RESTRICTED

Not for publication

SENSITIVITY OF PERIPHERAL VISION IN RELATION TO SKILL IN LANDING AN AIRPLANE

by

Miles A. Tinker

and

William S. Carlson

A report on research conducted at the University of Minnesota, Minneapolis, Minnesota, 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.

April 1943

CIVIL AERONAUTICS ADMINISTRATION

Division of Research

[illegible]

[illegible]

National Research Council

Committee on Selection and Training of Aircraft Pilots

Executive Subcommittee

C. W. Bray	J. C. Flanagan
D. R. Brimhall	H. M. Johnson
L. A. Carmichael	W. R. Miles
J. W. Dunlap	G. R. Wendt

M. S. Viteles, Chairman

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

April 26, 1943

Dr. Dean R. Brimhall
Director of Research
Civil Aeronautics Administration
Washington, D. C.

Dear Dr. Brimhall:

The attached report on Sensitivity of Peripheral Vision in Relation to Skill in Landing an Airplane, by M. A. Tinker and W. S. Carlson, is submitted by the Committee on Selection and Training of Aircraft Pilots with the recommendation that it be included in the series of technical reports published by the Division of Research, Civil Aeronautics Administration.

The authors were able to demonstrate that an individual's disjunctive reaction times were significantly related to variations in the size, illumination, and location of visually presented stimuli. Interactions of the sources of variation in the visual stimuli also seem to cause significant variance in the reaction times. It is not possible, however, on the basis of this study to estimate the magnitudes of the relationship among the variables.

It will be noted that the criteria of landing employed in this study, i.e., ratings on good and bad landings, proved to be inadequate for classification of student pilots. As a consequence, no significant conclusions or recommendations can be presented with respect to visually presented stimuli in selecting pilots or in improving performance in landing a plane. However, it is shown that reaction times do vary significantly with changes in peripherally presented stimuli. Further research, providing more exacting criteria of landings, will permit a further evaluation of the practical relation between such variables and performance in landing an airplane.

Very truly yours,

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

MSV:rm

EDITORIAL FOREWORD

The following report describes an experiment, undertaken early in the history of the Committee on Selection and Training of Aircraft Pilots, with the view of studying the relationship between visual functions and skill in landing an airplane. This investigation, which antedates the study by Tiffin and Bromer included in Report No. 10, reflects the early interest of the Committee, and particularly of Dr. Dean R. Brimhall, Director of Research, Civil Aeronautics Administration, in the relation between visual functions and pilot efficiency.

Unfortunately, difficulties experienced in obtaining accurate estimates of skill in landing a plane severely limit the contribution of this research project in solving an eminently practical problem in aviation. Nevertheless, the basic experimental methods warrant consideration in planning further research in this important area.

CONTENTS

SUMMARY

A disjunctive reaction time experiment was conducted to test the hypothesis that response time to peripherally presented stimuli might be related to skill in landing shown by pilots.

The subject fixated a faint spot of light projected on a ground glass screen. A second light source projected a ring, centered about the fixation point, with a break opposite one or another corner of the square screen. At a convenient level below the screen was placed a response lever. By an appropriate movement the observer could indicate in which quadrant the break in the ring of light appeared. A gravity tachistoscope placed between the projector and screen served as a shutter, exposing the ring for an interval of $\frac{1}{4}$ second. Releasing the shutter closed the circuit of a direct-current magnetic clutch which coupled a synchronous motor to a pointer. The correct response movement by the subject opened the circuit, disengaging the pointer.

Three sizes of rings, three sizes of breaks, and three levels of illumination were used. Three variables at three levels produce 27 combinations. The factorial design of the experiment allowed an investigation of all combinations, conserving all two-factor interactions and all main effects. The design also provided for recovery of three-fourths of the information on triple interaction, which was confounded with individual differences. A given subject was tested with 9 combinations so chosen that each variable appeared an even number of times at each of its three levels. Ten reaction times were measured for each combination in the experimental series. The ten values were averaged to yield 9 scores for each subject. The experimental series was preceded by a practice series. Forty-eight subjects took part in the study.

Results indicate that latency increases as the figure is presented in the visual periphery; that latency varies inversely with size of the break to be discriminated; and that changing illumination within a small range produces a statistically non-significant curvilinear effect on latency of response. Although significant two-factor and three-factor interactions are obtained, indicating that the effects of the variables are interdependent, a portion of these effects is probably due to the optical structure of the experimental situation. A marked practice effect was demonstrated within a single 50-minute session even though a pre-experimental series of responses was provided.

Reaction time was found to depend also upon the quadrant of the visual field in which the discrimination is to be made. Latency is significantly greater when the critical portion of the stimulus figure appears in the lower half of the visual field than when the upper half is involved, a finding which is contrary to previously reported results.

To evaluate the significance of visual reaction time in relation to landing skill a criterion was required. Five flight instructors ranked students in order of merit for landing and also filled out a portion of an early version of the <u>Ohio State Flight Inventory</u> for rating landings. Two student fliers judged high by both methods and two judged low by both were selected from each instructor's group. This procedure made up "high" and "low" groups each containing 10 students. (There is reason to believe that the ratings were poorly done.)

The "high" group made fewer response errors than the "low" group but the differences are not significant. None of the ratings on landing skill is related significantly to reaction time for peripheral discriminations. While the validity of the criterion is questionable, it does not appear likely that the type of reaction-time situation employed here would be highly differentiating even with better co-operation of the raters.

Flying conditions in training represent a considerable simplification over the situation encountered in commercial, civil, or military aviation. Differentiation by instructor's ratings depends on motivation and a number of other factors as well as on technical competence and aptitude. When subjects were rated on landing during check flights all were given ratings of 2 or 3 on a possible five-point scale. This high homogeneity renders differentiation on laboratory testing rather difficult.

It is suggested that a more exacting situation may be required on which to establish criteria of performance. A more complex task might differentiate when the simpler requirements fail. Night and formation flying are suggested as examples of the kind of task mentioned above. Further variations of experimental procedure are also recommended.

SENSITIVITY OF PERIPHERAL VISION IN RELATION TO SKILL IN LANDING AN AIRPLANE

INTRODUCTION

Airplane pilots encounter many situations where adequate maneuvering would seem to require effective use of peripheral or extra-foveal vision. Landing the airplane is perhaps one of the most important of such situations in which changes in the visual stimulus complex occur in rapid succession. Accurate perception of the entire pattern of visual stimuli must be followed by quick, accurate response to the important elements of the perceived pattern if skillful landings are to be achieved. Many of the cues in these changing stimulus patterns are to be found in the peripheral zones.

STATEMENT OF THE PROBLEM

There are two problems involved in the present study: (I) Analysis of the variability occurring in reaction time as the result of changes in distance from the fixation point, size, and illumination of a peripherally presented stimulus, and (II) a study of the relationship between quickness of reaction time to peripheral stimuli and skill in landing an airplane.

PROBLEM I

Disjunctive Reaction Time as a Function of Variations in the Stimulus Figure

Previous considerations of the problems involved in a study of visual fields led to the choice of the following three variables for investigation in a choice reaction-time situation:

(1) Distance of the stimulus object from the center of the visual field (fixation point).
(2) Size of the stimulating object.
(3) Brightness or intensity of the stimulus.

Use of a factorial experimental design made it possible to ascertain how much of the total variability exhibited by the reaction-time data could be assigned to each of these three experimental variables and to their interactions.

A. APPARATUS

The apparatus consisted of a ground-glass screen, 43 inches square, mounted on a frame so that the subject, standing behind the screen, could fixate in a normal manner on the center of the screen. An adjustable platform gave the subject proper height for a head rest which maintained his eyes 11 inches from the fixation point.

The choice-reaction-time apparatus consisted of a box 3 x 3 x 6.5 inches attached on the frame at the base of the glass screen. Four slits, one leading from each corner to the center, were cut in the top of the box. A vertical lever was mounted in the center of the box so that it could be pushed along any one of the four slits by the subject. When pushed along the slits this lever tripped a mercury switch, breaking the circuit and stopping the timer. When released it returned to its original position in the center of the box.

A free-falling gravity tachistoscope was constructed to produce a high degree of consistency in timing successive exposures. A shield with an adjustable opening was built into a fairly heavy frame which moved up and down in slots. This apparatus was 11 feet in height and was set 14 feet 4 inches from the ground-glass screen. Preliminary calibration with a standard electric timer gave absolute consistency of exposure within the limits of the timer[1] and the latency of the switches. This consistency was further checked within the limits of the measuring device for a series of 10 measurements by attaching a strip of wax-coated paper on the falling shutter in contact with a high-frequency electric vibrating marker run on a 60-cycle current. Accurate determination of exposure time which would maintain a relatively difficult visual task and at the same time make it possible for the subjects to discriminate nearly all of the stimuli gave a value of 250 milliseconds.

Reaction times were measured with a Standard Electric Timer, Model S-1, equipped with a D. C. Clutch. A G-M Laboratories D-C Power Supply Unit (Filtered Rectifier No. 2530-Z, 6 volts at 6 amperes output) was used in the clutch circuit of the timer. This circuit included a make mercury switch attached to the tachistoscope which was tripped simultaneously with the beginning of the exposure. In the same circuit were the break mercury switches in the subject's response box. Connection with the proper switch in the response box was achieved by the experimenter's moving the arm of a four-position selector switch to the appropriate connection prior to an exposure. Thus, the timer was activated by the fall of the tachistoscope shutter and stopped by the subject's movement of the lever into the proper corner of the response box.

The stimuli consisted of photographic negatives of wide, broken rings, one each mounted on a 2 x 2 inch projection slide.[2] Each ring had a break opposite one or another corner of the square slide. These rings were made in three sizes: as projected they were 23 1/2, 34 3/8, and 41 inches in diameter and subtended angles of 93°46', 114°56', and 123°34' respectively. Each diameter was tested with breaks in the circles subtending 2°12', 7°10', and 14°16' of visual angle. A Spencer lantern with a 100-watt S-11 bulb was used to project the slides. A set of polaroid discs attached to the lantern made it possible to vary the brightness level of the stimulus (.02, .05, and .10 footlamberts[3]).[4] By means of a second lantern, equipped also with polaroid discs, a faint spot of light was focused on the center of the screen as a fixation point.

B. LABORATORY PROCEDURE

In the test situation, then, the observer fixated binocularly the illuminated ring on a dark field. This ring was broken in one of four chosen positions on its circumference. Conveniently accessible to the subject was a manual response lever mounted on a spring swivel. The observer was to move the lever from its normal

---

[1]The timer recorded in 1/100 seconds.

[2]Landolt Rings, such as used by Ferree in many of his researches. Ferree, C. E. Rand, G., & Lewis, E. F. The effect of increase of intensity of light on the visual acuity of presbyopic and non-presbyopic eyes. Transactions of the Illuminating Engineering Society, 1934, 29, 296-313. (Editor's Note. The smallest ring (diam. = 23 1/2 inches) stimulates more of the foveal and parafoveal regions than the periphery.)

[3]Footlambert: "A unit of brightness, equal to the uniform brightness of a perfectly diffusing surface which emits or reflects one lumen per square foot." (From Warren, H. C. Dictionary of [illegible] Boston: Houghton Mifflin, 1934.)

[4]These values [illegible] the 114°56' ring because transmission properties of the screen [illegible] angle of incidence of the projection beam.

center position along any one of four guide slots which corresponded to the screen quadrant in which the break in the ring appeared.

On his arrival at the laboratory the subject was told the general nature of the experiment and shown the workings of the apparatus. Then, after he was placed in position behind the ground-glass screen, the following directions were read: "Look directly at the dim spot of light at the center of the screen. Shortly after I say 'ready' there will flash onto the screen for a brief instant a circle of light with a break in its circumference. The break will be opposite one of the four corners of the screen. As soon as you see where the break occurs, push the response lever into the corresponding corner of the response box. After the response is made, release your pressure on the lever and it will return to its original position at the center of the box. You will be given a rest at frequent intervals."

After two or three sample slides were presented to make sure that the subject understood the directions, a practice series of 9 slides was given. This was followed by the experimental series as described in the experimental design. Records were taken for both the practice and the experimental series. Rest periods were given after about every 15 responses. Each subject required approximately 50 minutes to finish the experiment.

The experimenter (Carlson or Tinker) operated the tachistoscope, changed slides and varied the brightness levels of the lights. An assistant set the four-point selector switch in the circuit with the response box and timer, and recorded the response times. The subject himself reset the switches in the response box after each response.

C. EXPERIMENTAL DESIGN

The experimental procedures for Problem I were factorially designed[5] in order to ascertain whether a significant portion of the observed variability in reaction time could be referred to one or more of the variables in the situation, or to any combination of them. To ascertain the significance of any portion of this variability assigned to a given causal agent, use is made of the ratio of the variance within groups to that between groups or among means.

Three sizes of circles (a_1, a_2, a_3) are combined with three sizes of break (b_1, b_2, b_3) under three levels of illumination (c_1, c_2, c_3). In all, therefore, 27 different combinations, or test objects, are derived. These 27 combinations can be arranged in three blocks[6] of 9, so chosen that within any one block all comparisons for main effects and interactions of two factors are conserved (see Table I). Of the four pairs of degrees of freedom for triple interaction,[7] three

---

[5]Full explanations of the factorial design of experiments can be found in standard references for the analysis of variance, e.g., Lindquist, E. F. Statistical Analysis in Educational Research. Boston: Houghton Mifflin, 1940.

[6]A 'block' refers to a table or column of data from a particular subject.

[7]Triple interaction refers to variance which would be due to the combined interaction of three variables. Only rarely are these higher order interactions significant.

pairs are conserved and one pair [illegible] with [illegible] (individual) differences.

By following the arrangement shown in Table I, it is possible to recover three-fourths of the information lost in confounding triple interaction and individual differences. The variables are given in the column headings and appear always in the order, ring size (a), break size (b), and brightness level (c). The three numbers indicate the levels of these variables for a particular combination, i.e. the first entry, $a_1b_1c_1$, means that the presentation included the smallest circle with the smallest break at the lowest illumination. The 12 vertical columns of Table I represent 4 sets of 3 blocks (individuals), i.e. each vertical column corresponds to the 4 subjects who were given the same presentation in the same order. Within each set, one pair of degrees of freedom for triple interaction is confounded with block (individual) differences. The information on any particular pair is therefore preserved in three sets of blocks, lost in one set, i.e., four different arrangements of the 27 combinations are given. In the first three vertical columns a fourth of the triple interaction is not separable from block differences; in the second three columns a fourth of the triple interaction is similarly confounded, but not the same part as in the first three. Each of the four pairs of degrees of freedom is thus confounded in turn.

Any one observer was tested with 9 combinations. For each stimulus arrangement 10 responses were averaged to yield a response time score. Because there were 48 subjects for whom records were required, they were tested in 12 groups of 4 observers each as shown in Table I.

Each of the 9 slides to be used with a given observer was prepared in two forms, differing only with respect to the quadrant in which the break was placed. The quadrant choices were randomized to avoid systematic effects from this source. The series of 18 slides so obtained was repeated five times to give a total of 90 responses, 10 for each combination. These groups of 10 yielded 9 composite scores for each observer.

---

[8] When the variance of the means is derived from two or more sources and no separation of the sources is possible making required comparisons interdependent, the degrees of freedom or sources of variation are said to be confounded. (Most experiments that are said to lack proper controls would involve confounding of two or more sources of variability.)

[9] After allowing time for instructions and practice series of slides, a 50-minute period permitted an experimental series of 90 responses. Since a presentation appeared 10 times in each session, this meant that only 9 of the 27 combinations could be employed with a given subject, and that in a group of three subjects all combinations could be tested. This design provides for 12 subjects in 4 groups of 3. Repetition of this sequence 4 times gives a total group of 48 observers. Ideally, all 48 observers should have been tested in each combination. This procedure, due to lack of time, was not feasible in the present study. The significant advantage of this type of design is that it becomes possible to obtain 10 instead of [illegible] responses for each subject for each given presentation, and permits the recovery of [illegible] of the information that would have been [illegible] in the complete experiment. This complete experiment would have provided all 27 combinations tested on every subject, but would have required [illegible] experimental labor and time

Outline of Factorial Design for Reaction-Time Experiment

Subjects	1-4	5-8	9-12	13-16	17-20	21-24	25-28	29-32	33-36	37-40	41-44	45-48
Variables	abc	abc	abc	abc	abc	abc	abc	abc	abc	abc	abc	abc
Combinations of Variables	111	113	112	111	112	113	111	113	112	111	112	113
	212	211	213	213	211	212	212	211	213	213	211	212
	313	312	311	312	313	311	313	312	311	312	313	311
	123	122	121	122	123	121	122	121	123	123	121	122
	221	223	222	221	222	223	223	222	221	222	223	221
	322	321	323	323	321	322	321	323	322	321	322	323
	132	131	133	133	131	132	133	132	131	132	133	131
	233	232	231	232	233	231	231	233	232	231	232	233
	331	333	332	331	332	333	332	331	333	333	331	332

D. DISCUSSION AND RESULTS (Problem I)

A summary of results appears in Table II. For each source of variation is given the appropriate degrees of freedom, sum of squares, and mean square or variance. The entries in the F column are the ratios of the several experimental variances to the residual or test variance. The latter is an estimate of the variability that the data would exhibit if the experimental variables had made no systematic contribution to the total variance. The F ratios are accompanied by the value required for significance at the .05 and .01 levels. The variance attributed to blocks is essentially a measure of individual differences. As previously noted, one-fourth of the variance due to triple interaction is included in block differences. The inflation of individual differences thereby produced may safely be neglected. In designs of this type block effects are commonly treated as a purely formal subdivision of the total sum of squares in order to reduce the residual variance preparatory to making tests of significance. It is used here as a convenient estimate of reliability. The subjects were available only for a single experimental session. Under these conditions it is customary to compute split-half correlations or related coefficients of internal consistency. A significant correlation means that differences between individuals are making a greater contribution to the scores than are the uncontrolled sources of variation. Precisely the same interpretation may be given to the block variance in this experiment.

The experimental variables are designated by the letters A, B, and C for circle size, break size, and level of illumination. There are 144 scores for each ring size, all variables and their combinations being equally represented at the three levels. The three totals yield the two degrees of freedom for estimating the variance associated with circle size. Because there is no reason to assume that the variable is linear, the sum of squares is given separately for the regression (A_r) and the deviation (A_d) components. Ring size is clearly linear in its effect. Within the limits of the range tested in this experiment one may say that increases in the visual angle subtended by the broken ring will produce proportional increases in response time. This result confirms an early finding by Hall and von Kries.[10] Visual acuity is also known to vary inversely with increasing dependence on peripheral vision. Data from an experiment by Hecht and Verrijp[11] suggest that the critical flicker frequency also varies inversely with the distance from the fovea at which the retina is stimulated. Although the effect of varying the retinal location depends on the level of illumination, the relationship between fusion frequency and stimulus position is evident at levels of illumination comparable with those used in this experiment.

Variations in the size of the break to be discriminated are associated with significant changes in response time. The effect is not strictly linear because the change in break size from 2° to 7° produces a greater effect than the change from 7° to 14°. It is, of course, well known that intensity discrimination varies with the size of the retinal area which is stimulated. But these break sizes are all above the threshold at the levels of illumination employed in this experiment. With very low brightness levels or very short exposure intervals, pre-experimental observers were brought down to threshold performance. Subjective reports suggest that under these conditions the ring is frequently seen as a closed annulus. Above the threshold the differences are in the time required for the discontinuity to appear.

Changes in brightness do not produce significant differences in response time when separately considered. In general, previous investigations agree that increased stimulus intensity is accompanied by shorter reaction times. The relation is certainly not linear, but its exact form varies widely from one experiment to another. Visual acuity also varies with intensity but the acuity curve for 2° test objects is well above normal threshold within the illumination limits here employed. Hecht[12], among others, is inclined to subsume the findings on acuity and intensity discrimination under one hypothesis. The effect of increasing the area of the stimulus field on intensity discrimination, and of increasing brightness on acuity are both interpreted

---

10 Hall, G. S. & Kries, J. v. Über die Abhängigkeit der Reaktions-zeiten vom Ort des Reizes. Arch. Anat. Physiol. Lpz. (Supl.), 1879, 1-10.

11 Hecht, S. & Verrijp, C. D. Intermittent stimulation by light: III. The relation between intensity and critical fusion frequency for different retinal locations. J. Gen. Physiol., 1934, 17, 251-268.

12 Hecht, S. A theory of visual intensity discrimination. J. Gen. Physiol., 1935, 18, 767-789.

TABLE II

Analysis of Variance of Reaction-Time Data

Source	Sum of Squares	Degrees of Freedom	Sum of Squares	Degrees of Freedom	Mean Square	F	5% Level	1% Level
Blocks	4374698.8	47	4374698.8	47	93078.7	10.7	1.4	1.6
A_r	1233104.3	1	1259848.8	2	629924.4	72.6	3.0	4.7
A_d	26744.5	1						
B_r	72810.1	1	106409.3	2	53204.7	6.1	3.0	4.7
B_d	28199.2	1						
C_r	4753.1	1	20856.0	2	10428.0	1.2	3.0	4.7
C_d	16102.9	1						
$A_r \cdot B_r$	54203.4	1	58465.9	4	14616.5	1.7	2.4	3.4
$A_r \cdot B_d$	191.5	1						
$A_d \cdot B_r$	225.0	1						
$A_d \cdot B_d$	3846.0	1						
$A_r \cdot C_r$	12320.0	1	166791.5	4	41697.9	4.8	2.4	3.4
$A_r \cdot C_d$	71734.8	1						
$A_d \cdot C_r$	24727.6	1						
$A_d \cdot C_d$	58009.1	1						
$B_r \cdot C_r$	47533.5	1	72546.4	4	18136.6	2.1	2.4	3.4
$B_r \cdot C_d$	11475.8							
$B_d \cdot C_r$	12348.8	1						
$B_d \cdot C_d$	1188.3	1						
A · B · C	210144.3	8	210144.3	8	26268.0	2.9	2.0	2.6
Residual variance		358	3113066.9	358	8975.0			
Total		431	9482827.9	431	22001.9			

as being due to an increase in the number of functioning retinal elements. If one assumes, furthermore, that the state of the visual apparatus at any given time is a function of the final stationary state for a specified set of conditions, this hypothesis might be adequate also for reaction-time data. Bartley's[13] finding that the latency of cortical response to visual stimulation varies with the area of the stimulus object is of interest in this connection. He noted also that a curve representing the area and response-time function shows an abruption at about 20° of visual angle. Size changes are more effective when small areas are involved. In this study, it will be recalled, the change from a 2° to a 7° break produced the

---

[13]Bartley, S. H. The time of occurrence of the cortical response as determined by the area of the stimulus object. <u>Am. J. Physiol.</u>, 1935, 110, 666-674.

greater portion of the total change occurring between 2° and 14°. Wilcox[14] has noted that acuity may vary curvilinearly with brightness. He used two dark bars on a lighted field and two illuminated bars on a dark field. For the first situation he obtains the usual rising curve for acuity plotted against illumination. Resolution in the second case rises to a maximum at something under 10 photons, falls thereafter. No satisfactory explanation has been offered for this finding, although the technique has been severely criticized on the ground that it differs from the conventional way of measuring acuity. The broken rings used in this experiment present a comparable test situation. Although the illumination changes are too small to produce a statistically significant effect, the response-time changes observed follow the same trend found by Wilcox at comparable illuminations. That some sort of glare phenomenon involving entoptic stray light or retinal halation can reduce acuity is to be expected. But a curvilinear relationship is not so readily explained.

The interaction variance for circle and break size leaves tenable the hypothesis that their effects on response time are additive. The meaning of the interaction between circle size and illumination is ambigous. In all probability, a portion of this variance is an indirect measure of the effect of illumination. It is due to changes in the amount and distribution of the transmitted light with variations in the angle of incidence at the screen. This angle changes from one circle size to another. The break-size and illumination interaction, though not significant, is conspicuous in one part of the original data. The effect appears when large circles are considered. With these figures in dim light, break-size changes seem to be effective. And light changes are effective if the breaks are small. The triple interaction is derived from the interdependence of the two-factor interactions. Although it can be given a statistically straightforward interpretation, its experimental meaning is uncertain. Nevertheless, it is further evidence that even the most carefully conducted studies of single variable relations to visual processes are likely to encourage over-simplified hypotheses.

The residual or test variance is made up of intra-individual differences, including practice effects, motivational changes, fatigue, and fluctuations of attention. It includes also experimental errors due to apparatus variations, unequal difficulty of the response quadrants, and related sources. Some of these factors probably make no important contribution to the total variability, others contribute appreciably. The results of various calibration procedures make it clear that the variable error of instrumental recording does not exceed two per cent of the mean response time. Variability introduced because the response quadrants are of unequal difficulty is probably important. It will be recalled that of the ten responses made by the subject to a given test object, five were made in one randomly chosen quadrant, five in another quadrant randomly chosen from the remaining three. The data from each observer could, therefore, be arranged to provide 9 differences between quadrants by comparing two groups of five responses obtained under similar conditions except for the response quadrant. Six distributions of differences were obtained, the comparisons being quadrants 1 - 2, 1 - 3, 1 - 4, 2 - 3, 2 - 4, and 3 - 4. Whether the relative difficulty within a given quadrant pair is obtained by

---

[14] Wilcox, W. W. The basis of the dependence of visual acuity on illumination. Proc. Nat. Acad. of Sci., 1932, 18, 47-56.

direct comparison or inferred from other comparisons the results are almost identical. In arbitrary units, the relative difficulties of the four quadrants of the visual field are: upper left, 0; upper right, 9.2; lower right, 27.5; lower left, 46.4 (most difficult). All differences except that between the two upper quadrants are significant beyond the .01 level. The interpretation of the differences between stimulation in the left and right visual fields is complicated by the fact that binocular vision was permitted. The results do, however, agree with previous findings for horizontal differences but not for vertical. Woodworth[15] following Hall and von Kries says that supra-foveal stimulation yields shorter reaction times than infra-foveal stimulation. Significant differences in the reverse direction were obtained in this experiment. The comparison is, of course, made with reservation because the present data have been obtained from a choice-reaction task instead of the simple reaction to which Woodworth refers.

Practice effects are also responsible for an appreciable contribution to the residual variance. An estimate is obtained by comparing the response times for each subject on the first eighteen responses with the performance on the fifth series for the same test objects. In spite of the fact that preliminary practice was given, the mean difference between the first and fifth response-series totals is 1178 ms., with a σ of 204 ms. This difference exceeds the .01 level of significance.

E. SUMMARY AND CONCLUSIONS
(Problem I)

The effects of three variables on the latency of a visually controlled disjunctive reaction were studied in a factorially designed experiment. The stimulus objects were illuminated figures similar to the Landolt broken ring. Observers were to indicate by an appropriate movement of a manual response lever the quadrant of the visual field (translucent screen) in which the break appeared.

Results indicate that latency increases as the figure is presented in the visual periphery, that latency varies inversely with the size of the break to be discriminated, and that changing illumination within a small range produces a statistically non-significant curvilinear effect on latency of response. Although significant two-factor and three-factor interactions are obtained, a portion of these effects is probably due to the optical structure of the experimental situation.

A marked practice effect was demonstrated within a single 50-minute session even though a pre-experimental series of responses was provided.

Reaction time depends also upon the quadrant of the visual field in which the discrimination is to be made. Latency is significantly greater when the critical portion of the stimulus figure appears in the lower half of the visual field than when the upper half is involved, a finding which is contrary to previously reported results.

---

[15]Woodworth, R. S. Experimental Psychology. New York: Henry Holt, 1938, p.327.

PROBLEM II

Comparison of Reaction Time with Landing Skill

After establishing the relationship between the experimental variables and response time, the next step was to look for a correlation with the criterion, skill in landing. This criterion was derived from the ratings by flight instructors on two types of scales and from the ratings by federal examiners of the selected aspects of the student pilot's skill in landing his airplane.

A. THE RATING INSTRUMENTS

Three types of ratings of the pilot's skill in landing were obtained, two from flight instructors and one from the C.A.A. flight examiner.

1. Each instructor placed his students in rank order, giving a rank of one to the student who showed the greatest skill in landing, a rank of two to the next best student, etc. There were five instructors, each having 8 to 10 students.

2. Each instructor rated his students on a rating scale for landings. This scale was slightly modified from one appearing in an early form of the Ohio State Flight Inventory.[16] The scale appears in Appendix A.

3. The C.A.A. flight examiner rated students on a 5-point scale for landings, ranging from (1) excellent, through (5) poor.

Ratings for four separate landings were made on the Ohio State Flight Inventory, two on each of two days. Ratings on all students were obtained when they reached the same stage in their training course - just prior to making their cross-country flights. In view of the general lack of high validity of ratings it was decided to treat them in the following manner: The ratings on the various sections of the Ohio State Flight Inventory were given letter grades of A, B, or C to indicate good, fair, or poor skill in the various aspects of landing. Then, from inspection of the total pattern, each individual pilot was assigned a rating of 1, 2, or 3 (good, fair, or poor).

The rankings and the ratings by instructors were both considered in selecting a group of 10 students considered best and another group of 10 considered poorest. Since each flight instructor had about 10 students, two high and two low students were picked from each sub-group. Only those students were selected who were consistently high or low on both (1) the rank order, and (2) the ratings on the Ohio State Flight Inventory scale.

Various comparisons (see discussion of results) on speed of reaction were made between the group rated high and the group rated low. The fact that the 48 students were not trained by the same instructor imposed a limitation on the effective selection of criterion groups. Unless good and poor prospects were equally distributed among the five instructors the criterion groups derived from the composite ratings were not fully efficient. A student considered good in one group may have seemed only mediocre in another. The simplest

[16] The early version of the Ohio State Flight Inventory is described in: Edgerton, H. A. & Walker, R. Y. Criteria of Flight Competence. Progress report, September 1940. Washington, D. C.: National Research Council Committee on Selection and Training of Aircraft Pilots, 1940. (Copy in Committee files.)

solution (though impossible in the present study) would have been to have the subjects all rated by competent observers who were not directly connected with the training program. A second difficulty was that although the ratings were made at the time when all students had reached the same stage of training, they had not all progressed at the same rate. In effect, therefore, the individual differences are kept at a minimum by rating a common achievement level rather than keeping the amount of practice constant.

B. TREATMENT OF RESULTS

The ten fliers rated as best in landing and the ten rated as poorest were compared with respect to reaction-time responses. When a student failed to make the correct response (missed items), the highest score made by that student for the same stimulus at some other place in the series was assigned. This procedure was followed because it was necessary in the computations to have a numerical score for every item.

In the comparisons to be made between the "good" and the "poor" landers a distinction is made between the preliminary or practice series and the main or experimental series.

C. RESULTS

Table III presents a comparison of the means for the good and poor landing groups in the practice and the main series:

TABLE III

	Mean R.T. main series	Mean R.T. practice series	Mean misses
Good Landers	59.2	67.7	3.1
Poor Landers	56.0	77.0	5.5
t values (signif.)	0.68	7.10	1.22

The t-values in the above table suggest that two of the differences in means are due to chance[17] while the difference in mean reaction time in the practice series is significant (t-value - 7.10). However, if <u>one</u> atypical case is eliminated from the series, the t-value of 7.10 drops to 0.49 signifying a chance difference only.

Similarly the mean variance is significantly different for the two groups, but with the elimination of the atypical case the difference becomes a chance one.

---

[17]The t-value for these comparisons of means would have to be 2.101 to be significant at the 5% level, or 2.878 at the 1% level.

When given their flight test by the C.A.A. flight examiner, the students were rated on a 5-point scale. Actually they were all given a "2" or a "3" rating on landing. The data for these two groups ("2" and "3") are analyzed and presented in Table IV.

TABLE IV

	Mean R.T. main series	Mean R.T. practice series	Mean Misses main series	Mean Misses practice series
The "2" group	55.3	74.1	2.68	1.42
The "3" group	57.1	74.3	4.13	0.94
t-values (signif.)	0.67	chance difference	1.20	1.18

Here also the differences are not significant and are unstable. It should be noted, however, that in three of the four comparisons those rated as poor had slightly longer reaction times and made slightly more misses than those rated as good.

C. DISCUSSION AND CONCLUSIONS

In all the comparisons, when a single atypical case is eliminated the differences become insignificant and unstable. This indicates that reaction time to peripheral stimuli does not differentiate between good and poor landing skill. Such a conclusion, however, must rest on the assumption that the criterion of landing skill is a valid one. The writers are now very skeptical about the ratings made by the flight instructors[18] and by flight examiners. Apparently the instructors are unable to give a valid rating even with the help of the scale used. One is not justified, therefore, on the basis of these data in concluding that quickness of response to peripheral stimuli cannot diagnose degrees of skill in landing. The crucial test awaits a more adequate criterion of landing skill.

D. EXTENSIONS OF THE PROBLEM

There is a possibility that speed of reaction to peripheral stimuli may have an application in at least one other aspect of flying, namely formation flights, both during daylight and at night or twilight. Obviously here is a situation where quick and accurate response to stimuli appearing off in the periphery of the visual field is perhaps crucial. The technique as described here should be

---

[18] The writers were surprised by what may be interpreted as a lack of cooperation on the part of some flight instructors. They went through the motions of aiding us but obviously the affair was not taken very seriously. In fact, one instructor stated that he was too busy to give more than part of the rating requested.

[illegible] were for measuring reaction time to peripheral stimuli. It would be necessary to work in collaboration with officers of the government at some large government aviation field where flight formation is taught if this problem were to be investigated. The criterion would have to be in terms of actual performance during flight in formation. A more complex task such as formation flying might differentiate students where the simpler requirements of landing fail.

An improvement in differentiation might be obtained by complicating the experiment. A possible illustrative variation of this sort could be discrimination of movement rather than of stationary inhomogeneity in the visual periphery. Also, the reaction time itself might be further complicated.

APPENDIX A

A sample of the scale used in this study for rating landings (slightly modified from an early form of the scale in the Ohio State Flight Inventory).

[illegible]

Lat. B[illegible] [illegible]	Glide		During Roll		[illegible] Glide		[illegible] Roll	
	1st	2nd	1st	2d	1st	2d	1st	2d
Constant	()	()	()	()	()	()	()	()
Not more than 5° r. or l. deviation	()	()	()	()	()	()	()	()
Not more than 10° r. or l. deviation	()	()	()	()	()	()	()	()
More than 10° r. or l. deviation	()	()	()	()	()	()	()	()
Continuous fluctuation of less than 10°	()	()	()	()	()	()	()	()
Continuous fluctuation of more than 10°	()	()	()	()	()	()	()	()

LEVEL OFF

() () Starts to level off proper distance from the ground
() () Starts to level off 5 ft. above proper distance
() () Starts to level off 5 ft. below proper distance
() () Levels off more than 5 ft. above proper distance
() () Levels off on the ground
() () Pushes stick forward to resume glide within 10 ft. of the ground

LANDING
() () Three-point landing
() () Stalls ship 3 ft. off ground, "pancakes"
() () Stalls ship 5 ft. off the ground
() () Stalls ship 10 ft. off the ground
() () Wheel landing
() () No drifting or skidding on landing
() () Slight drift or skid on landing
() () Drifts or skids on landing
() () Recognizes bad landing
() () Pulls out of bad landing and makes recovery
() () Fails to pull out of bad landing

PRECISION
() () Lands on spot
() () Lands on spot or within 100 ft. beyond
() () Lands on spot or within 200 ft. beyond
() () Over-shoots spot landing more than 500 ft.
() () Under-shoots spot landing less than 50 ft.
() () Over-shoots spot by more than 300 ft.
() () Under-shoots spot by more than 50 ft.
() () Over-shoots field
() () Under-shoots field

CONDITIONS
Into wind________________
Cross wind_______________
Down wind________________
Gusty air________________
Forced landing___________
Wind velocity____________