of investigation and two different apparatuses.

EXPERIMENT 1.—COMPARISON OF CONTROLS A AND B

Apparatus

A convenient apparatus for carrying out experiments 1 and 2 was provided by the SAM Pedestal Sight Manipulation Test, shown in figure 18.1. The construction, calibration, and operation of this test have been described by Judson S. Brown (1) and reliabilities have been reported by Moncreif H. Smith (3) in publications from the AAF School of Aviation Medicine. This test was originally designed as a selection device for B-29 gunners. It provides a realistic and standardized series of pursuit-course attacks requiring tracking, ranging, and triggering by the operator. For each series of attacks there can be obtained a time score for tracking correctly, for ranging correctly, and for triggering correctly, and also a time score for any combination



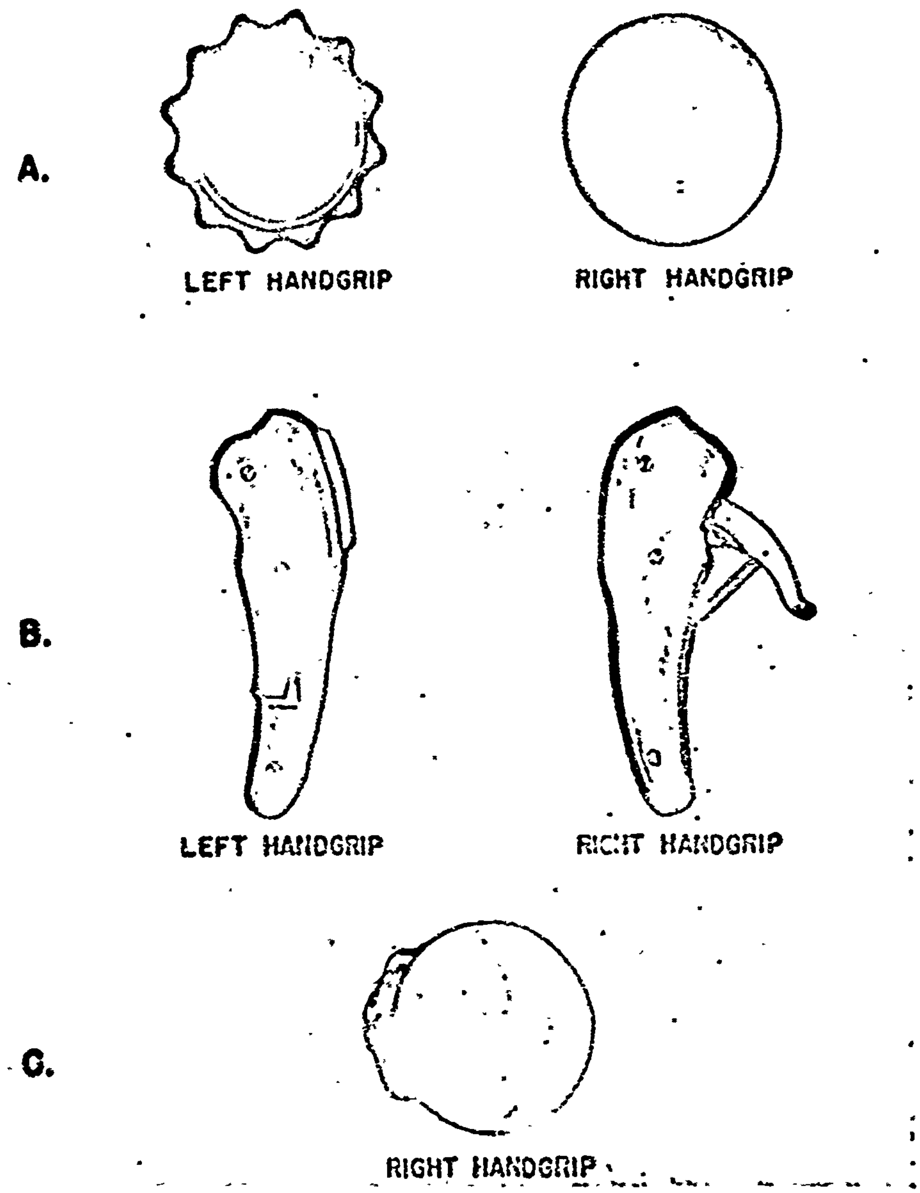


Figure 18.2.—Original handwheel controls (A) and two experimental controls (B and C) used in Experiments 1 and 2.

of these functions. Independent scores can also be obtained for azimuth and elevation tracking. The criteria of correctness was defined in terms of angular degrees of divergence of the line of sight from the target. It is believed that the scores provided by the apparatus are as good criteria of gunnery performance as any currently available.

Description of Controls

The standard hand controls for the B-29 Pedestal Sight have already been described. A view of these controls is shown in figure 18.2. The standard control will be referred to as control A for the purposes of this report. In experiment 1, this control was compared with control B which is shown in figure 18.2. The latter control was designed originally by Mr. Richard L. Solomon of Project AC-94 with the assistance of Mr. Nicholas Yakimovich of the Research Division, Laredo Army Airfield (4). The new control was designed to fit the hand comfortably. Triggering was accomplished by the thumb of the left hand on a trigger which was a part of and moved with the control. Ranging was accomplished by applying pressure with the fingers to a spring-loaded trigger-like extension on the right-hand control. As the target came nearer and increased in size, the operator of this control squeezed his fingers together, moving the trigger-like extension in toward the rest of the handle. The two sets of hand controls were easily interchangeable on the test apparatus.

Experimental Procedure

Two groups of subjects were used, one group beginning with the hand control A, the other with the hand control B. A counterbalanced experimental testing sequence was followed in order to equalize practice effects. Each group operated the hand control A for two 10-minute test periods and the hand control B for two similar periods, each test period occurring on a separate day. The order of testing for one group (N=9) was ABBA and for the other (N=10) was BAAB. The groups consisted chiefly of enlisted men. None had any previous experience with the apparatus or in the use of the B-29 sight.

Each 10-minute trial included four successive 2½-minute series of eight different attack courses (half beginning on the right and half on the left). The test was carefully calibrated prior to each day's testing in accordance with standard procedures. Azimuth and elevation tracking scores were obtained in minutes and hundredths of minutes on target. "On target" was defined as the time during which the center dot of the reticle was within approximately 10 mils vertically or horizontally of the center of the target. Ranging within approximately 6 mils was scored as correct. On each trial two scores were obtained simultaneously: a score giving the amount of time during which the gunner was simultaneously sighting and ranging correctly; and a score indicating the amount of time he was simultaneously sighting, ranging, and triggering correctly. Each person was instructed to trigger only while he was tracking and ranging correctly. Since a high proportion of the group thought they were tracking and ranging correctly nearly all of the time, they held the trigger down almost continuously.

Results

The results obtained in this experiment are summarized in table 18.1 which gives the scores obtained on controls A and B expressed as mean percent of time on the target, the standard deviations of these scores, the differences between scores obtained by the two hand controls, and the significance of these differences. Scores are presented for each of

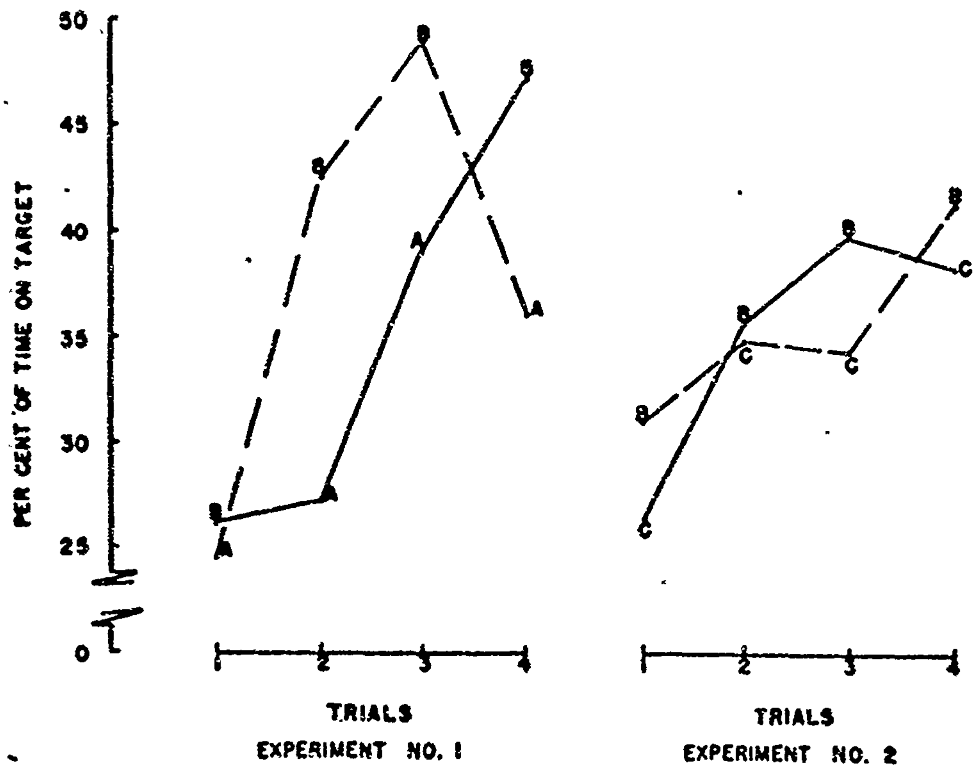


FIGURE 18.3.—Comparison of combined tracking, ranging, and triggering performance—experiment 1. control A versus control B; experiment 2, control B versus control C.

the two groups of subjects and for the groups combined. Data for combined tracking, ranging, and triggering are shown graphically in figure 18.3 for each of the four trials of each experimental group. It can be seen from table 18.1 and from figure 18.3 that there was considerable learning during the trials and that the order of testing had a significant effect on the results for the two subgroups. These results indicate that for the combined tracking and ranging scores, the new hand control B is markedly superior to the old type hand control A. The differences are significant at the 1-percent level of confidence. For combined tracking, ranging, and triggering, the new-type hand controls are even more markedly superior.

For all of the data combined it will be noted that the percentage of time during which subjects sighted, ranged, and triggered correctly with hand control A was 32.1 percent while the percentage of time with hand control B was 41.6 percent. This superiority of approximately one-fourth more time-on-target is not only of statistical significance but, from a practical point of view, is highly significant in indicating the amount of improvement that can be gained by designing controls in relation to human abilities.

TABLE 18.1.—*Mean scores expressed as percent of time on target, standard deviations, and* t *ratios for hand controls A and B*

[N=19]

Score	Experimental group	N	Hand control (A)		Hand control (B)		Mean of difference	SD of difference between paired scores	*t*
			Mean percent time on	SD	Mean percent time on	SD			
Tracking and ranging	ABBA	9	40.4	4.6	47.9	5.6	7.4	4.2	5.0
	BAAB	10	41.4	11.4	43.0	10.2	1.6	5.2	.9
	Combined	19	40.9	8.9	45.4	8.6	4.5	5.6	3.4
Tracking, ranging, and triggering	ABBA	9	30.6	6.2	46.1	5.3	15.5	5.1	8.6
	BAAB	10	33.6	13.2	37.2	12.0	3.6	7.2	1.5
	Combined	19	32.1	10.6	41.6	10.2	9.6	8.6	4.7

EXPERIMENT 2.—COMPARISON OF CONTROLS B AND C[1]

Description of Controls

In this investigation the new-type hand control B, which was found in the preceding study to be significantly superior to the old-type hand control A, was compared with a third control which for the purposes of this study was designated control C. This control is shown in figure 18.2. The left-hand grip was the same for controls B and C. The right-hand control C, which is primarily a ranging control, was patterned after a design proposed by K. U. Smith of Project AC-94 in collaboration with Frank Crossley of the Research Division, Laredo Army Airfield, (2).

With the new hand control C, the movement required for framing is distinct from and independent of the tracking movements in azimuth and elevation. The finger and wrist movements required for framing do not change appreciably as the sight travels in azimuth or elevation. However, the hand accomplishing the framing moves with the azimuth and elevation movements of the sight and serves to stabilize the tracking. In the original design proposed by Project AC-94, this ranging control was to have been operated by the left hand and the tracking and triggering control by the right hand. However, since the B-29 Pede-

[1] Mrs. Sue Diggles assisted the author in the testing of subjects and the analysis of data.

stal Sight is so designed that ranging can be accomplished only by the right hand, the control was tried out in this study on the right hand.

Apparatus and Procedure

The apparatus used in experiment 2 was the same as that employed in experiment 1 and the procedure was similar. In experiment 2, a total of 42 rated pilots were used as subjects. The scoring system was arranged so that an independent ranging score was obtained, in addition to a score which indicated the time during which simultaneous tracking and ranging was accomplished correctly, and the time during which simultaneous tracking, ranging, and triggering was accomplished correctly. Subjects were divided into two groups and a counterbalanced experimental testing sequence was employed as in experiment 1. Subjects were given the 10-minute test trials on four successive days.

Results

The results of experiment 2 are summarized in table 18.2 in which are given the scores for each experimental group and for combined groups. It will be noted from an examination of this table that

TABLE 18.2.—*Mean scores expressed as percent of time on target, standard deviations, and t ratios for hand controls B and C*

[N=42]

Score	Experimental group	N	Hand control (B) Mean percent time on	Hand control (B) SD	Hand control (C) Mean percent time on	Hand control (C) SD	Mean of difference	SD of difference between paired scores	t
Ranging only	BCCB	21	52.6	9.5	50.2	9.0	2.4	6.9	1.6
	CBBC	21	50.6	7.0	46.8	8.1	3.8	9.7	1.8
	Combined	42	51.6	8.3	48.5	8.7	3.1	8.5	2.3
Tracking and ranging	BCCB	21	42.6	8.2	39.4	8.0	3.2	6.2	2.3
	CBBC	21	41.7	9.6	37.0	7.4	4.7	8.8	2.4
	Combined	42	42.1	7.6	38.2	7.8	4.0	7.6	3.4
Tracking, ranging, and triggering	BCCB	21	36.4	9.8	34.6	7.4	1.8	6.7	1.2
	CBBC	21	37.0	6.4	33.0	7.9	4.0	7.3	2.4
	Combined	42	36.7	8.2	33.8	7.6	2.9	7.1	2.6

the hand control B was in all cases superior to the hand control C when all data were combined. The ranging score with control B was superior at the 2-percent level of confidence. Combined tracking and ranging scores were significantly higher and combined tracking, ranging, and triggering scores were also much higher for control B. These differences were significant at better than the 1-percent level.

The retest reliability of the scores is indicated by a coefficient of 0.40 for the B condition and one of 0.50 for the C condition for the data of the 42 subjects.

By comparing the data in table 18.1 for hand control A [illegible] in table 18.2 for hand control C, it will be noted that perf[illegible] control C was intermediate between that on controls A and B [illegible] data suggest therefore that ranging control C is somewhat better than the original control A employed on the B-29 Pedestal Sight but not as good as control B.

In considering the results of the comparisons made in experiment 2, it should be remembered that control C was initially designed to be used by the left hand. It is possible that if control C had been operated by the left hand in this experiment, tracking and framing would have been accomplished with somewhat greater accuracy. The comparison between controls B and C is believed justifiable since in both cases the right hand did the ranging.

GENERAL CONCLUSIONS

The results of experiment 2 as well as the results of experiment 1 indicate that changes in the design of the controls employed in operating a gun sight can result in marked changes in the accuracy with which the sight can be used. Since statistically significant results were obtained in both experiments with relatively small numbers of cases, it can be concluded that the apparatus and procedures employed in these studies provide a reliable method for the systematic investigation of problems in the design of controls for equipment in which the operator is required to keep a moving object centered on a set of cross hairs.

SUMMARY

1. This study was undertaken as part of a comprehensive research program on the design of gun sight controls for most efficient use by the operator. The study dealt specifically with controls for the B-29 Pedestal Sight.

2. In experiments 1 and 2, the tracking, framing, and triggering performance of subjects was recorded when using two new-type hand controls as well as when using the original hand controls.

3. The resulting data indicate that changes in the design of controls can result in marked changes in the accuracy with which a gun sight can be used. Both new-type controls were found to be superior to the original controls on the B-29 Pedestal Sight.

REFERENCES

1. Brown, J. S. The S. A. M. Pedestal Sight Manipulation Test Model B (CM824B) for the selection of B-29 gunners. *Report No. 1, Project No. 352, AAF School of Aviation Medicine*, Randolph Field, Texas, 10 Oct. 1945.

2. Smith, K. U. Psychological factors in the operation of flexible gunnery equipment. *Informal Memorandum No. 10, Report No. 2*, Applied Psychology Panel, N. D. R. C., 20 Aug. 1945.

3. Smith, M. H. The S. A. M. Pedestal Sight Manipulation Test, Model B (CMS24B) for the selection of B-29 gunners. *Report No. 2, Project 352, AAF School of Aviation Medicine*, Randolph Field, Texas, 18 Jan. 1946.
4. Solomon, R. L. An experimental study of the efficiency of sighting performance with modified sight controls on the B-29 Pedestal Sight. *O. S. R. D. Report No. 5558, Project AC-94*, Applied Psychology Panel, N. D. R. C., 10 Sept. 1945.

CHAPTER NINETEEN

The Effect of Anoxia On Visual Illusions

Walter F. Grether, John T. Cowles, and Richard E. Jones[1]

INTRODUCTION

A variety of environmental influences, such as long periods of work, reduced oxygen supply, noise, and extremes of humidity and temperature are generally assumed to impair human efficiency. But in most cases experimental attempts to measure this impairment in terms of decrements in performance have been relatively unsuccessful. When placed in a test situation, human beings apparently are able to resist the environmental influences and achieve an approximately normal test score. This is particularly true when the environmental influences being studied cannot be concealed from the experimental subjects. Moreover, most experiments have been concerned with the subject's ability to perform a motor or intellectual task, with emphasis upon speed and accuracy. In such tests the subjects are well acquainted with the criteria on which their performance is being scored.

During the recent war, a German research worker, W. Ehrenstein,[2] reported a technique for measuring fatigue by means of visual illusions. He tested shipyard workers at the beginning and end of a 9-hour day of hard physical labor, and found that the extent of the illusions was greater at the end of the day. This technique is of special interest because it involves perceptual judgments of equality in which there are no right or wrong responses in the usual sense. Thus, the subjects would not be expected to conceal the effects of fatigue by concentration of effort during the test. Such a test, if sufficiently discriminative and correlated with more direct measures of human efficiency, has considerable potential application.

The experiments reported here were stimulated by the German experiment mentioned above. As a starting point it was decided to

[1] The first author exercised general supervision over experiment 1 and prepared this report. The second author did most of the detailed planning and test administration for experiment 1. The third author assisted in experiment 1 and carried out experiment 2.

[2] Fitts, P. M. German applied psychology during World War II. *The American Psychologist*, 1946, 1, 151–161.

test the effect of mild anoxia on visual illusions, since this condition was much easier to produce in the laboratory and control experimentally than was fatigue, and since the symptoms of anoxia in several respects resemble those of fatigue. It was thought that if clear-cut results were obtained for anoxia, the technique would merit further study in relation to other environmental influences.

EXPERIMENT 1.—THE EFFECT UPON FOUR VISUAL ILLUSIONS OF ANOXIA PRODUCED BY EXPOSURE TO A PRESSURE ALTITUDE OF 15,000 FEET

Purpose

The purpose of this experiment was to compare the magnitude of the illusory effects on four visual illusions under the following two conditions: (1) a pressure altitude of 15,000 feet with supplementary oxygen supplied by mask (no anoxia); and (2) a pressure altitude of 15,000 feet with subjects breathing through a mask but not receiving supplementary oxygen (anoxia).

Apparatus

The four visual illusions used in this experiment are shown in figure 19.1. These illusions were drawn on white cardboard, with the dimensions as shown in the figure, and placed in a stand on a table top. The angle of the stand was such that the illusory figures were approximately perpendicular to the line of sight, and the figures were viewed at arms length.

The first of these illusions (A in fig. 19.1) was the so-called Ehrenstein illusion in which a vertical and horizontal line were compared in length. In this apparatus the length of the vertical line was adjusted by the subject by means of a sliding card which extended beyond the top of a wooden frame surrounding the figure. The task of the subject was to adjust the vertical line until he judged it to be equal to the horizontal line in length. The second illusion (B in fig. 19.1), named for Poggendorf, required the subject to aline the two ends of an interrupted diagonal line. This alinement was accomplished by adjusting a slide which extended beyond the frame at the side. In both of these illusions a scale at the back indicated to the experimenter the magnitude of the illusory error in the subject's adjustment.

The remaining two illusions were of the reversible-figure type and were scored in terms of the number of reversals indicated by the subject who operated an electric counter by means of a hand switch.

A small altitude chamber was suitably instrumented for conducting this experiment. Two subjects' positions were provided at either side of a small table on which there were stands for supporting the illusions. The experimenter's station was at one end of the table and

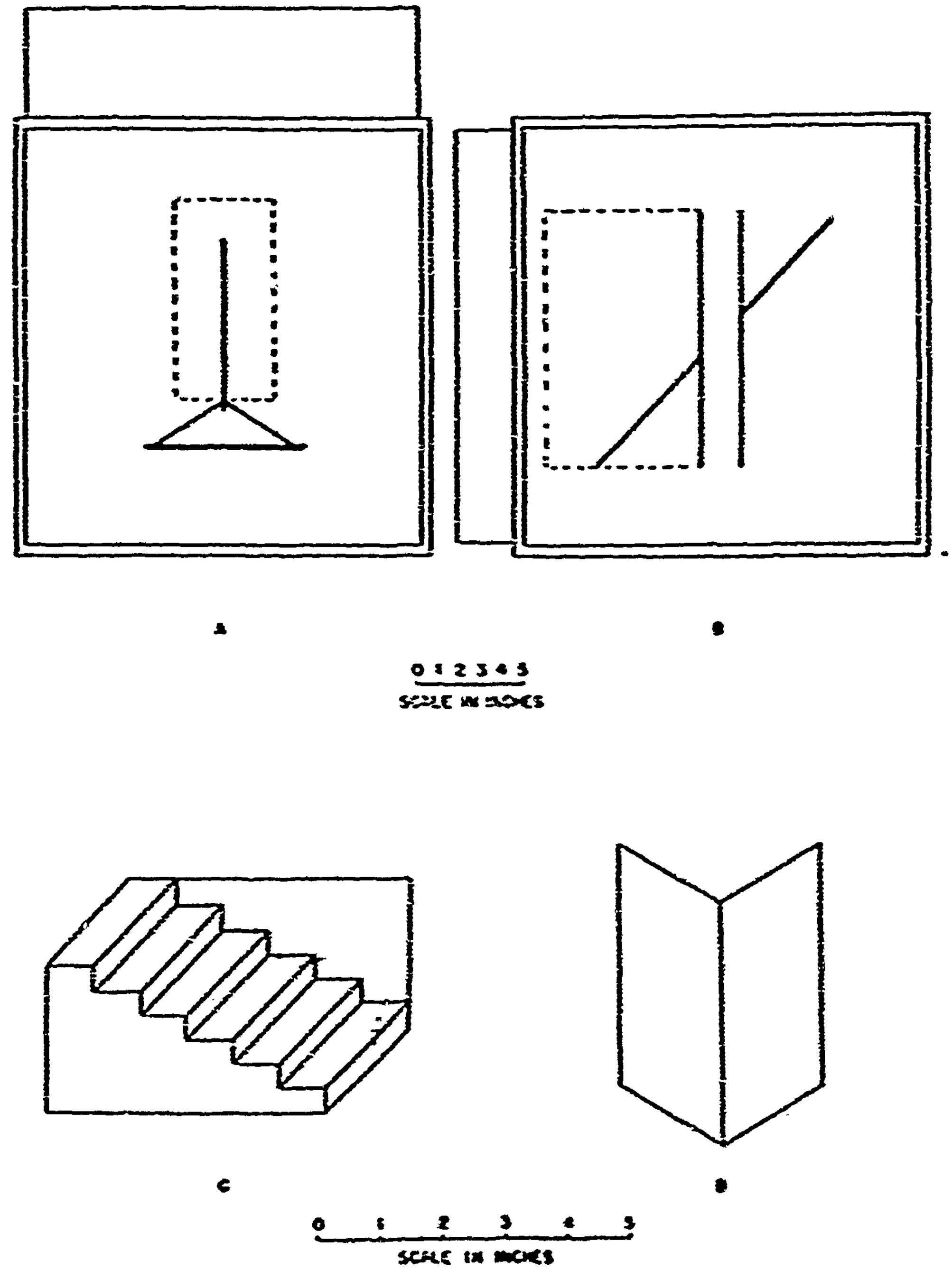


FIGURE 19.1.—Four illusory figures used in study of effect of anoxia upon visual illusions.

included a microphone for communicating with the subjects and with the chamber operator on the outside. The chamber operator also was provided with a microphone, and all four of the men, that is, the two subjects, the experimenter, and the chamber operator, wore headsets. Both the subjects and the experimenter wore demand-type oxygen masks, but a valve was provided by which the experimenter could cut off the supplementary oxygen supply to the subjects without their knowledge.

Procedure

Each experimental run was begun by fitting the subjects with oxygen masks, and then seating them at their stations in the altitude chamber. During the simulated ascent to 15,000 feet the subjects were given general instructions and were told in general terms the purpose of the experiment. All subjects were told that they would receive no oxygen, but that it was necessary for them to wear masks since other subjects would receive oxygen, and since it was necessary to keep the experimental conditions constant in other respects for all subjects. Only one of the subjects apparently suspected that he had received oxygen during part of the test run.

After the pressure altitude of 15,000 feet had been reached, a 10-minute adaptation period was introduced before administration of the illusions was begun, in order to insure a stable level of anoxia. Ten trials were given on each of the adjustable illusions, A and B, beginning alternately with the slide set to full deflection at either side of the equality point. The subjects were allowed to move the slides back and forth until apparent equality had been achieved. The reading on the scale at the back of the illusion was recorded for each trial. Each of the reversible illusions was viewed for 1½ minutes with the subjects indicating reversals by hand switches which operated counters. For figure C, the staircase illusion, the subjects were told to press the switch down and hold it down as long as the staircase appeared to be seen from below and to release the switch whenever the staircase appeared to be seen from above. For figure D, the book illusion, the subjects were instructed to push the switch down and hold it down as long as the back of the book appeared to be away from them with the pages open as if for reading. The switch was to be released whenever the back of the book appeared to be toward the subject. On the reversible illusions, therefore, the counter score indicated only one-half the total number of reversals.

During the first administration one-half of the subjects received no oxygen, and the remainder received oxygen. Immediately following completion of the tests the oxygen valve was either closed or opened, depending upon the conditions for the first test administration. Ten minutes of rest was then interposed to allow physiological adaptation to the changed condition, following which the administration of the tests was repeated.

In addition to the counterbalancing of the oxygen versus no-oxygen conditions, eight different test sequences were used to counterbalance for possible interference, learning, and fatigue effects, making a total of 16 different experimental sequences. Two subjects were tested in each sequence, making a total of 32 subjects, all of whom were adult males.

Results

The results of this experiment are presented in table 19.1 which gives the mean illusory effects, the standard deviations for the oxygen

TABLE 19.1.—*Summary of data on effect of anoxia produced by a pressure altitude of 15,000 feet upon four visual illusions*

[Experiment 1. N=32]

	Test conditions			
	Figure A		Figure B	
	Oxygen	No oxygen	Oxygen	No oxygen
Mean illusion (inches)	1.50	1.47	0.316	0.312
SD (inches)	.65	.53	.16	.23
r	+0.72		+0.32	
t	.31		.08	
	Figure C		Figure D	
	Oxygen	No oxygen	Oxygen	No oxygen
Mean illusion (reversals)	19.6	20.4	25.2	24.3
SD (reversals)	11.4	11.8	13.7	15.6
r	+0.75		+0.93	
t	.55		.78	

versus no-oxygen conditions, the coefficients of correlation, and the t value for the two sets of results. For figures A and B the mean illusion is given in inches. For figures C and D the mean illusion is given in number of reversals, which was double the counter readings. It will be noted that the results are virtually equal for the oxygen and no-oxygen conditions for all four of the illusions. Moreover, all of the t values are very small, indicating that the results have no statistical significance. The results indicate that the degree of anoxia that results at a pressure altitude of 15,000 feet has little, if any, effect on the magnitude of visual illusions.

EXPERIMENT 2.—THE EFFECT OF PROLONGED ANOXIA PRODUCED BY A PRESSURE ALTITUDE OF 10,000 FEET UPON TWO VISUAL ILLUSIONS

Purpose

The purpose of this experiment was to determine the effect upon the magnitude of two visual illusions of prolonged exposure to a pressure altitude of 10,000 feet without supplementary oxygen.

Apparatus and Procedure

An experiment which was being run by the Physiology Branch of the Aero Medical Laboratory provided an opportunity to measure the effect of prolonged mild anoxia on the magnitude of visual illusions. In this experiment, subjects remained in the altitude chamber at a pressure altitude of 10,000 feet for a period of approximately 8 hours without supplementary oxygen, and without wearing oxygen masks. Only the two adjustable illusions, A and B, used in experiment 1 were used in this experiment, and each test consisted of 20 instead of 10 trials. As in the first experiment the alternation in initial position of the slide at either side of the apparent equality point was used.

The illusions were administered to 10 subjects under 4 different conditions as follows: (1) at the beginning of the simulated flight, but after the 10,000 foot pressure altitude had been reached; (2) after 5 hours at altitude; (3) after 7 hours at altitude; and (4) at ground level pressure in the chamber on the morning of the day following the flight. Tests 1 and 4 may be considered as no-anoxia conditions, and tests 2 and 3 as mild-anoxia conditions. Under conditions 3 and 4, however, influences other than anoxia were probably also acting upon the subjects. The subjects had eaten no breakfast on the day of the test and were provided with one package of K ration for consumption at any time they chose during the chamber run. During the latter part of the flight most of the subjects were quite hungry and also somewhat bored with their experience. They had, however, not been required to perform physical exercise or work in the chamber and thus could hardly be described as physically fatigued.

Results

The results of experiment 2 are presented in table 19.2 for the four experimental conditions. It will be noted in the table that all of the differences between means are relatively small and lack statistical significance. We may conclude from these results that the conditions acting upon the subjects in this experiment had no effect upon the magnitude of the visual illusions used.

TABLE 19.2.—*Summary of data on effect of prolonged anoxia upon two visual illusions*

[EXPERIMENT 2. N=10]

	Test conditions					
	Figure A					
	1	2	3	4	1 and 4	2 and 3
Mean illusion (inches)	1.25	1.25	1.16	1.01	1.13	1.21
SD (inches)	.65	.64	.55	.72	.68	.65
r					+.98	
t					1.78	
	Figure B					
Mean illusion (inches)	0.44	0.43	0.40	0.42	0.43	0.41
SD (inches)	.15	.22	.19	.18	.12	.15
r					+.07	
					.51	

SUMMARY AND CONCLUSIONS

1. In experiment 1 a total of 32 subjects were tested on four visual illusions in the altitude chamber at a pressure altitude of 15,000 feet. Two of these illusions were geometrical illusions adjusted to the point of apparent equality by the subjects. The remaining two illusions were of the reversible figure type. A comparison was made between the illusory effects obtained when the subjects were receiving no supplementary oxygen (anoxia) and when the subjects were receiving supplementary oxygen (no anoxia).

2. In experiment 2 a total of 10 subjects were tested on only the first 2 (adjustable) illusions under 4 experimental conditions: (a) immediately after reaching a pressure altitude of 10,000 feet; (b) after 5 hours at 10,000 feet; (c) after 7 hours at 10,000 feet; and (d) on the following day at ground-level pressure.

3. The results of these experiments indicate that the illusory effects studied in this investigation are not influenced by mild anoxia. It is possible, of course, that the extent of the illusions might be influenced by more severe anoxia or by other environmental conditions.

CHAPTER TWENTY

Effect of Increased Positive Acceleration (G) on Ability to Read Aircraft Instrument Dials[1]

Melvin J. Warrick, Ralph E. Nelson, and Douglas W. Lund[2]

INTRODUCTION

When a force is applied to a body to accelerate it (change its speed or direction of motion), the body in turn, because of its inertia, exerts an opposite force. This force may be expressed as multiples of the acceleration of gravity—32.2 feet per second per second. Under the normal influence of gravity a body exerts a force of 1 G. A body accelerated at the rate of 64.4 feet per second per second exerts a force of 2 G, etc. This force may be in any direction with reference to the human body, being always equal and opposite to the accelerating force. Thus when a car is accelerated forward, its occupants experience a force backwards against the seats. When a car is turned to the left, the occupants exert a force to the right.

In aircraft, G forces are incurred with any change of speed or direction, particularly in flight maneuvers. In a turn of radius 1,000 yards at 375 mph, a force of slightly more than 3 G is produced (2). If a pilot is seated, as in a conventional aircraft, this force acts to "pull" the body and its components "downward" toward the feet (positive G). The limbs become virtually heavier and the blood tends to leave the head because of the inability of the heart to overcome the virtual increase in the weight of the blood.

The human centrifuge produces G forces in exactly the same manner as does the turning aircraft. A rotating boom carries a seat

[1] This chapter is based upon research findings reported in Headquarters AMC, Engineering Division, Memorandum Report TSEAA-694-10.

[2] The first author was responsible for the design of the experiment, for the equipment, for the procedures used, and for the writing of this report. The second author was responsible for the construction, installation, and maintenance of the apparatus, and for the test administration. The third author, Chief of the Acceleration Unit, contributed the use of the centrifuge and contributed to the writing of the introduction of this report. Ray Whitney, the centrifuge operator, by the generous contribution of his time and experience, materially facilitated the experiment.

carriage around in a horizontal circle. The carriage is free to pivot about a horizontal axis tangent to the circle of rotation of the boom. As the speed of rotation of the boom increases, the G force increases and the bottom of the carriage swings outward and upward toward the horizontal. Thus the G force is always directed toward the subject's feet.

It has been demonstrated (1, 2) on the human centrifuge that as G is increased, the subject may experience a dimming of a relatively bright light at about 3.5 G. As G is increased, peripheral vision is lost and at about 5 G foveal vision is lost and "blackout" occurs. At this point the subject is still conscious and able to respond to a gross auditory stimulus. If the G force is reduced, recovery from blackout takes place within a matter of seconds. At about 1 G above the blackout level unconsciousness occurs. Return to consciousness may take about 30 seconds, followed by a period of a minute or so of disorientation and confusion. It should be noted that the maximum effects of increased G are usually felt within 5 to 8 seconds after the onset of the force and that the body can tolerate rather large values of G if they are of short enough duration (3).

The techniques of locating the points where visual dimming, loss of peripheral vision, and blackout occur have been used to determine the effectiveness of various types of anti-G (4) suits which have been found quite effective in raising the G tolerance (2, 4).

Although it is obvious that under conditions of excess G a person's ability to pilot an aircraft would be seriously hampered, it has not been demonstrated that under relatively low G forces his ability is impaired. It was the purpose of this experiment to determine whether or not such a simple, practical function as reading aircraft instruments is impaired by relatively low G.

APPARATUS AND PROCEDURES

An instrument-reading test was assembled containing eight rows of the nine common aircraft instrument dials used in the AAF classification test, Table and Dial Reading Test, CP-622-A. The dials were mounted on a 16- by 18-inch piece of bristol board, four rows numbered 1 through 4 on one side and four rows numbered 5 through 8 on the reverse side (see fig. 20.1). Each row presented similar instruments in the same order but with different readings.

Immediately above each dial was presented a number which corresponded to or was markedly different from the reading of the dial. Thus the test was essentially a true-false test of whether or not the dial read the same as the number above it. This number was intentionally drawn quite large in hopes that it would remain clearly readable under conditions of moderate G. The test items were se-

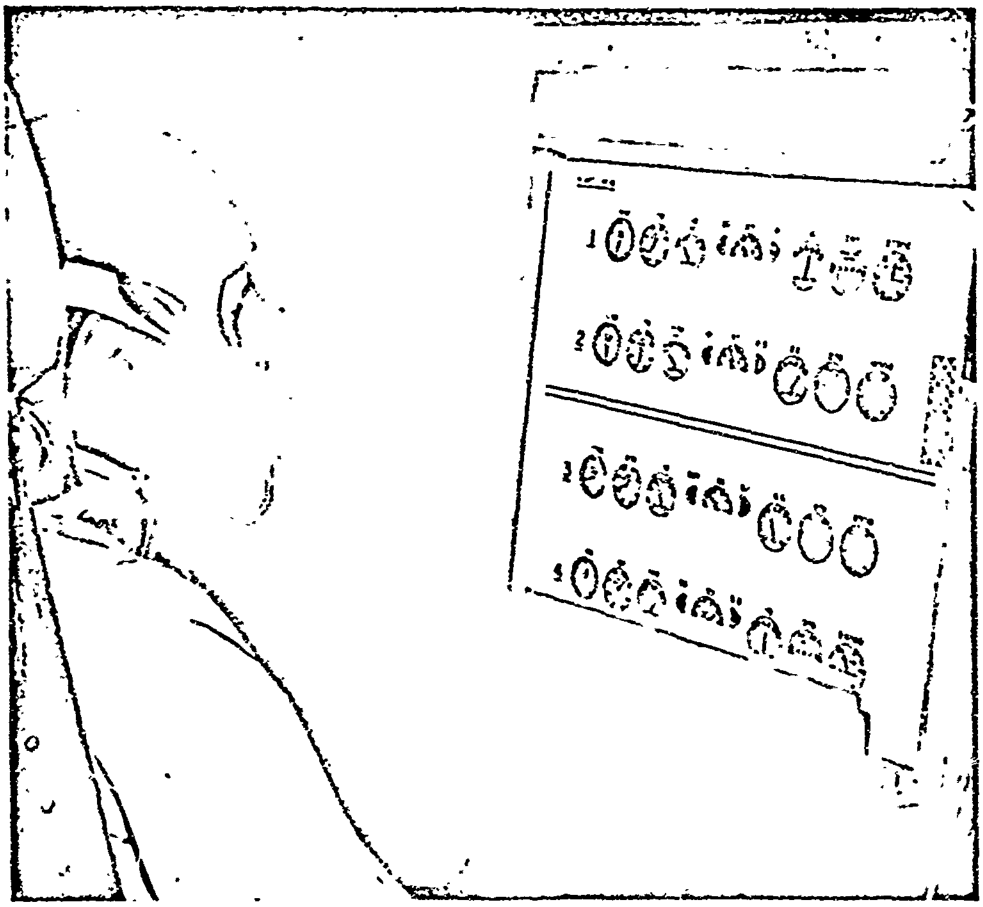


FIGURE 20.1.—Test apparatus used on the centrifuge to determine the effect of G on the ability to read aircraft-instrument dials.

lected with care to equalize the difficulty of reading the comparable instruments of each row and at the same time avoid a systematic pattern of right and wrong items. The misleads were chosen to avoid any such debatable misleads as an erroneous sign, a misplaced decimal point, or an incorrect number of zeros. If the answer was right it was as nearly correct as the authors could read the dials under ideal conditions. If the answer was wrong it was in error by at least one scale division.

The test panel was held on the centrifuge carriage in front of the subject by a special wooden frame which allowed the panel to be removed or turned over readily. The wooden frame also acted as a windshield partially protecting the subject from the wind blast encountered in rotation of the centrifuge and provided a mounting for the reading lights illuminating the test.

The reading lights were mounted at the top and bottom of the frame to provide maximum illumination of the test panel and minimum illumination of the surroundings. It was felt that extraneous illumination of either the walls or floor would be distracting to the subject and perhaps cause dizziness. Without visual cues the subject was not acutely conscious of his abnormal position, nearly horizontal, nor of his circular motion. The reading lights were controlled by the test administrator who during the test run was located in an adjacent control room. Accurate timing of the effective test runs was obtained by use of a synchronous motor timer which, under control of the operator, kept the reading lights on for exactly 30 seconds.

The subject was allowed to adjust the seat to bring the test panel holder into what he considered a comfortable reading position under the circumstances. Actually the subjects invariably left the seat in its normal position so that the test was about 18 inches from the eye, perpendicular to the line of vision, with the center slightly below eye level.

Since it was known that increased G interferes with motor activity, it was deemed necessary to provide as simple and convenient a method as possible for the subject to respond to avoid the possibility that difficulty in responding would bias the results. It seemed that a vocal response would be least likely to be affected by increased G and at the same time provide a convenient method for the administrator to record the responses accurately. An interphone system was installed and a microphone mask and head-set helmet provided for the subject. Corresponding equipment was provided for the administrator. The subject's microphone was always open so that he could communicate with the administrator at any time, while the administrator, in turn, could give the necessary directions to the subject at the appropriate times. The use of an interphone communication system was also of value in reducing the apprehension of the subject, for he knew that at any time he could request that the run be terminated.

Although some static was encountered in the interphone system incident to the rotation of the centrifuge, there were no reports by either subject or administrator of an inability to understand clearly any of the instructions or responses. The volume of the interphone amplifier was maintained at a rather abnormally, but not distractingly, high level to counteract the possibility that auditory acuity might be impaired under increased G.

As a safey precaution, an observer was stationed at the center of the centrifuge who could in an emergency stop the centrifuge. Between test runs, when the reading light was off, a small light behind the subject provided sufficient illumination for the observer to notice if the subject became sick or slumped forward because of unconsciousness. Such effects were not anticipated nor did they arise at the moderate G levels used.

The safety light served another very useful function during the interval between tests, when the reading light was off, by providing sufficient illumination for the subject to keep the test panel located and, at the proper time, turn it over. Care was taken in positioning the safety light to make certain that the illumination of the test dials was insufficient to allow a subject to read any items prematurely. It is also very likely that by providing this ambient lighting any effect that G might have on the ability to dark or light adapt was minimized.

It was quite obvious that, at best, there were many factors incident to riding the centrifuge other than a simple increase in G. Unless all these factors could be accurately controlled, equated, and measured, it would seem inadvisable to compare directly results obtained from the centrifuge with results from other test conditions. Thus, to obtain a best approximation of the results to be expected under normal G with all the distraction factors present, a very slow speed of rotation of the centrifuge was used as the control condition. Thus, strictly speaking, the question became that of determining the effect of increasing G on the centrifuge from a very low positive value to a moderate positive value.

It should be noted here that although the distraction factors incident to riding the centrifuge were somewhat equated, a few subjects reported that they noticed an increased wind against the eyes, that the static in the interphone increased, and that a slight momentary loss of orientation occurred at the higher G when the head was moved rapidly in shifting from the end of one row of test dials to the beginning of the next. However, it was believed that these factors were not potent enough to affect the results seriously.

From what was previously known about human behavior under increased G it was felt that 3 G was the maximum safe G level that could be tolerated without definite symptoms of visual dimming. An examination of unpublished records of 151 centrifuge subjects revealed that about 9 percent reported some dimming at or slightly below 3 G. One and one-half G (½ G above normal) was selected as the control condition primarily because of the ease with which it could be obtained and maintained. Thus the experimental comparison was made between the two G conditions, 1½ and 3 G.

In order to equalize the effect of any differences in difficulty of the test items and to counterbalance for learning, the subjects were divided into two groups of equal size, one group following a sequence of 1½, 3, 1½, 3 G, and the other group following the alternate sequence of 3, 1½, 3, 1½ G. Each group had the same test in the same order of test items. Thus the first group had at 3 G the same items that the second group had at 1½ G.

An optimum test interval of 30 seconds was established in preliminary investigations. No subject was able to complete two rows of instruments in that time interval; however, most subjects were able to complete one row. Furthermore, it was believed that surely within 30 seconds the maximum effect of increased G would be obtained, but that longer periods would not only make the administration more difficult but might place undue strain on the subjects. Since it took about 15 seconds to accelerate the centrifuge to the desired G value and to instruct the subjects between test runs, and about 30 seconds to turn the panel over, the subjects were under conditions of increased G for a maximum of about 4 minutes including starting and stopping. There is no evidence that physical harm would or did accrue from such G experiences.

Thirty-four rated military pilots were used in this study. The subjects had had no previous experiences on the centrifuge, but, of course, had experienced varying amounts of G in flight maneuvers. It was assumed that since the subjects were on flying status their physical conditions, particularly their eyes, met the physical standards established by the AAF. None of the subjects wore glasses during the test although the use of glasses or goggles might have eliminated the distraction, if any, of wind against the eyes.

The day prior to the test proper the subjects were individually indoctrinated concerning the centrifuge. Essentially the same material was covered as is discussed in the introduction to this report, with, however, a shift in emphasis considered desirable to minimize any apprehension incident to the forthcoming unique physical experience. The procedure for reading the dials and for responding was explained in detail. The subject was instructed to read silently the number appearing above the dial, then to read the dial, then determine whether or not the dial was reading the same as the number above it, and finally to respond verbally with the name of the instrument and either "true" or "false." The subject was instructed that he could use the common name identifying the instrument and not necessarily the name printed on the dial.

The subject was then given a sample test which was read and responded to in the prescribed manner. At this time the subject also was instructed how to turn the panel over so as to be able to read the items on the back of the panel. Mistakes, which were few, were explained and corrected, and the limits of accuracy required explained in detail. Following the indoctrination and the sample test, the subject was given a practice test session which duplicated in every way the final test. This was considered desirable since in such a very new experience the subject might be so distracted as to forget, or be confused about, the test procedure.

The practice test procedure exactly duplicated the procedure used with the test proper except that after the practice test any errors were clarified and in the case of a few individuals assistance was required during the test run. A description of the procedure followed in the test proper is given in the following paragraphs.

The subject was seated on the centrifuge and instructed to adjust the seat so that he would be able comfortably to read test material placed on the test panel holder. He was strapped loosely in the seat as a precautionary measure against the remote possibility that he might become unconscious and in slumping forward strike his head on the test panel. The safety light was adjusted to provide minimum illumination of the test dials. The subject put on and adjusted his helmet, containing the headset, and his mask, carrying the microphone. The mask was adjusted so as not to impede either mouth or nose breathing. The test administrator checked the interphone system to determine that it worked properly both ways.

The test procedure was then reviewed with the subject. He was given the sample test panel and again asked to read it, respond in the proper fashion, and turn it over in the prescribed manner. By the second day, during the test proper, the subjects seemed to understand the procedure thoroughly.

The reading light was then turned out, leaving only the safety light for illumination, and the test panel was placed in position. The administrator retired to his post and took up communication with the subject by interphone.

The subject was not informed of the sequence of G conditions he would experience. Actually the sequence was the same as that on the previous day's practice test. As the centrifuge accelerated to the operating speed, the administrator said, "When the light comes on start reading at line one; when you finish line one, read line two." Just before turning on the reading light and starting the 30-second test run, the administrator reiterated, "Start reading at line one."

These instructions were repeated for each of the four test runs in references to line one, three, five, and seven. At the end of the second 30-second run the centrifuge was maintained at or slowed to 1½ G and the subject was instructed in detail how to turn the test panel over. When he acknowledged that he had completed this operation the instructions were given concerning line five and the test continued. No subject evidenced any difficulty in following the instructions during the test proper.

The administrator recorded each response as correct or incorrect according to a predetermined correct key. Fortunately there were no omits. The subjects were not told their scores at the various G levels or how they compared with other subjects.

RESULTS

In table 20.1 are presented the means and standard deviations of the R–W (rights minus wrongs) scores, the number of errors, and the number of attempts at 1½ and 3 G. The significance of the difference within each pair is indicated by the t values. Those significant at or above the 1-percent level of confidence are so indicated.

TABLE 20.1.—*Effect of increased positive G on ability to read aircraft instruments dials*

[N=34]

	G force	Mean	SD	r	t[1]
R-W scores	1½	12.63	4.09	0.44	[2] 2.75
	3	10.50	4.13		
Errors	1½	3.47	1.69	.38	[2] 3.54
	3	4.71	2.06		
Attempts	1½	19.59	3.40	.88	1.36
	3	19.97	3.09		

[1] These data are based on the totals obtained from two 30-second tests per individual at each G level.
[2] Significant at the 1-percent level of confidence.

In computing these statistics the raw data from all the 1½ G runs for each individual were combined and all the raw data from the 3 G runs for each individual were combined. All 34 subjects were then considered as a group in which it was hoped that the influences of learning, order of tests, and difficulty of test items were equitably counterbalanced. The computations of the various sample parameters were then performed in the standard manner. Right-minus-wrong scores were used, not to correct for "guessing," but to provide a better index of the proportion of correct responses than would be obtained if a percentage figure were used since the number of attempts varied considerably.

Inspection of table 20.1 indicates that the subjects made significantly more errors in reading simulated aircraft dials under conditions of 3 G than under conditions of 1½ G. This finding is perhaps particularly significant since the task is a relatively simple one as compared to the task of manipulating an aircraft under conditions of increased G. How much the performance of other perceptual tasks would be impaired if the task were compounded or how well an individual would eventually adjust to increased G are problems for further study.

The subjects attempted to read slightly more dials at 3 G than at 1½ G. Although this difference is not statistically significant, it perhaps indicates that extraneous factors were not biasing the results or perhaps that the subjects' critical abilities were impaired to the extent that they did not recognize or were less concerned with wrong responses.

The data from the practice test were recorded and analyzed in the same manner as those of the test proper. Analysis of these data, although not conclusive, supports the results of the test proper.

It seems quite clear from the results that ability to read the dials was impaired; however, there is no evidence indicating why this phenomenon occurred. At least two hypotheses can be offered pending further investigations; vision may be impaired, or the interpretation of visual stimuli may be impaired. In any case the common factor is probably anoxia caused by a reduction in the blood supply to the eyes or to the brain.

SUMMARY AND CONCLUSIONS

1. The purpose of this study was to determine whether the ability to read aircraft instrument dials is impaired under conditions of moderately low G. Thirty-four rated military pilots were required to read printed simulated instrument dials under conditions of 1½ and 3 G as produced by the human centrifuge. It was found that the subjects made significantly more errors under conditions of 3 G than they did under conditions of 1½ G.

2. Since the ability to read simulated aircraft dials accurately was decreased under conditions of 3 G as compared to conditions of 1½ G, it is concluded that moderate G impairs the ability to read aircraft instruments.

REFERENCES

1. AAF Manual No. 25-2, *Physiology of Flight*, March, 1945.
2. AAF Technical Report No. 5433, *Design and Use of Anti-G Suits and Their Activating Values in World War II.* March, 1946 (restricted).
3. Armstrong, H. G., and Heim, J. W. Effect of acceleration on the living organism. *J. Avn. Med.*, December, 1938.
4. Maison, G. L., and Maaske, C. A., and others. Acceleration and the G suit. *Air Surgeon's Bull.*, 2, January, 1945.

CHAPTER TWENTY-ONE

Summary and Evaluation

PAUL M. FITTS

SUMMARY OF RESEARCH FINDINGS

The studies of human efficiency in operating equipment and the investigations of perceptual and motor capacities relating to equipment design, reported in this volume, have yielded findings of practical importance to psychologists and engineers concerned with the relationship of man and machine.

It has been possible to draw conclusions regarding several important visual display problems. The conditions under which quantitative data should be displayed in tabular or in graphic form have been specified; several factors have been shown to be important for the design of scales employed in measuring and plotting air navigation courses; a number of features have been found to contribute significantly to speed and accuracy of clock reading; the spacing of scale divisions and the size and number of intermediate numerical markings on instrument dials have been shown to be major determinants of reading accuracy; the direction and plane of movement of controls and indicators have been shown to be an important factor in determining human efficiency in initiating appropriate control movements; and the coordinate system and type of directional stabilization used with radar-scope presentations have been found to be significantly related to speed and accuracy of interpretation.

Capacities of individuals for discriminating pressure, shape, and location have been studied, and practical findings obtained that are of wide interest because of their significance for understanding the proprioceptive and tactual control of behavior and because of their implication for maximizing the cues available to the man who is operating machine controls.

Investigations of the efficiency of different types of adjustive reactions have shown that there are important differences between the accuracy of the hands and feet in controlling a visual indicator; that there are significant differences in the accuracy of control toward and away from the body as compared with left-right control movements; and that successive clockwise or counterclockwise adjustments are made with speed and accuracy superior to a series of alternating adjustments. In a following-pursuit task (tracking a target with a gun sight) it has been found that the shape and mode of action of hand controls has a very important influence on the accuracy of performance.

In two studies of the effects of unusual environmental conditions upon perceptual abilities, it was found that mild anoxia has little or no influence upon the extent of illusory visual judgments, and that the amount of positive acceleration is significantly related to the accuracy of dial reading.

FUTURE RESEARCH

Investigations of the kind reported in this volume will continue to be emphasized by the Psychology Branch. In the future, moreover, it is hoped to establish some of the theoretical principles underlying engineering psychology. The number of theoretical problems in this field is large. Information is needed, for example on such questions as how men orient themselves in three-dimensional space, what sensory cues are relied on in making precise connective and adjustive control movements, and how interpretation of single instrument displays is influenced by the total pattern of surrounding objects in the workplace.

During the coming year it is planned to undertake several studies at a more complex level than any carried out thus far. It is expected that equipment will be available in the laboratory for simulating with a high degree of validity some of the tasks of the pilot and of the radar operator and for obtaining detailed quantitative records of performance. It is also expected that an air-borne psychological laboratory will be available for obtaining reliable records of pilot reactions and flying proficiency.

A number of university psychology laboratories have initiated research on equipment-design problems under contract with the Air Matériel Command. These universities should make a significant contribution to engineering psychology. The research at universities will complement the work of the Psychology Branch, and make it possible to investigate certain problems more intensively and more systematically than is possible at Wright Field.

One university contract calls for systematic study of different types of visual and auditory warning devices. The effectiveness of different stimuli is being studied in a situation where the environment and the task given the subject are carefully controlled, and also in a Link trainer where the subject is engaged in a complex and changing problem. Two related university projects involve the study of factors influencing the ability of individuals to interpret dials, scales, and computers, and the ability to read various kinds of verbal materials, including numerals and letters, quickly and accurately. Factors such as style of numerals, illumination, noise, and vibration will be studied systematically. Another university laboratory is studying orientation problems in relation to the design of various types of equipment used by the pilot, radar operator, and navigator in keeping track of ground position, heading, and attitude. Two other university projects are concerned chiefly with control-design problems and questions of motor abilities. Design of controls for use in the prone and supine positions and study of the effect of general body position on motor abilities is the topic of one project. The other project concerns the design of controls to provide most effective tracking by the gunner, and the frequency analysis of tracking errors in relation to such factors as the mass of the instrument being positioned, angular speed of the target, control principle employed, and muscle groups used. Several additional university contracts are concerned with more specific problems.

It is believed that future research both at Wright Field and at universities will be characterized by increasingly efficient methodology as techniques of investigation in this new field are refined, as experimental variables come to be understood better, as questions of interaction effects, population differences, and learning are settled, and as better apparatus is developed for studying equipment design problems.

EVALUATION

The scope of engineering psychology appears broader, the psychological problems more numerous, and the applications more significant than when the research program was initiated a year ago. It has been found, for example, that such a familiar instrument as a clock is misread over 30 percent of the time when a poor dial design is used, and that the difference in number of errors when a good and a poor dial is used is almost tenfold. As another example, research data indicate that the efficiency of the defensive fire-control system of a large bomber—a system that has cost millions of dollars to produce—can be increased by approximately 25 percent simply by changing the shape and mode of action of the hand grips used by the gunner.

For the past 200 years man has been adjusting constantly to the new devices of a machine age. The importance of designing machines in relation to man's own capacities for perceiving and reacting is finally coming to be recognized. The application of psychological research techniques to this problem is an exciting challenge.

Since the scope of this new field of psychology is so broad, and the need for basic psychological information so widespread, it is hoped that in the future many psychologists will undertake research on equipment-design problems.

APPENDIX A

Names of Personnel Who Were Assigned to the Psychology Branch of the Aero Medical Laboratory at Some Time Between 1 September 1945 and 1 October 1946

Name	Military status at time of separation or as of Oct. 1, 1946	Months of service with the Psychology Branch		
		Military	Civilian	Total
Aborn, Murray	S/Sgt	1	0	1
Abrams, George D	Corp	7	0	7
Allenstein, Morton B	Sgt	2	0	2
Bakalas, Joseph	Sgt	0	5	5
Bedworth, Sally J		0	2	2
Bieler, Robert E	Corp	7	0	7
Boory, John F	S/Sgt	3	0	3
Brick, Jay R	Maj	5	0	5
Brown, Judson S	Maj	0	6	6
Brown, Walter T	Sgt	1	0	1
Carter, Launor F	Maj	0	7	7
Cerf, Arthur Z	S/Sgt	2	0	2
Christensen, Julien M	First Lt	6	3	9
Connell, Shirley C		0	2	2
Cowles, John T	First Lt	0	2	2
Cowles, Mary E	Sgt	0	7	7
Diggles, Sue S		0	8	8
Ducody, Michael M	Sgt	5	0	5
Eklof, Charles R	Corp	2	0	2
Fitts, Paul M	Lt. Col	11	2	13
Gagne, Robert M	First Lt	1	0	1
Glnn, Vernon H	First Lt	5	0	5
Grether, Walter F	Maj	6	7	13
Jenkins, William O		0	11	11
Johnson, Albert P	Maj	5	0	5
Jones, Richard E	Capt	8	0	8
Joseph, Robert T	S/Sgt	4	0	4
Korinek, George	Capt	5	0	5
Lehman, Thomas R	Sgt	4	0	4
Lyon, Wolcott N	Corp	4	0	4
McKewen, Peter S	Corp	7	0	7
Meshew, Robert L	Sgt	1	0	1
Milton, John L	First Lt	8	0	8
Muehlbauer, Helmut	Pvt	5	0	5
Murray, Norman L	Capt	6	0	6
Nelson, Ralph E	Capt	6	0	6
Neston, William	Sgt	1	0	1
Parish James W	Capt	3	0	3
Pepitone, Albert	Sgt	4	0	4
Robbins, Irving	S/Sgt	3	0	3
Roettele, Robert J	Sgt	4	0	4
Scardino, Robert	First Lt	1	0	1
Shakely, John A	Corp	1	0	1
Showalter, Ralph E	Second Lt	2	0	2

Name	Military status at time of separation or as of Oct. 1, 1946	Months of service with the Psychology Branch		
		Military	Civilian	Total
Smith, Raymond B.	Sgt.	4	0	4
Sowder, Wanda L.		0	5	5
Stauffer, Neil P.	Sgt.	3	0	3
Strong, Evelyn		0	5	5
Terry, Jeannette		0	5	5
Thomas, Frances		0	1	0
Van Saun, H. Richard	Capt.	0	7	7
Wagner, Jerome	Sgt.	1	0	1
Wall, Patricia A.		0	13	13
Warrick, Melvin J.	First Lt.	0	7	7
Webb, Wilse B.	First Lt.	1	0	1
Young, Katharine D.		0	8	8

Name Index

Subject Index