A TEST OF DECISION TIME: RELIABILITY AND "GENERALITY"

by

Leon Festinger
Seymour Wapner

A report on research conducted at the University of Rochester, Rochester, New York, by means of a grant-in-aid from the National Research Council Committee on Selection and Training of Aircraft Pilots from funds provided by the Civil Aeronautics Administration.

September 1945

CIVIL AERONAUTICS ADMINISTRATION
Division of Research
Report No. 48
Washington, D. C.

National Research Council

Committee on Selection and Training of Aircraft Pilots

Executive Subcommittee

M. S. Viteles, Chairman

E. C. Andrus	J. C. Flanagan
C. W. Bray	W. E. Kellum
D. R. Brimhall	W. R. Miles
J. W. Dunlap	P. J. Rulon

National Research Council

1945

LETTER OF TRANSMITTAL

NATIONAL RESEARCH COUNCIL

2101 Constitution Avenue, Washington, D. C.
Division of Anthropology and Psychology

Committee on Selection and Training of Aircraft Pilots

September 28, 1945

Dr. Dean R. Brimhall
Asst. to Administrator for Research
Civil Aeronautics Administration
Room 5835, Commerce Building
Washington 25, D. C.

Dear Dr. Brimhall:

Attached is a report entitled "A Test of Decision Time: Reliability and 'Generality'", by Leon Festinger and Seymour Wapner. This report has been approved by the Committee on Selection and Training of Aircraft Pilots for transmittal to the Civil Aeronautics Administration with the recommendation that it be published in the Technical Series, Division of Research, Civil Aeronautics Administration.

Of particular interest is the fact that the research described in this report bears simultaneously upon a basic problem of psychological theory and a problem of great practical interest in the selection and supervision of flight personnel. The development of this research project has been followed closely by representatives of the Army Air Forces who have also expressed an interest in additional work in the same research area.

Cordially yours,

Morris S. Viteles, Chairman
Committee on Selection and
Training of Aircraft Pilots
National Research Council

MSV:rm

Rec 10-11-45

EDITORIAL FOREWORD

The attached report grows out of an interest in determining the feasibility of predicting the success of a pilot in making prompt decisions in conflict situations.

The materials used in the study were taken from earlier research on two "personality tests," proposed for use in selecting candidates for flight training, conducted by the staff of the Statistical Unit, Committee on Selection and Training of Aircraft Pilots, located at the University of Rochester. The study was made possible through the cooperation of 51 male students in the Navy V-12 program and 75 female students at the University of Rochester who served as subjects. Dartmouth College contributed to the investigation through the loan of a timing apparatus. The research was conducted by Dr. Leon Festinger and Dr. Seymour Wapner.

Acknowledgments are made to Col. J. C. Flanagan, Army Air Forces; to Dr. P. J. Rulon, member of the Committee on Statistics; and to Dr. Raymond Franzen, Consultant to the Committee on Selection and Training of Aircraft Pilots, for guidance in the statistical treatment of the data.

CONTENTS

Aircraft pilots are frequently faced with the necessity of making decisions in conflict situations, especially when engaged in aerial combat. Techniques for quantitative evaluation of the difficulty with which trainees resolve conflict situations might therefore improve the selection of those individuals likely to make the best pilots as well as the best fighters.

The present report describes two studies concerned with this problem. Part I describes the development of a visual line test which provides a measure of the increase in decision time in a conflict situation. Part II describes a study of the "generality" of the decision-time measure, using the test discussed in Part I and three other tests.

In the visual line test the subject was shown two vertical lines and required to decide whether one line was longer or shorter than the other. Conflict between alternatives was provided by using stimuli which differed by a small amount or were equal in length.

The subjects in Part I consisted of 54 men in the Navy V-12 training program at the University of Rochester. Each subject was given 5 practice trials and a total of 40 comparison trials in the test situation. The time taken to respond and the kind of response for each of the 40 trials were recorded.

The measures obtained were: (1) the average of the decision times for discriminations based on comparisons of 5 1/2" and 6 1/2" variables with the 6" standard stimulus (base time), (2) the difference between the base time and the average of the decision times for the 6 1/8" comparison lines ($D_{6\ 1/8"}$), (3) the difference between the base time and the average of the decision times for the 5 7/8" comparison lines ($D_{5\ 7/8"}$), (4) the difference between the base time and the average of the decision times for the 6" comparison lines ($D_{6"}$), (5) the average of the measures 2, 3, 4 (D_A), (6) the "constant error" score based on the number of times "longer" was the response out of 10 equal comparisons ($\underline{CE}$), and (7) the errors for the trials in which there was an actual difference between comparison and standard stimuli (Number of Errors).

In the analysis of the findings uncorrected odd-even reliability coefficients were secured for base time and all D measures. The highest reliability (.93) was obtained for base time and the lowest (.66) for $D_{5\ 7/8"}$. Intercorrelations among all decision-time measures ranged from .70 to .96. The correlation coefficients for the five D measures with the constant error were all close to zero, while the relationships between the number of errors and each of the five D measures were low and negative.

Part II of the study was designed to study the "generality" of decision time, that is, to determine whether decision times were specific to the particular conflict situation or were indicative of some characterological

aspect of the individual. Four tests were selected to provide situations with various degrees of divergence in what appeared on a priori grounds to be relevent characteristics. These tests were: (1) the visual line test (described in Part I), (2) the tactual line test, (3) the tactual angle test, and (4) the word test.

These four tests, measuring decision time in conflict situations, were administered to 75 women students at the University of Rochester. Thirty trials were given for each of the tests, except for the word test in which 40 trials were administered.

Measures obtained from the test data were: (1) the difference between the average time to make a decision on difficult discriminations and the average time to make a decision on easy discriminations (decision time), (2) the average time taken to respond on difficult discriminations (conflict time), (3) the average time taken to respond on easy discriminations (base time), (4) the difference between conflict time and estimated base time on the regression line (decision-time residual), and (5) the errors occurring on all tests, excluding the word test (Number of Errors).

Two examiners were trained to administer the tests before the experiment was begun. The t values based on the differences between the mean decision times of each of the four tests for the two examiners were not significant, indicating the absence of examiner differences in the study.

Measures of test reliability were obtained by splitting the series of trials for each test into halves, and correlating the first half against the second half. Reliability coefficients for conflict time, decision time, decision-time residuals, and base time were calculated for each of the tests. The reliabilities for conflict time range from .84 to .93; those for decision time range from .73 to .86; those for decision-time residuals from .33 to .76; and for base time from .81 to .92.

In accordance with the original purpose of the study which was concerned with "generality" of decision time from one test to another, intercorrelations among the four tests were obtained for conflict time, decision time, decision-time residuals, conflict time holding base time constant by means of second order partials, and base time. The intercorrelations among base times and among conflict times are highest (ranging from .52 to .83); the intercorrelations among decision times are next highest (ranging from .52 to .69); the intercorrelations among residuals and among second order partials are lowest (ranging from .17 to .44). The small range of correlation coefficients within any of the correlation matrices appears to indicate that the degree of "generality" is about the same for the four situations within the limits of the experiment. Such findings do not suggest any decrease in "generality" as the pairs of situations become more and more divergent.

As pointed out in the study, it might be desirable to use two of the tests in combination, with resultant increase in reliability and "generality."

Coded scores on two of the tests (decision time for the visual line test and decision time for the tactual line test) were added and correlated against the sum of the coded scores for decision times from the tactual angle and word tests. This correlation was fairly high (.73), suggesting an increase over "generality" among single tests.

The outcome of this research has been the development of tests which appear to be fairly reliable indices of the way people react to conflict situations and which are suitable for actual field validation against criteria such as passing or failing flight training, success in aerial combat, and success as a bomber pilot or fighter pilot.

A TEST OF DECISION TIME: RELIABILITY AND "GENERALITY"

INTRODUCTION: HISTORICAL AND THEORETICAL

When flying an airplane the pilot is constantly faced with the necessity of making decisions between two or more possible courses of action. Most types of flying require the pilot to be able to make quick yet accurate decisions. A pilot who makes "snap" decisions may be undesirable if he lacks proper judgment. On the other hand, a pilot who is too cautious and cannot make decisions quickly enough may also be undesirable.

The first part of the present study is devoted to the construction of a test designed to measure the speed with which individuals make decisions in a conflict situation. The second part of the study explores the extent to which this characteristic of individuals is "general" from one situation to another.

Conflict and Decision Time. The whole area of conflict is a relatively new field of investigation for experimental psychology. Even of more recent origin is the realization that measures like those of VTE (vicarious trial and error), relative frequency of choice, and time to make a choice are related to the degree of conflict.

The earlier data on decision time come from experiments where the main emphasis was on psychophysical problems rather than on conflict. Thus experiments by Fernberger and his group[1] centered to a large extent around the problem of whether judgment time was longer when two, or three categories of judgment were allowed. That is, if judgments were allowed only in categories of, say, "longer" and "shorter," how did the judgment time in general compare with the judgment time if the additional response of "equal" was allowed?

Another problem concerned the effect on judgment time of increasing or decreasing the size of the equal category. If the subject were given instructions to the effect that he should not use the "equal" category unless absolutely necessary (thus restricting the size of the "equal" category), how would the judgment time compare with that when no instructions were given concerning the "equal" category, or an instruction was given to use the judgment of "equal" liberally?

Concomitant with the above problem was the question of how the judgment times in the different categories of response compared with one another. That is, were judgment times in the "equal" category higher or lower than judgment times in the other categories?

---

[1] Fernberger, S. W. The effect of the attitude of the subject upon the measure of sensitivity. Amer. J. Psychol., 1914, 25, 538-543.

Fernberger, S. W., Glass, E., Hoffman, I., and Willig, M. Judgment times of different psychophysical categories. J. exp. Psychol., 1934, 17, 286-293.

Fernberger, S. W., and Irwin, F. W. Time relations for different categories in judgment in the "absolute method" in psychophysics. Amer. J. Psychol., 1932, 44, 505-525.

The investigators in this field made no attempt to obtain the functional relationships between judgment time and magnitude of stimulus difference.

Kellogg[2] and George[3] contributed valuable experiments in this field and did make attempts to plot the curves of judgment time against magnitude of stimulus difference. Much controversy, however, existed with respect to the problems mentioned above.

Many more experiments have been done on a wide variety of problems involving the measurement of judgment time, reaction time, and latency time. These experiments, while undoubtedly very valuable in the field of prime interest to the experimenter, are relatively valueless for shedding light on decision time as related to conflict. The experiments mentioned here and in a few other places in the course of this introduction are the only ones that the writers feel have contributed to the understanding of decision time in its present context.

The first experiment on decision time in which its role as indicative of conflict was first clearly presented, was done by Cartwright.[4] Cartwright was endeavoring to settle the problems mentioned above, which were sources of open disagreement in psychophysics. Cartwright evolved a psychological theory of conflict from which he concluded that it was not the number of categories of response which was important for the judgment time. The crucial thing was, rather, how close any particular stimulus pair was to a boundary between different categories of response. Let us take as an example a situation where a subject can give a response of either "shorter," "equal," or "longer." As the magnitude of stimulus difference approaches the boundary between where the subject would respond "shorter" and where the subject would respond "equal," the decision time increases to a maximum. This boundary will also correspond to the 50% relative frequency point of judgments of "shorter." The same relationship will hold for the boundary between judgments of "equal" and judgment of "longer." In other words, these boundary points are points of maximum conflict between alternatives. Where only two categories of response are allowed, there is only one such maximum point.

This clear statement of conflict as being the determinant of increases and decreases in decision time was followed up by Barker,[5] who published

---

[2]Kellogg, W. N. The time of judgment in psychometric measures. Amer. J. Psychol., 1931, 42, 65-86.

[3]George, S. S. Attitude in relation to psychophysical judgment. Amer. J. Psychol., 1917, 28, 1-37.

[4]Cartwright, D. The relation of decision-time to the categories of response. Amer. J. Psychol., 1941, 54, 174-196.
Cartwright, D. Decision-time in relation to the differentiation of the phenomenal field. Psychol. Rev., 1941, 48, 425-442.

[5]Barker, R. B. An experimental study of the resolution of conflict by children. See Studies in personality, contributed in honor of Lewis M. Terman. New York: McGraw-Hill, 1942.

the results of an experiment in which decision time and a measurement of VTA were used to estimate the amount of conflict engendered in children by various types of choices.

Experiments on decision time or judgment time have not confined themselves to the perceptual psychophysical field. The experiment mentioned above by Barker dealt with conflict of choice between different liquids. Various liquids were arranged in order of preference by means of the paired comparison technique. Decision times could then be plotted against an abscissa of increasing difference in preference. The results showed that as the difference in preference became smaller the decision time increased.

Wells,[6] and Dashiell[7] did experiments involving esthetic judgments, also by the method of paired comparisons, in which judgment times were taken. Both of these experiments show that the decision times increased as the alternatives drew closer together in terms of preference.

The next step forward as regards the theory of decision time in its relation to conflict was made by Cartwright and Festinger.[8] This article elaborates a mathematically stated theory of the relationship between decision time, relative frequency of judgment and degree of conflict. Festinger[9] presents the results of experiments which support the predictions and derivations of the theory to a very large extent.

The main results from the previous experiments and their relations to the decision-time theory which are pertinent to our present work are as follows:

1. In any choice situation, as the relative frequency of choice approaches 50%, the decision time approaches a maximum.

2. The magnitude of the decision time will be a function of the restraint with which the individual approaches the choice situation. Thus, instructions which affect this restraint can alter the magnitude of the decision time.

---

[6] Wells, H. M. The phenomenology of acts of choice; an analysis of volitional consciousness. Brit. J. Psychol., Monogr. Suppl., 1927, 4, No. 11.

[7] Dashiell, J. F. Affective value-distance as a determinant of esthetic judgment-times. Amer. J. Psychol., 1937, 50, 57-67.

[8] Cartwright, D., and Festinger, L. A quantitative theory of decision, Psychol. Rev., 1943, 50, 595-621.

[9] Festinger, L. Studies in decision: I. Decision-time, relative frequency of judgment and subjective confidence as related to physical stimulus difference. J. exp. Psychol., 1943, 32, 291-306.
Festinger, L. Studies in decision: II. An empirical test of a quantitative theory of decision. J. exp. Psychol., 1943, 32, 411-423.

3. As the restraint with which the individual approaches the situation increases, and the decision time increases correspondingly, there will be a decrease in the number of errors made in the judgments.

All the above derivations and many more have been corroborated from the results of several experimenters.[10] The importance of decision time as a measure of conflict can be accepted.

The time it takes an individual to make a choice in a conflict situation is undoubtedly compounded of at least two things. This can best be made clear by indicating that even if there were no choice to be made, some time would still be consumed by the subject in perceiving, acting and speaking. The time taken to make a decision is made up, not only of the time it takes to resolve the conflict, but also in part of the time it takes the individual to act or react. In order to make the measure of decision time less dependent upon extra-conflict factors, the time to make a choice minus the action time or base time, will be used throughout the experiments.

PART I: CONSTRUCTION OF DECISION-TIME TEST

Apparatus and Procedure. In this test the subject was shown two vertical lines and was required to make a decision as to whether one line was longer or shorter than the other. Conflict between the alternatives was provided by using stimuli which differed by a small amount or were equal in length. Difficulty of decision was measured by the time taken to make a decision.

An exposure apparatus was designed which permitted easy adjustment of length of lines and accurate measurement of the time between presentation of the stimuli and the subject's response.

The exposure apparatus consisted of a black metal panel (see Fig. 1). Two vertical apertures 1/2" in width and 10 3/4" long were cut in the front of the apparatus. The size of the visible aperture could be changed from behind the apparatus by means of metal slides. Lowering the slides decreased the length of line. Raising the slides increased the length of line. The slides were calibrated to read off exact lengths of line to be exposed to the subject. The smallest possible aperture that could be made was 1 3/8" and the largest possible aperture was 10 3/4". In the actual test the shortest line used was 5 1/2" and the longest line was 6 1/2".

Behind the apertures there were white cardboard strips which fit closely to the front face of the apparatus so that the subject actually saw two vertical white lines of 1/2" width.

---

[10] Kellogg, W. N. Op. cit. (Referred to in Footnote 2.)
Johnson, D. M. Confidence and speed in the two-category judgment. Arch. Psychol., 1939, 241, 1-52.
Cartwright, D. Op. cit. (1st reference referred to in Footnote 4.)
Festinger, L. Op. cit. (1st reference referred to in Footnote 9.)
Festinger, L. Op. cit. (2nd reference referred to in Footnote 9.)

The apertures were kept closed from the subject's view between trials by means of the scissors-shutter shown in Figure 1, while any adjustments in length of lines to be exposed were being made by the experimenter. The shutter opened almost instantaneously by means of spring action.

The apparatus was wired to an electric time clock so that opening the shutter closed a mercury switch which started the clock. This clock was also wired to a voice key and holding relay so that the clock kept running until the subject's verbal response into a microphone activated a voice key, thus breaking the circuit. A diagram of the circuit is shown in Figure 2.

In the test situation the subject sat on a chair seven feet away from and directly in front of the exposure apparatus which was placed on a table 30 inches high. A microphone was adjusted in front of the subject's mouth so that it did not obstruct his view of the exposure apparatus (Figures 3 and 4).

After the microphone was adjusted and the subject seated in a position along the midline between the two stimuli, he was given the following instructions:

> "This is a test of visual acuity. You are to judge whether the line on your left is longer or shorter than the other. Sometimes the lines may appear nearly equal to you, but actually they are never equal, and so always make a judgment of 'longer' or 'shorter.'
>
> "Keep your mouth directly in front of the microphone, and speak in a loud clear voice. Keep your head in the same position. Before each trial I will say 'ready' so that you can get set for the discrimination."

Five practice trials were given. The first three were easy discriminations and the last two, difficult ones. Then the actual test started.

After each trial as soon as the subject responded the shutter was closed and the lengths of line reset for the next trial. A total of 40 comparisons were given to the subject. In each comparison one of the lines (standard stimulus) was 6" in length. Table 1 gives a summary of the various lengths of comparison lines used and the number of times each was presented.

TABLE 1

SUMMARY OF COMPARISONS PRESENTED IN DECISION-TIME TEST

Length of Comparison Line	Number of Times Presented
5 1/2 in.	5
6 1/2 in.	5
5 7/8 in.	10
6 1/8 in.	10
6 in.	10
Total	40

The order of presentation of stimuli was arranged to insure that a particular size of comparison line was on the left side half the time and on the right side half the time; that each group of comparison lines of given size was distributed through the 40 trials; and that there were no systematic groupings of the order in which one size of comparison line followed another on successive trials. The actual order or presentation is shown in the attached sample data sheet (Figure 5).

Subjects. The test was administered to 51 men in the Navy V-12 training program at the University of Rochester.

Measures Employed. The experimenter recorded the time taken to respond and the response (longer or shorter) for each of the forty trials. From these data, the following measures were obtained:

(a) Base Time. The decision times for the easy discriminations, i.e., those with the comparison line at 5 1/2" and 6 1/2", were averaged to give the base time.

This measure represents the time required to make a decision in this experimental situation when there is little or no conflict between the alternatives. Thus it constitutes a basic value from which the decision times in conflict situations should be measured.

(b) $D_{6\ 1/8"}$. This is the difference between the base time and the average of the decision times for the 6 1/8" comparison lines.

This and the next measure give the decision times for a moderately difficult discrimination which probably represents a fairly high degree of conflict. The base time is subtracted to yield a measure of the increase in decision-time due to the conflict.

(c) $D_{5\ 7/8"}$. Computed as in (b) for the 5 7/8" comparison lines.

(d) $D_{6"}$. Computed as above for 6 " comparison lines.

This measure represents decision time under a maximum degree of conflict for this situation. The decision time here should be higher than for any of the preceding cases.

(e) D_A. The average of measures (b), (c), and (d).

This represents a composite measure of the increase in decision time from no-conflict to conflict.

(f) CE. A "Constant Error Score" was derived from the number of times "longer" was the response in the 10 equal comparisons. Those subjects who gave 5 responses of "longer" out of the 10 equal comparisons were said to have a constant error of "0". Those with 4 or 6 "longer" responses were given a constant error score of "1." Those with 3 or 7 "longer" responses were given a constant error score of "2," and so on.

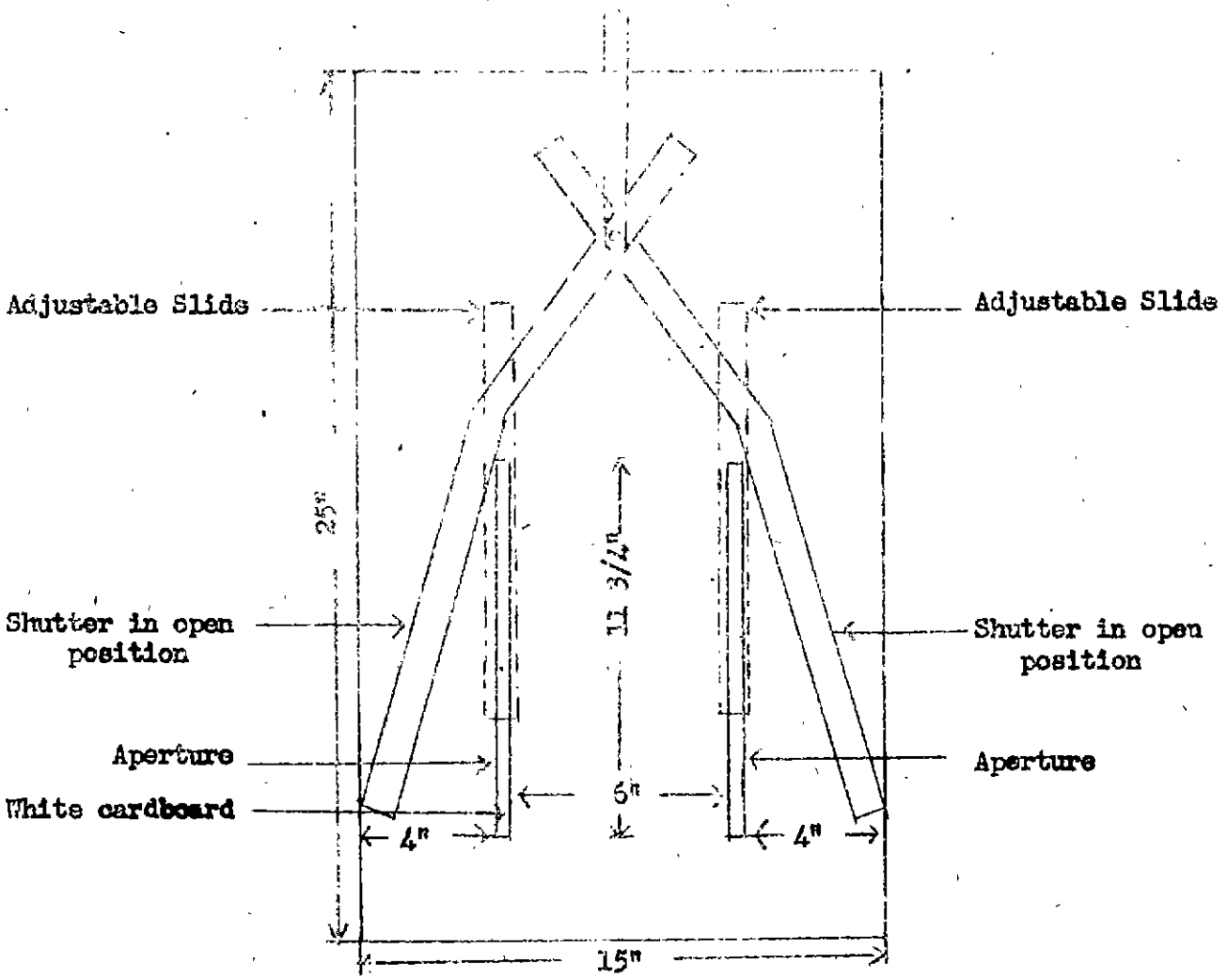


FIGURE 1

VISUAL LINE EXPOSURE APPARATUS

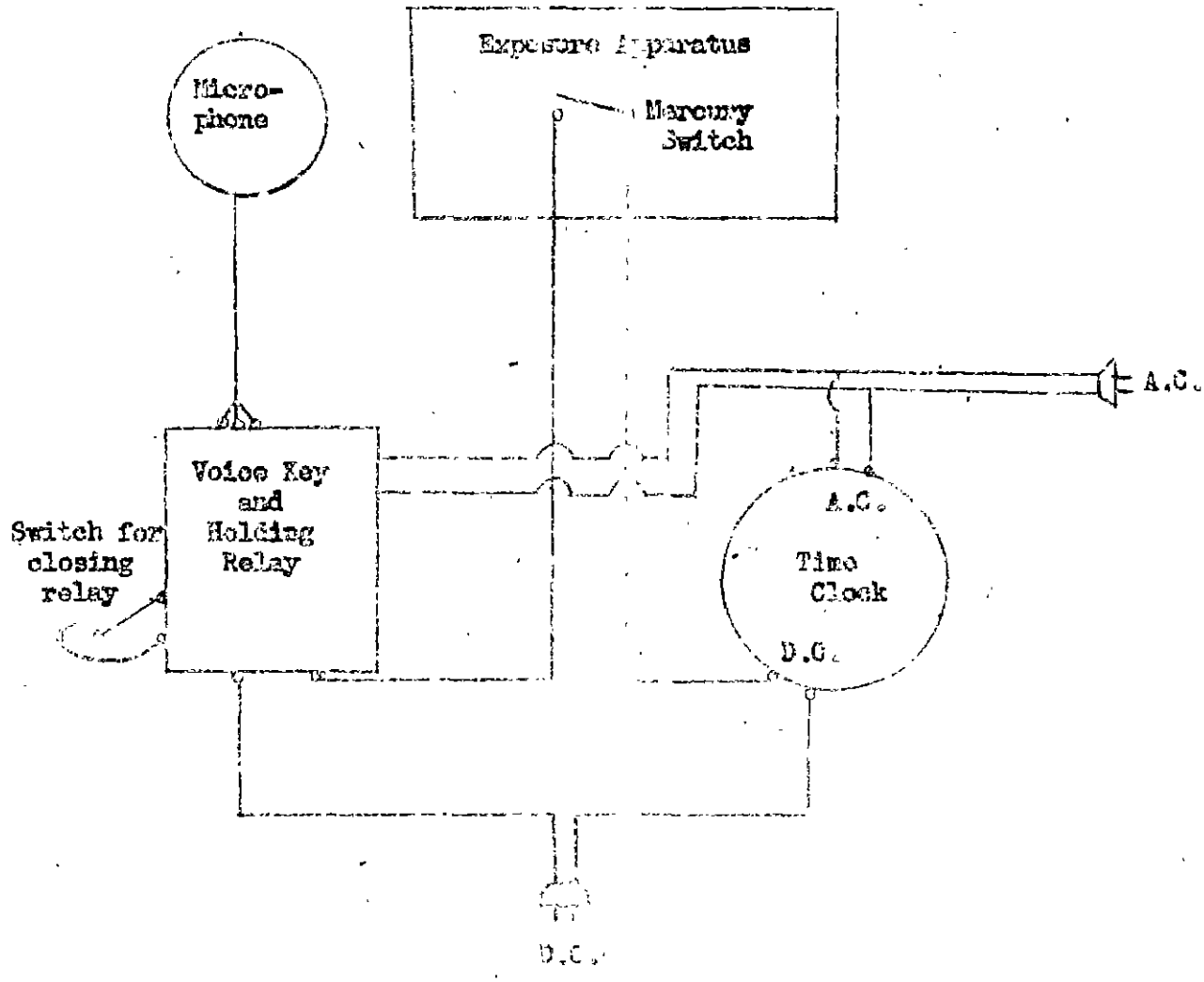


FIGURE 2

ELECTRICAL CIRCUIT FOR DISTINCTION-TIME APPARATUS

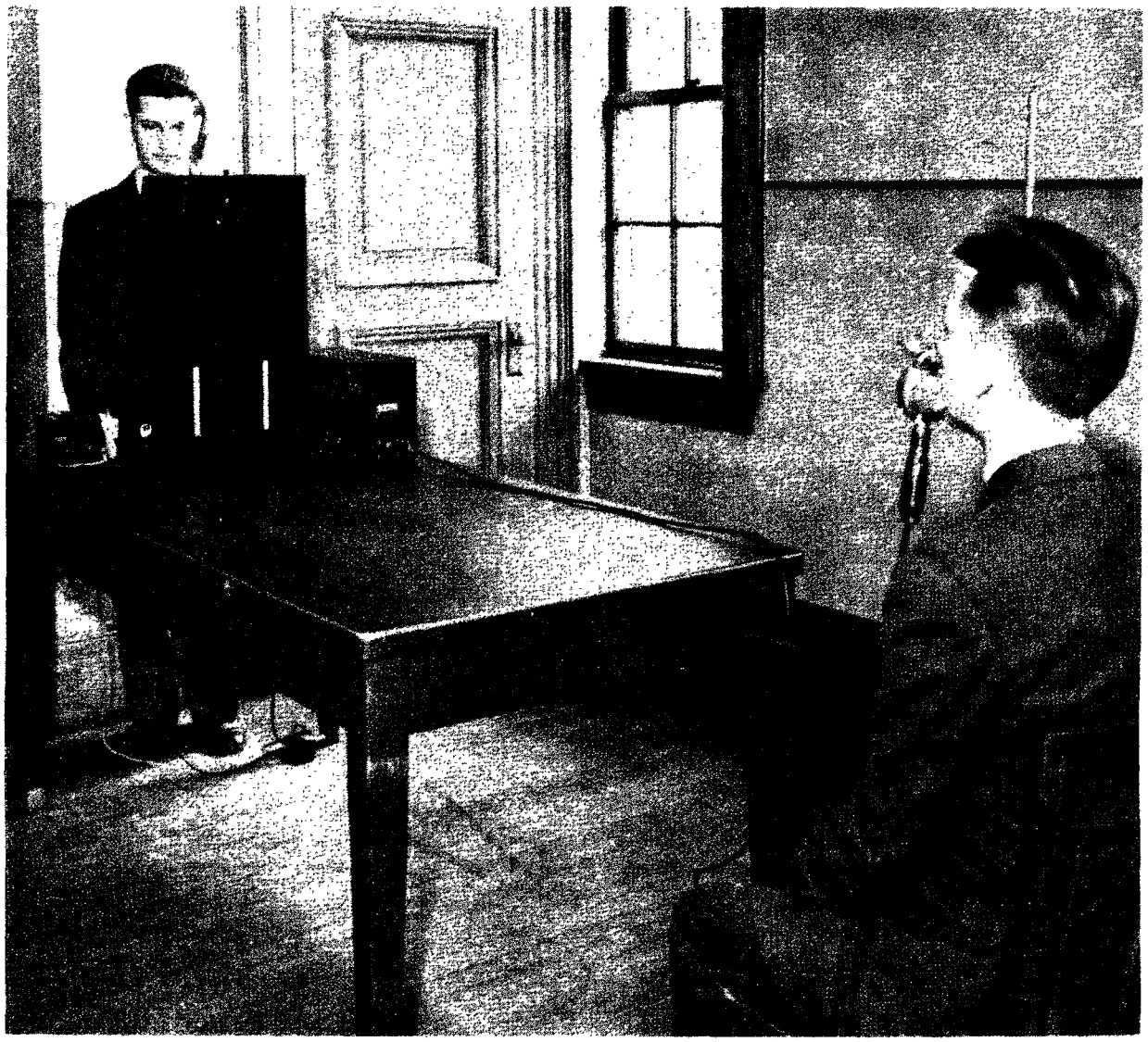


FIGURE 3

VISUAL LINE TEST

(FRONT VIEW)

FIGURE 4

VISUAL LINE APPARATUS
(REAR VIEW)

Stimulus Light	Stimulus Time	Decision Time	Response
L	5		
R	5 1/8		
R	5 1/2		
[illegible]	5 1/8		
[illegible]	5		
L	5 1/8		
[illegible]	5		
[illegible]	5 7/8		
[illegible]	5 1/2		
R	5		
L	5 1/2		
L	5 7/8		
R	6 1/8		
L	6 1/2		
L	6 1/8		
R	6		
L	5 7/8		
R	6 1/2		
R	5 1/2		
L	6 1/8		

Stimulus Light	Stimulus Time	Decision Time	Response
R	5 7/8		
L	6		
L	6 1/8		
L	5 7/8		
L	6		
R	6 1/8		
R	5 1/2		
L	6 1/8		
R	5 7/8		
L	5 1/2		
L	6 1/2		
R	6		
L	5 7/8		
R	6 1/8		
R	6		
L	6 1/8		
R	6 1/2		
L	5 7/8		
L	6		
R	6 1/8		

FIGURE 5

SAMPLE DECISION-TIME DATA SHEET

This measure was employed principally to estimate whether constant errors in the situation would be related to the decision-time results. If such were the case then care would have to be taken to minimize constant errors.

(g) Number of Errors. The number of errors for the 30 trials in which there was an actual difference between the stimuli was calculated.

In a previous study (Cartwright and Festinger,[11] Festinger[12]) it was predicted that the more quickly conflicts were resolved the greater would be the tendency to make incorrect choices. The number of errors was therefore recorded in the present study to see whether the same results would be forthcoming. The measure serves as a partial check on the meaningfulness of the decision-time measures.

Results. Odd-even reliability coefficients for the first five measures and correlations for some of the measures are shown in Table 2.

The uncorrected reliabilities for the base time and for all the D measures are high. The highest reliability (.93) is found for the base time; the lowest (.66) occurs for $D_{5\ 7/8"}$.

The correlations between the Number of Errors and the five measures of decision time are all negative although of not very great magnitude. This result was expected on theoretical grounds.

No relationship is apparent between any of the measures and indications of Constant Error. The importance of demonstrating that this measure is not related to any other measure of decision time or to the total number of errors made has already been indicated.

PART II: "GENERALITY" OF DECISION TIME

The Question of "Generality." Some understanding of the extent of "generality" of a measure is desirable when the measure presumably reflects a characterological aspect of the individual. The measure of decision time probably depends upon the cautiousness and restraint with which different individuals approach choice situations and the difficulty which different individuals have in resolving conflicts.

It may be regarded as axiomatic that, no matter what aspect of the individual we are measuring, as the situations in which we measure become more and more divergent (divergent in respect to relevant variables) the correlations between different situations will become smaller and smaller. At one end of this continuum of similarity between one situation and another is the case where we try to duplicate the same situation twice. We usually call

---

[11] Cartwright, D., and Festinger, L. Op. cit. (Referred to in Footnote 8.)

[12] Festinger, L. Op. cit. (2nd reference referred to in Footnote 9.)

TABLE 2

VISUAL LINE TEST

CORRELATIONS AND RELIABILITIES
N = 51

	Constant Error	Number of Errors	Mean	Sigma	Reliability
Base Time	.00	-.32	.69	.16	.93
D_6	-.20	-.34	.52	.63	.88
$D_{5\ 7/8}$	.01	-.25	.26	.22	.66
$D_{6\ 1/8}$	-.13	-.24	.25	.23	.74
D_A	-.15	-.33	.34	.33	.91
Constant Error	-	.04	2.3	1.5	-
Number of Errors			1.5	1.6	-

this "reliability." Theoretically, then, such a reliability coefficient gives us the "generality" of our test when the identical situation is duplicated. Of course, this "generality" that we obtain in this manner is dependent upon the precision of our measurement, the extent to which we have succeeded in duplicating the situation, and the extent to which the person is the same from one time to the next. Any "generality" coefficient obtained between two different situations will depend upon the extent to which the situations diverge in relevant characteristics and the extent to which the individual remains constant.

Let us go now to the other extreme of our "generality" continuum. It is certainly possible to find two situations so greatly divergent and dissimilar in very relevent characteristics that we would obtain no "generality" between the two situations. These situations could be dissimilar because of the instructions given in the situation, because of the different confidences with which the individual approached the two situations, because of the different characters of the two tasks, and many other things which might add to the picture.

Even something like I.Q. could probably be made to yield zero "generality" by changing the situation from a competitive to a noncompetitive one; from a friendly atmosphere to a nonfriendly atmosphere; from a highly motivated to a poorly motivated one. Variations in testing procedures from one time to the next would also act to decrease "generality."

The best one can do then in answering the question of "generality" is to restate the question so as to ask "What is the extent of the dropping off of our generality correlation coefficients as the situations become more and more divergent?" Even this question is to a large extent unanswerable because usually we have no *a priori* way of measuring the extent of divergence between our situations. We are thus driven to the expediency of selecting certain arbitrary situations which we suspect differ from each other in varying amounts and where we can arrange the differences between tests in an ordered series. We may then proceed to obtain our "generality" correlation coefficients and observe the degree to which our correlations decrease as the situations become more dissimilar. We may then draw conclusions in terms of our preconceived notions about the extent of the differences between the situations involved.

In the present experiment we have attempted to be somewhat systematic in our choice of tasks. We have selected a total of four tests so that there are two which are similar with regard to the nature of the judgment made (judgment of length of line visually and tactually); two which are similar with regard to the sense modality involved in making the judgment (judgment of length of line and judgment of steepness of slope, both tactually); and one which seemed widely different from any of the other three tests; namely, intellectual judgment concerning the meaning of words.

We thus have four tests which we may be pretty sure present some range of divergence. The absolute magnitude of our correlations among the tests and the rapidity or slowness with which the correlation decreases as our situations become more and more divergent will mainly determine what our conclusions are with respect to the "generality" of the decision-time measure.

The instructions given to the subject would undoubtedly also have some effect on the "generality." In all four of the situations our instructions have been the same. No mention has been made about either the speed with which the subject should react or the accuracy the subject should try to achieve. It is, of course, possible to insert into the instructions statements telling the subject to react as quickly as possible, or statements telling the subject that accuracy is by far the most important thing, or it is even possible to insert the somewhat equivocal instruction that speed and accuracy are equally important. We have chosen a type of instruction which fails to mention things like speed and accuracy so that the attitude which the subject assumes is left free. It was felt that the type of attitude which the subject sets for himself when he approaches such a situation is one of the important aspects of his behavior in conflict situations.

There are some data already in the literature on the extent to which changes in instructions or changes in the reality of the situation will affect the decision time. The data are by no means entirely adequate to answer the questions regarding relationship between different types of instructions, but they will at least be indicative of trends and suggestive of ideas.

Barker[13] presented individuals with choices in two different types of situations; a real situation and a hypothetical situation. In the real situation they were confronted with a choice between two liquids, and the one they chose they had to drink. In the hypothetical situation they were confronted with a choice of two liquids and were asked to choose the one they would drink if they had to drink one of them. In other words, between these two situations we have the difference between a real choice and a "make-believe" choice. As one might expect, the decision time increases from the hypothetical to the real choice. The correlation between decision-time in the real and in the hypothetical situations for 19 subjects is .69.

This correlation of .69 is the only data on hand which can give us some idea of the extent of "generality" which might obtain when the nature of the choice to be made remains exactly the same and the only thing which differs is the attitude of the subject toward the choice.

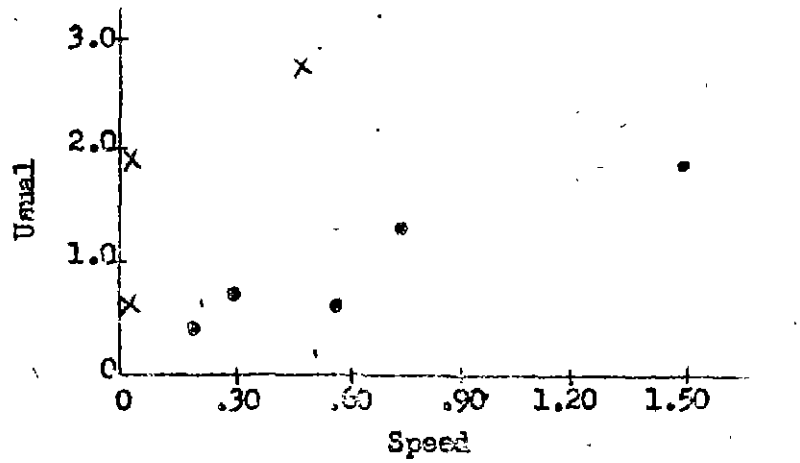
There are preliminary data[14] which suggest that correlations of like magnitude would result if we compared individuals under instructions where neither speed nor accuracy is emphasized (these instructions are referred to as "usual" instructions) with the same individuals under instructions which tell them that speed is the main thing. In an experiment by Festinger the same subjects were used under three different types of instructions, namely, instructions emphasizing speed, instructions emphasizing accuracy, and "usual" instructions.[15] Data are also available from Johnson who used his subjects under the same three conditions of instruction. Only five subjects were used in Festinger's experiment and three in Johnson's. The scatter-diagrams of the interrelations among the three conditions of instruction are presented in Figure 6 for the subjects from both experiments. The subjects from Festinger's experiment are represented by dots. The subjects from Johnson's experiment are represented by crosses. The actual data are presented in Table 3. The relationship between decision time under "usual" instructions and under "accuracy" instructions seems quite good. The other relationships seem relatively poor.

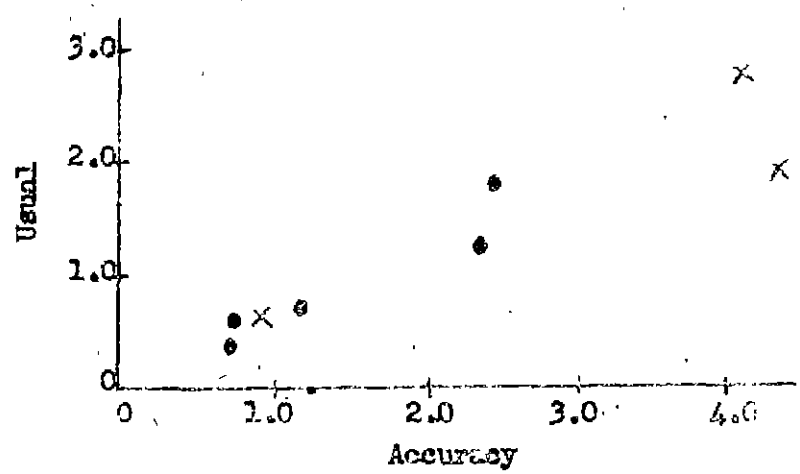
---

[13] Barker, R. B. Op. cit. (Referred to in Footnote 5.)

[14] Festinger, L. Op. cit. (1st reference referred to in Footnote 9.)
Johnson, D. M. Confidence and speed in the two-category judgment. Arch. Psychol., 1939, 241, 1-52.

[15] The specific instructions used by Festinger in the "usual" condition were: "When the room is darkened, you will see two vertical lines illuminated for a period of time. You are to make a judgment as to whether the line on the right is longer or shorter than the line on the left. If the line on the right is shorter, you will say 'shorter'; if the line on the right is longer, you will say 'longer.' Do not say any other words before the judgment of longer or shorter. When you give your judgment, enunciate the word clearly and distinctly and fairly loudly." Festinger, L. Ibid.





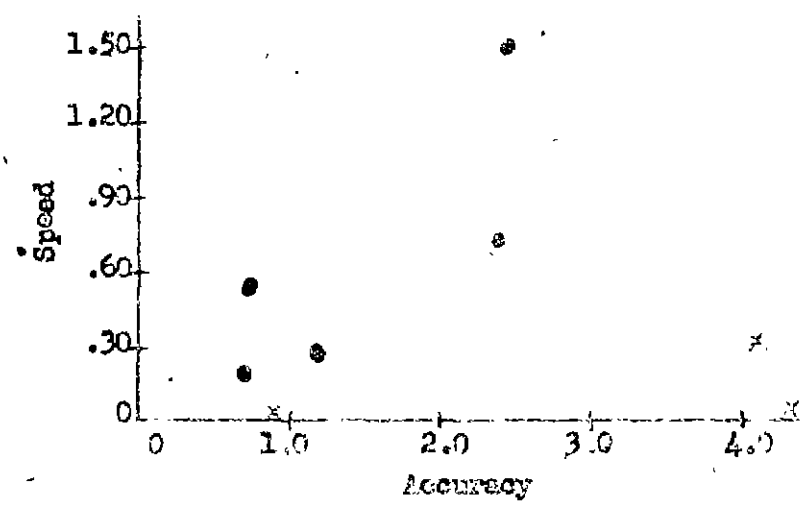


FIGURE 6

RELATIONSHIPS AMONG RECEIVER TIMES (IN SECONDS) USED

DECISION TIME UNDER THREE CONDITIONS OF INSTRUCTION

Subject	Usual	Accuracy	Speed	
H. S.	.64	.73	.56	Data from Festinger
B. W.	1.85	2.46	1.50	
G. S.	1.30	2.37	.74	
J. B.	.40	.74	.19	
N. P.	.69	1.14	.28	
K.	2.78	4.12	.47	Data from Johnson
V.	1.97	4.40	.01	
Z.	.65	.90	.02	

APPARATUS AND PROCEDURE

Apparatus for Visual Line Test. The procedure for the visual line test was exactly the same as described in Part I, except for changes in the number of trials, which will be described below.

Apparatus for Tactual Line Test. The object of this test was to have the subject, while blindfolded, compare lengths of lines. The comparisons were to be made by running the index fingers back and forth along the lines. Figure 7 shows a diagram of the apparatus used for presentation of the tactual lines. This apparatus consisted of a baseboard on which was mounted horizontally a long calibrated metal bar. A stationary vertical stop bar divided the horizontal bar into two equal parts. Variation in horizontal length of line could be accomplished by adjustment of two movable stops at either end of the horizontal bar. In front of the horizontal bar was a finger rest bar on which the subject rested his two index fingers between trials. When the subject was seated before the apparatus, he was given the following directions for the tactual line test:

> "In this experiment your task is to compare two lengths of line while you are blindfolded. You are to judge whether the horizontal line on your left is longer or shorter than the other. If the line on your left is longer, say 'longer.' If the line on your left is shorter, say 'shorter.' Sometimes these lines may appear nearly equal. Actually they are never equal so always make a judgment of 'shorter' or 'longer.' Be sure to speak in a loud clear voice so that you will activate the microphone. Always keep your fingers pressed down upon the finger rest bar in between trials. Before each trial, I will say 'ready' so that you can get set for the next discrimination. When I say 'go' place your index fingers on the horizontal lines. Make the judgment by moving your fingers back and forth along them, and then return your fingers to the rest bar. You will be given five practice trials before we begin the actual experiment."

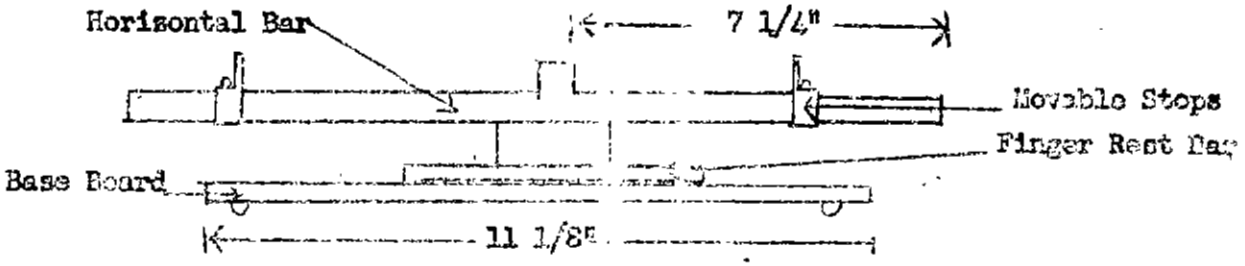


FIGURE 7

TACTUAL LINE EXPOSURE APPARATUS

Figure 8 shows a photograph of a subject in process of making the judgment on the tactual lines test.

Apparatus for Tactual Angle Test. Figure 9 shows a photograph of a subject in the process of making a judgment on the tactual angle test.

The apparatus consisted of a baseboard upon which were mounted two metal bars hinged at the bottom so that they could be rotated through any desired angle with the horizontal, as shown in Figure 10.

The procedure during this test was similar to the procedure during the tactual line test. The subject, again blindfolded, was required to compare the steepness of slope of the two lines. This was to be done by running the index fingers up and down the sloping bars. The directions given to the subject for this test were:

> "The procedure for this test is exactly the same as for the preceding one except that here you will be judging the steepness of slope. If the left line is steeper, say 'steeper.' If the left line is less steep, say 'less steep.' Be sure to speak in a loud voice. You will again be given five practice trials before we begin the actual experiment."

Apparatus for Word Test. In the word test the subject was required to judge which of two words best fitted a sentence or phrase. The apparatus used is shown in Figure 11. The permanently opened slit in the face of the apparatus was used for exposing the sentence or phrase. The sentences and words were printed on a long strip of white cloth attached to rollers in back of the apparatus so that changing the stimuli for the next trial was accomplished by winding the roller. The directions given to the subject were as follows:

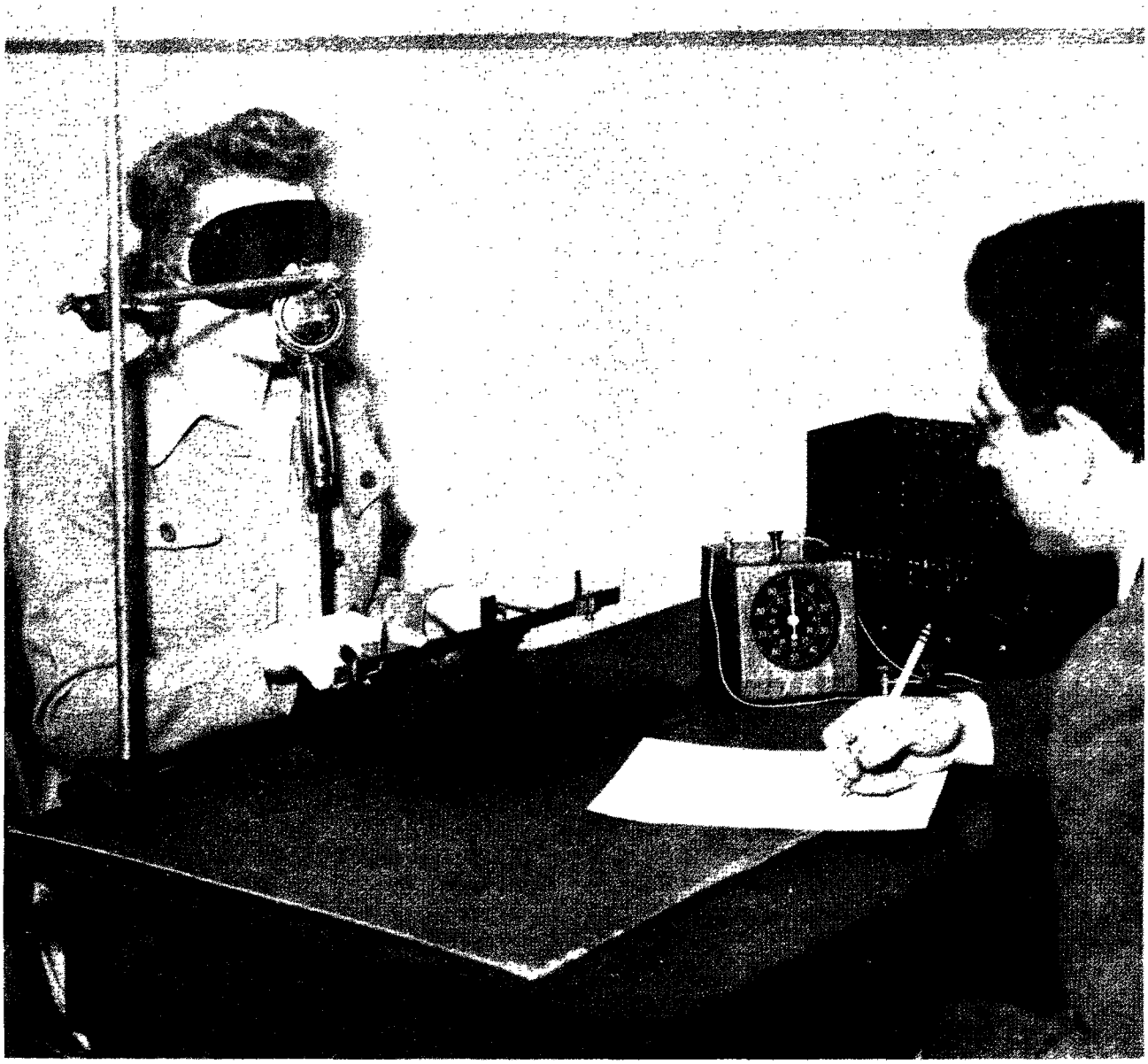


FIGURE 8

TACTUAL LINE APPARATUS

FIGURE 9

TACTUAL ANGLE TEST

"In this experiment I will expose a sentence or a phrase. After you have read the sentence or phrase you will be shown two words. You will be required to choose which of the two words fits the sentence or phrase more closely. You will do so by saying the word which fits more closely in a loud clear voice. The alternative words will not be presented until you give a signal that you are ready to see them. You will have three practice trials."

Figures 12 and 13 show photographs of a subject taking the word test.[16]

Timing Methods. Each apparatus was wired to an electric time clock so that closing a switch started the clock. This clock was also wired to a voice key and holding relay so that the clock kept running until the subject's verbal response into the microphone activated a voice key, thus breaking the circuit. A diagram of the electrical circuit for all four tests is shown in Figure 14. The electric clock was calibrated to one-hundredths of a second.

The method employed for stopping the clock was constant for all tests, namely, the response into the microphone. This action ended the decision period for a particular trial.

In each test the starting of the clock coincided approximately with the time when the subject started to perceive the stimuli. There were, however, some differences among the tests in the manner of starting the clock.

In the visual line test the clock was started by a mercury switch which tipped when the shutters were opened. In both the tactual line and the tactual angle tests, the clock started when the subject removed his fingers from the rest bar. In the word test, the clock was started when the shutter was opened.

Summary of Comparison Stimuli Used for Each Test. Thirty trials for each test were given in the visual line test and the two tactual tests. Table 4 presents, for the visual line test and the two tactual tests, the dimensions of the comparison stimuli and the number of times each such comparison stimulus was presented in the series of thirty trials.

The order of presentation of stimuli was arranged to insure that particular comparison stimuli were on the left side half the time and on the right side half the time. The trials for each magnitude of comparison stimulus were distributed equally throughout the 30 trials. The trials were arranged so as to control the order in which comparison stimuli followed one another. The actual order of presentation of the trials for the visual line and the two tactual tests is shown in Figure 15, which is a copy of the sheet on which the data were collected.

---

[16] If at any time during the experiment the subject raised any question concerning such matters as procedure, etc., the experimenter replied "It's up to you."

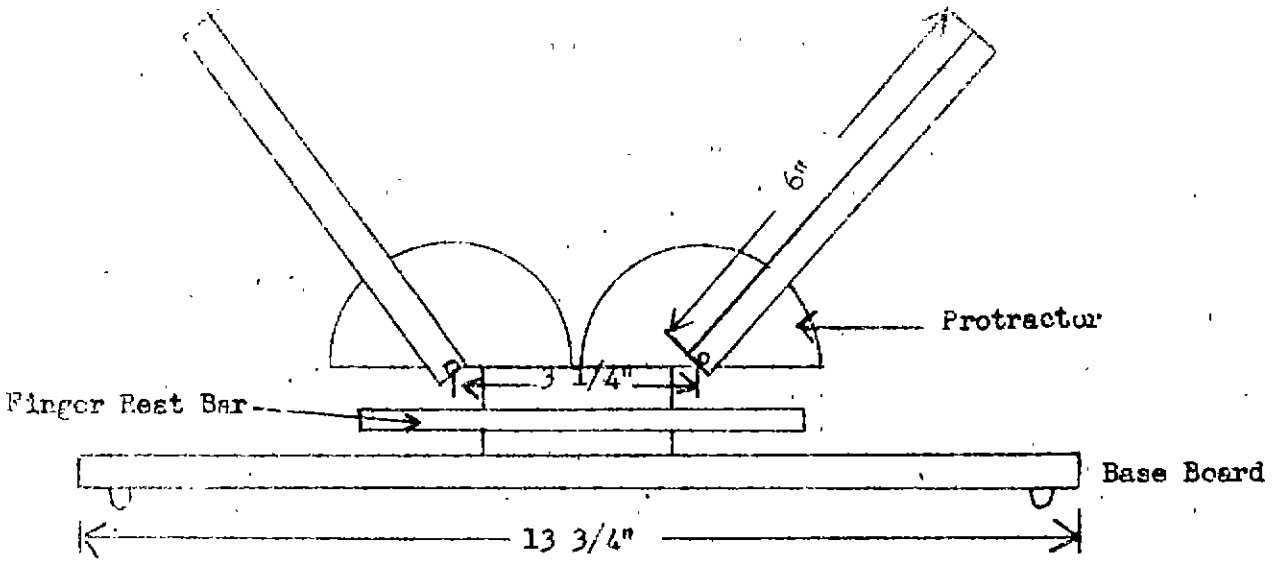


FIGURE 10

TACTUAL ANGLE EXPOSURE APPARATUS

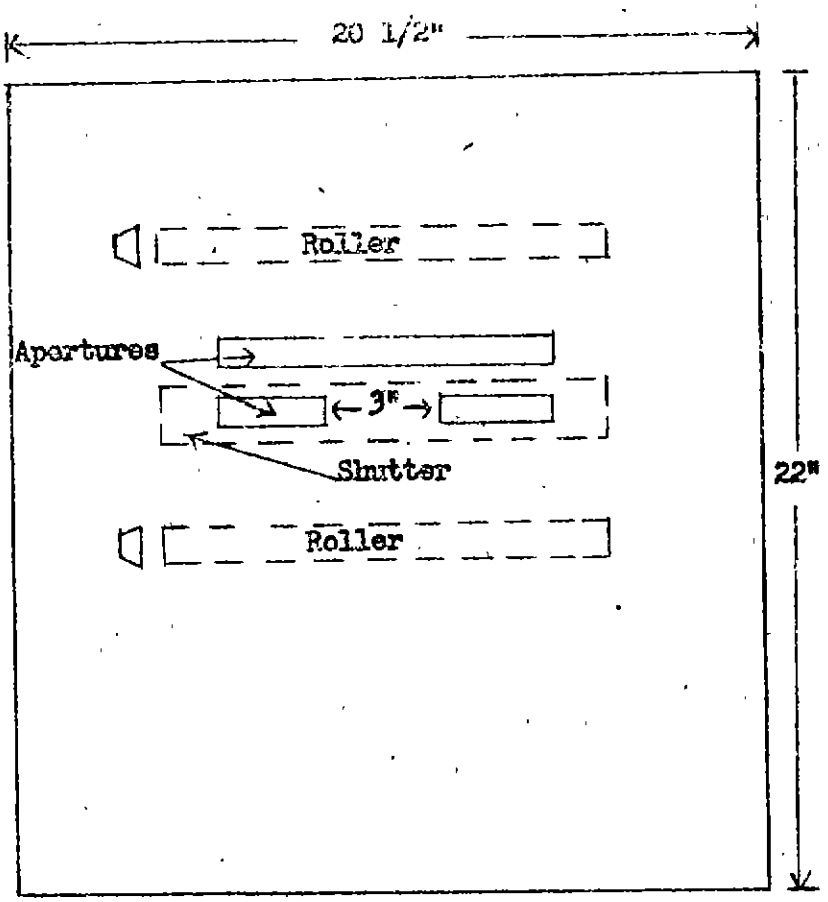


FIGURE 11

WORD TEST EXPOSURE APPARATUS

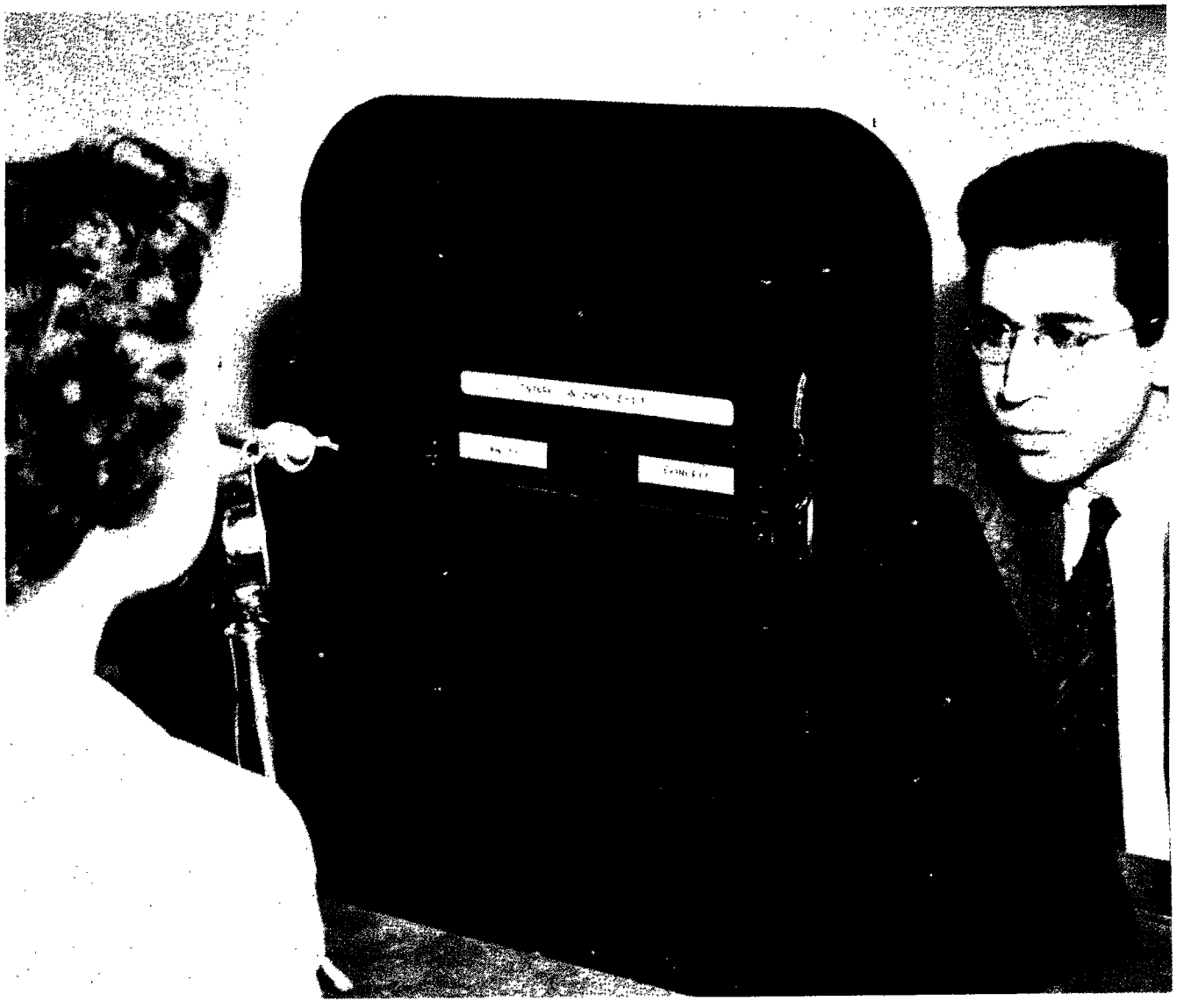


FIGURE 12

WORD TEST
(FRONT VIEW)

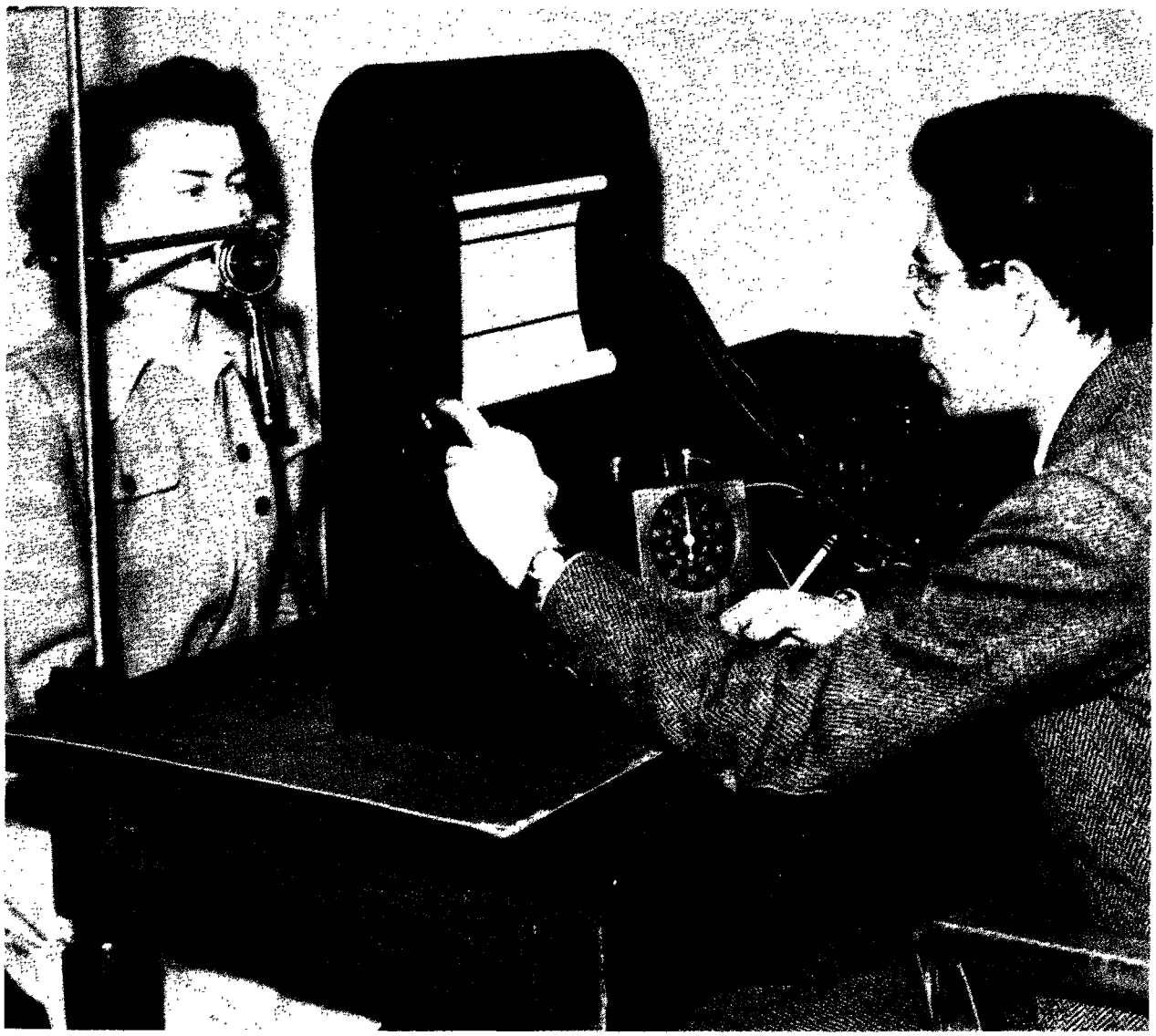


FIGURE 13

WORD TEST
(REAR VIEW)

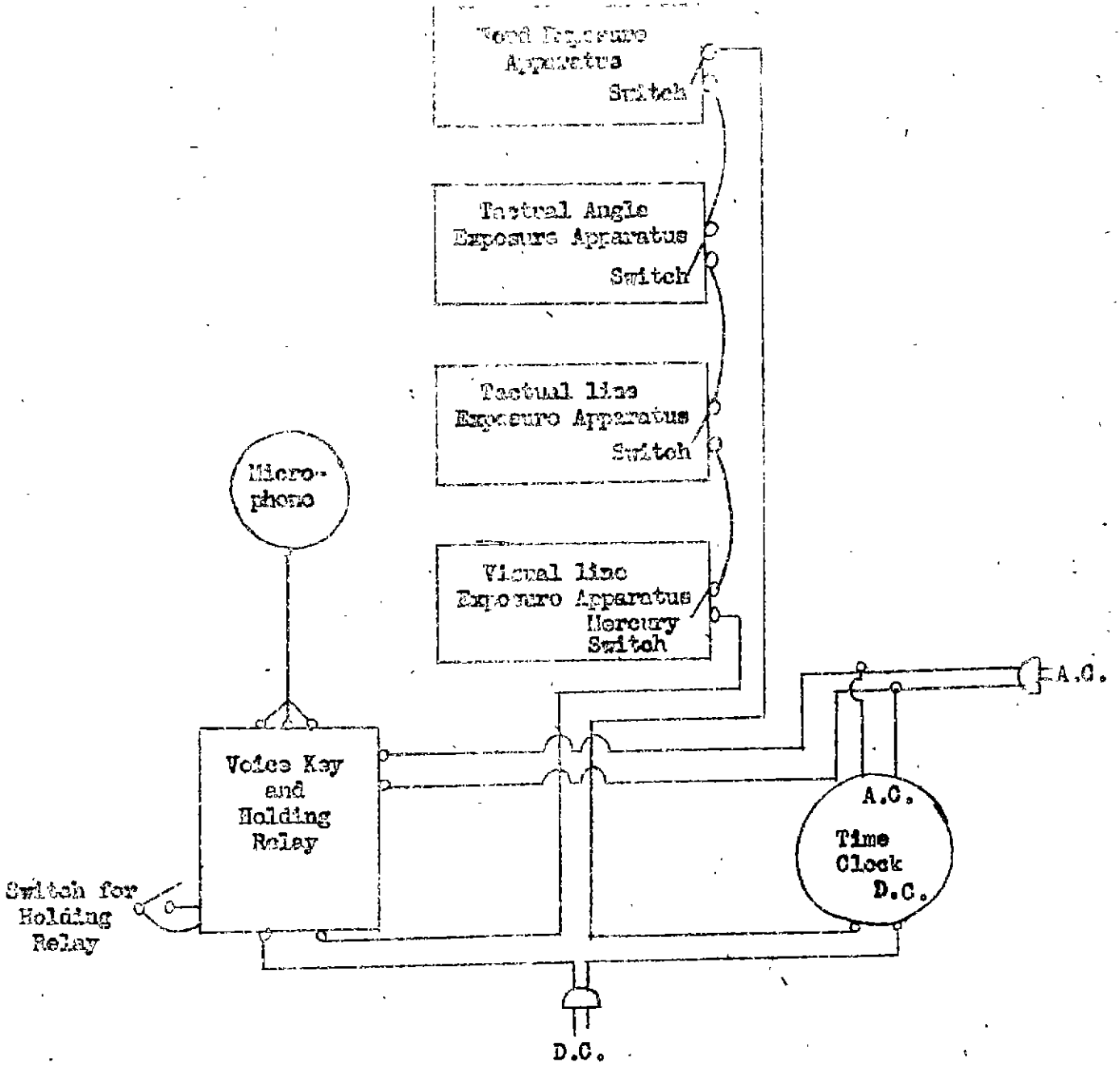


FIGURE 14

ELECTRICAL CIRCUIT

Subject:

Experimenter:

Date:

	Stim.	Resp.	D. T.		Stim.	Resp.	D. T.		Stim.	Resp.	D. T.
R*	6"			R	4 1/2"			R	140°		
L	5 7/8"			L	4 1/4"			L	145°		
L	6 1/2"			L	5 1/2"			L	120°		
R	6 1/8"			R	4 3/4"			R	135°		
L	6"			L	4 1/2"			L	140°		
R	5 1/2"			R	3 1/2"			R	160°		
R	5 7/8"			R	4 1/4"			R	145°		
L	6 1/8"			L	4 3/4"			L	135°		
R	6 1/2"			R	5 1/2"			R	120°		
L	6"			L	4 1/2"			L	140°		
L	6 1/8"			L	4 3/4"			L	135°		
R	5 7/8"			R	4 1/4"			R	145°		
L	6"			L	4 1/2"			L	140°		
R	5 1/2"			R	3 1/2"			R	160°		
R	6 1/2"			R	5 1/2"			R	120°		
L	5 7/8"			L	4 1/4"			L	145°		
R	6"			R	4 1/2"			R	140°		
L	6 1/2"			L	5 1/2"			L	120°		
L	6"			L	4 1/2"			L	140°		
R	6 1/8"			R	4 3/4"			R	135°		
L	5 1/2"			L	3 1/2"			L	160°		
R	6"			R	4 1/2"			R	140°		
R	5 7/8"			R	4 1/4"			R	145°		
L	6 1/8"			L	4 3/4"			L	135°		
R	6"			R	4 1/2"			R	140°		
L	6 1/2"			L	5 1/2"			L	120°		
L	5 1/2"			L	3 1/2"			L	160°		
R	6 1/8"			R	4 3/4"			R	135°		
L	5 7/8"			L	4 1/4"			L	145°		
R	5 1/2"			R	3 1/2"			R	160°		

*Position refers to position of standard for subject

FIGURE 15

DATA COLLECTION SHEET

TABLE 4

SUMMARY OF SIZES OF COMPARISON STIMULI FOR EACH TEST

	Number of Times Presented	Visual Line	Tactual Line	Tactual Angle
Easy	5	5 1/2"	3 1/2"	160°
	5	6 1/2"	5 1/2"	120°
Difficult	6	5 7/8"	4 1/4"	145°
	6	6 1/8"	4 3/4"	135°
	8	6"	4 1/2"	140°
Size of Standard		6"	4 1/2"	140°

In the word test, 40 trials were given, 10 of which were easy discriminations and 30 of which were difficult discriminations. In the difficult discriminations the two words were synonyms. In the easy discriminations the two words had quite different meanings. (These words were selected from a longer list after some preliminary work.) The list of phrases and words is given below in the order presented to the subjects. The easy discriminations are marked with an asterisk.

	Phrase	Words
1.	interest in one's self	vanity - conceit
2.	to depart in a hurry	escape - flee
3.	to spring forward	leap - jump
4.	an even surface	flat - level
5.	a property of a good cutting tool	sharp - keen
6.	to do something quickly	haste- hurry
7.	to talk to an audience	speech - lecture
8.*	to proceed leisurely	walk - run
9.	to keep out of view	conceal - hide
10.*	to be energetic	active - lazy
11.	a line of light	ray - beam
12.	moderate degree of heat	warm - tepid
13.	to set free from danger	save - rescue
14.*	piece of office furniture	oven - desk
15.	not easily moved	rigid - firm
16.	to plead	beg - implore
17.	something to be aware of	threat - menace
18.*	done for amusement	play - work
19.*	a connected thought	word - sentence
20.	to want	desire - wish
21.	unadorned	simple - plain
22.	abounding in money	wealthy - rich
23.*	a form of punctuation	period - paragraph
24.	a request which must be obeyed	command - order
25.	to commence	begin - start
26.	ungraceful in movement	awkward - clumsy
27.*	a resting position	sit - jump
28.	free from injury	safe - secure

	Phrase	Words
29.	to assist	help - aid
30.*	instrument played with a bow	piano - violin
31.*	something to write with	chair - pencil
32.	temporarily unwell	sick - ill
33.	a sudden fear	frightened - scared
34.	below	under - beneath
35.	a collection of houses	town - village
36.	early evening	twilight - dusk
37.*	season of warm weather	winter - summer
38.	a narrative	story - tale
39.	to give an account of	tell - relate
40.	a speech difficulty	stammer - stutter

The easy discriminations in the word test were distributed through the 40 trials, with the correct word occurring on each side equally often.

Occasionally during the administration of the tests, trials would have to be discarded because the subject's response was not loud enough to activate the microphone. In the visual line test, the tactual line test, and the tactual angle test such trials were repeated later in the series. In the word test such trials could of course not be repeated because the subject had already seen the words and would anticipate them in another presentation. For this reason 40 trials rather than 30 were used in the word test. The average number of trials omitter per subject on the word test was 3.55. The average number of trials omitted for the easy comparisons was .67. The average number of trials omitted for the difficult comparisons was 2.88. The frequency distribution of number of trials omitted is given below:

Number Omitted	f
12	1
11	0
10	0
9	1
8	7
7	4
6	5
5	5
4	10
3	9
2	11
1	17
0	5
Total	75

Subjects and Order of Presentation. The four tests were administered to 75 women students enrolled at the University of Rochester. The order of presentation of the four tests was the same for all subjects. The visual line test was presented first and was followed by the tactual line test, the tactual angle test and finally the word test. The entire procedure lasted about 1 hour and 15 minutes. The two authors administered all the tests.

MEASURES EMPLOYED

Decision Time. This is the principal measure derived from the data. This measure is the difference between the average time it takes to make a decision on the difficult discriminations and the average time it takes to make a decision on the easy discriminations. The quantity is positive when the difficult discriminations take longer than the easy discriminations. The decision time represents the increase in time which is presumably due to the presence of conflict. Thus, for each test we come out with a decision-time score representing the extent to which the time per trial has been prolonged when conflict is present between the alternatives.

Conflict Time. The conflict time is the average time taken to respond on the difficult discriminations. This measure represents the time to react plus the lateness in time due to conflict.

Base Time. The base time is the average time taken to respond on the easy discriminations.

This measure represents the time required to make a decision when there is little or no conflict between the alternatives. Thus it constitutes a basic value from which the decision time in conflict situations should be measured.

Decision-Time Residual. This measure is the difference between the conflict time and the estimated base time on the regression line. The effect of this is to make the measure completely independent of its own base time.

Number of Errors. This measure is merely the number of errors occurring on a test. It was obtained for only three tests since no errors could be counted for the word test. Since any response could be called correct for those trials in which the comparison stimulus was equal to the standard, the maximum possible numbers of errors is twenty-two.

In Part I it was found that the number of errors correlated negatively with decision time. These correlations were about the magnitude of -.3 (based on 51 cases).

This result corroborated one of the theoretical predictions mentioned in the introduction. The number of errors will be used here again to check on this relationship.

RESULTS

Examiner Differences. The two examiners who did all of the testing underwent a fairly prolonged period of training before the actual experiment started. The training was in the nature of discussions of the procedure, observation of each other administering the tests, and subsequent discussions. Any necessary changes in procedure were made during this preliminary period.

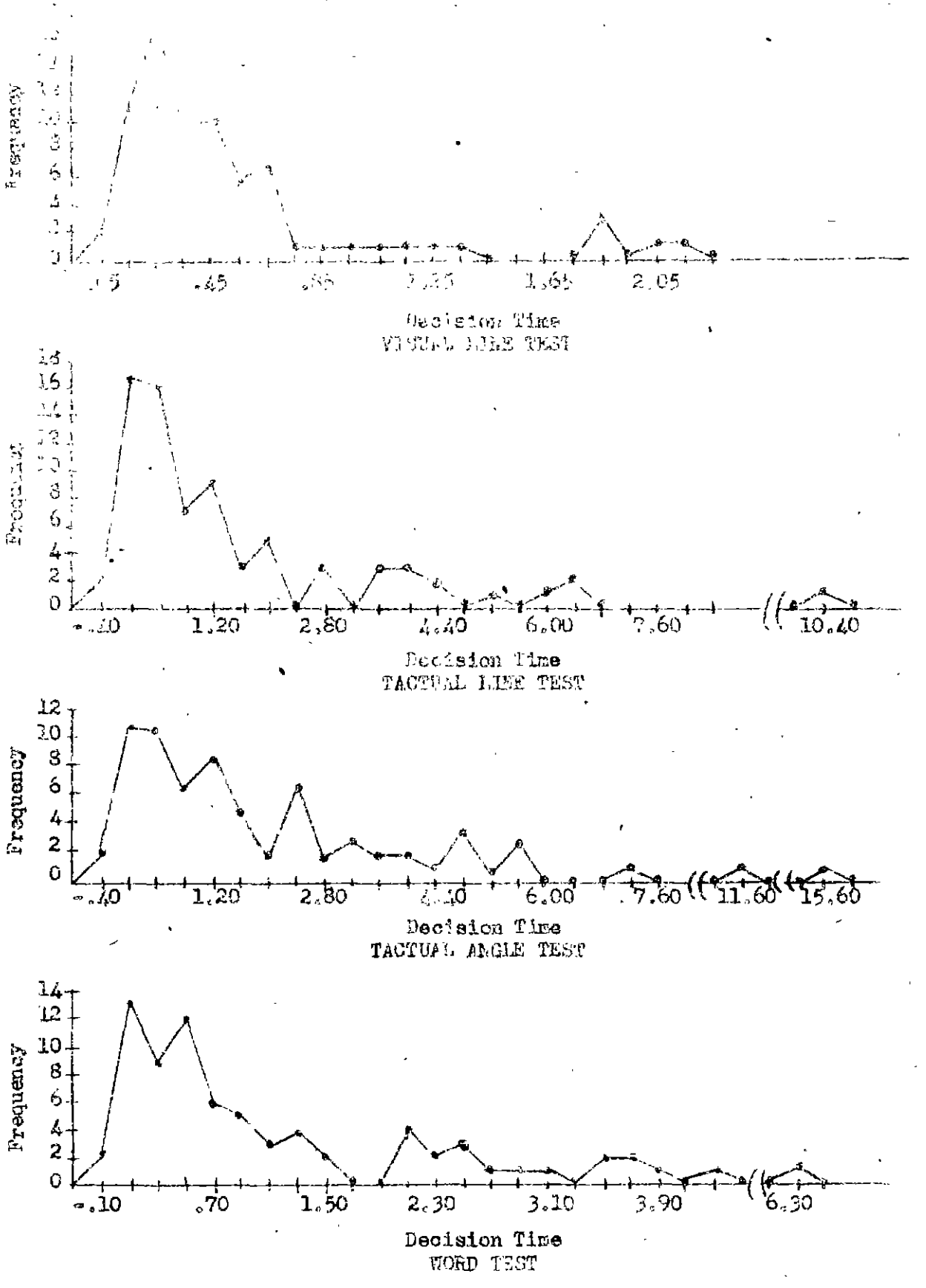


FIGURE 16

DISTRIBUTIONS OF DECISION-TIME MEASURES (IN SECONDS)

As a check [illegible] that differences between examiners would be found [illegible] of the fact that such examiner differences [illegible] studies, the results on examiner differences will be presented. Here differences are assessed only with respect to the measure of decision time. Table 5 shows the mean decision time for the subjects tested by each examiner on each of the 4 tests, the difference between the means for each of the 4 tests, the t value for each of the 4 differences, and the P value for each of the 4 t tests. Each t test is for 73 degrees of freedom. In no case does a t value even approach significance. The lowest P value was .23, which would mean that a t of this magnitude or greater would be obtained by chance about once out of every four times. We may certainly conclude from this analysis that there are no examiner differences in the present study. Consequently, the subsequent data will be presented for the group as a whole.

Distribution of Decision-Time Measures. Figure 16 shows the frequency distributions of the decision-time measures for each of the 4 tests.[17] It may be noted that in each case the distribution is markedly skewed. There is, in general, a piling up of scores toward the low end of the distribution and then a gradual tailing out to higher scores.

TABLE 5

EXAMINER DIFFERENCES ON DECISION-TIME MEASURES

	N	Visual Line	Tactual Line	Tactual Angle	Words
Mean Examiner 1*	39	.46	1.73	1.99	1.02
Mean Examiner 2*	36	.59	1.17	2.26	1.31
Difference in Means		-.13	.56	-.27	-.29
t		1.192	1.133	.428	1.000
P		.23	.26	.67	.32

*Examiner 1 = S. W.
Examiner 2 = L. F.

[17] The time intervals along the abscissas are not uniform for all curves, having been selected to facilitate comparison of the shapes of the curves. It will be noted that negative scores occur on three of the tests.

Number of Errors. On three of the tests, visual line, tactual line, and tactual angle, it was possible to count up the number of errors the subject made. It will be remembered that we were able to predict theoretically that greater decision time should be accompanied by fewer errors. The correlation between number of errors and decision time was -.22 for visual line test, -.20 for the tactual line test, and -.11 for the tactual angle test. None of these correlations is high, although the one for the visual line test just reaches the 5% level of significance. The fact that negative correlations occur on each of the three tests, and that a negative correlation of about -.3 was found in Part I seems to be fairly strong evidence that a negative relationship really exists between magnitude of decision time and number of errors.

It is interesting to note that no such relationship exists between number of errors and base time. The correlations are .01 for visual line test, -.07 for tactual line test, and .02 for tactual angle test. In the previous study mentioned above a correlation of -.3 was obtained for the visual line test. The best conclusion one can make from the present data is that there is no relationship between number of errors and the magnitude of base time.

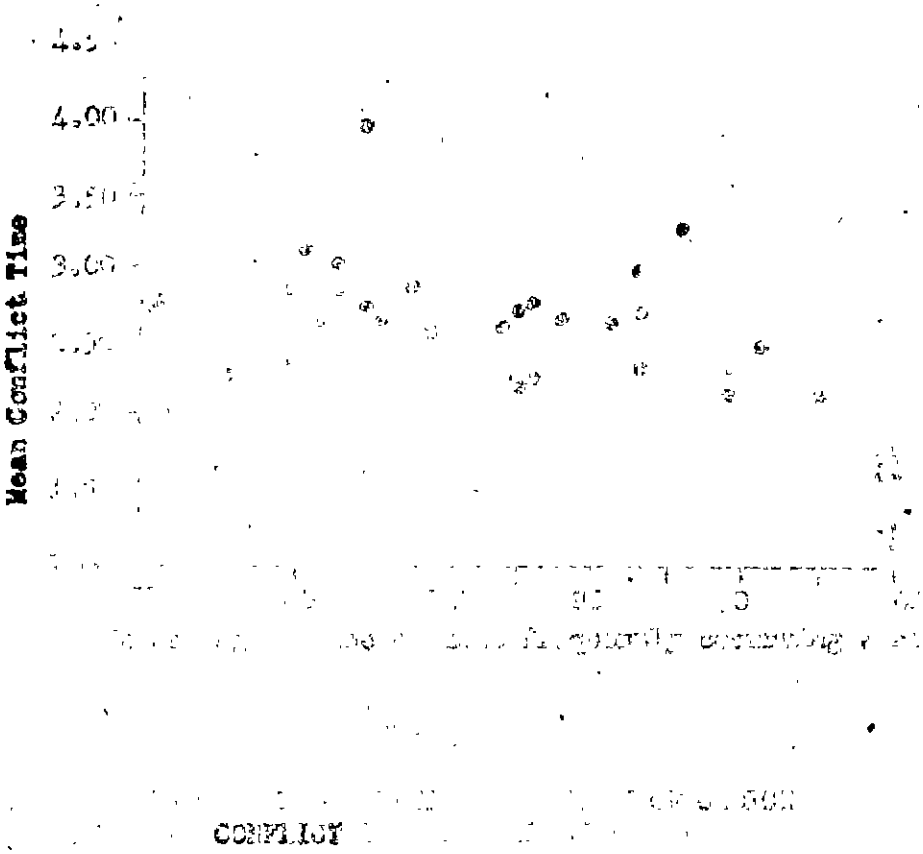
Analysis of the Word Test. On the words test, in contrast to the three others, it was not possible to measure gradations in the degree of conflict between the alternatives on an a priori basis. We could only distinguish the easy discriminations as a group from the difficult discriminations as a group. While the easy discriminations probably comprise a rather homogeneous group where no conflict is present, the difficult discriminations undoubtedly present a wide range of intensity of conflict.

We do have a measure of the degree to which the subjects agree in choosing the same word. For example, on the easy discriminations all the subjects chose the same word, except for 8 of the word pairs where there was one disagreement. The per cent of subjects choosing the same word varied from 51% to 95% on the difficult choices.

If we can assume that the degree to which a group of people agree among themselves, as to the meaning of a word, is indicative of the degree of uncertainty within the individual, then the percentage of subjects choosing the same word could be used as a measure of the degree of conflict engendered by the pair. There should be a greater degree of conflict for each individual on those pairs where there was more disagreement within the group as a whole.

We may make this clear by a more specific statement of the proposed analysis. For each pair of words, we tabulated the number of times each word was chosen. The relative frequency of choice of the more frequent of the two words was then calculated. This gives us a score which can range from .50 to 1.00. A score of .50 would indicate that on this pair an equal number of people selected each word. A score of 1.00 would indicate that for this pair everyone chose the same word. (It may be noted here that the latter occurred only for the "easy" discriminations

[illegible] the relative frequency of choice for the [illegible] and the mean conflict time for the group as a whole. As the relative frequency approaches 50%, the conflict time becomes higher and higher.



It should be remembered that the relative frequency is a measure of disagreement within the group. Each individual's response is a conflict response. The average time for the group, on the other hand, is a measure of the degree of conflict within the individual. It is interesting, of course, to note that the 10 easy discriminations are a separate and distinct group in the lower right-hand corner of the figure. A fairly wide gap exists between these 10 "easy" discriminations and the closest of the other 30 discriminations in terms of conflict time.

A product-moment correlation was calculated omitting the 10 "easy" discriminations. This correlation is equal to -.31. For N = 30, a correlation of .36 is necessary for significance at the 5% level.

Reliability. An evaluation of the tests must be based in part on the reliability of the tests. In a previous study the reliability of fixation time on the visual line test, calculated by the odd-even method, was .91 the reliability of the brightness test .91. These reliabilities were based on 30 trials of difficult discriminations and 10 trials of easy discriminations

The reliabilities for the present study were calculated by splitting the series of trials into two halves. The scores for the first half were correlated against scores on the second half. This method may be expected to yield lower reliabilities than the odd-even method because any progressive changes occurring in the course of the series of trials would tend to lower the correlation. These reliabilities are based upon 20 difficult discriminations and 10 easy discriminations.

Table 6 presents the uncorrected reliabilities for conflict time, decision time, decision time residuals, and base time for each of the tests. It may be seen that the reliabilities for conflict time and base time are uniformly high, ranging from .81 to .93. The reliabilities for decision time are slightly lower but quite satisfactory with the possible exception of the tactual line test which yields a reliability of only .53. The reliabilities for the decision time residuals are again lower as might be expected. They range from .33 to .76. For two of the tests, visual line and tactual angle, the reliability is still quite adequate.

TABLE 6

UNCORRECTED RELIABILITIES
(1st half vs. 2nd half)

Test	Conflict Time	Decision Time	Decision Time (Residual)	Base Time
Visual line	.84	.78	.67	.86
Tactual line	.93	.53	.33	.92
Tactual angle	.90	.83	.76	.90
Words	.92	.86	.40	.81

Intercorrelations among Tests. To determine the extent of "generality" or communality of decision time among the various test situations, we are faced with somewhat of a problem. Our basic measure (the one that has psychological meaning) is the measure of decision time. This measure is obtained, however, by subtracting base time from conflict time, these last two being the ones directly determined in the investigation. The subtraction of the base time does not serve the purpose of holding the base or reaction time constant completely. This is not a problem if one views this measure as the basic thing in which one is interested. The fact that it is not independent of the base time can be viewed as merely a reflection of the empirical fact that there is relationship between base time and decision time. In other words, some factors which are responsible for variation in base time are also responsible for variation in decision time.

A critic might, however, desire to know in addition what the intercorrelations are among the conflict times when the base times are held statistically constant; that is, when all variation in the base times are removed and base time is made independent of decision time. An argument for this procedure might be as follows: If one were to set up an ideal experiment, it might be desirable to select a group of people, all of whom had the same base time so that the conflict time measure for this group of people would be a clear uncomplicated measure of the increased time occasioned by the presence

of conflict. Since we were unable to do this, partial or semi-partial correlations may be used to approximate this condition. The semi-partial correlations, or correlations among residuals makes the conflict time independent of the base time on its own test only. The second order partial correlations makes the conflict time independent of both base times in any correlation.

One should, of course, expect the magnitude of the intercorrelations to decrease in the following order:

1. The intercorrelations among conflict times will be highest since they are actually compounded of decision time plus base time.

2. Intercorrelations among the decision times will be next in order of magnitude.

3. The semi-partial and partial correlations will be lowest in order of magnitude since a whole group of factors are held constant which ordinarily affect variation in all the tests.

The intercorrelations among the base times will also be presented for reference.

In all tables of intercorrelations, with the exception of the second order partials, the intercorrelations corrected for attenuation are presented in parentheses. The corrections were made using the reliabilities corrected by the Spearman Brown Prophecy formula.

The major emphasis should be put on the uncorrected correlations and the corrected correlations should be regarded only as auxiliary and interpreted with many reservations for the following reasons:

1. The reliabilities on which the corrections for attenuation are based are viewed by the authors as only very rough approximations, since split-half reliabilities were used which were corrected by the Spearman Brown Prophecy formula.

2. The correction for attenuation assumes that errors of measurement are uncorrelated among the tests. Since a good part of the unreliability is, no doubt, due to progressive changes in the form of more rapid decision time as the test progressed, there is good reason to believe that this assumption is violated.

3. In any use to which the tests may be put, some unreliability of the tests will remain a factor and, therefore, the correlations corrected for attenuation seem somewhat lacking in importance.

4. Since the correlations are based on relatively few cases the random errors may be expected to be quite large. Such random errors can produce very great distortions in such corrections.

Base-Time Intercorrelations. It will be recalled that the average of the times taken to respond on the 10 "easy" discriminations is called the base time. The intercorrelations among the four tests on this measure are presented in Table 7. It can be seen from inspection of Table 7 that the intercorrelations for base time range from .52 to .80. All of these correlations are significant at the 1% level. As might be expected, these data indicate that there is a fairly high degree of "generality" for the measure of the time to respond in a no-conflict situation.

TABLE 7

BASE TIME INTERCORRELATIONS*
(N = 75)

	Visual Line	Tactual Line	Tactual Angle	Words
Visual line	-	.52 (.55)	.60 (.64)	.62 (.68)
Tactual line		-	.80 (.84)	.57 (.62)
Tactual angle			-	.64 (.70)
Words				-
Mean	.84	6.05	3.66	1.47
Sigma	.20	3.94	1.96	.38

* The correlations in parentheses are corrected for attenuation using corrected reliabilities.

Intercorrelations among Conflict Times. Table 8 presents the intercorrelations of conflict time among the four tests. These correlations range from .57 to .83 averaging about .66.

TABLE 8

INTERCORRELATIONS AMONG CONFLICT TIMES*
(N = 75)

	Visual Line	Tactual Line	Tactual Angle	Words
Visual line	-	.57 (.61)	.65 (.71)	.62 (.67)
Tactual line		-	.83 (.87)	.66 (.69)
Tactual angle			-	.66 (.69)
Words				-
Mean	1.36	7.51	5.78	2.63
Sigma	.59	5.41	4.08	1.53

* The correlations in parentheses are corrected for attenuation using corrected reliabilities.

These intercorrelations represent "generality" of a compounded measure where time to react is added in with increase in time due to conflict.

The intercorrelations between all the base times and all the conflict times can be found in Appendix A. These correlations together with those presented in Tables 7 and 8 give the complete intercorrelation matrix for those measures directly determined in the investigation.

Decision-Time Intercorrelations. As indicated previously the more basic data here are the intercorrelations of the decision-time measures among the tests. In Table 9 these intercorrelations are presented together with the means and sigmas for each of the tests.

The correlation coefficients range from .52 to .59. All of the correlations are significant at the 1% level. The small range for these correlation coefficients is surprising because it indicates that the degree of "generality" is about the same regardless of how divergent one test situation is from any other test situation within the limits of our experiment.

TABLE 9

DECISION-TIME INTERCORRELATIONS*
(N = 75)

	Visual Line	Tactual Line	Tactual Angle	Words
Visual line	-	.52 (.66)	.59 (.66)	.59 (.66)
Tactual line		-	.57 (.72)	.56 (.70)
Tactual angle			-	.57 (.62)
Words				-
Mean	.52	1.46	2.12	1.16
σ	.46	1.93	2.61	1.26

* The correlations in parentheses are corrected for attenuation using corrected reliabilities.

Correlations were also computed between base time and decision time for each of the four tests. It is recognized that those correlations have a somewhat spurious element in them. This spurious element will have an effect in a negative direction since, when base time is large, decision time will tend to be small and when base time is small, decision time will be large. The correlations are presented, however, merely for the purpose of showing that there is some positive relationship between base time and decision time. Since the obtained correlations are positive and fairly high, we can be pretty certain of our conclusions because we obtain those high positive correlations in spite of a spurious negative element. These correlations are: for the visual line test .55; for the tactual line test .66; for the tactual angle test .58; and for the word test .79.

Intercorrelations Using Decision-Time Residuals. These correlations represent the intercorrelations among the conflict times when each conflict time is made independent of the base time on the same test. This independence is accomplished by subtracting from each conflict time the value of the base time which lies on the regression line. The formula used in the computation is:

$$r_{\text{between residuals}} = \frac{r_{x_1x_2} - r_{x_1y_2}r_{x_2y_2} - r_{y_1x_2}r_{x_1y_1} + r_{y_1y_2}r_{x_1y_1}r_{x_2y_2}}{\sqrt{\left(1 - r^2_{x_1y_1}\right)\left(1 - r^2_{x_2y_2}\right)}}$$

(Where x represents conflict time, and y represents base time)

This same result is also accomplished by setting the standard deviation of the conflict time equal to 1 and the standard deviation of the base time equal to the correlation between base time and conflict time. Thus defining the relationship between base time and decision time as 0. Decision time is then independent statistically from base time. These correlations are presented in Table 10.

TABLE 10

INTERCORRELATIONS USING RESIDUALS*
(N = 75)

	Visual Line	Tactual Line	Tactual Angle	Words
Visual line	-	.32 (.47)	.39 (.47)	.44 (.70)
Tactual line		-	.17 (.25)	.18 (.34)
Tactual angle			-	.32 (.48)
Words				-

*The correlations in parentheses are corrected for attenuation using corrected reliabilities.

The correlations range from .17 to .44. It still appears that "generality" among the tests are all about the same. The fluctuations within the matrix can undoubtedly be attributed to chance.

Intercorrelations Using Second Order Partials. The previous section, it will be remembered, dealt with intercorrelations among conflict time when each conflict time was independent of its own base time and not of any of the others. The intercorrelations among the tests using second order partial correlations make the intercorrelations independent of both base times involved in any single correlation. They are presented in Table 11. They may be seen to be slightly higher than the intercorrelations among the residuals. This increase is practically negligible and is probably a chance affair. The correlations here range from .19 to .44.

TABLE 11

INTERCORRELATIONS USING SECOND ORDER PARTIALS
(N = 75)

	Visual Line	Tactual Line	Tactual Angle	Words
Visual line	-	.33	.42	.44
Tactual line		-	.19	.22
Tactual angle			-	.34
Words				-

We have thus demonstrated that even when all sources of variance common to base time and decision time are held constant there is still an appreciable degree of communality among the tests. These sources of variance which have been held constant in this manner may be equally important in the evaluation of the individual as those sources of variation which are still operating. Whether or not in an actual testing situation one should leave in these sources of variation (use the decision-time measure) or eliminate the sources of variance (use residuals) is a practical problem the solution of which depends on the exigencies of the situation involved.

The correlations among the tests for decision-time have shown a fair degree of relationship. In a practical testing situation one might wish to use two tests in conjunction, since such a combination would probably be more reliable and more "general." One such correlation was calculated. The decision time for the visual line test was added to the decision time for the tactual line test. This combination was correlated with the combination of decision time for the tactual angle and word tests. The combinations were simple additions of coded scores. In order to have approximately equal weighting of each test in a combination, those tests whose standard deviations were nearer alike were paired.[18]

This correlation (visual line plus tactual line versus tactual angle plus words) turns out to be .73 for decision time. When decision time residuals are used the correlation is .53. This represents an appreciable increase over "generality" among single tests. This increase might be due to increased reliability, and/or measurement over a wider area.

---

[18]Since the additions were in terms of coded scores, the standard deviations of these coded scores were considered. These coded standard deviations were one-half the actual standard deviations for the two tactual tests. There was no change for the other tests.

SUMMARY AND CONCLUSION

Pilots are constantly faced with the necessity of making decisions in conflict situations particularly during combat. If tests and techniques could be designed to measure the difficulty with which trainees resolve conflict situations, the refinement of predictive batteries might be extended to improve the selection of those most likely to succeed in aerial combat. Laboratory studies, reported by other investigators, have been principally concerned with the determination of the functional relationship between magnitude of decision-time and intensity of conflict. The present study represents an attempt to focus attention upon the study of decision-time in relation to problems of pilot selection.

Part I of the study in this report describes a test, designed to measure the length of decision-time, which was also employed with three other tests in research presented in Part II. In the original test the subject was shown two vertical lines and instructed to decide whether one line was longer or shorter than the other. Conflict between the alternatives was provided by using stimuli which differed by small amounts or were equal in length. Odd-even reliability coefficients ranged from .93 for the base time to .66 for $D_{5\ 7/8"}$ (the difference between the base time and the average of the decision-times for the 5 7/8" comparison lines). Correlation coefficients of such magnitude appear to indicate fairly satisfactory reliability.

Part II was designed to study more specifically the "generality" of decision-time, i.e., whether decision times were specific to the particular conflict situation or were indicative of some characterological aspect of the individual. This problem was attacked by determining the intercorrelations among decision-times obtained in four conflict situations differing in degree of similarity. The situations, designed to yield pairs of tests with various degrees of divergence in what appeared on *a priori* grounds to be relevant characteristics, were (a) the visual line test, (b) the tactual line test, (c) the tactual angle test, and (d) the word test.

Measures of test reliability were obtained by splitting the series of trials for each test into halves, and correlating the first half against the second half. Reliability coefficients for conflict time, decision time, decision-time residuals, and base time were calculated for each of the tests. The reliabilities for conflict time range from .84 to .93; those for decision time range from .53 to .86; those for decision-time residuals from .33 to .76; and for base time from .81 to .92.

In accordance with the original purpose of the study which was concerned with "generality" of decision time from one test to another, intercorrelations among the four tests were obtained for conflict time, decision time, decision-time residuals, conflict time holding base time constant by means of second order partials, and base time. The intercorrelations among base times and among conflict times are highest (ranging from .52 to .83); the intercorrelations among decision times are next highest (ranging from .52 to .59); the intercorrelations among residuals and among second order partials

are lowest (ranging from .17 to .44). The small range of correlation coefficients within any of the correlation matrices appears to indicate that the degree of "generality" is about the same for the four situations within the limits of the experiment. Such findings do not suggest any decrease in "generality" as the pairs of situations become more and more divergent.

The present study has served the purpose of providing instruments which yield reliable means of the way people react characteristically to conflict situations. The instruments are suitable for field validation against such criteria as passing or failing flight training, success in aerial combat, and success as a bomber pilot or a fighter pilot.[19]

---

[19] The question has been raised relative to the correlation of these tests with intelligence measures. The present study was not designed to answer this question. It so happens, however, that scores are available for the 75 women subjects on the American Council on Education examination, which had been administered previously by the University of Rochester. These correlations are as follows: visual lines, -.12; tactual lines, -.06; tactual angles, -.32; words, -.12.

APPENDIX A

CORRELATIONS BETWEEN BASE TIME AND CONFLICT TIME

CORRELATIONS BETWEEN BASE TIME AND CONFLICT TIME

Base Time	Conflict Time: Visual Line	Tactual Line	Tactual Angle	Words
Visual Line	.76	.52	.59	.47
Tactual Line	.53	.96	.84	.63
Tactual Angle	.55	.76	.85	.60
Words	.61	.57	.62	.87