RESTRICTED Not for publication

STUDIES OF PREDICTORS OF ACHIEVEMENT IN LEARNING TO FLY

Four Early Investigations of Aptitude and Personality Tests in Relation to Flying

Reports on research conducted by means of grants-in-aid from the National Research Council Committee on Selection and Training of Aircraft Pilots from funds provided by the Civil Aeronautics Administration.

March 1944

CIVIL AERONAUTICS ADMINISTRATION
Division of Research
Report No. 27
Washington D. C.

National Research Council

Committee on Selection and Training of Aircraft Pilots

Executive Subcommittee

M. S. Viteles, Chairman

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

National Research Council

1944

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

March 13, 1944

Dr. Dean R. Brimhall
Director of Research
Civil Aeronautics Administration
Washington 25, D. C.

Dear Dr. Brimhall:

Attached is a report entitled Studies of Predictors of Achievement in Learning to Fly, embodying the results of experiments conducted independently by John P. Foley, Jr.; Thelma Hunt, E. Lowell Kelly and A. P. Johnson, and William M. Lepley. This report 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 issued by the Division of Research, Civil Aeronautics Administration.

These experiments represent four of the earliest selection research studies sponsored by the Committee on Selection and Training of Aircraft Pilots. The methods and data presented in these studies are of interest, not only in terms of the positive results obtained, but also in presenting the negative findings. The report of negative findings may help prevent fruitless repetition of experimental work on tests and measures which have already been tried out and discarded because of their unpromising character. The studies again emphasize the need for extended research on criteria of flight proficiency.

Cordially yours,

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

MSV:bs

EDITORIAL FOREWORD

In 1939 and 1940, the Committee on Selection and Training of Aircraft Pilots was first faced with the problem of selecting and classifying applicants for flight training. It seemed well to include, in the early phases of the research program, exploratory studies of existing instruments possessing a reasonable degree of "face validity." Such studies, therefore, were set up at several universities.

Presented in this report are the results of four of these exploratory investigations. These studies made use of a wide variety of the available physiological and psychological instruments with the hope that some of them would prove of predictive significance in the selection and training of pilots.

Certain limitations must be noted in connection with these experiments: The tests were administered, for the most part, to "in-training" students who already had been pre-selected on the basis of the physiological and psychological standards represented in the C.A.A. flight physical examination. They were also selected on the basis of certain other criteria, such as grades, class, interest, etc., established by the local universities. The investigations were conducted at different centers and on different populations. The results are therefore not directly comparable either from experiment to experiment or from these experiments to studies on the present candidates for flight instruction. The number of cases involved in the studies is not large, limiting the conclusiveness of the results. Moreover, the criteria of flight performance available at the time of these investigations were not standardized or refined and are not comparable to the criteria now available.

In spite of these limitations, these studies were productive in delineating promising areas of investigation which were further exploited and have produced positive and helpful results described in various of the earlier reports in this series. The biographical inventory, used so extensively in the selection of pilots, is a good example of one positive outcome from an initial exploratory study. The studies also suggested the requirements for more intensive work in the field of criteria, and helped to produce the Ohio State Flight Inventory, the Purdue Rating Scale and other developments for improving ratings of flight performance.

The exploratory studies were of value not only in producing positive leads but also in indicating areas which appeared to offer little promise of fruitful outcomes for continued investigation. The studies are published not only by reason of their positive findings, but because they may also help investigators avoid the repetition of work which has already been done and not extended because of its unpromising character.

CONTENTS

INDEX OF VARIABLES INVESTIGATED IN EACH OF THE FOUR STUDIES

Measures of Flight Competence	Part
C.A.A. Examiner's Rating	III
Chief Instructor's Over-all Rating	III
Composite Rating (Flight Rating Index, Chief Instructor, and C.A.A. Examiner)	III
Flight Instructor's Daily Ratings from Students' Log Books	I,II
Flight Rating Index (Composite Instructor's Rating)	III
Flight Test Scores for Pilot Certificates	I
Grades in Navigation and Meteorology	III
Graphic Rating for General Aptitude for Aircraft Operation (Chief Instructor)	IV
Graphic Rating on 9 Aspects of Flight Performance	IV
Paired-Comparison Rank of Student Pilots by Instructors	I,II
Pass-Fail	II
Rating of General Confidence in Student as a Pilot Made by Operator (Chief Instructor) of Field	IV
Ratings of Local Committee on Selection	IV
Ratings of Relative Competence (Superior vs. Inferior Pilots)	II
Time in Air	I

Physiological and Psychological Measures	
A.C.E. Psychological Examination	III
Age	I
Bernreuter Personality Inventory	I,II,IV
Blood Pressure (Systolic, Standing, and Reclining)	IV
Coordination and Balance Tests	IV
Ataxiameter (Eyes Closed)	IV
Ataxiameter (Eyes Open)	IV
Differential Pressures (Squeeze)	IV
Finger-to-Finger Coordination	IV
Foot Span	IV
Hand Separations	IV
Horizontal Linear Space	IV
Knee Bend	IV
Reproduction of Standing and Leaning Postures	IV
Step to Line	IV
Walk to Line	IV
Depth Perception (Howard-Dohlman Test)	III
Dexterity Test (Number of Pins in Proper Hole)	III
Finger Dexterity Test	III
Flack-Woolham Index	III
Hearing Loss (Audiometric Test)	III
Height	I
Humm-Wadsworth Temperament Scale	III
Hunter Test of Social Attitudes	I,II
Iowa Placement Test in Mathematics	III
Medical Examiner's Composite Rating	IV
Minnesota Spatial Relations Boards (Four Form Boards)	III

SUMMARY

In the beginnings of any large scale research program undertaken with the view to predicting success or failure (or competence) in such a task as flying, two problems appear immediately: First, there must be some means of evaluating the task itself, i.e., the case of aircraft pilots, methods must be available by which the "good" flyers can be distinguished from the "poor" ones. Secondly, a battery of tests or measures must be assembled which will (with as high a degree of assurance as possible) predict performance on the task, i.e., distinguish between those who do well and those who do poorly in learning to fly.

During the early phases of research on pilot selection, measures already in use for evaluation flight performance were accepted as criterion scores for the task of flying. Such measures as instructor's day-to-day ratings of the student pilots, inspector's ratings of performance during flight tests, and "passing" or "failing" in flight training were among the most prevalent ones. To the extent that these measures are unstable and unstandardized any program aimed at accurate prediction of ultimate performance is seriously handicapped.

One of the first steps in an exploratory study in the field of pilot selection is to choose from the existing tests those which possess a high degree of "face validity," i.e., those which, on the basis of content or makeup, "seem" to bear a direct association with the task of flying, particularly with the aspects being used for evaluating the task. Those instruments are then subjected to exploratory experimentation to provide an estimate of the degree to which they will predict flight performance. The procedure is to administer all of the selected psychological and physiological tests to a group of flight students who are in training at the time or who are just about to begin their training. After various stages of their training are completed, it is then possible to compare their ratings in flight performance with their scores on the various tests to determine if there is any direct relationship between the two variables.

Such preliminary investigations are usually undertaken with small groups of subjects and are apt to be loosely designed or poorly controlled. For these reasons, definite conclusions often cannot be drawn. Such studies serve only as a quick means of sorting out the most promising tests and measures for further and more exacting experimental study or as a basis on which to devise new and more applicable instruments. They also provide experimental data on certain measures of flight performance, and may thus suggest more accurate bases for the evaluation of flight performance itself.

This report presents the results of four early investigations of this type, sponsored by the Committee on Selection and Training of Aircraft Pilots. These studies are exploratory in nature. They were undertaken with the view of evaluating certain of the existing physiological and psychological tests in terms of their association with current criterion measures of flight performance.

Various measures of flight performance were used in these studies, including: (1) passing or failing in flight training, (2) day-to-day ratings by flight instructors, (3) flight test ratings by inspectors, and (4) paired-comparison rankings of students toward the end of their training.

Among the psychological tests investigated were: (1) personality tests, (2) intelligence tests, (3) attitudes tests, (4) interests inventories, (5) scholastic tests or indices, etc.

The physiological measures evaluated include: (1) respiratory measures, (2) cardiovascular measures, e.g., heart rate, blood pressure, etc., (3) physical fitness indices, (4) visual and auditory measures, (5) over-all medical ratings, and (6) such data as height, weight, and age.

Other varieties of tests were also analyzed in these studies, namely: (1) measures of bodily steadiness and body sway, (2) "tension," (3) manual dexterity, (4) depth perception, (5) reproduction of postures, (6) reaction time, etc.

No attempt is made in this summary to present all the findings of the four studies. General conclusions can, however, be briefly stated. The studies almost uniformly demonstrated that, in spite of "face validity," the physiological and psychological tests or measures studied had little or no relationship to the task of flying, as here evaluated. It was not possible to distinguish between the "good" and the "bad" pilots, or the "passers" and the "failers," on the majority of these tests. Pilot populations, however, were shown to score differently on certain of the tests from groups of students who were not taking flight training. It should be pointed out that some of the psychological tests did show an appreciable association with one of the ratings of flight competence. The usefulness of this finding is limited, however, because of the fact that these ratings are of questionable reliability and validity.

Another, and perhaps the most important generalization which could be drawn from these studies, was that none of the different measures of flight performance employed in these studies provided a satisfactory estimate of success in flying. What is needed, therefore, is a more intensive study of the criteria of flight competence.

As was pointed out earlier in this summary, the generalizations made on the basis of these studies must be accepted with caution. First, because of the small number of individuals involved in the studies, and secondly, for the reason that when larger populations are used and more refined measures of flight performance itself are isolated it is possible that many of the psychological and physiological tests investigated in these experiments will show a significant degree of association with the criterion.

PART I

AN OBJECTIVE MEASURE OF BODILY STEADINESS IN RELATION TO FLIGHT PERFORMANCE

by

JOHN P. FOLEY, JR.

The George Washington University

AN OBJECTIVE MEASURE OF BODILY STEADINESS IN RELATION TO FLIGHT PERFORMANCE

INTRODUCTION

In a paper by J. G. Jenkins[1] it was pointed out that one of the major areas of research in aviation was the investigation of the psycho-physiological responses of student pilots. It was the purpose of this experiment to study one of the responses falling in this area, namely, the bodily steadiness of student pilots.

Specifically, this investigation included an analysis of the pre- and post-flight measures of bodily steadiness of a representative group of 10 student pilots. The data from these experiments were analysed with two aims in view: (1) to investigate the immediate influence of flight lessons upon the measures of bodily steadiness, and (2) to study the relationship between bodily steadiness and (a) a criterion of flying success, (b) an objective index of "tension," and (c) scores on standard psychological tests.

APPARATUS

Measures of bodily steadiness were taken from the "Wabblemeter," an instrument developed by Moss and Brown [2] for the purpose of measuring bodily steadiness as an indication of riding comfort. The apparatus, as employed in this study, consists of a platform, 15" x 12 3/4", on which the subject stands. The platform rests upon a central fulcrum, and loss of balance or unsteadiness is measured by the extent to which the platform is moved or "wabbled" within a given period of time during which the subject attempts to keep it as nearly level and balanced as possible. The platform can be locked into a stationary horizontal

---

[1]Jenkins, J. G. Report at round table on selection and training of civilian airplane pilots (unpublished). Eleventh annual meeting, Eastern Psychological Assoc., Atlantic City, N. J. April 5, 1940.

[2]The Wabblemeter was originally developed by F. A. Moss, of the George Washington University, and R. W. Brown, Director of Engineering Laboratories, Firestone Tire and Rubber Co., Akron, Ohio, under the auspices of the Riding Comfort Subcommittee of the Research Committee of the Society of Automobile Engineering. Dr. F. A. Moss furnished the Wabblemeter used in this investigation.

Anon. Review of riding-qualities research. S.A.E. Journal, 1931, 28, No. 5, 577.

Brown, R. W. Measuring riding qualitites. S.A.E. Journal, 1931, 28, No. 5, 577-580.

(a) Moss, F. A. Development of the Wabblemeter for measuring bodily steadiness as an indication of riding comfort. S.A.E. Journal, 1931, 28, No. 5, 580-581.

(b) _____ Further experimental work with the new Wabblemeter. S.A.E. Journal, 1931, 29, No. 3, 243-246.

(c) _____ New riding-comfort research instruments and Wabblemeter applications. S.A.E. Journal, 1932, 29, No. 4, 182-184.

position before and at the termination of the test period by means of two brakes manipulated by the experimenter. The following characteristics and advantages of the present Wabblemeter may be mentioned:

1. Automatic recording. The [illegible] work by a lever arrangement, the movements of the platform being recorded mechanically and automatically on two Veeder counters. This instrument thus eliminates the necessity for electrical connections or for oil, which were necessary in some of the cruder and earlier machines.

2. The machine integrates all the movements made when the subject stands on the platform. Moss points out[3] that two of the earlier machines were based on the all-or-none principle the record of wabble being maximum if sufficient to record at all. Differences in records, therefore, occurred only with differences in the numbers of wabbles. In the present instrument, however, the exact recording count depends upon the extent of the wabble. All movements are recorded according to extent.

3. Compactness. The instrument is very small, making for ease of shipment and transportation, and is portable. It is self-contained, there being no extraneous batteries or other equipment necessary.

4. The instrument is easily set up and easily used.

5. It is practically noiseless in operation.

6. It is not influenced by temperature changes, or by other extraneous environmental factors. Moss points out that lack of this quality was a distinct disadvantage in some of the earlier machines, especially in the oil-pumping type.

The Wabblemeter was devised as an objective measure of bodily steadiness and fatigue which would be reliable and free from subjective error. Initial attempts to obtain fatigue measures tended to be confined to the known physiological tests of fatigue. Of these, the carbon-dioxide combining power of the blood appeared to be most accurate, although it was abandoned because of the large sample (5-10 cc.) of blood required. After much experimental work, Moss concluded that bodily unsteadiness was an excellent criterion of fatigue. Preliminary investigations indicated that simple tests, such as of eye-hand coordination,[4] were unsatisfactory, since fatigue tended to lower the steadiness of the body as a unit. It was thus felt that if the effect of fatigue was to be measured from only one portion

---

[3]Moss, F.A., op. cit. (a), p. 580.

[4]An example of such an attempt to measure steadiness is found in Swaps, A. and Brandenburg, G.C. The steadiness meter and its use in a preliminary investigation of riding qualities. S.A.E. Journal, 1931, 28, No. 5, 582-583.

of the body, it should be from the legs.[5] This led to the conception of the wabblometer, an instrument upon which the subject stands, and which records in two planes the ability of a person to maintain perfect vertical equilibrium. Early models necessitated the application of a stencil to the charted record on a graphic recording machine; another, the "oil-pumping" variety, was scored in terms of the amount of oil pumped. The present and most recent instrument records mechanically, and has now been developed into a relatively compact, accurate form.[6]

The wabblometer has been previously used by Moss in a series of road tests with automobile passengers.[7] He reports progressive increase in wabblometer score on long-distance trips, the increase bearing a positive relationship to distance travelled but varying with differing conditions of the road and other more extraneous driving conditions.

The only previous application of wabblometer tests to airplane flights was made by Moss in the summer and autumn of 1931.[8] With the cooperation of the Ludington Lines,[9] wabblometer tests were made on an unreported number of airline pilots and passengers before and after flights of more than 200 miles between Newark, N. J. and Washington, D. C. Moss reports that "the pilots tested showed an average increase of 42% in wabblometer record after the flight, the least increase shown by any pilot tested being 16%, and the greatest, 73%. The passengers tested showed an average increase of 77%, the least being 24%, and the greatest, 161%.....Pilots, as a group, showed a considerably more steady normal of before-flight records on the wabblometer than did passengers. The passengers showed a much higher percentage of increase in wabble after the flight than did the pilots. The increase in wabblometer readings for the passengers is somewhat greater than is the increase for an equal distance covered by usual means of travel. This does not hold true, however, in the case of pilots. This might suggest a factor of importance in producing unsteadiness after airplane riding, other than fatigue from vibration and lack of smoothness in operation of the plane, which is absent in the pilot accustomed to flying; that is, a factor akin to 'air-sickness'."[10]

Moss concludes his article with the following three suggested applica-

---

[5]Brown, R.W., op. cit., p. 579.

[6]The Wabblometer, originally made by the Pioneer Apparatus Co., Brooklyn, N.Y., was manufactured by The American Instrument Co. of Washington, D.C.

[7]Moss, F.A., op. cit., (b) and (c).

[8]Moss, F.A., op. cit., (c). The data are unfortunately not reported in full in the original article; the present resume is complete with respect to published accounts of the use of the wabblometer in airplane research as of 1941.

[9]Now Eastern Air Lines.

[10]Moss, F.A., op. cit., (c), p. 183.

"(1) Testing of workers of [illegible] to be compared with other modes of travel.

"(2) Testing of condition of pilots before flight; for example, if a pilot, whose normal record is about a 20 reading of the wabblemeter, shows a reading of 35, his fitness for flight on that particular day should be carefully checked.

"(3) Testing of recruits for aviation training, after a study has been made of the correlation between steadiness as measured by the wabblemeter and good flying."

It will be noted that these suggestions are directly in line with the present investigation.

It should be emphasized that although Moss contends that the present instrument is directly a measure of bodily steadiness and indirectly an indication of fatigue, considerable caution should be exercised in such interpretations. It is highly probable that other factors, such as the subject's willingness to cooperate (i.e., to maintain as nearly perfect balance as possible) enter in to influence the actual wabblemeter score. It should be emphasized, however, that the diagnostic and prognostic value of the wabblemeter as a test of flying aptitude is in no way impaired by the complexity of such contributing factors, since the validity of the instrument for this purpose is determined solely by the extent to which it correlates with an independent criterion of flying ability.

PROCEDURE [12]

The exact procedures followed in taking the wabblemeter records may be summarized under the following outline headings:

1. Posture on the instrument. In a preliminary series of experimental sessions (prior to the first instruction flight), the subject was instructed regarding the position to be taken on the instrument. He was asked to stand upright and in a natural and relaxed posture, with feet placed symetrically on either foot-rest as far forward as possible, with toes flush with the front of the platform. He was asked to fixate the horizon at all times, and was instructed to make every effort to balance at all times.

2. Method and duration of timing. All records were timed with a stop-watch which was started by the experimenter at the moment he released the brakes which had previously immobilized the platform. At the end of a 1-minute test period,[11] the brake levers were closed.

---

[11]This 1-minute test period, as in the case of other aspects of the procedure, is suggested by Moss, F.A., op. cit., (b).

[12]The data for this study were collected and recorded by Mr. J. T. Wilson assistant in psychology, to George Washington University (now Lt. j.g., USNR).

3. Absence of distraction. During the test proper, an effort was made to prevent all conversation with the subject or with others close to the subject. An attempt was also made to face the subject in such a direction that people and other objects would not be moving about in his visual field.

4. Times of test. The pre-flight test was made just prior to the student's entering the plane, and at a place outside of the propeller stream. The post-flight test was made as soon as the student had left the plane. In fact an effort was made to prevent him from leaving the plane until time for his test. The same procedure was followed, as nearly as possible, in testing the instructor before and after his first daily flight.

5. Recording of data. Before each test period, the wabblemeter readings on both counters were recorded. These were subtracted from the final readings following flight, the difference indicating the increase or decrease in wabble following flight. All readings were recorded on a special Data Sheet (see Figure 1) as soon as possible, and the remainder of the sheet was filled out in full from the Student Pilot Log Book.

There are a number of variables which must be considered and controlled in securing data on the wabblemeter or similar instrument. Among these are such factors as diurnal variation, sex differences, age differences, the effect of weight, the effect of height, and practice effects. Since the readings on each of the subjects participating in this study were taken at various times of the day over a period of 3-4 months, diurnal factors are effectively eliminated. Data on this group of 9 male and 1 female subjects were treated separately, before combining, so that possible sex differences in wabblemeter performance would not mask the actual results. The use of pooled data, in so far as the present group of 10 subjects may be considered representative of the total C.A.A. college population, would adequately take into consideration any possible effects due to age, weight, or height. Since none of the subjects had been tested on the wabblemeter before, and since all had completed the same ground course and were just beginning flight instruction, practice effects in performance on the instrument would be approximately equal in the case of all subjects. In fact, were the wabblemeter to prove useful in pilot selection, such practice effects would be an intrinsic part of the actual data if the latter were to be taken over a period of time, and consequently should not be eliminated except with respect to their differential operation.

As indicated above, the wabblemeter data consisted of pre-flight and post-flight records as well as differences (increase or decrease in wabble) for each student during Stages A, B, and C of flight training. Wabblemeter data were also obtained on 16 flights made by the instructor. The amount of time spent in the air on each flight was also recorded. In addition to the wabblemeter records, the following data were obtained for interpretive and evaluative purposes:

1. Physical data. Data such as sex, height, weight, and age were obtained from the Student Pilot Log Book.

Student _____ Age _____ Group _____ Stage _____

Daily Flight Record:

Make and Model Ship Piper Cub

Identification Number NC 2500Y

Flight Number _____

Time Up _____

Time Down _____

Time In Air _____

Instr's. Rating _____

Cumulative Flight Record:

Previous Flight Time _____

Time This Flight _____

Total Flight Time _____

Remarks _____

Pre-Flight Wabblometer Record:

	Front	Back
Final	_____	_____
Initial	_____	_____
Difference	_____	_____

Sum of 2 Differences _____ (A)

Post-Flight Wabblometer Record:

	Front	Back
Final	_____	_____
Initial	_____	_____
Difference	_____	_____

Sum of 2 Differences _____ (B)

(B) minus (A) ____ - ____ = _____ Increase following flight.

Figure 1.

2. "Tension" data.[13] These data, indicating "excessive landing tension," were secured by means of the Fries Flight Analyzer. The "tension index," furnished by Dr. A.C. Williams as the result of his research on "excessive landing tension," was defined as "tension on the stick at landing resulting in a deflection of the recording pen on the Flight Analyzer greater than 0.2 of an inch."

3. Psychological test data.[14] These data were furnished by Dr. Thelma Hunt. Scores on the following tests were recorded: Pressey Interest-Attitude Tests, Hunter-Test of Social Attitudes, Bernreuter Personality Inventory, and the Strong Vocational Interest Blank.

4. Criterion data. The criterion data employed in this study were: (a) mean of the flight instructor's daily ratings from the Student Pilot Log Book, (b) paired-comparison rank of the 10 student pilots by the instructor near the completion of flight training, and (c) flight test scores for pilot certificates.

RESULTS

The present group of 10 subjects, 9 male and 1 female, range in age from 19 to 25, with a mean age of 20.5 and S.D. of 1.69 (see Table 6). The mean height is 5' 10.2", with S.D. of 1.75; and the mean weight is 155.6 lbs., with S.D. of 13.77. The sole female subject falls well within the male range in all physical measures. In general, the physical data would seem to indicate that the subjects are typical of the usual student population. It should be emphasized at the outset that the present group is too small to justify detailed statistical treatment of the data. The present study should be considered primarily as exploratory and preliminary; data on additional cases are needed before conclusive results can be ascertained.

In Tables 1-3 inclusive will be found the wabblometer and tension data, together with the instructor's ratings, on each of the 10 subjects for Stages A, B, and C of flight training, respectively. In the case of each subject, the mean wabblometer records have been reported for pre-flight and post-flight readings, as well as for the difference (increase or decrease) in wabble following flight. Such differences are further broken down into those obtained with flights of less than 45' and those following

---

[13] The writer is indebted to Dr. A.C. Williams, Department of Psychology, University of Maryland, for making these data available. For further details regarding the tension index, see Williams, A.C., Automatic recording of muscular tension during flight (Progress report, September 1, 1940) Washington, D.C.: N.R.C. Committee on Selection and Training of Aircraft Pilots, September, 1940. (Copy in Committee files.)

[14] The writer is indebted to Dr. Thelma Hunt, Department of Psychology, The George Washington University, for making these data available. For further details regarding the psychological test results, see Part II of this report.

TABLE 1 - WABBLEMETER DATA, TENSION DATA, AND INSTRUCTOR'S DAILY RATINGS (IN LOG BOOK) ON 10 STUDENT PILOTS IN STAGE A OF FLIGHT TRAINING

SUB. NO.	Flights	Records	PRE-FLIGHT			POST-FLIGHT			DIFFERENCE			DIFFERENCE WITH RESPECT TO TIME IN AIR						TENSION	INSTRUCTOR'S RATING (IN LOG BOOK)		
												LESS THAN 45'			MORE THAN 45'						
			MEAN	S.D.	RANGE	MEAN	S.D.	RANGE	MEAN	S.D.	RANGE	MEAN	S.D.	RANGE	MEAN	S.D.	RANGE	DATA	MEAN	S.D.	RANGE
1	16	11	105.8	10.0	90 to 123	94.9	9.7	82 to 108	-10.9	14.9	-31 to +13	-10.9	14.9	-31 to +13	-	-	-	3	3.8	0.6	3 to 4.5
2	18	7	80.9	12.0	65 to 102	81.7	4.1	73 to 86	+0.9	14.8	-29 to +20	+0.9	14.8	-29 to +20	-	-	-	1	3.4	1.8	3 to 5
3	17	9	103.6	13.5	87 to 128	89.4	6.4	81 to 102	-14.1	14.7	-34 to +15	-14.1	14.7	-34 to +15	-	-	-	8	3.4	0.5	3 to 4.5
4	16	10	124.3	16.6	101 to 152	109.1	19.6	83 to 142	-15.1	11.9	-29 to +6	-15.1	11.9	-29 to +6	-	-	-	4	3.8	0.6	3 to 4.5
5	17	8	142.4	26.6	110 to 202	133.1	16.7	107 to 153	-9.3	17.7	-50 to +4	-9.3	17.7	-50 to +4	-	-	-	10	3.7	0.7	3 to 5
6	15	8	81.9	6.0	71 to 93	95.0	8.7	77 to 107	+13.1	8.0	-3 to +22	+13.7	8.4	-7 to +22	+3.0	0.0	one flight	6	3.6	0.7	3 to 4.5
7	17	8	113.4	19.3	91 to 146	106.7	22.9	73 to 135	-6.6	10.1	-22 to +8	-6.6	10.1	-22 to +8	-	-	-	5	3.8	0.7	3 to 5
8	19	8	153.1	14.2	135 to 176	134.0	14.8	100 to 155	-19.1	14.6	-37 to +6	-19.1	14.6	-37 to +5	-	-	-	7	3.8	0.8	3 to 5
9	24	10	82.9	12.1	63 to 104	81.6	8.0	70 to 95	-10.4	8.6	-22 to +7	-10.4	8.6	-22 to +7	-	-	-	2	3.9	0.0	3 to 4.0
10	30	13	176.1	16.4	137 to 198	166.9	20.6	136 to 208	-6.9	19.6	-45 to +33	-3.9	16.7	-35 to +33	-43	0.0	one flight	9	4.0	0.8	3 to 5
AVE.	18.9	9.2	119.3	15.6		111.4	14.3		-8.3	14.3		-8.8	13.9		-17.0				3.7	0.7	
S.D.	4.4	1.7	31.3	5.4		25.8	6.5		5.7	3.5		9.0	3.2		26.0				0.2	0.2	

TABLE 2 - WABBLEMETER DATA, TENSION DATA, AND INSTRUCTOR'S DAILY RATINGS (IN LOG BOOK) ON 10 STUDENT PILOTS IN STAGE B OF FLIGHT TRAINING

SUB. NO.	Flights	Records	PRE-FLIGHT MEAN	PRE-FLIGHT S.D.	PRE-FLIGHT RANGE	POST-FLIGHT MEAN	POST-FLIGHT S.D.	POST-FLIGHT RANGE	DIFFERENCE MEAN	DIFFERENCE S.D.	DIFFERENCE RANGE	DIFFERENCE WITH RESPECT TO TIME IN AIR: LESS THAN 45' MEAN	LESS THAN 45' S.D.	LESS THAN 45' RANGE	MORE THAN 45' MEAN	MORE THAN 45' S.D.	MORE THAN 45' RANGE	TENSION DATA	INSTRUCTOR'S RATING (IN LOG BOOK) MEAN	S.D.	RANGE
1	8	4	84.8	4.5	81 to 92	77.5	5.6	72 to 83	-7.5	4.3	-11 to 0	-7.5	4.3	-11 to 0				3	2.8	0.3	2.5 to 3
2	11	3	61.7	1.7	60 to 64	63.7	7.4	54 to 72	+2.0	7.0	-6 to +11	+2.0	7.0	-6 to +11				2	3.5	0.8	3 to 4
3	10	6	105.5	11.9	84 to 121	85.2	5.5	61 to 110	-20.3	19.8	-45 to +9	-20.3	19.8	-40 to +9				4	3.2	0.9	3 to 4
4	10	2	93.0	0.0	93 to 93	101.0	12.0	89 to 133	+8.0	10.2	-4 to +20	+8.0	10.2	-4 to +20				5	3.0	0.0	3 to 3
5	11	5	114.3	13.3	103 to 135	113.3	7.4	104 to 122	-1.0	9.1	-13 to +9	-1.0	9.1	-13 to +9				10	3.0	0.0	3 to 3
6	20	2	81.5	5.5	76 to 87	88.0	2.0	85 to 90	+6.5	9.3	-1 to +14	+6.5	9.3	-1 to +14				1	3.0	0.0	3 to 3
7	5	3	139.0	11.8	124 to 153	127.0	3.0	122 to 130	-12.0	12.2	-23 to +5	-12.0	12.2	-23 to +5				7	3.0	0.0	3 to 3
8	8	2	140.5	6.5	134 to 147	137.0	3.0	134 to 140	-3.5	3.5	-7 to 0	-3.5	3.5	-7 to 0				8	3.0	0.0	3 to 5
9	8	1	84.0	0.0	84 to 84	69.0	0.0	69 to 69	-15.0	0.0	-15 to -15	-15.0	0.0	-15 to -15				6	3.0	0.0	3 to 3
10	10	5	187.7	12.6	176 to 200	180.0	15.5	[illegible]	-7.7	5.4	-13 to -2	-7.7	5.4	-13 to -2				9	3.0	0.0	3 to 3
AVE.	9.4	2.9	[illegible]	[illegible]		102.2	10.1		7.0	11.3		-7.0	11.4						3.0	0.8	
S.D.	[illegible]	[illegible]	[illegible]	[illegible]		[illegible]	[illegible]		9.2	5.9		9.2	5.9						0.2	0.4	

TABLE 5 - WABBLEMETER DATA, TENSION DATA, AND INSTRUCTOR'S DAILY RATINGS (IN LOG BOOK) ON 10 STUDENT PILOTS IN STAGE C OF FLIGHT TRAINING

SUB. NO.	Flights	Records	PRE-FLIGHT			POST-FLIGHT			DIFFERENCE			DIFFERENCE WITH RESPECT TO TIME IN AIR						TENSION	INSTRUCTOR'S RATING (IN LOG BOOK)		
												LESS THAN 45'			MORE THAN 45'						
			MEAN	S.D.	RANGE	MEAN	S.D.	RANGE	MEAN	S.D.	RANGE	MEAN	S.D.	RANGE	MEAN	S.D.	RANGE	DATA	MEAN	S.D.	RANGE
1	34	13	74.2	12.0	60 to 97	67.6	9.8	52 to 84	-6.8	10.2	-27 to +15	-6.3	11.9	-27 to +15	-7.8	3.6	-14 to -8	4	2.9	0.4	2 to 3.5
2	30	7	74.4	10.7	62 to 82	71.4	5.2	65 to 82	-3.0	8.6	-21 to +10	-6.5	0.5	-6 to -7	-1.6	9.8	-14 to +10	1	3.2	0.8	3 to 3.5
3	27	12	87.9	6.5	75 to 102	82.5	9.1	71 to 107	-5.4	9.1	-17 to +18	-4.5	6.1	-13 to +4	-5.9	10.3	-17 to +18	5	3.0	0.4	2.5 to 3
4	42	13	109.5	13.6	92 to 143	99.7	9.0	89 to 118	-11.2	13.5	-37 to +2	-11.6	14.0	-37 to +2	-9.0	10.0	-19 to +1	8	3.0	0.0	3 to 3
5	33	11	102.8	5.4	95 to 114	104.0	9.7	90 to 119	+1.2	10.2	-14 to +18	+2.3	9.5	-6 to +16	+0.8	9.6	-14 to +13	7	3.1	0.6	3 to 3.5
6	32	19	77.0	8.2	58 to 93	73.1	10.1	57 to 97	-4.1	8.4	-17 to +19	-10.4	10.2	-17 to +10	-2.1	7.8	-11 to +19	3	3.1	0.5	3 to 3.5
7	26	10	140.0	16.0	117 to 170	121.0	18.7	93 to 146	-19.0	26.6	-64 to +23	+11.5	16.3	0 to 23	-26.6	24.2	-64 to +12	8	3.3	0.2	3 to 3.5
8	26	9	100.3	21.3	72 to 141	85.3	16.5	64 to 120	-13.9	10.2	-24 to +10				-13.9	10.2	-24 to +10	9	3.4	0.2	3 to 3.5
9	29	5	80.4	6.2	73 to 90	79.6	4.1	76 to 86	-0.8	7.3	-15 to +6	+1.0	0.0	1 to 1	-1.3	8.1	-15 to +6	2	3.7	0.5	3.5 to 4
10	30	3	189.3	19.6	175 to 217	170.7	23.7	141 to 199	-18.7	42.1	-76 to +24	-76.0	0.0	-76 to -76	+10.0	14.0	-4 to +24	10	3.2	0.7	3 to 3.5
AVE.	30.9	10.2	96.4	12.8		89.0	11.8		-7.3	15.0		-7.8	10.8		-7.2	12.3			3.9	0.3	
S.D.	5.3	4.3	35.4	5.4		30.3	6.0		6.7	10.4		24.1	6.7		9.2	5.3			0.7	0.3	

flights [illegible] are the flights long enough to make this analysis [illegible] worthy of consideration. All wobblometer entries in these and subsequent tables are in terms of total number of wobbles recorded on the two counters during the 2-minute test period.

It will be noted that there are wide individual differences in both pre-flight and post-flight wobblometer records in all three stages of flight training. The same individual also shows considerable variability from flight to flight, although the overlapping in ranges from one subject to another is not excessive. In fact, the subjects near the two extremes, e.g., Subjects 2 and 1 as compared with subjects 10 and 7, show little or no overlapping in range of wobblometer scores. This result is also apparent in the combined data for all three stages of training, presented in Table 5, although the overlapping of individual ranges is naturally greater in this case owing to the larger number of readings for each subject.

Inspection of Tables 1-3 will reveal that in general the 10 subjects tend to maintain the same relative positions throughout the three stages of flight training studied, in terms of wobblometer performance. The individual records thus do not appear to be differentially affected by "practice." Thus, Subjects 2, 6, 4, and 1 usually rank among the first 4 in terms of low wobblometer scores, whereas subject 10 shows the most wobbles throughout the three stages of flight training, followed by subjects 7, 8, and 5.

There is a marked absolute practice effect, however, which is manifested in both pre- and post-flight indices. Thus the pre-flight mean for the group as a whole decreases successively from 119.29 to 110.13 to 96.39 for Stages A, B, and C, respectively, and the corresponding post-flight means are 111.40, 103.24, and 89.06. Inspection of Tables 1-3 will reveal that these decreases are not statistical artifacts, but are borne out in most of the individual cases. Thus the mean pre-flight scores for Subjects 1, 3, 6, and 9 and the mean post-flight scores for Subjects 1, 3, 4, 5, and 6 show consistent reductions throught Stages A, B, and C training. In most of the other cases there is also an over-all decrease, although the trend is not entirely consistent throughout the three stages. Only in the case of Subjects 7 and 10 does there seem to be indication of an increase in unsteadiness with practice, and this increase is not completely consistent in the post-flight means. It might be pointed out that Subjects 7 and 10 are among the unsteadier subjects, Subject 7 ranking ninth and Subject 10 tenth in both pre- and post-flight wobblometer scores. Further experimentation with additional subjects, however, is necessary in order to determine if the practice effect, or decrease in wobblometer score, is greater in the case of the steadier subjects. Finally, it is interesting to note that the increase in steadiness during training, manifested by most of the subjects, was achieved in spite of the longer flight periods in Stage C.

Perhaps the most outstanding fact to be noted from Tables 1-3 is that the subjects as a group tended to show a decrease in wobble following flight, in contrast to the results reported by Mons[15] for automobile as well as

---

[15] Mons, F.A., op. cit., (b) and (c).

TABLE 4

WABBLEMETER DATA ON SIXTEEN FLIGHTS MADE BY INSTRUCTOR
(all flights less than 45 minutes)

N		PRE-FLIGHT			POST-FLIGHT			DIFFERENCE			TIME IN AIR (flights less than 45 min.)		
Fl.	Rec.	Ave.	S.D.	Range	Ave.	S.D.	Range	Ave.	S.D.	Range	Ave.	S.D.	Range
16	16	107.1	12.3	83 to 129	99.5	11.0	84 to 128	7.0	11.2	-25 to +28	17.6	11.2	+2 to -28

commercial airline pilots and passengers. Although the ranges of the individual subjects indicate that both positive and negative differences between pre- and post-flight records are usually shown by all subjects in each stage of training, nevertheless 8 of the 10 subjects in stage A, 7 in stage B, and 9 in stage C show mean decreases in wabble following flight. This general result will be noted more readily from Table 5, which presents the wabblometer data for all three stages of flight training combined. Only 1 of the 10 subjects (Subject 6) shows an over-all mean increase in wabble following flight, and this increase (+1.41) is very small. From Table 4 it will be noted that the instructor also shows this decreased wabble or "exhilaration effect," with a difference of -7.56 between pre- and post-flight means.

It is difficult to account for the discrepancy between the present results and those previously reported by Moss. This difficulty is further augmented by the fact that Moss failed to report his aeronautics data in complete and quantitative form. The present results, if substantiated in further experimentation, cannot be attributed to the original "nervousness" before a flight on the part of a student beginning to fly for the first time, since the decrease in post-flight wabble is shown in all stages of flight training, and is evidenced long after the initial emotionalized attitudes have disappeared. It is possible, of course, that the periods of flight were not long enough to induce the bodily unsteadiness and fatigue necessary to produce an increased post-flight reading. The Moss airplane data, it will be recalled, involved a flight of approximately 200 miles, or more than 1½ hours. Inspection of Table 5, however, shows that the preponderance of mean post-flight decreases in wabble is just as great following flights of more than 45' as in the case of flights less than 45'. In both cases, 9 of the 10 subjects show decreases, and the respective mean decreases for the group as a whole are 7.80 with S.D. of 7.60 in the case of the shorter group of flights, and 7.50 with S.D. of 7.70 in the case of those longer than 45'. It should be pointed out, however, that the longer flights rarely exceeded 1 hour, or 1 hour and 15' at the most, so that it is at least possible that longer flight periods might give opposite results. Finally, it might be suggested that the decrease in post-flight wabble may be correlated with the "lift" which many of the subjects, as well as others with flying experience, introspectively reported. In fact several experienced pilots,[16] not knowing the results of the present investigation, have told the writer that they would expect a reduction in wabble following flight. Whether these results are indicative of such an "exhilaration effect," in which case Moss's data are open to question, or whether subsequent experimentation will indicate that the "exhilaration effect" is limited to re-

---

[16] e.g., Prof. Norman Ames, a World War pilot, and Director of The George Washington University C.A.A. Student Pilot Training Program.

TABLE 5 - VARIABILITY DATA
(Covering training period as a whole)

SUB. NO.	Flights	Records	PRE-FLIGHT			POST-FLIGHT			DIFFERENCE			DIFFERENCE WITH RESPECT TO TIME IN AIR					
												LESS THAN 45'			MORE THAN 45'		
			AVE.	S.D.	RANGE	AVE.	S.D.	RANGE	AVE.	S.D.	RANGE	AVE.	S.D.	RANGE	AVE.	S.D.	RANGE
1	58	28	88.1	10.4	60 to 123	79.7	9.3	52 to 108	-8.5	11.7	-3 to +15	-8.6	12.6	-27 to +15	-7.8	3.6	-14 to -3
2	59	17	74.8	10.4	60 to 102	74.3	5.3	54 to 86	-0.5	11.3	-29 to +20	-0.1	11.7	-29 to +20	-1.6	9.8	-14 to +10
3	54	27	97.0	20.5	75 to 128	85.4	3.6	61 to 110	-11.6	14.1	-48 to +18	-14.1	15.3	-48 to +25	-5.9	10.3	-17 to +18
4	68	25	114.1	14.4	92 to 152	103.6	14.4	83 to 142	-11.2	12.6	-37 to +20	-25.9	12.8	-37 to +20	-9.0	10.0	-19 to +1
5	61	22	119.0	17.3	95 to 202	115.9	12.5	90 to 153	-2.9	13.3	-50 to +18	-5.0	14.7	-50 to +18	+0.8	9.6	-14 to +13
6	57	29	78.7	7.9	58 to 96	80.1	9.5	57 to 107	+1.4	8.2	-17 to +22	+3.6	8.4	-17 to +22	-1.2	7.6	-11 to +19
7	51	21	129.7	16.9	91 to 170	116.4	19.2	73 to 146	-13.2	20.0	-64 to +23	-5.1	11.2	-23 to +23	-26.6	24.2	-64 to +12
8	53	19	126.8	17.4	72 to 176	111.5	14.9	64 to 155	-15.0	11.9	-37 to +10	-16.0	13.1	-37 to +5	-13.9	10.2	-24 to +10
9	61	16	82.2	10.1	73 to 104	80.2	6.3	69 to 95	-7.7	7.9	-22 to +7	-9.8	7.9	-22 to +7	-1.3	8.1	-15 to +6
10	70	19	181.3	16.7	137 to 217	169.3	20.4	136 to 203	-8.9	23.4	-76 to +33	-9.1	15.3	-76 to +33	-7.7	11.4	-43 to +24
AVE.	59.2	22.3	107.6	13.5		100.1	13.0		-7.7	14.0		-7.8	12.5		-7.5	12.1	
S.D.	5.0	4.5	30.8	31.0		27.5	5.3		5.2	4.4		7.6	2.5		7.7	5.3	

latively short flying periods, only to be determined by further investigation.[17]

Attention should be called to the fact that the rank-order correlation between pre- and post-flight wabblemeter scores for all stages of training combined is slightly over +.95. This serves as an indirect indication of the rather high reliability of the wabblemeter measures.

Let us now turn to a consideration of the relationship between wabblemeter scores and certain other experimental variables, one of the chief of which is the criterion or valid measure of flying ability. As will be seen from Table 6, recourse can be made to any of three different types of available criteria, in the absence of a more adequate measure.[18] One of these is the subject's score on the flight test given at the completion of training prior to the granting of the pilot certificate. The individual differences in flight test scores are small, however, 5 of the 10 subjects making a score of 80 and 3 making a score of 85, so that this measure does not furnish the differentiation in flying ability necessary for criterion use. The use of this measure is also negated by the fact that different amounts of training were administered to the various subjects prior to their flight tests, and by the fact that all of the subjects were brought up to approximately the same point in flying ability. The second criterion of flying ability is represented by the mean of the instructor's daily ratings on a 5-point scale, recorded following each flight in the Student Pilot Log Book. Inspection of the original ratings on each subject revealed that there was a pronounced central tendency of judgment, all of the students usually receiving ratings of 3, with occasional 4's toward the later stages of training. This central tendency of judgment is further indicated by the mean ratings for each of the 10 subjects, reported in Table 6, all of the means falling between 3.14 and 3.50.

Thus recourse must be made to the third type of criterion for the best available index, viz., to the paired-comparison rank of the 10 subjects by the instructor near the completion of flight training. We need not here discuss the merits of the paired-comparison method of ranking. Suffice it to say that such a method did serve to differentiate clearly between the 10 subjects. (See Table 6.) If we omit separate consideration of each of the three stages of flight training, owing to the small number of determinations, and consider only the relationship between the final ranks and the respective

---

[17] Another possible functionally differentiating feature of the Moss and present investigations is the degree of cabin enclosure, the completely air-tight cabins of the commercial air-lines being somewhat different in this respect from the small cabins of the Piper Cubs, which are rather roughly enclosed with a transparent plastic. It is doubtful, however, whether the resulting greater "ventilation" in the Cub could account for the "exhilaration effect" found in the present investigation.

[18] It is hoped that work by other investigators in concurrent projects, devoted to the development of a more adequate criterion, will yield such a measure for use in subsequent investigations.

TABLE 6 - CRITERIA, PHYSICAL, AND TENSION DATA
(Covering training period as a whole)

SUB. NO.	PAIRED COMPARISON RANK	CRITERIA				PHYSICAL DATA				EXCESSIVE TENSION ON STICK
		INSTRUCTOR'S RATING (IN LOG BOOK)			FLIGHT TEST SCORE (%)	SEX	HEIGHT	WEIGHT	AGE	
		AVE.	S.D.	RANGE						
1	1	3.2	0.4	2 to 4.5	65	M	5'10.5"	135	19	2
2	2	3.3	0.9	3 to 5	80	M	6'0"	165	25	1
3	3	3.1	0.8	2.5 to 4.5	80	M	6'1"	165	21	5
4	4	3.3	0.4	3 to 4.5	85	M	5'9.8"	147	20	6
5	5	3.3	0.4	3 to 5	83	M	5'11"	150	19	8
6	6	3.2	0.4	3 to 4.5	85	M	5'11"	160	21	3.5
7	7	3.4	0.5	3 to 5	80	M	5'9"	140	21	7
8	8	3.6	0.6	3 to 5	80	M	5'11"	185	19	9
9	9	3.4	0.6	3 to 4.5	80	M	5'8"	149	20	3.5
10	10	3.5	0.7	3 to 5	78	F	5'8"	160	20	10
AVE.		3.3	0.5		81.6		5'10.2"	155.6	20.5	
S.D.		0.4	0.3		2.3		1.8	13.8	1.7	

wabblometer scores, we obtain rank order correlations between final paired-comparison rank and pre-flight and post-flight wabblometer scores of .53 and .65 respectively. These correlations are positive but not exceptionally high, especially in the light of the possible use of the wabblometer as a predictive instrument. It should be noted, however, that these correlations are lowered primarily by one subject, Subject 9, who is rated ninth by the instructor but whose pre- and post-flight wabblometer scores give him a relatively high steadiness ranking.

Finally, it should be pointed out that there is no reliable indication of a correlation between decrease in wabblometer score following flight[19] and the criterion rating ($r = .22$). It might also be mentioned that the instructor's mean pre- and post-flight wabblometer scores as well as the decrease following flight (Table 4) are approximately identical to the respective mean scores for the group of 10 subjects (Table 5).

Since the present group of students was also serving as subjects for Dr. Williams in his study of excessive landing tension,[20] it is interesting to investigate the relationship between wabblometer scores and excessive landing tension measured by the Friez Flight Analyzer. The tension index employed may be defined as tension on the stick at landing resulting in a deflection of the recording pen on the Flight Analyzer greater than 0.2 of an inch. Only the rankings of the subjects with respect to this tension index are reported in the accompanying tables. The tension rankings for the various subjects in stages A, B, and C of flight training are reported in Tables 1-3, respectively, and the final tension rankings on all stages of training combined are reported in Table 6. The rank of 1 indicates least amount of excessive landing tension.

Since the interest of this study is not in the tension data per se, but rather in their relationship to wabblometer scores, and since there is considerable variability in the tension rankings of certain subjects from stage to stage, the combined data above should be considered. The correlation between mean pre-flight wabblometer scores and tension rankings is +.92, and that between post-flight wabblometer scores and tension rankings is +.95. These correlations are exceptionally high, and indicate clearly that the steadier subjects showed the least amount of excessive landing tension on the stick, whereas the unsteady subjects had the highest landing tension records. It will be noted that these correlations are considerably higher than the correlations between mean wabblometer scores and the criterion (+.53 and +.65 for pre- and post-flight indices, respectively), reported above. They are also higher than the correlation between the final tension rankings and the criterion, which is +.66. The relationship between mean

---

[19] Assigning a rank of 1 to the highest negative difference between pre- and post-flight wabblometer scores.

[20] Williams, A. C., op. cit.

TABLE 7 - PSYCHOLOGICAL TEST DATA

SUB. NO.	PRESSEY INTEREST-ATTITUDE		HUNTER SOCIAL ATTITUDE		BERNREUTER												STRONG VOCATIONAL INTEREST BLANK
					B1-N		B2-S		B3-I		B4-D		F1-C		F2-S		
	DEC-ILE	RANK	CEN-TILE	RANK	CEN-TILE	RANK	CEN-TILE	RANK	CEN-TILE	RANK	CEN-TILE	RANK	CEN-TILE	RANK	CEN-TILE	RANK	Occupations Scored "A" - "B+"
1	30	7	90	10	24	6	83	2	26	6	71	6	33	4	75	9	Office worker, sales manager, real estate salesman, production manager, purchasing agent.
2	20	5	90	9	22	8	65	4	24	7	81	3.5	33	4	60	5	Math. teacher, English teacher, production mgr., personnel mgr. accountant, musician, author, physical dir., office work.
3	20	4	42	3	22	7	38	9	28	5	95	2	33	4	32	3	Production mgr., purchasing agt., real estate salesman.
4	30	6	32	2	40	4	51	5	40	3	66	7	46	6	66	6.5	Sunday-school teacher, YMCA sec'y, musician, office worker, personnel manager.
5	10	1	78	8	29	5	51	6	17	8	74	5	50	7	54	4	Engineer, chemist, prod. mgr.
6	10	2	58	5	69	1	46	7	67	1	19	10	84	10	66	6.5	Math. teacher, chemist, personnel mgr., musician.
7	60	8	82	7	18	9	17	10	14	9	81	3.5	26	2	15	2	Lawyer, sales mgr., real estate salesman, life ins. salesman, advertising man.
8	10	3	48	4	3	10	90	1	2	10	98	1	1	1	72	8	Sunday-school teacher, YMCA phy. dir., personnel mgr. accountant, sales mgr., office worker, life ins. salesman.
9	90	10	16	1	42	3	42	8	32	4	20	9	70	9	8	1	Physician, math. teacher, chemist, personnel mgr., psychologist, musician, YMCA physical dir.
10	80	9	85	3	63	2	71	3	64	2	57	8	57	8	81	10	Nurse, housewife, office worker, secretary.

wabblemeter scores and final tension ratings is very significant both from a theoretical and a practical or applied point of view. If confirmed by future samplings, such correlations might easily have considerable predictive value.

Lastly, the relationship between wabblemeter scores and scores on the various psychological tests[21] may be considered. The following psychological test scores were available on the present group of subjects: Pressey Interest-Attitude Tests, Hunter Test of Social Attitudes, Bernreuter Personality Inventory, and Strong Vocational Interest Blank. In Table 7 will be found the decile or centile scores for each subject on each test except the Strong Vocational Interest Blank, in which these occupations for which the individual's interest score was A or B are reported. The rank order for each subject on the various tests is also reported, a rank of 1 being assigned to the "most mature" score on the Pressey Interest-Attitude Tests, to the "most liberal" on the Hunter Test of Social Attitudes, and to the "most neurotic," "most self-sufficient," "most introverted," "most dominant," "most self-confident," and "most social," respectively, on the six scales of the Bernreuter Personality Inventory.

In Table 8 are reported the correlations between mean wabblemeter scores, both pre-flight and post-flight, and the scores on each of the psychological tests. It will be noted that all correlations are low, none being greater than .28. No correlations are given for the Strong Vocational Interest Blank since the results on that test are not reported in quantitative form, but inspection of Tables 5 and 7 fails to reveal any relationship between type of occupational interest and wabblemeter score. Thus the psychological test results show no relationship to bodily steadiness as measured by the wabblemeter, a fact which suggests that the wabblemeter measures a fairly specific motor characteristic, rather than a more general personality tendency. This specificity of wabblemeter score, when considered together with its relatively high correlation with criterion ratings and its exceedingly high correlation with the more elaborate and expensive tension indices, suggests its value as a measure of psycho-physiological changes associated with flight and as an indicator of flying aptitude, and points to the desirability of further work on a larger sampling.

---

21 For detailed consideration of the psychological test results and for a discussion of the possible use of such personality, attitude, and interest inventories in the selection of student pilots, the reader is referred to the section of this report by Thelma Hunt.

TABLE 8

CORRELATIONS BETWEEN MEAN PRE- AND POST-FLIGHT WABBLEMETER SCORES AND SCORES MADE BY THE GROUP OF 10 STUDENT PILOTS ON PSYCHOLOGICAL TESTS

	PRESSEY INTEREST-ATTITUDE	HUNTER SOCIAL ATTITUDE	BERNREUTER					
			B1-N	B2-S	B3-1	B4-D	F1-C	F2-S
Mean Pre-flight Wabblemeter	±.16	±.04	-.19	-.09	-.27	-.26	-.26	.25
Mean Post-flight Wabblemeter	±.11	-.10	-.03	.11	-.16	-.12	-.08	.05

SUMMARY AND CONCLUSIONS

The present investigation was undertaken with a view to investigating measures of bodily steadiness of a representative group of 10 student pilots. The results were analyzed to show (1) the influence of the flight lesson upon the measures of bodily steadiness, (2) the relation between steadiness scores and a criterion of flying success, (3) the relationship between steadiness and an objective index of "tension," (4) the association between steadiness and scores on standard psychological tests (the Pressey Interest-Attitude Tests, the Hunter Test of Social Attitudes, the Bernreuter Personality Inventory, and the Strong Vocational Interest Blank), and (5) the relationship of changes in steadiness to length of flight.

Steadiness was measured by means of the "Wabblemeter." Scores were obtained from this instrument by having the subject stand on a platform which is balanced on a central fulcrum and measuring the extent to which the platform was moved or "wabbled" within a given period of time during which the subject attempted to keep it level. Steadiness measures were secured during both pre-flight and post-flight period.

"Tension" scores were procured from another experiment in which these same subjects had participated. The amount of "tension" was measured, in that study, by the extent to which the pilot "squeezed" a specially constructed stick while landing the plane.

Presented below are the general results of this investigation. Definite conclusions could not be drawn because of the size of the sample and limitations in the experimental design.

1. There were wide individual differences in both pre- and post- flight "Wabblemeter" scores, and the same individual showed considerable variability from flight to flight.

2. The 10 subjects maintained the same relative position in measures of steadiness, through three stages of flight training.

3. There was a marked absolute decrease (practice effect) in the steadiness scores with training. This decrease was present in both pre- and post-flight measures.

4. Pre-flight steadiness correlated .95 with the post-flight scores, indicating little if any differential effect from the flight.

5. Flight test scores and instructor's day-to-day ratings were found to be of no value as criterion measures.

6. Pre-flight steadiness scores correlated .53 with a paired-comparison ranking of the students made by the instructor near the end of training, and .92 with a "tension" measure. (The steadier subjects tended to be less tense.)

7. Post-flight steadiness scores correlated .65 and .95 respectively with these same indices.

8. Activity in "Wakefulness" correlated with the pursuit-rotor performance ranking to the extent of .2[illegible].

9. The correlation between the "tension" scores and the paired-comparison criterion was .66.

10. Correlations between the scores on the psychological tests and the pre- and post-flight steadiness scores ranged from -.28 to +.16.

PART II

PERSONALITY, INTEREST, AND ATTITUDE INVENTORIES IN THE SELECTION OF CIVILIAN PILOTS

by

THELMA HUNT

The George Washington University

PERSONALITY, INTEREST, AND ATTITUDE INVENTORIES IN THE SELECTION OF CIVILIAN PILOTS

THE PROBLEM

This project represents an analysis of four psychological tests administered to three student pilot groups and to a normal control group. The tests used were: (1) the Bernreuter Personality Inventory, (2) the Pressey Interest-Attitude Test, (3) the Hunter Social Attitudes Test, and (4) the Strong Interest Blank. The three student pilot groups tested include 35 George Washington University students, 43 Maryland University students and 54 North Carolina students (University of North Carolina and North Carolina State College). A small control group of 22 George Washington University students expressed no intention of participating in the student pilot training course and matched for age, sex, and status in school with 22 of the student pilots was given the Bernreuter, Pressey, and Hunter Tests. This small control group did not deviate significantly from the college norms published for the tests used. Since the published norms represent a wider geographical and institutional scattering they have been used as "normals" in comparing the average student pilot records obtained in the three schools participating in this study.

This preliminary report includes studies of (1) comparison of records on the tests of successful student pilots with normal control records, (2) comparison of records of superior student pilots with normal control records, and (3) study of relationship between test records and ratings of relative competence of certain student pilots.

COMPARISON OF SUCCESSFUL STUDENT PILOTS WITH NORMAL CONTROL RECORDS

Tables 1 to 15 present the distributions for the student pilots who completed the course successfully and for the normal. Significant observations from these tables will be indicated according to test.

1. The Bernreuter Personality Inventory. All the student pilot groups made lower average Neurotic Tendency scores than the normal (Table 1) Median score for the combined groups of successful student pilots is equivalent to the 35th percentile for the normal. Among the three student pilot groups the North Carolina group is significantly above the George Washington and Maryland groups. How much of this is due to sectional differences cannot be ascertained from data on hand at present. All the student pilot groups are, on the average, less introverted than the normal as shown by the lower Introvert scores (Table 2). Again, the North Carolina group deviates less from the norm than the other two student pilot groups. The composite student pilot group is more dominant than the normal group as shown by the average Dominance-Submission score (Table 3), although the North Carolina group does not deviate significantly from the normal. Average student pilot scores in the remaining qualities for which the Bernreuter Test was scored (Self-sufficiency, Self-confidence, and Sociability) do not differ

[illegible] among the [illegible] (Tables 6, 7, [illegible])

2. The Pressey Interest-Attitude Test. This is a questionnaire for investigating the maturity of interest and attitudes. Average of the successful pilot group shows greater maturity of interest and attitude than for the normal control group. As shown in Table 9 the average score for pilots is minus 6, for the normal is plus 6. (The higher the score on this test the greater the immaturity.) Of the three pilot groups, the George Washington group shows highest maturity, a finding probably related to the fact that George Washington is an urban university attracting for that reason a somewhat more mature group of students.

An analysis of the four subtests of the Pressey Interest-Attitude Test (Tables 8-11) shows significant difference between pilots and non-pilots in two of the four, a slight difference in one, and no difference in one. The two showing the greatest differentiation are the subtests representing judgments or opinions of wrongness of various things and the one representing expressions of like or admiration for various types of persons. The subtest representing expressions of interest in various things showed a smaller differentiation. No differentiation between the successful pilot groups and the norm occurred in the subtest representing expressions of worries or fears.

3. Hunter Social Attitudes Test. Tables 12 to 18 present the records on a social attitudes test for the successful pilot groups. Positive scores indicate liberalism; negative scores, conservatism. In general, social attitudes of the successful pilot group are neither more conservative nor more liberal than the normal. Median scores are respectively minus 5 and minus 4 (Table 12). The test can be broken down into measures of specific attitudes, including attitudes in the following categories: Negro, War, Economics and Labor, Social Life, Government, and Religion (Tables 13 - 18). In only three of these is the difference in Median between the total pilot and normal group more than 2 points. These three are Negro, Economics and Labor, and Religion (Tables 13, 15, and 18). In all these the difference is less than 4. The differences are in the direction of more liberalism in attitude toward the negro and toward religion and greater conservatism in attitude toward labor and economic problems.

4. Strong Interest Blank. The Strong Interest Blanks for the pilot groups have been scored only for the nine occupations of Group I of the possible scorings worked out by the author of the test. These give the distributions of ratings shown in Table 19. A ratings signify that the subject possesses the interest pattern of those successful in the occupation, B plus ratings that he probably does, B and B minus ratings are "doubtful," and C plus or C ratings signify that the subject does not possess the interest pattern of successful people in the occupation. Among the successful pilots the largest percentage of A ratings is for the occupation of engineer. Twenty-seven percent have A ratings; and forty-one percent have A or B plus ratings. A control group rated at George Washington showed only 7 percent with A or B plus ratings. Very differentiative on the low end of the ratings are the occupations of minister, certified public accountant, and mathematician. For the occupation of minister, 80 percent

of the successful pilot group have ratings in the "no interest group." The George Washington Control Group show only 53 percent with these low ratings. For the occupation of certified public accountant, 84 percent of the successful pilot group show "no interest" ratings. For the control group there are only 46 percent with such ratings. For mathematician, the successful pilot group shows a high percentage of low ratings, but the same is found in the control group, so that the student pilot skewness of distribution is probably less significant.

SPECIAL STUDY OF SUPERIOR STUDENTS

For the George Washington and Maryland University Groups, a list was obtained of those students selected on the basis of superior flying ability for the advanced training course, and those dropped for any reason from the course, before completion. Of those to whom the various personality, attitude, and interest tests were administered, four of the George Washington Group were selected for the advanced course and nine of the Maryland Group. Only two students to whom the tests were administered (one in each of the two schools) were dropped.

Individual test records for these students are presented in Tables 20, 21, and 22. In Table 23 the standing of the superior group is compared with that of the whole group of pilots completing the course and with normal non-pilot college ratings.[1]

On the whole the Bernreuter Personality Inventory and the Pressey Interest-Attitude Test seem to be more valuable than the other two in differentiating the superior group of student pilots. As a group the superior student pilots are characterized by low Neurotic Tendency, high Self-sufficiency, low Introversion, high Dominance, high Self-confidence, and below average Sociability. These tendencies are evident in comparing them with the total pilot group and are even more evident in most of the traits in comparing them with a normal non-pilot college group. A marked difference in maturity of emotions, interests, and attitudes is shown by the Pressey Test. The average maturity rating for the superior group of student pilots is 90 as compared with 69 for the whole group of pilots and 50 for college students in general.

The superior group of student pilots is on the average differentiated in a direction of greater liberalism in attitude toward government and economic and labor problems, and somewhat more conservatism in attitude toward social laws and conventions. Attitudes toward the negro, war, and

---

1 It was originally planned to make a special study of the failure students but the number of failure students available was too small to warrant any deductions as to value of the test in predicting failure. One of the failures (N.W.T.) probably should not be counted as a real failure. He failed for lack of time to complete the course in the required time from interference of extraneous factors. It is interesting, perhaps, to note that the other failure student does exhibit ratings in general on the Bernreuter and Pressey Tests which are the reverse of the trends in the superior group.

[illegible] Occupational Interest ratings [illegible] a predominance of [illegible] among the high ratings, and of Minister, Certified Public Accountant, and Mathematician among the low ratings, but the percentage of these ratings does not differ significantly from the whole pilot group.

CORRELATION STUDY OF TEN STUDENTS OF ONE INSTRUCTOR WHO WERE RATED IN FLYING ABILITY BY THE METHOD OF PAIRED COMPARISONS

The students of one flight instructor at the University of Maryland (ten students in number) were rated by the instructor in flying ability by the method of paired comparisons. Although the number included in the group is too small to allow the drawing of reliable conclusions, and the ratings of flying ability may be subject to criticisms as to accuracy, the data are presented for possible suggestive modes of approach in future attacks upon the problem. Rank order correlations between these ratings and the ratings on the Bernreuter Personality Inventory, the Pressey Interest-Attitude Test, and the Hunter Social Attitudes Test are presented in Table 24. Only a few of the qualities show significant relationship in this small group to the personal ratings of flying ability by the instructor. These qualities are: Self-sufficiency, Sociability, Maturity of Emotions-Interests-Attitudes, and Attitude toward Economic and Labor Problems. The relationships are in the direction of association of flying ability with high Self-sufficiency, low Sociability, high Maturity, and Liberalism in attitude toward Economic and Labor Problems. These relationships were shown in more striking fashion in the comparison of the significantly superior pilots with the total groups of student pilots.

TABLE 1

DISTRIBUTION OF SCORES ON BERNREUTER B1-N
(Neurotic Tendency)

Scores	Normal Group	G.W.U. Pilot Group	Md. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
160 up	1				
150	-				
140	1	1			1
130	1	-			-
120	-	-			-
110	1	-	1		1
100	1	-	-		-
90	-	2	-		2
80	1	-	-		-
70	1	-	-		-
60	1	-	-	2	2
50	2	1	-	-	1
40	3	-	1	-	1
30	3	1	-	1	2
20	3	1	-	2	3
10	3	-	-	2	2
0	3	-	-	2	2
-10	4	1	-	1	2
-20	2	-	2	2	4
-30	4	-	1	5	6
-40	4	-	-	2	2
-50	4	-	-	4	4
-60	4	1	3	2	6
-70	4	1	1	1	3
-80	6	2	-	2	4
-90	6	-	5	2	7
-100	5	3	1	5	9
-110	8	3	4	2	9
-120	4	1	1	4	6
-130	4	2	6	3	11
-140	4	-	2	1	3
-150	2	2	2	1	5
-160	2	1	2	-	3
-170	2	3	3	2	8
-180	6	5	4	5	14
TOTAL	100	31	35	53	123
Q3	-3	-58	-82	-22	-39
Median	-67	-105	-115	-73	-97
Q1	-117	-161	-146	-116	-138

TABLE :

DISTRIBUTION OF SCORES ON BERNREUTER B3-I
(Introversion-Extroversion)

Scores	Normal Group	G.W.V. Pilot Group	Mi. V. Pilot Group	M.C. Pilot Group	Total Pilot Group
160 up					
150					
140					
130					
120					
110		[illegible]			1
100		[illegible]			[illegible]
90	1	[illegible]			[illegible]
80	1	[illegible]			[illegible]
70	2	[illegible]			[illegible]
60	2	[illegible]	1		2
50	2	[illegible]	[illegible]	1	1
40	3	[illegible]	[illegible]	1	1
30	5	2	1	2	5
20	5	[illegible]	[illegible]	2	2
10	5	[illegible]	[illegible]	1	3
0	6	[illegible]	3	2	5
-10	6	1	1	7	9
-20	8	[illegible]	[illegible]	6	7
-30	7	[illegible]	2	3	5
-40	8	4	2	4	10
-50	8	[illegible]	4	5	9
-60	7	3	4	4	11
-70	7	3	4	3	10
-80	6	1	6	6	13
-90	5	2	5	1	13
-100	3	[illegible]	[illegible]	2	4
-110	2	3	3	2	8
-120	1		3	1	4
-130					
-140					
-150					
-160					
-170					
-180					
TOTAL	100	31	39	53	123
Q3	8	-17	-39	-6	-12
Median	-30	-62	-64	-34	-51
Q1	-63	-86	-82	-66	-79

TABLE 3

DISTRIBUTION OF SCORES ON BERNREUTER B4-R
(Dominance-Submission)

Scores	Normal Group	U.W.U. Pilot Group	M.U. Pilot Group	N.C. Pilot Group	Total Pilot Group
160 up	2	3	1	1	5
150	1	2	2	1	5
140	3	-	2	1	3
130	4	2	1	3	6
120	4	1	1	-	2
110	4	-	3	1	4
100	5	3	5	4	12
90	4	1	6	5	12
80	6	3	4	1	8
70	6	1	3	3	7
60	6	-	4	3	7
50	6	4	1	2	7
40	6	2	1	4	7
30	6	4	1	4	9
20	4	-	-	5	5
10	5	-	-	6	6
0	6	-	-	1	1
-10	5	-	2	2	4
-20	4	1	1	2	4
-30	2	1	1	2	4
-40	2	-		-	-
-50	2	2		1	3
-60	2	1		1	2
-70	1				
-80	-				
-90	1				
-100	1				
-110	-				
-120	1				
-130	1				
-140					
-150					
-160					
-170					
-180					
TOTAL	100	31	39	53	123
Q3	90	123	104	95	106
Median	47	75	74	46	74
Q1	0	37	67	17	32

TABLE 4

DISTRIBUTION OF SCORES ON BERNREUTER B2-S
(Self-sufficiency)

Scores	Normal Group	G. W. U. Pilot Group	Md. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
160 up					
150	1				
140	1				
130	2				
120	2	1		1	2
110	2	2		2	4
100	2	-	2	1	3
90	3	1	1	1	3
80	4	2	1	1	4
70	6	1	3	2	6
60	6	3	2	2	7
50	7	1	3	1	5
40	7	2	4	1	7
30	7	6	2	2	10
20	8	1	3	5	9
10	8	1	7	3	11
0	8	1	3	5	9
-10	6	1	3	6	10
-20	5	3	2	4	9
-30	4	-	2	4	6
-40	3	-	-	1	1
-50	3	1	-	2	3
-60	2	2	1	1	4
-70	1	2		1	3
-80	1			-	-
-90	1			6	6
-100				-	-
-110				1	1
-120					
-130					
-140					
-150					
-160					
-170					
-180					
TOTAL	100	31	39	53	123
Q3	62	67	57	34	56
Median	25	36	25	1	19
Q1	-8	-11	6	-26	-13

TABLE 5

DISTRIBUTION OF SCORES ON BERNREUTER F1-C
(Confidence in Self)

Scores	Normal Group	G.W.U. Pilot Group	Ld. U. Pilot Group	N. C. Pilot Group	Total Pilot Group
160 up	1				
150	1		1		1
140	-	2	-		2
130	1	-	-	1	1
120	1	1	-	-	1
110	1	-	1	1	2
100	1	-	-	-	-
90	2	1	-	-	1
80	1	-	-	1	1
70	2	-	-	2	2
60	1	-	-	-	-
50	2	-	-	2	2
40	2	-	-	3	3
30	2	1	1	2	4
20	2	1	2	2	5
10	4	-	-	-	-
0	4	1	1	5	7
-10	4	1	1	4	6
-20	4	2	-	2	4
-30	4	-	1	1	2
-40	4	-	-	1	1
-50	4	2	2	4	8
-60	4	1	3	5	9
-70	4	1	2	3	6
-80	6	4	3	-	7
-90	5	1	2	1	4
-100	5	1	5	3	9
-110	4	1	1	2	4
-120	4	1	2	-	3
-130	2	1	3	1	5
-140	4	2	3	1	6
-150	4	4	2	1	7
-160	3	-	1	-	1
-170	2	-	-	3	3
-180	5	2	2	2	6
TOTAL	100	31	39	53	123
Q3	3	-7	-19	24	1
Median	-60	-74	-88	-35	-59
Q1	-112	-131	-128	-86	-111

TABLE 6

DISTRIBUTION OF SCORES ON BELL-METTER F2-S
(Sociability)

Scores	Normal Group	G. V. U. Pilot Group	Mi. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
150 up					
150					
140		1			1
130		-			-
120		-			-
110	1	-			-
100	1	-			-
90	1	-		1	1
80	1	1			1
70	2	-	1	-	1
60	2	1	1	-	2
50	2	1	-	-	1
40	3	1	1	-	2
30	4	-	-	3	3
20	4	2	2	1	5
10	6	2	3	2	7
0	6	2	4	5	11
-10	7	1	4	3	8
-20	8	1	3	1	5
-30	8	3	3	2	8
-40	7	3	1	3	7
-50	7	-	1	6	7
-60	6	2	4	7	13
-70	4	-	1	4	5
-80	4	-	3	3	6
-90	4	3	2	4	9
-100	4	-	3	1	4
-110	2	1	1	1	3
-120	2	3	-	4	7
-130	1	1	1	-	2
-140	1	-		-	-
-150	-	-		-	-
-160	1	1		1	2
-170	1	1		1	2
-180					
TOTAL	100	31	39	53	123
Q3	10	21	6	-4	4
Median	-27	-25	-22	-49	-38
Q1	-64	-84	-71	-76	-77

TABLE 7

DISTRIBUTION OF TOTAL SCORES ON PRESSEY INTEREST-ATTITUDE TEST

Scores	Normal Group	C.W.U. Pilot Group	Md. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
80 and up			1		1
75-79	1		"		"
70-74	1		"		"
65-69	1		"	1	1
60-64	1		"	1	1
55-59	2		"	1	1
50-54	2		1	"	1
45-49	3		"	"	"
40-44	4		1	1	2
35-39	3	2	2	"	4
30-34	4	"	"	2	2
25-29	5	"	1	2	3
20-24	6	1	"	"	1
15-19	6	1	"	3	4
10-14	7	3	1	4	8
5-9	8	1	2	4	7
0-4	7	1	5	3	9
-5 to -1	8	3	7	6	16
-10 to -6	9	2	3	10	15
-15 to -11	5	4	4	6	14
-20 to -16	5	5	1	2	8
-25 to -21	3	3	2	1	6
-30 to -26	3	5	3	2	10
-35 to -31	3	"	2	3	5
-40 to -36	1	1	2	1	4
-45 to -41	1		"	"	"
-50 to -46			"	1	1
-55 to -51			1		1
-60 to -56					
-65 to -61					
-70 to -66					
-75 to -71					
TOTAL	100	32	39	54	125
Q3	27	5	4	12	8
Median	8	-13	-4	-5	-6
Q1	-8	-22	-22	-12	-17

TABLE 8

DISTRIBUTION OF SCORES ON PART I
OF PRESSEY INTEREST-ATTITUDE TEST

Scores	U.S.N. Pilot Group	Marine Pilot Group	U.S. Pilot Group	Total Pilot Group
30 and up	2	1	9	12
28-29	-	1	1	2
26-27	-	1	2	3
24-25	2	2	-	4
22-23	-	1	3	4
20-21	3	1	2	3
18-19	1	-	3	3
16-17	2	1	3	5
14-15	1	1	3	5
12-13	1	5	3	9
10-11	2	2	1	5
8-9	3	3	4	10
6-7	3	2	4	9
4-5	1	5	3	9
2-3	4	3	5	12
0-1	2	3	3	8
-2 to -1	3	4	3	10
-4 to -3	2	3	1	6
-6 to -5	1	2	1	4
-8 to -7	2			2
-10 to -9				
-12 to -11				
-14 to -13				
-16 to -15				
-18 to -17				
-20 to -19				
-22 to -21				
-24 to -23				
-26 to -25				
-28 to -27				
-30 to -29				
Below -30				
TOTAL	32	39	54	125
Q3	14	14	23	17
Median	5	5	13	9
Q1	0	0	4	2

TABLE V

DISTRIBUTION OF SCORES ON PART II
OF PRESSEY INTEREST-ATTITUDE TEST

Scores	W.U. Pilot Group	Md. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
30 and up			2	2
28-29			1	1
26-27			-	-
24-25		1	-	1
22-23		1	-	1
20-21		-	-	-
18-19		1	-	1
16-17		-	-	-
14-15		-	-	-
12-13	1	-	-	1
10-11	1	1	-	2
8-9	-	2	1	3
6-7	3	3	5	11
4-5	-	3	3	6
2-3	6	3	4	13
0-1	7	7	16	30
-2 to -1	9	10	13	32
-4 to -3	3	5	4	12
-6 to -5	2	1	4	7
-8 to -7		-	1	1
-10 to -9		1		1
-12 to -11				
-14 to -13				
-16 to -15				
-18 to -17				
-20 to -19				
-22 to -21				
-24 to -23				
-26 to -25				
-28 to -27				
-30 to -29				
Below -30				
TOTAL	32	39	54	125
Q3	3	5	3	3
Median	0	0	0	0
Q1	-2	-2	-2	-2

TABLE 10

DISTRIBUTION OF SCORES ON PART III
OF PRESSEY INTEREST-ATTITUDE TEST

Scores	G.W.U. Pilot Group	Md. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
30 and up			1	1
28-29		2	-	2
26-27		-	-	-
24-25	1	1	-	2
22-23	1	1	1	3
20-21	1	-	-	1
18-19	-	1	1	2
16-17	2	2	2	6
14-15	-	1	1	2
12-13	-	1	4	5
10-11	-	4	2	6
8-9	4	4	8	16
6-7	4	2	5	11
4-5	2	3	8	13
2-3	3	5	6	14
0-1	1	2	1	4
-2 to -1	2	1	2	5
-4 to -3	5	2	2	9
-6 to -5	4	5	5	14
-8 to -7	-	1	3	4
-10 to -9	1	-	1	2
-12 to -11	-	-	1	1
-14 to -13	-	1		1
-16 to -15	1			1
-18 to -17				
-20 to -19				
-22 to -21				
-24 to -23				
-26 to -25				
-28 to -27				
-30 to -29				
Below -30				
TOTAL	32	39	54	125
Q3	8	11	9	9
Median	3	5	5	5
Q1	-3	-1	-1	-3

TABLE 11

DISTRIBUTION OF SCORES ON PART IV
OF PRESSEY INTEREST-ATTITUDE TEST

Scores	G.W.U. Pilot Group	Md. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
30 and up				
28-29			1	1
26-27			-	-
24-25			1	1
22-23			-	-
20-21			-	-
18-19			-	-
16-17		1	-	1
14-15		-	-	-
12-13		-	-	-
10-11		1	1	2
8-9		1	-	1
6-7		2	-	2
4-5		-	-	-
2-3		-	-	-
0-1		-	-	-
-2 to -1		-	-	-
-4 to -3	1	3	2	6
-6 to -5	1	1	2	4
-8 to -7	1	1	3	5
-10 to -9	-	1	-	1
-12 to -11	3	1	3	7
-14 to -13	2	5	8	15
-16 to -15	3	-	6	9
-18 to -17	2	2	3	7
-20 to -19	5	2	6	13
-22 to -21	7	3	2	12
-24 to -23	-	3	1	4
-26 to -25	2	2	4	8
-28 to -27	1	3	3	7
-30 to -29	1	3	3	7
Below -30	3	4	5	12
TOTAL	32	39	54	125
Q3	-14	-8	-13	-13
Median	-19	-19	-16	-19
Q1	-22	-27	-25	-25

TABLE 12

DISTRIBUTION OF TOTAL SCORES FOR HUNTER SOCIAL ATTITUDES TEST

Scores	Normal Group	G.W.U. Pilot Group	Md. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
80 and up	1	3	1		4
75-79	1	-	-		-
70-74	1	-	-		-
65-69	2	-	1		1
60-64	1	1	-	1	2
55-59	2		-	-	-
50-54	2	1	-	-	1
45-49	1	-	-	1	1
40-44	-	-	1	1	2
35-39	2	-	-	3	3
30-34	1	1	2	3	6
25-29	3	3	1	2	6
20-24	4	-	2	2	4
15-19	5	1	-	2	3
10-14	5	1	3	3	7
5-9	7	2	2	3	7
0-4	6	2	3	2	7
-5 to -1	9	2	1	5	8
-10 to -6	8	2	5	5	12
-15 to -11	7	4	2	5	11
-20 to -16	5	1	5	5	11
-25 to -21	5	-	4	2	6
-30 to -26	5	4	2	2	8
-35 to -31	6	1	1	3	5
-40 to -36	3	1	2	2	5
-45 to -41	2	1	-	-	1
-50 to -46	2		1	1	2
-55 to -51	1				
-60 to -56	2				
-65 to -61	1				
-70 to -66					
-75 to -71					
TOTAL	100	31	39	53	123
Q3	16	27	12	19	18
Median	-4	-2	-8	-4	-5
Q1	-22	-17	-21	-17	-19

TABLE 7

DISTRIBUTION OF SCORES ON PART I (REVISED)
SOCIAL ATTITUDES TEST

Scores	U.W.U. Pilot Group	Mid. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
30 and up	2	1		3
28-29	-	2	1	3
26-27	-	3	-	3
24-25	1	2	-	3
22-23	-	-	1	1
20-21	1	-	1	2
18-19	-	-	3	3
16-17	1	1	4	6
14-15	2	2	2	6
12-13	-	-	2	2
10-11	-	1	1	2
8-9	2	2	4	7
6-7	3	-	4	7
4-5	2	1	1	3
2-3	1	1	6	8
0-1	3	4	5	12
-2 to -1	3	1	1	5
-4 to -3	1	2	7	10
-6 to -5	4	2	2	7
-8 to -7	1	3	1	5
-10 to -9	1	3	1	5
-12 to -11	1	3	-	4
-14 to -13	2	2	1	5
-16 to -15	-	1	3	4
-18 to -17	-	2	2	4
-20 to -19	-	-		-
-22 to -21	-	-		-
-24 to -23	-	-		-
-26 to -25	1	-		1
-28 to -27	1	1		2
-30 to -29				
Below -30				
TOTAL	31	39	53	123
X	8	15	12	13
Median	0	0	3	1
Q1	-6	-10	-8	-6

TABLE 24

DISTRIBUTION OF SCORES ON PART II (WAR)
SOCIAL ATTITUDES TEST

Scores	G.W.U. Pilot Group	Md. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
30 and up				
28-29				
26-27				
24-25				
22-23		1		1
20-21		-		-
18-19	1	-		1
16-17	-	-		-
14-15	-	-		-
12-13	1	1	2	4
10-11	2	1	-	3
8-9	4	1	3	8
6-7	1	4	5	10
4-5	1	2	6	9
2-3	3	3	13	19
0-1	6	3	5	14
-2 to -1	3	3	4	10
-4 to -3	3	5	3	11
-6 to -5	1	3	-	4
-8 to -7	2	4	4	10
-10 to -9	1	2	4	7
-12 to -11	2	2	3	7
-14 to -13		1	1	2
-16 to -15		-		-
-18 to -17		3		3
-20 to -19				
-22 to -21				
-24 to -23				
-26 to -25				
-28 to -27				
-30 to -29		-		
Below -30				
TOTAL	31	39	53	125
Q3	8	4	4	5
Median	1	-3	2	1
Q1	-3	-8	-4	-6

TABLE [illegible]

DISTRIBUTION OF SCORES ON PART III (ECONOMICS AND LABOR)
SOCIAL ATTITUDES TEST

Scores	C.A.A. Pilot Group	M.C. Pilot Group	N.C. Pilot Group	Total Pilot Group
30 and up				
28-29	1			1
26-27	-			-
24-25	-			-
22-23	1			1
20-21	-			-
18-19	1	1		2
16-17	1	-	1	2
14-15	1	-	-	1
12-13	-	2	4	6
10-11	-	2	4	6
8-9	3	-	3	6
6-7	3	1	3	7
4-5	1	3	4	8
2-3	-	5	3	8
0-1	3	4	4	11
-2 to -1	4	4	6	14
-4 to -3	4	5	6	15
-6 to -5	3	3	4	10
-8 to -7	3	3	4	10
-10 to -9	1	2	4	7
-12 to -11	-	-	2	2
-14 to -13	-	-	3	3
-16 to -15	1	1	-	2
-18 to -17			-	-
-20 to -19			-	-
-22 to -21			1	1
-24 to -23				
-26 to -25				
-28 to -27				
-30 to -29				
Below -30				
TOTAL	32	39	53	123
Q3	2	3	7	6
Median	-3	-1	-1	-1
Q1	-5	-6	-5	-5

TABLE 1c

DISTRIBUTION OF SCORES ON PART IV (SOCIAL LIFE)
SOCIAL ATTITUDES TEST

Scores	G.W.U. Pilot Group	Md. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
30 and up				
28-29				
26-27				
24-25				
22-23				
20-21				
18-19				
16-17	1			1
14-15	1	1		2
12-13	1	4	4	9
10-11	3	1	2	6
8-9	6	4	4	14
6-7	5	8	6	19
4-5	8	5	14	27
2-3	-	5	4	9
0-1	1	6	9	16
-2 to -1	3	5	4	12
-4 to -3	1		4	5
-6 to -5	-		1	1
-8 to -7	-		1	1
-10 to -9	-			-
-12 to -11	1			1
-14 to -13				
-16 to -15				
-18 to -17				
-20 to -19				
-22 to -21				
-24 to -23				
-26 to -25				
-28 to -27				
-30 to -29				
Below -30				
TOTAL	31	39	53	123
Q3	9	7	6	8
Median	6	5	4	5
Q1	4	1	0	1

TABLE 17

DISTRIBUTION OF SCORES ON PART V (GOVERNMENT)
SOCIAL ATTITUDE TEST

Scores	G.W.U. Pilot Group	Md. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
30 and up				
28-29				
26-27				
24-25				
22-23				
20-21				
18-19	1			1
16-17	-			-
14-15	-			-
12-13	-		1	1
10-11	1		1	2
8-9	1		-	1
6-7	2	1	1	4
4-5	2	2	4	8
2-3	3	1	3	7
0-1	2	4	6	12
-2 to -1	4	3	8	15
-4 to -3	1	4	9	14
-6 to -5	3	9	7	19
-8 to -7	3	6	5	14
-10 to -9	2	7	3	12
-12 to -11	3	1	4	8
-14 to -13	2	1	1	4
-16 to -15	1			1
-18 to -17				
-20 to -19				
-22 to -21				
-24 to -23				
-26 to -25				
-28 to -27				
-30 to -29				
Below -30				
TOTAL	31	39	53	123
Q3	3	-2	0	0
Median	-2	-5	-3	-4
Q1	-9	-8	-6	-8

TABLE 18

DISTRIBUTION OF SCORES ON PART VI (RELIGION)
SOCIAL ATTITUDES TEST

Scores	G.W.U. Pilot Group	Md. U. Pilot Group	N.C. Pilot Group	Total Pilot Group
30 and up				
28-29				
26-27				
24-25				
22-23				
20-21				
18-19				
16-17				
14-15				
12-13				
10-11		1		1
8-9	1	-	1	2
6-7	-	1	1	2
4-5	3	-	-	3
2-3	2	1	2	5
0-1	1	2	2	5
-2 to -1	2	3	6	11
-4 to -3	8	3	7	18
-6 to -5	4	5	10	19
-8 to -7	3	9	10	22
-10 to -9	3	7	6	16
-12 to -11	2	5	4	11
-14 to -13	1	1	3	5
-16 to -15	1		1	2
-18 to -17				
-20 to -19				
-22 to -21				
-24 to -23				
-26 to -25				
-28 to -27				
-30 to -29				
Below -30				
TOTAL	31	39	53	123
Q3	-1	-4	-3	-3
Median	-4	-7	-6	-6
Q1	-8	-9	-9	-9

TABLE 2

DISTRIBUTION OF RATINGS ON STRONG INTEREST BLANK

Rating	G.W.U. Pilot Group	U. of Md. Pilot Group	N.C. Pilot Group	Total Pilot Group
		Occupation Physician		
A	2	2	4	8
B plus	1	3	3	7
B	6	2	5	13
B minus	2	8	7	17
C plus	7	2	7	16
C	8	17	24	49
		Occupation H. S. Math.-Physical Science Teacher		
A	6	6	13	25
B plus	6	7	10	23
B	5	7	12	24
B minus	2	8	4	14
C plus	2	3	7	12
C	5	3	4	12
		Occupation Engineer		
A	6	8	16	30
B plus	2	5	8	15
B	5	6	6	17
B minus	4	2	11	17
C plus	2	4	6	12
C	7	9	3	19
		Occupation Lawyer		
A	3	2	1	6
B plus	2	2	0	4
B	3	1	3	7
B minus	10	7	8	25
C plus	5	10	12	27
C	3	12	26	41
		Occupation Minister		
A	1	0	0	1
B plus	0	0	2	2
B	1	1	2	4
B minus	6	5	4	15
C plus	3	1	3	7
C	15	27	39	81

TABLE 19 (Cont'd)

DISTRIBUTION OF RATINGS ON STRONG INTEREST BLANK

Rating	G.W.U. Pilot Group	U. of Md. Pilot Group	N.C. Pilot Group	Total Pilot Group
	Occupation Social Science High School Teacher			
A	4	5	5	14
B plus	1	5	1	7
B	5	1	10	16
B minus	5	8	4	17
C plus	4	5	10	19
C	7	10	20	37
	Occupation Certified Public Accountant			
A	1	2	0	3
B plus	0	1	0	1
B	1	0	1	2
B minus	5	4	3	12
C plus	9	8	6	23
C	10	19	40	69
	Occupation Mathematician			
A	0	0	0	0
B plus	1	1	0	2
B	1	2	1	4
B minus	2	2	5	9
C plus	4	5	6	15
C	18	24	38	80
	Occupation Chemist			
A	3	9	10	22
B plus	4	1	9	14
B	5	2	3	10
B minus	6	5	11	22
C plus	0	6	8	14
C	8	11	9	28

TABLE 20

INDIVIDUAL PERCENTILE RECORDS OF SUPERIOR STUDENTS AND FAILURE STUDENTS

Student	Bernreuter Test						Pressey Test	
	B1-N	B2-S	B3-I	B4-D	F1-C	F2-S	Score	Sm. Age
SUPERIOR STUDENTS[2]								
J.D.D.	23	77	5	89	19	60	86	22
R.L.K.	94	8	90	53	97	55	20	17
W.C.R.	3	90	3	98	5	69	99	22
J.H.W.	5	96	6	97	5	88	69	21
W.F.M.	21	65	24	81	33	60	75	21½
L.W.P.	69	46	67	19	85	66	92	22
A.S.A.	14	29	11	79	22	10	95	22
G.F.B.	88	7	88	13	96	48	99	22
J.O.H.	41	92	39	77	33	69	74	21½
J.L.L.	10	77	8	84	10	60	90	22
G.T.S.	6	31	4	84	16	12	92	22
F.A.S.	14	64	13	98	17	70	70	21
H.W.	56	15	53	54	82	4	7	16
Medians	21	64	13	81	22	60	90	22
FAILURE STUDENTS[3]								
G.W.	94	7	94	17	98	74	31	18½
N.W.T.	20	87	24	60	33	50	91	22

---

2 Included in the group of "Superior Students" are those recommended from the George Washington and Maryland groups for the advanced course. These students were recommended upon the basis of superior work in the initial course.

3 Included in the group of "Failure Students" are those dropped from the training course before completion.

TABLE 21

INDIVIDUAL PERCENTILE RECORDS OF SUPERIOR STUDENTS AND FAILURE STUDENTS

Student	Hunter Social Attitudes Test						
	Parts						
	I	II	III	IV	V	VI	Total
SUPERIOR STUDENTS							
J.D.D.	6	5	20	40	44	28	5
R.L.K.	59	50	90	22	99	2	36
W.C.R.	30	20	32	82	11	36	19
J.H.W.	46	90	75	10	94	51	60
W.F.M.	72	85	55	55	75	68	74
L.W.F.	79	78	73	7	81	12	57
A.S.A.	45	99	78	23	81	79	73
G.F.B.	51	33	37	69	75	97	55
J.O.H.	45	53	64	25	15	51	27
H.L.L.	75	38	85	6	30	23	37
G.T.S.	45	7	90	22	88	7	27
F.A.S.	5	72	88	85	93	78	40
H.W.	4	15	30	52	Inc.	85	Inc.
Medians	45	50	73	25	78	51	38
FAILURE STUDENTS							
G.W.	98	51	88	23	93	9	65
N.W.T.	58	70	15	50	11	34	39

[illegible]

Student	Occupations with High Ratings (A)	Occupation with Low Ratings (C)
	SUPERIOR SCIENCE	
J.D.G.	None	Math.-Physical Science Teacher Minister Social Sc. H.S. Teacher Mathematician
R.J.K.	Engineer	Minister Social Sc. H.S. Teacher C.P.A. Mathematician
W.C.H.	Math.-Ph. Sc. Teacher Engineer Chemist	Lawyer Minister C.P.A.
J.H.W.	None	Physician Minister Mathematician
W.F.M.	Math.-Ph. Sc. Teacher	Physician Lawyer Minister Mathematician
L.W.P.	Math.-Ph. Sc. Teacher Chemist	C.P.A.
A.S.A.	Engineer	Minister Social Sc. H.S. Teacher C.P.A. Mathematician
C.F.B.	None	Lawyer Minister C.P.A. Mathematician
J.O.H.	Chemist	None
J.L.L.	Physician Engineer Chemist	Minister Social Sc. H.S. Teacher

TABLE 22 (Cont'd)

G.T.S.	None	Physician Engineer Minister C.P.A. Mathematician Chemist
F.A.S.	Social Sc. H.S. Teacher C.P.A.	Physician Engineer Minister Mathematician Chemist
H.W.	Math.-Ph. Sc. Teacher Engineer Chemist	Lawyer Minister C.P.A.

FAILURE STUDENTS

G.W.	Lawyer C.P.A.	Math.-Ph. Sc. Teacher Mathematician
N.W.T.	Chemist Physician Math.-Ph.Sc. Teacher Engineer	Lawyer Minister Social Sc. H.S. Teacher C.P.A.

AVERAGE RATINGS OF THREE GROUPS ON PARTS OF
THE STRONG, KUDER, AND HUNTER TESTS

Trait Rated	Average Percentile ratings: Superior Student Pilot Group	All Student Pilots Completing Course	Normal Non-pilot College Group
Neurotic Tendency	21	34	50
Self-sufficiency	64	46	50
Introversion	13	35	50
Dominance	82	66	50
Self-confidence	22*	50	50
Sociability	80**	42	50
Maturity of Emotions, Interests, Attitudes	90	69	50
Attitude toward Negro	45***	46	50
Attitude toward War	50	54	50
Attitude toward Economic and Labor Problems	73	65	50
Attitude toward Social Laws and Conventions	25	38	50
Attitude toward Government	78	60	50
Attitude toward Religion	51	51	50

* Low rating signifies high Self-confidence.

** High rating signifies low Sociability.

*** High ratings signify Liberalism; low ratings, Conservatism.

TABLE 24

RANK ORDER CORRELATIONS OF INSTRUCTOR'S RATINGS OF FLYING ABILITY WITH VARIOUS TEST SCORES

Test	Rho.
With Neurotic Tendency	-.19
With Self-sufficiency	.45
With Introversion	-.05
With Dominance	-.09
With Self-confidence	-.16*
With Sociability	.53**
With Maturity of Emotions-Interests-Attitudes	.49
With Liberalism toward Negro	-.08
With Liberalism toward War	-.11
With Liberalism toward Economic and Labor Problems	.25
With Liberalism toward Social Laws and Conventions	.04
With Liberalism toward Government	-.01
With Liberalism toward Religion	.02

* High test scores signify low Self-confidence.

** High test scores signify low Sociability.

SUMMARY AND CONCLUSIONS

This study presents an intensive analysis of four psychological tests: (1) the Bernreuter Personality Inventory, (2) the Pressey Interest-Attitude Test, (3) the Hunter Social Attitudes Test, and (4) the Strong Vocational Interest Blank. The instruments were administered to a total of 132 student pilots in training at the George Washington University, Maryland University, the University of North Carolina, and North Carolina State College.

Scores on these tests were analyzed to show: (1) the comparison of the successful student pilots' records with the normal control records, (2) the comparison of records of superior student pilots with the normal control records and records for the total group of pilots, and (3) the relationship between the psychological measures and ratings of relative flight competence.

Within the limitations of this investigation, the following tentative conclusions may be drawn:

1. All successful student pilot groups made lower average neurotic tendency scores (B_1N) on the Bernreuter Inventory than did the normal control group.

2. All successful student pilot groups were on the average less introverted than the normal groups as measured by the Bernreuter Inventory (B_3I).

3. The composite successful pilot group is more dominant than the normal (B_4D--Bernreuter Inventory).

4. Scores for the successful pilots and normals did not differ appreciably on the factors of Self-sufficiency (B_2S), Self-confidence, (F_1C), and Sociability (F_2S).

5. On the Pressey Interest-Attitude Test, the composite successful pilot group showed greater maturity of interest and attitude than the normal control group.

6. Of the four sub-tests on the Pressey instrument, two (opinions of strangers and expressions of like and dislike for types of persons) showed appreciable difference between the successful pilot and non-pilot groups; one (interest in various things) showed a slight difference between the two groups; and one (expression of worry or fears) showed no difference.

7. The measures on the Hunter Social Attitudes Test indicate that the successful pilot group and the normals showed no difference in liberalism and conservativism (means of -5 and -4 respectively).

8. In the six sub-tests of the Hunter test (attitude toward the negro, toward war, toward economics and labor, towards social life, and toward government and religion) there was little difference. On tests of attitude

toward the negro, economics and labor, and religion, differences of 2-4 points in the median scores were found. Differences in the other tests were even smaller.

9. The Strong Vocational Interest Blank was scored for nine occupations (Table 19). Among the successful pilots the largest percentage of "A" ratings (signifying that the subject possesses the interest patterns of those successful in the occupation) is for the occupation of Engineer with only a small percentage of a control group of students at The George Washington University choosing in this manner. For the occupation of Minister, 80% of the successful pilot group have rating in the "no interest category" while the control group showed only 52%. For the occupation of Certified Public Accountant, 84% of the successful pilot group show "no interest" ratings while only 46% of the controls did likewise. For Mathematician there is very little difference in the two groups.

10. Analysis of the scores of the "superior" group of pilots as compared with all pilots (washouts and less superior ones) and with normal non-pilot students showed the superior pilots to be characterized by low neurotic tendency, high self-sufficiency, low introversion, high dominance, high self-confidence, and below average sociability. They differ slightly from the total pilot population and markedly from the non-pilot Groups in these traits. A marked difference is shown in maturity of emotions, interests, and attitudes on the Pressey test for the three groups (ratings of 90, 69, and 50 for the superior, total pilot, and control groups respectively). The superior pilot group is on the average more liberal toward government and economic and labor problems, and more conservative toward social laws and conventions than the others. No differences were found among the three groups with regard to attitude toward the negro, war, or religion.

The occupational ratings give a predominance of high ratings to Engineering for the superior group and low ratings for Minister, Certified Public Accountant and Mathematician. There were no significant differences between the superior pilot group and the total pilot groups in these ratings.

11. When rank-order correlations were computed between a paired-comparison rating of 10 of the student pilots (the pilots who are discussed in Part I of this report) only self-sufficiency, sociability, maturity of emotions, interest, and attitudes, and attitude toward economic and labor problems demonstrated any relationship to the criterion (correlations of .45, .53, .49, and .25 respectively). All other correlations ranged between -.19 and .04.

Editor's Note: Tests of the significance of differences between groups are implied in the above report, although they are not presented. It should be observed, however, that statistically "significant" differences among the groups studied may be related to a number of uncontrolled factors of selection.

PART III

ANALYSIS OF TEST DATA ON 1938-39 PURDUE C.P.T. STUDENTS

by

E. LOWELL KELLY
and
A. P. JOHNSON
Purdue University

ANALYSIS OF TEST DATA ON 1938-39 C.A.A. STUDENTS[1]

INTRODUCTION

With the organization of the Experimental Civilian Pilot Training Program at Purdue University in the Spring of 1939, the Coordinating Committee in charge of this program felt the need for selective measures in addition to the regular physical examination in selecting applicants. Accordingly, a research program was designed to study the process of selection with the aim of finding some tests and measures which might aid in the future selection of student pilots. A series of tests and measures was assembled and administered to students participating in the C.P.T. Program for 1939. The results of these tests were then correlated with pilot success, at the conclusion of the training program, to determine whether any of the test performances were related to success in flying and might thus serve as selective devices in future programs.

At the time that this study was made, two measures of flying ability were available: (1) the instructor's day-to-day flight rating of his students, and (2) an over-all rating made by the Chief Instructor. After the preliminary analysis of these data was completed, the C.A.A. Flight Examiner's rating became available and it seemed desirable to expand the treatment of the results using this new criterion.

The present report is a presentation of the results of the complete analysis made of the tests and measures given to the students participating in the C.P.T. Program for 1938-39.

TESTS AND MEASURES USED

The testing program for this investigation included administering thirty-four different tests; these, together with other scores already available, made a total of thirty-nine measures available for each student pilot. The administration of these tests was conducted by three different divisions of Purdue University and the following discussions are grouped according to the divisions administering them. All test data as used in determining their relationship with flying ability are presented in Appendix A, Table A-1. In some cases it was necessary to transmute the original data into a different form in order that a correlational analysis might be made; the details of such transmutations of the data will be given with the discussion of the tests so treated. The original form of such data is given in Appendix A, Table A-2.

---

1 The writers wish to acknowledge the cooperation of the following people: W. A. Hiestand, C. L. Morgan, and S. Meikle of the local Purdue Committee; G. Satter, E. Ewart, L. Tuning, and James Wright, Research Assistants; and Captain L. I. Aretz, Chief Flight Instructor.

A. Tests Administered by the Personnel [illegible][2]

1. Minnesota Spatial Relations Boards.[3] This test consists of a series of four form boards into which the subject is to place fitted blocks in the shortest possible time. The score is the total time for the four boards.

2. Humm-Wadsworth Temperament Scale.[4] This is a pencil and paper inventory yielding seven different measures. In interpreting scores on this test it is necessary to consult the manual of the test. Either a very high score or a very low score may be indicative of maladjustment. However, for the purposes of this study the meaning of the score is not important since the correlation might be significant regardless of the direction of the score.

3. Revised Minnesota Paper Form Board.[5] This test belongs to a series of tests of which the Minnesota Spatial Relations Boards mentioned above are a part. It is designed to test one phase of mechanical aptitude.

4. Otis Intelligence Test, Higher Form Self-Administering.[6] This is a pencil and paper intelligence test in which the scores represent items correct and not I.Q.'s.

5. Scholastic Index.[7] This measure was secured from the Purdue University Registrar's Office and represents the cumulative scholarship average. The larger indices indicate higher scholastic averages.

6. Personality Rating.[7] This measure was secured from the forms issued by the Personnel Service. The measure represents the composite of a group of ratings by different individuals. The higher the score, the better the rating.

7. Orientation Tests.[8] These measures were secured by the Personnel Service from the Division of Educational Reference and represent percentile rankings of the students on the orientation tests given upon entering the University. The battery of tests from which these scores were obtained is made up of the A. C. E. Psychological Examination, the Iowa Placement Test in Mathematics, and the Purdue

---

2 Under the direction of Dr. A. P. Johnson, Purdue University, Lafayette, Indiana.

3 Paterson, Elliott, Anderson, Toops, Heidbreder - Minnesota Mechanical Ability Tests. The University of Minnesota Press, Minneapolis, 1930, pp. 311-312.

4 Humm, Doncaster and Wadsworth, Guy Jr. "The Humm-Wadsworth Temperament Scale," reprint from the American Journal of Psychiatry, 1935, 92, No. 1.

5 Paterson, Minnesota Tests, Op.cit., p. 310. Quasha, W. H. and Likert, R. "The Revised Minnesota Paper Form Board Test," The Journal of Educational Psychology, 1937, 20, 197-204.

6 Otis, A. S. "Manual of Directions" (World Book Company, 1922).

7 Further information may be obtained from the Personnel Office or the Registrar's Office, Purdue University, Lafayette, Indiana.

8 See Division of Educational Reference, Purdue University, Lafayette, Indiana.

Placement for ... English. The better ... the higher the score the better the performance.

B. Measures secured by the Biology Department.[9]

8. Vital Capacity.[10] This is a measure of lung capacity in cubic centimeters; the higher the score the greater the capacity.

9. Flack-Wodham Index.[11] This is an index of physical fitness and is a composite of three different measures: 1) expiratory force as measured by a mercury U tube in millimeters, 2) persistence measured in seconds of holding a column of mercury at 40 millimeters, and 3) breatholding in seconds. The formula for obtaining the index as given by Schneider is:

$$\text{F-W Index} = \frac{\text{Per} \times \text{Exp} \times \text{Br}}{100 \times \frac{(\text{Age in years})}{(4)} \times 1.807}$$

According to Schneider the average index for a group of boys was found to be .45. It will be noted that this is a great variance with the data listed in Table 1.

10. Schneider Index.[12] This index is a measure of physical fitness which is used to test applicants for the Army Air Corps. According to Armstrong, an index of 18 to 14 is considered excellent; 13 to 11 very good; 10 to 9 fair; 8 doubtful; and below 8 unsatisfactory.[13]

11. Nystagmus Test.[14] This is a measure of the time in seconds necessary for the eyes to return to normality after the subject has been rotated rapidly in a hunched over, sitting position.

12. Pulse Rate Change.[15] This is a measure of the change in the pulse rate resulting from the rotation of the subject in the Nystagmus test. Since in some instances there was an increase in pulse rate, while in others there was a

---

9 Under the direction of Dr. W. A. Hiestand, Purdue University, Lafayette, Indiana.

10 Schneider, E. C. Physiology of Muscular Activity. W. B. Saunders, Philadelphia, 1933, p. 320.

11 Ibid. Schneider cites the original reference; Cripps, L. D. "The Application of the Air Force Physical Efficiency Tests to Men and Women," Special Report Ser. No. 14, 1924, Medical Research Council, London.

12 Schneider, op. cit., p. 321.

13 Armstrong, H. G. Principles and Practice of Aviation Medicine. Williams and Wilkins Company, Baltimore, 1939, p. 40.

14 Ibid., page 122.

15 Ibid., page 122.

decrease, it seemed desirable to make all of the measures positive and this was accomplished by adding "15" to each measure. (The original data are presented in Appendix A, Table A-2.)

C. Tests given by the Division of Education and Applied Psychology.[16]

13. Dexterity Test.[17] In this test the scores represent the number of pins placed in the proper holes in a board in 30 seconds. Two trials were given for each hand and two trials for both hands together. For correlational purposes these scores were summated to give one total score.

14. Finger Dexterity Test.[18] This test is similar to the one mentioned above except that the score represents time in seconds to complete the first half of the board and time in seconds for the second half of the board. These two scores were summated for correlational purposes.

15. Tweezer Dexterity Test.[19] This test is similar to the one mentioned above except that tweezers were used to manipulate the pins. The scores were again time scores for the first and second halves of the board.

16. Morgan Tracing Test.[20] This test consists of a series of 15 styli graduated in size which are to be drawn along a tracing board having electrical contact on the sides -- a contact constitutes an error. Four trials were given with each stylus; a maximum of four contacts was allowed with each stylus. The score represents the styli number weighted by the number of contacts made with each stylus. The lower the score the better the performance.

17. Simple Sound Reaction Time.[21] In this test, simple reaction time to a single sound stimulus was measured to the nearest 1/200 of a second.

18. Simple Light Reaction Time.[22] Simple reaction time to a single light stimulus was measured to the nearest 1/100 of a second.

19. Light-choice Reaction Time.[23] Reaction time was measured to the

---

16 Under the direction of Dr. C. L. Morgan, Purdue University, Lafayette, Indiana.

17 Bingham, W. VanDyke. Aptitudes and Aptitude Testing. Harper Brothers. New York, 1937, p. 281. (The Dexterity Test is a modification of the Finger Dexterity Test.)

18 Ibid. p. 281.

19 Ibid. p. 284.

20 Devised by Dr. C.L. Morgan.

21 Woodworth, R.S. Experimental Psychology, Henry Holt, New York, 1938, Ch. XIV, pp. 298-339, gives a complete discussion of the measurement of reaction time.

22 Ibid. pp. 298-339.

23 Ibid. pp. 298-339.

nearest 1/120 of a second to a light stimulus requiring a choice in response.

20. Steadiness Test.[24] This instrument is a metal plate pierced with a series of holes which grow progressively smaller. The subject inserts a metal stylus into the holes one at a time. Every contact with the sides of the hole is registered electrically and counts as an error. The steadiness score represents the number of the first hole in which more than two contacts were made. The higher the score (the smallest hole in which more than two contacts with the periphery are made) the better, i.e., the steadier the movements of the hand.

21. Whipple Tracing Test.[25] This test provides another measure of steadiness. The subject inserts a metal stylus in a groove between two graduated metal plates. Contact with the sides of the groove (the metal plates) is recorded electrically and counts as an error. The score represents the average distance in millimeters traced before contact for five trials. Measurements were made of both right hand and left hand with movements away from and toward the body.

22. Whipple Tapping Test.[26] Scores represent number of taps made by a metal stylus on metal plates. These taps are recorded electrically. Two 15-second trials for the right hand and for the left hand were given.

23. Percent Loss or Gain for Understanding Speech (audiometer).[27] This is a test of hearing ability as measured by the Western Electric 6A Audiometer. Auditory Acuity is measured in decibels for various frequencies for each ear.[28] Seven different measures are given for each ear. West gives a technique whereby one single score for "understanding speech" can be found.[29] Scores were computed by this technique. Table A-2 (Appendix A) presents the loss (shown by the minus sign) or gain of hearing in percent. For correlational purposes it was necessary to convert these measures to positive values; this was done by adding 12 to each measure (see Appendix A, Table A-1, for converted data).

---

24 Whipple, G.M. Manual of Mental and Physical Tests, Warwick & York, Baltimore, Maryland, 1924, Pt. I, p. 156.

25 Ibid. p. 151.

26 Ibid. p. 151.

27 Western Electric 6A Audiometer, Instruction Bulletin No. 841 (Western Electric Company, New York).

28 Fletcher, Harvey, Speech and Hearing, D. Van Nostrand Company, New York, 1929, Ch. VI, Methods of Testing Acuity of Hearing, p. 198.

29 West, Kennedy, and Carr The Rehabilitation of Speech, Harper and Brothers, New York, 1937, p. 397.

24. Depth Perception as measured by the Howard-Dohlman Tests.[30] Five measurements were made from the front and five from the back. The score represents the average error of the ten trials in millimeters; the smaller the score the more accurate the perception.

25. Stereopsis.[31] This is a measure of the ability to achieve three dimensional vision dependent upon coordination of the eyes as measured by the telebinocular. The score represents the highest percent level at which no errors were made.

26. Telebinocular Tests.[32] This instrument measures various functions important in vision. The first test is used merely to indroduce the subject to the instrument. Results were interpreted into P(ass) or F(ail). It will be noted that only five of the subjects failed any of the tests. Since the student pilots were selected on a basis of physical fitness with considerable attention given to vision, it is to be expected that there would be few failures on this test.

THE CRITERIA OF FLYING ABILITY[33]

27. Flight Instructor's Rating (Flight Rating Index). As one measure of flying ability a composite index was obtained from the ratings given the student pilots by their flight instructors in accordance with the CAA training requirements. Student pilots were rated during each period (usually a half-hour long) of the three stages of flight training on a five-point scale ranging from a rating of "1", excellent, to a rating of "5", poor. The index was obtained by summating the weighted ratings throughout the training period and dividing this sum by the total number of ratings given. It will be noted that the total range for the flight rating index is quite limited, being from 1.33 to 2.65 (see Appendix A, Table A-3).

28. Chief Instructor's Over-all Rating. As another measure of flying ability an over-all rating of the pilots was received from the Chief Instructor. This rating was made after the close of the training period and represents the Chief Instructor's opinion of the student's flying ability when all factors are taken into consideration. It will be noted (a) that the Chief Instructor, Captain Aretz, felt capable of listing the ten best pilots in order, (b) that of a group of 31 average pilots, only the best and the poorest could be differentiated, and (c) that the nine poorest pilots are listed in order. For purposes of analysis it was felt that the best of the average pilots belonged with the good group and that the poorest of the average pilots belonged with the poor group (see Appendix A, Table A-3).

---

30 Armstrong, Aviation Medicine, op. cit., pp. 77-78.

31 Keystone Diagnostic Series (Card, Instructions for Use of Stereometric Slides), Keystone View Company, Meadville, Pennsylvania.

32 Betts, E.A. The Prevention and Correction of Reading Difficulties. Row, Peterson & Company, Evanston, Illinois, 1936, p. 323.

33 The reliabilities for these criteria are presented in Table 1.

29. C.A.A. Examiner's Rating. At the time that the student pilots were examined for their private license, the C.A.A. Flight Examiner rated them as to their proficiency in flying. This rating was secured from the C.A.A. in Washington, D. C. with the hope that it might show a higher degree of relationship with the tests than the other criteria.

30. Composite Rating. Since each of the three pilot ratings discussed above was made by a different judge, it seemed reasonable to assume that a combination of these ratings might yield a composite rating which would be more valid than any single measure of flying ability. Each of the three ratings was transmuted to sigma scores. In the case of the Chief Instructor's rating, this was done by Hull's Transmutation technique.[34] (See Appendix A, Table A-3, for sigma scores of the Chief Instructor's rating.) The sigma scores of the three ratings were then summated and averaged, yielding a single composite index of flying ability (see Appendix A, Table A-3).

31. Grades in Navigation and Meteorology. At the time that the C.A.A. Examiner's rating was obtained, examination grades made by the student pilots on (a) Navigation, (b) Meteorology, and (c) the Civil Aeronautics Regulations were also obtained. These grades were based upon written tests taken at the time of application for a private license. (See Appendix A, Table A-3.)

RESULTS

A. The Reliabilities of the Various Tests

One of the more important things to know about a test is its reliability, i.e., to determine whether it measures consistently whatever it measures. There are two general methods by which one can determine the reliability of a test. One method is to give the test two times and correlate the first performance with the second; naturally the more agreement, the more consistently the test measures whatever it measures, and the higher the correlation. The other technique is a statistical manipulation having the effect of splitting the test into two halves. In case of a pencil-and-paper test this may be achieved by scoring the test on odd numbered items and again on even numbered items. From these two measures on the same test a correlation may be determined. In such a procedure we actually have the correlation between halves of the test; but we are interested in determining the reliability for the total test. The Spearman-Brown formula[35] gives us a mathematical technique whereby we can estimate the reliability of the whole test knowing the reliability of any fraction of it.

The reliabilities of the tests used in this study are shown in Table 1, page 68. In some instances it was felt that there was no necessity of computing reliability coefficients of the tests as adequate ones were supplied with the test. In other cases the data were of such a nature as to make it impossible

---

34 Guilford, J.P. Psychometric Methods. McGraw-Hill, New York, 1936, p. 248.

35 Guilford, J.P. Psychometric Methods. McGraw-Hill, New York, 1936, p. 418.

TABLE 1

RELIABILITIES OF VARIOUS TESTS AND CRITERIA

Test No.	Tests & Criteria	r	P.E.
	I. Reliabilities Computed from the Test Data:*		
1.	Minn. Spatial Relations Boards	.74	.04
13.	Dexterity Test (Sum of odd vs. sum of even trials)	.80	.04
14.	Finger Dexterity (Score on 1st half of board vs. 2nd half)	.87	.03
15.	Tweezer Dexterity (Score on 1st half of board vs. 2nd half)	.78	.04
17.	Sound Reaction Time (odd vs. even trials)	.88	.02
18.	Light Reaction Time (odd vs. even trials)	.76	.05
19.	Light-choice Reaction Time (odd vs. even trials)	.68	.06
21. a.	Whipple Tracing: right hand, toward (ave. of trials 1 and 2 vs. ave. of trials 4 and 5)	.20	.18
22. a.	Whipple Tapping: right hand (trial 1 vs. trial 2)	.71	.06
b.	Whipple Tapping: left hand (trial 1 vs. trial 2)	.88	.02
24.	Depth Perception (ave. of trials 1 and 4 vs. ave. trials 2 and 5)	.76	.07
27.	Flight Rating Index (index for odd vs. even half hours)	.98	.00
	II. Reliabilities Supplied by the Authors of the Tests:		
2. a.	Humm-Wadsworth Scale: Normal	.82	.04
b.	" " " Hysteroid	.85	.02
c.	" " " Cycloid Manic	.73	.03
d.	" " " Cycloid Depressed	.88	.01
e.	" " " Schizoid Autistic	.88	.01
f.	" " " Schizoid Paranoid	.70	.03
g.	" " " Epileptoid	.75	.02
3.	Revised Minnesota Paper Form Board	.85	.01
4.	Otis Intelligence Test, higher form	.92	.01
6. a.	Personality rating (odd vs. even items)	.86	.03
b.	Personality rating (rating vs. rerating)	.91	.01
7. a.	ACE Psychological Examination	.94	--
b.	Iowa Placement Tests in Mathematics	.87	--
c.	Purdue Placement Tests in English	.95	--
	III. Tests of Such Obviously Low Reliability as Not to Merit Computations:		
20. a.	Steadiness Test: right hand		
b.	Steadiness Test: left hand		
21. b.	Whipple Tracing: right hand; away		
c.	" " : left hand; toward		
d.	" " : left hand; away		

IV. Tests For Which No Reliability Could be Computed:**

8.	Vital Capacity	12.	Change in Pulse Rate
9.	Flack-Goolham Index	16.	Morgan Tracing
10.	Schneider Index	23.	Audiometer (hearing)
11.	Nystagmus Test	25.	Stereopsis

26. Telebinocular Tests of Vision

* The reliabilities in this section have all been and stepped up by means of the Spearman-Brown Prophecy Formula.

** Reliability was not computed for the Scholastic Index, nor for Criteria 28-31.

to compute the reliability of the test. In a few cases the reliability was obviously so low as not to merit computation. In those tests in which the reliability was computed it was necessary to score the test on alternate items or performances.

B. Intercorrelations of the Various Criteria

Since each rating is presumably based upon the same consideration of the student's flying ability, it is reasonable to assume that there would be fair degree of correlation between the various criteria. If a high degree of relationship exists, then we have good evidence of the reliability (objectivity) of our ratings. Following this assumption, correlations were computed between the three different measures of flying ability. As will be seen in the following table (Table 2), the correlations are disappointingly low.

TABLE 2

INTERCORRELATIONS BETWEEN RATINGS OF FLYING ABILITY

	Flight Instructor (27)	CAA Examiner (29)
Chief Instructor's Rating (28) (sigma scores)	+.216	+.245
CAA Examiner's Rating (29)	+.375	

It would appear that each rating is based upon factors different in large part from those considered in the other ratings. As a check upon this point it was felt advisable to determine the highest degree of correlation which could be achieved when the best possible combination of two ratings was used to predict the 3rd; this was done by computing the multiple correlation coefficients (Table 3). It can be seen that even the best possible combination of rating, $R_{29(27,28)}$, yields a correlation which is unsatisfactorily low. This point will be further discussed later in the report.

TABLE 3

MULTIPLE R'S OF CORRELATIONS BETWEEN RATINGS

(27) Flight Instructor's Rating
(28) Chief Instructor's Over-all Rating
(29) C.A.A. Examiner's Rating

$R_{27(29,28)} = +.40$
$R_{28(27,29)} = +.28$
$R_{29(27,28)} = +.41$

C. Correlations of Tests and Measures with the Criteria

As has been stated, the major purpose of this study was to find some tests and measures which might be of value in aiding in the selection of student pilots for future flight training programs. To determine whether any of the tests and measures used might be of value in building a predictive battery, it was necessary to determine the correlation existing between them and the criteria. Coefficients of correlation were computed between all of the tests and measures (except the Telebinocular Tests) and the various criteria (see Table 5).

It will be remembered that in some of the tests a high score was considered as indicative of poor performance; also in the case of the Flight Instructor's Rating a high numerical rating was considered poor. Because of this fact it was necessary to reflect the signs of some of the correlations. In all cases positive correlations listed in the table are to be interpreted as meaning that a good performance on the test is related to a good rating of flying ability.

In studying the table of correlation coefficients it will be seen that none of them shows a high degree of relationship. By following Fisher's[36] test of the significance of "r" we can determine the possibility that the "r" obtained may have occurred by chance. According to this interpretation of the significance of "r", there are only 5 chances in 100 that a correlation as high as .285 would occur in a sample of 50 if no true correlation existed in the population. Similarly, there is only 1 chance in 100 that a correlation as high as .368 would so occur -- this might be termed a very significant correlation. Table 4 shows the "r's" which approach these figures. Of approximately 100 correlations computed, 7 were significant at the 5 percent level and 2 at the 1 percent level.

TABLE 4

CORRELATIONS REACHING TWO LEVELS OF SIGNIFICANCE

Test or Measure		Flight Instructor (27)	CAA Examiner (29)	Composite Rating (30)
Humm-Wadsworth, Cycloid Depressed	(2,d)	.054	-.301*	-.176
" " Schizoid Paranoid	(2,f)	-.275	-.213	-.397**
" " Epileptoid	(2,g)	-.057	-.287*	-.260
Minnesota Paper Form Board	(3)	-.179	.297*	-.164
Scholastic Index	(5)	-.099	-.410**	-.265
Iowa Placement Tests in Math.	(7,b)	-.253	-.297*	-.210
Light-choice Reaction Time	(19)	.229	-.343*	.107

* Significant at the 5% level ** Significant at the 1% level

36 Guilford, J. P. Psychometric Methods. McGraw-Hill, New York, 1936, page 335. See also Table K, pages 548-549 for significant values of "r".

CORRELATIONS (and P^{**}) BETWEEN TESTS OR MEASURES AND FOUR VARIOUS CRITERIA

Test No.	Tests or Measures	Flight Instructor (r)	CAA Examiner (r)	Composite Rating (r)	Overall by Chief Instructor (χ^2)**
1.	Minn. Spatial Relations Boards	-.197	.013	-.199	0.82
2.a.	Humm-Wadsworth; Normal	.035	.231	.229	3.14
b.	" " Hysteroid	-.099	-.255	-.119	0.44
c.	" " Cycloid Manic	.205	-.172	.032	1.47
d.	" " Cycloid Depressed	.054	-.301	-.176	0.94
e.	" " Schizoid Autistic	.128	-.225	.007	1.28
f.	" " Schizoid Paranoid	-.275	-.213	-.397	3.30
g.	" " Epileptoid	-.057	-.287	-.260	3.03
3.	Minnesota Paper Form Board	-.179	.297	-.164	2.77
4.	Otis Intelligence, higher form	-.134	.062	-.035	1.69
5.	Scholastic Index	-.099	-.430	-.265	0.74
6.	Personality Rating	-.134	-.149	-.171	4.54
7.a.	ACE Psychological Examination (%ile)	-.119	-.027	-.139	1.15
b.	Iowa Placement Test in Math. (%ile)	-.253	-.297	-.210	0.75
c.	Purdue Placement Test in English (%ile)	.075	-.019	-.105	5.00
8.	Vital Capacity	.020	.119	.111	1.02
9.	Flack-Woolham Index	.140	.006	.070	1.38
10.	Schneider Index	-.116	-.157	-.230	1.60
11.	Nystagmus Test	.166	.190	.330	4.05
12.	Pulse-rate change*	-.140	.073	-.063	1.18
13.	Dexterity Test (summation)	.051	-.143	-.044	0.59
14.	Finger Dexterity	.164	.200	.204	0.29
15.	Tweezer Dexterity	.068	.062	-.066	1.49
16.	Morgan Tracing (meaningless score)	-.180	.197	.037	2.56
17.	Simple Sound Reaction Time	-.147	.049	-.112	1.60
18.	Simple Light Reaction Time	.044	-.016	-.083	2.50
19.	Light-Choice Reaction Time	.229	-.343	.107	5.31
20.a.	Steadiness Test; right hand	.060	.006	-.004	1.61
b.	" " left hand	.127	.029	.067	0.21
21.a.	Whipple Tracing; right hand, toward	.100	-.010	.111	0.06
b.	" " right hand, away	.046	.094	.004	
c.	" " left hand, toward	.023	-.195	-.083	
d.	" " left hand, away	-.072	-.096	-.029	
22.a.	Whipple Tapping, right hand	.168	.209	.176	
b.	" " left hand	.113	.143	.131	
23.	Hearing λ (Audiometer)	.253	.198	.158	1.18
24.	Depth Perception	-.057	.104	.048	2.51
25.	Percent Stereopsis	-.057	-.138	-.042	0.42
28.	Chief Instructor's Rating (Sigma score	.216	.245		
29.	CAA Examiner's Rating	.375			
		C.A.R. (31,c)	NAV & MIT (31,a) & (31,b)		
(4)	Otis Intelligence, higher form	.308	.181		
(5)	Scholastic Index	-.190	.280		
(7.a)	ACE Psychological Examination (%ile)	.002	.004		

*Note: Scores above 15 indicate positive acceleration of pulse rate following exercising; scores below 15 indicate negative acceleration.

** With 2 degrees of freedom, Chi-squared must equal 5.991 for the relationship to be significant 95% chances in 100

None of the tests correlates significantly with the flight instructor's rating (one, Schizoid-Paranoid, approaches significance), six of them correlate significantly with the CAA Examiner's rating; and one correlates significantly with the composite rating.

It was thought that the combining of the flight instructor's rating the CAA Examiner's rating, and the Chief Instructor's over-all rating might give a composite rating which would show more relationship to the test than did any one of the three other ratings. Actually the order of the obtained "r's" with this criterion is quite similar to those obtained with the other two criteria. In several instances the size and/or the direction of the correlation with the composite rating is different from that of the other two ratings. Actually the effect of combining the three ratings to attain a composite rating seems to have been one of cancelling the value of the other ratings in several instances (see correlations with Light-Choice Reaction Time, Table 4). This is explained by the fact that the direction of relationship is not always the same with each of the criteria.

Since the rating by the Chief Instructor was not, strictly speaking, a continuous distribution, the Chi-Square test was used to determine the presence or absence of a significant relationship between these ratings and the test scores.[37] For the purpose of this study it was necessary to plot a six-fold table to insure enough cases in each of the cells. It will be noted that none of the obtained Chi-squares equals 5.991, i.e., none of the relationships are significant at the 5% level.[38]

DISCUSSION OF FINDINGS

On reviewing the data presented in the last section, the following facts seem to stand out as significant:

1. Since each of the three criterion ratings is ostensibly a measure of the same thing, ability to fly an airplane, it seemed reasonable to assume that they would show a fairly high degree of intercorrelation. Yet, as is shown in Table 2, no one of the criterion ratings correlates highly with any other one. The fact that they do not intercorrelate highly does not dismiss the possibility that one of these ratings is a good criterion. Lack of high intercorrelation merely indicates that they are not (as was presumed) measuring the same thing. It should also be pointed out that high intercorrelations among the criterion measures would not definitely establish their value as criteria of flight proficiency. The selection of criteria must be finally determined on some basis other than their correlations with other criteria.

2. While the reliability of the flight rating index was found to be very high, .98, yet the range is extremely limited, being from 1.33 to 2.65. The ratings tend to center around the lower part of the possible distribution, i.e., every student pilot tends to receive either a "good" or an "above average" rating These two facts seem to indicate a tendency for the raters to think in relative terms for the amount of training taken rather than in absolute terms and conse-

---

37 Guilford, op. cit., page [illegible]

38 Guilford, op. cit., [illegible]

quently to give the student pilot approximately the same rating for each period of instruction.

3. Of all the correlations between the tests and the criteria, only 7 were found to be of statistical significance and six of these were with one criteria, the CAA Examiner's rating. These correlations, however, do not necessarily establish either the validity of the tests or the validity of the CAA Examiner's ratings. They merely indicate that for this particular study the CAA Examiner's rating is probably the most useful of the criteria employed. It must also be remembered that the correlations between these criterion measures were extremely low indicating that each of them was assessing a different characteristic. A criterion against which a predictive test battery is to be validated would surely need to be a great deal more valid than any of those employed in this study.

4. One would assume that the group participating in the C.P.T. program was a very select group. The very fact that they have successfully surmounted the hurdles of the educational system to become college students would indicate this. Add to the selective factor of the educational system the fact that they had to pass a rigorous physical examination to qualify for the training and we have further support of our assumption. It is interesting to note a further supporting factor from the test results: it was found that in the Minnesota Spatial Relations Form Boards no Purdue C.A.A. student pilot scored below the 60th percentile for engineering students and that the average C.A.A. score fell at the 95th percentile for engineering students at the University of Minnesota.[39]

5. Out of a total battery of 39 psychological and physiological tests purporting to measure some characteristic related to or essential to flying, only seven different tests showed a higher degree of relationship to any of the criteria than might have occurred by chance.

6. Academic achievement as shown by grade point average appears to be negatively related to flying ability. This is shown by the correlation of -.41 between scholastic index and the C.A.A. Flight Examiner's Rating (see Table 4). While one cannot be certain that the true correlation is as high as the one obtained, yet it does seem safe to assume that the relationship is not a positive one. Evidently the practice of selecting student pilots on the basis of scholastic ability is not justifiable on the grounds that good "book students" make good "flying students."

The correlations between mathematical ability as measured by the Mathematics Orientation Test and the three criteria of flying ability indicate little relationship between this capacity and ability to learn to fly. All of the correlation coefficients are low negative. Measures of intelligence also show zero correlation with criteria of flight success.[40]

---

39 Miller, D.C. and Dean, Ross. A Report on the C.A.A. Testing Program, on file in Mr. A.P. Johnson's office, Personnel Office, Purdue University, Lafayette, Indiana.

40 Editor's Note. It must be pointed out that the unreliability of the criteria used in this study may be responsible for the apparent lack of relationship between the flight ratings and these measures. Other studies tend to show a positive relation between mathematics, intelligence, and the ability to learn to fly.

CONCLUSIONS

1. The lack of a valid measure of flying ability is clearly indicated by the results of this study. With the imperative need for selecting only those applicants who will become successful pilots and for discriminative measures of pilot proficiency it is necessary that future research be directed toward the construction of adequate criteria.

2. There is no justification for the selection of student pilots from among those making superior marks in college courses. The correlation between intelligence and flying ability is about zero[41] but scholastic standing is somewhat negatively correlated with success in flying. If college marks are to be used as a selective factor, the best student pilots are likely to come from those making average rather than superior academic grades.

3. No conclusion can be drawn as to the efficacy of the tests and measures used in this study in aiding in the selection of future student pilots. The fact that very few tests correlated significantly with any one measure of flying ability is nullified by the corresponding fact that no measure of flying ability correlated highly with any other measure of flying ability. It would seem that the predictive value of these tests can only be determined at some later date when a more valid measure of flying ability has been devised.

---

41 Editor's note: The fact that two years of college were required for acceptance for flight training at the time of this study probably accounts for the low correlation between intelligence and flying ability. As the educational standards are relaxed, selection on the basis of intelligence tests should become increasingly important.

APPENDIX A

ORIGINAL DATA

TABLE A-1

TEST SCORES OF C.P.T. STUDENTS

Pilot No.	Minn. Spat. Rel	Humm-Wadsworth Temperament Scale N	H	M	D	A	P	E	Minn. Pap. Form Board	Otis Intell.	Schol. Index	Personal Rating
1.	871	52	29½	41	35	16	16	25	46	43	3.80	84.2
2.	890	46	26	31	27	20	23	13	40	43	2.15	78.7
3.	939	58	23	29	21	11	9	17	53	54	4.59	84.9
4.	708	56	43	24	25	24	16	20	50	55	3.74	82.4
5.	749	48	42	45	41	35	26	19	48	61	5.80	----
6.	802	48	38	26	37	17	34	21	43	56	4.60	92.5
7.	820	45	49	24	29	22	32	31	49	41	4.13	91.9
8.	650	35½	33	44	55	27	21	21	57	54	5.25	85.1
9.	867	52	17	24	15	19	16	20	43	61	3.77	72.5
10.	890	42	35	46	58	28	20	24	48	56	4.29	87.4
11.	715	47	32	34	36	38	18	16	50	58	4.08	84.8
12.	1007	40	46	38	46	33	33	19	42	53	3.74	77.4
13.	798	49½	29	23½	25	15½	28½	27	54	57	4.02	85.5
14.	596	27	56	43	59	30	43	22	56	55	4.85	95.1
15.	891	26	54	23	61	35	38	16	51	48	3.81	78.2
16.	766	50½	36	19	20	25	25	18	57	50	3.90	----
17.	807	51	25	34	33	30	17	31	48	54	4.59	82.2
18.	703	57	30	34	45	51	13	22	52	56	3.84	85.1
19.	943	51	45	25	24	22	19	16	39	58	3.48	84.4
20.	732	26	58	32	54	36	24	20	53	52	4.46	77.3
21.	759	44	25	36	33	26	16	29	53	56	4.54	83.2
22.	882	45	15	23	24	12	14	22	--	--	3.97	82.3
23.	759	39	35	47	43	19	22	27	48	62	3.76	75.0
24.	688	42	35	29	38	26	27	28	62	57	3.97	85.0
25.	762	40	53	40	46	24	26	22	56	49	3.91	79.0
26.	810	38	55	23	23	14	32	20	49	56	3.94	87.5
27.	848	40	45	27	36	37	45	22	44	47	4.06	78.0
28.	701	51	35	27	31	26	28	29	48	58	3.69	83.9
29.	832	46	18	30	25	45	19	17	54	57	4.42	83.3
30.	814	58	19	18	15	19	24	12½	51	47	4.10	82.1
31.	729	43	37	36	30	15	21	30	51	48	3.91	81.5
32.	786	32	58	46	56	24	28	31	53	52	4.24	86.9
33.	712	61	20	12	9	14	11	19	54	59	4.76	86.0
34.	897	42	35	25	27	31	23	16	40	46	3.10	75.0
35.	745	19	53	31	49	38	36	25	51	60	3.77	75.2
36.	651	56	20	29	24	19	13	24	50	57	4.72	85.6
37.	673	44	32	36	45	40	26	18	64	64	4.27	----
38.	634	44	53	30	31	29	25	17	58	61	4.17	86.7
39.	802	56	32	25	19	6	9	25	54	62	5.37	87.6
40.	657	44	51	36	34	44	20	25	59	56	4.66	77.9
41.	829	50	38	29	26	19	19	15	38	51	3.80	85.6
42.	706	38	54	35	48	41	26	21	49	53	4.06	78.8
43.	697	50	39	28	23	17	23	17	53	52	4.48	84.7
44.	742	47	52	25	12	31	27	14	46	55	2.68	92.0
45.	778	61	37	37	29	14	13	11	45	52	3.62	81.3
46.	800	32	26½	32	42½	36	29	19½	49	52	5.20	88.1
47.	681	48	41	28	31	17½	19	14½	56	55	4.26	89.0
48.	913	52	23½	33	29	31	15	19	--	--	3.72	90.0
49.	811	44	56	41	41	35	34	17	53	61	3.94	84.5
50.	827	42	43	26	29	29	34	21	52	43	3.79	80.8

TABLE A-I (cont.)

TEST SCORES OF C.F.T. STUDENTS

Pilot No.	Psychol. %ile	Math. %ile	English %ile	Vital Capacity	Flack-Woolham Index	Schneider Index	Nystagmus	Pulse Rate Change	Dexterity (Summation)	Finger Dexterity	Tweezer Dexterity	Morgen Tracing	Sound R.T.	Light R.T.
1.	12	11	6	4400	504.43	9	21	21	128	458	440	211	19	27
2.	12	12	9	5500	179.01	7	26	15	109	460	340	129	20	23
3.	79	92	69	5050	302.95	13	29	17	125	430	377	166	20	24
4.	81	25	—	4350	89.86	11	12	15	103	486	487	129	23	25
5.	97	--	--	4400	39.19	7	22	13	125	473	301	102	18	22
6.	84	75	60	5250	342.80	11	22	15	117	528	297	114	18	24
7.	15	59	17	4600	121.01	11	13	16	97	523	339	60	20	26
8.	83	84	72	4750	122.37	13	20	13	115	440	419	72	20	23
9.	82	90	70	5050	147.97	13	11	13	106	447	485	142	18	23
10.	48	26	21	5250	230.39	8	19	15	107	500	365	180	23	30
11.	88	71	95	4200	140.11	12	12	15	115	455	430	124	18	22
12.	67	59	73	4750	58.70	13	20	11	110	470	380	152	20	[illegible]
13.	47	26	47	4250	110.70	14	14	5	106	395	442	101	24	24
14.	32	25	49	5000	148.80	13	30	15	121	473	364	86	17	23
15.	59	54	35	4050	140.14	12	11½	15	105	481	399	73	21	24
16.	50	--	--	4650	107.78	13	12	15	119	509	382	182	18	22
17.	69	86	70	4750	105.67	12	30	15	103	495	451	142	27	29
18.	65	33	49	3900	89.21	11	14	11	112	442	390	88	18	24
19.	55	45	53	4600	111.60	11	14	15	98	484	450	225	14	21
20.	70	--	40	3750	38.82	13	11	15	122	427	321	74	18	23
21.	--	--	--	4500	259.19	12	12	7	106	469	438	155	19	22
22.	90	73	96	3250	85.37	7	13	21	110	486	385	154	18	24
23.	74	89	76	4450	114.75	7	12	15	125	471	---	179	23	24
24.	71	78	53	4400	73.50	12	28	17	117	384	363	44	20	24
25.	82	81	41	5000	263.34	6	24	13	110	458	429	126	26	25
26.	69	73	55	3950	66.39	8	20	15	96	482	424	102	21	28
27.	75	90	38	4250	73.31	10	23	15	98	645	485	270	17	23
28.	52	--	20	4650	312.66	7	14	17	101	492	446	213	26	28
29.	89	90	95	4750	244.58	14	13	15	114	484	400	124	18	23
30.	66	39	75	4000	100.70	10	13	18	106	455	359	15	17	23
31.	61	70	56	4350	106.92	10	19	15	117	459	347	204	25	26
32.	33	68	7	4750	209.43	8	22	8	112	448	437	102	23	23
33.	81	81	46	4000	166.10	11	13	26	100	445	342	44	26	28
34.	44	34	54	4250	112.05	11	14	15	104	518	336	45	18	25
35.	94	98	77	4300	82.67	14	23	15	97	512	427	74	19	23
36.	68	43	83	5450	273.00	12	18	9	115	484	368	129	23	28
37.	94	99	86	3550	153.75	12	19	15	113	398	374	100	19	24
38.	58	61	14	4250	53.60	10	23	9	116	440	444	60	16	22
39.	49	73	49	4750	91.52	15	20	9	114	419	380	57	21	22
40.	64	72	54	4950	172.26	9	14½	11	118	410	420	215	19	23
41.	26	72	25	4250	183.15	15	17	15	104	479	328	88	19	23
42.	64	--	--	4200	223.09	10	8	16	120	442	352	215	25	22
43.	50	56	24	4550	34.56	9	14	16	109	456	429	60	21	27
44.	62	62	37	5100	191.13	11	11	15	134	427	317	155	22	22
45.	38	36	17	3750	96.57	12	17	13	120	459	366	114	15	23
46.	87	98	80	3950	90.75	11	18	--	110	470	532	74	23	29
47.	65	72	24	4250	105.97	15	10	13	112	387	320	88	21	28
48.	53	94	47	4050	81.84	7	8	13	98	486	515	15	25	28
49.	45	42	18	4200	178.04	14	26	30	108	416	424	88	23	26
50.	68	37	16	4500	83.48	10	17	15	122	402	321	—	25	25

TEST SCORES OF C.P.T. STUDENTS

Pilot No.	Light RT Choice	Steadiness Test Right	Steadiness Test Left	Whipple Tracing To	Whipple Tracing Away	Whipple Tracing To	Whipple Tapping Away	Whipple Tapping Right	Whipple Tapping Left	Audiometer	Depth Percep.	% Stereopsis	Telebinocular Test Trial	Both eyes	Left eye	Right eye	Stereopsis	Vert. Imbal.	Far Fusion	Lat. Imb Far	Lat. Imb N.	Near Pt.	Ametropia-N	Ametropia-Far	Color-blind
1.	50	5	5	9	16	7	3	157	152	15.74	10.9	70	P	P	P	P	P	P	P	P	P	P	P	P	P
2.	49	6	4	14	19	19	16	144	114	28.85	11.2	80	P	P	P	P	P	P	P	P	P	P	P	P	P
3.	47	6	4	8	10	17	10	137	108	15.79	8.6	80	P	P	P	P	P	P	P	P	P	P	P	P	P
4.	45	7	5	16	14	17	21	103	94	19.96	10.0	40	P	P	P	P	P	P	P	P	P	P	P	P	P
5.	51	6	5	11	9	14	5	112	117	16.36	11.5	60	P	P	P	P	P	P	P	P	P	P	P	P	P
6.	30	6	5	12	11	11	8	118	129	13.28	9.3	60	P	P	P	P	P	P	P	P	P	P	P	P	P
7.	55	7	6	8	11	16	20	108	96	20.53	22.7	10	P	P	P	P	P	P	P	P	P	P	P	P	P
8.	46	5	7	8	13	16	16	148	137	13.82	22.4	80	P	P	P	P	P	P	P	P	P	P	P	P	P
9.	52	5	4	13	10	18	16	90	98	15.42	12.0	70	P	P	P	P	P	P	P	P	P	P	P	P	P
10.	52	9	8	12	8	17	9	125	122	16.50	7.3	30	P	P	P	P	P	P	P	P	P	P	P	P	P
11.	53	5	4	12	18	18	18	120	98	18.89	8.6	80	P	P	P	P	P	P	P	P	P	P	P	P	P
12.	51	7	7	12	19	9	14	149	107	11.93	11.6	40	P	P	P	P	P	P	P	P	P	P	P	P	P
13.	51	8	7	13	22	7	13	94	89	24.55	11.8	40	P	P	P	P	P	P	P	P	P	P	P	P	P
14.	53	5	5	13	9	15	9	162	114	19.93	7.9	30	P	P	P	P	P	P	P	P	P	P	P	P	P
15.	53	7	6	9	11	16	16	131	89	1.35	13.1	20	P	P	P	P	P	P	P	P	P	P	P	P	P
16.	56	6	5	17	18	18	15	157	142	20.05	7.3	60	P	P	P	P	P	P	P	P	P	P	P	P	P
17.	51	8	7	2	14	14	13	131	115	2.01	14.5	50	P	P	P	P	P	P	P	P	P	P	P	P	P
18.	46	7	6	17	8	15	15	74	80	27.03	9.3	30	P	P	P	P	P	P	P	P	P	P	P	P	P
19.	45	4	4	12	11	9	16	131	101	16.99	9.6	70	P	P	P	P	P	P	P	P	P	P	P	P	P
20.	50	5	4	16	13	17	20	126	114	18.54	13.3	10	P	P	P	P	P	P	P	P	P	P	P	P	P
21.	43	6	5	18	11	13	14	113	103	14.71	7.3	90	P	P	P	P	P	P	P	P	P	P	P	P	P
22.	51	6	4	7	8	13	11	124	130	13.63	8.6	20	P	P	P	P	P	P	P	P	P	P	P	P	P
23.	53	7	5	14	19	19	16	144	111	16.31	13.3	30	P	P	P	P	P	P	P	P	P	P	P	P	P
24.	52	7	5	11	15	11	17	115	91	16.63	11.3	50	P	P	P	P	P	P	P	P	P	P	P	P	P
25.	51	7	5	20	19	23	24	131	106	12.83	13.0	70	P	P	P	P	P	P	P	P	P	P	P	P	P
26.	58	6	5	16	11	16	19	131	99	0.73	6.1	70	P	P	P	P	P	P	P	P	P	P	P	P	P
27.	59	7	4	10	19	19	21	79	80	29.61	10.2	40	P	P	P	P	P	P	P	P	P	P	P	P	P
28.	47	8	7	11	6	8	15	133	110	15.94	14.9	—	P	P	P	P	P	P	P	P	P	P	P	P	P
29.	51	8	8	13	10	17	10	116	105	10.33	16.9	80	P	P	P	P	P	P	P	P	P	P	P	P	P
30.	—	6	4	7	15	14	18	151	108	14.77	15.7	60	P	P	P	P	P	P	P	P	P	P	P	P	P
31.	55	9	6	13	12	23	15	102	111	14.69	10.0	80	P	P	P	P	P	P	P	P	P	P	P	P	P
32.	54	6	5	11	13	14	21	126	82	15.10	8.8	40	P	P	P	P	P	P	P	P	P	P	P	P	P
33.	65	7	6	14	18	22	12	111	109	21.16	4.8	20	P	P	P	P	P	P	P	P	P	P	P	P	P
34.	50	9	9	11	12	9	10	148	123	16.68	11.4	70	P	P	P	P	P	P	P	P	P	P	P	P	P
35.	41	5	5	19	10	14	19	134	126	12.57	12.8	70	P	P	P	P	P	P	P	P	P	P	P	P	P
36.	49	5	4	11	10	14	18	165	131	17.20	11.2	60	P	P	P	P	P	P	P	P	P	P	P	P	P
37.	52	4	3	19	18	23	22	116	117	13.81	13.7	50	P	P	P	P	P	P	P	P	P	P	P	P	P
38.	37	5	4	16	15	15	21	136	97	13.62	11.7	55	P	P	P	P	P	P	P	P	P	P	P	P	P
39.	45	5	5	12	19	19	18	125	129	15.03	13.5	80	P	P	P	P	P	P	P	P	P	P	P	P	P
40.	50	4	3	16	18	21	18	149	119	14.20	14.4	70	P	P	P	P	P	P	P	P	P	P	P	P	P
41.	49	8	7	10	13	8	15	125	127	18.64	14.6	60	P	P	P	P	P	P	P	P	P	P	P	P	P
42.	49	5	5	20	12	22	18	121	128	5.26	7.0	30	P	P	P	P	P	P	P	P	P	P	P	P	P
43.	63	5	5	14	2	21	22	109	88	22.13	6.7	90	P	P	P	P	P	P	P	P	P	P	P	P	P
44.	58	6	5	12	13	18	13	114	107	7.17	11.7	20	P	P	P	P	P	P	P	P	P	P	P	P	P
45.	47	9	7	13	8	14	19	154	129	9.36	8.4	80	P	P	P	P	P	P	P	P	P	P	P	P	P
46.	47	6	7	8	13	15	20	105	90	19.81	15.4	90	P	P	P	P	P	P	P	P	P	P	P	P	P
47.	64	8	5	77	8	19	17	134	96	16.65	13.4	30	P	P	P	P	P	P	P	P	P	P	P	P	P
48.	58	6	5	15	17	19	18	89	90	12.29	10.4	70	P	P	P	P	P	P	P	P	P	P	P	P	P
49.	54	7	5	9	16	18	15	98	95	14.95	9.5	80	P	P	P	P	P	P	P	P	P	P	P	P	P
50.	45	8	4	12	13	17	16	141	106	21.69	13.3	—	P	P	P	P	P	P	P	P	P	P	P	P	P

TABLE A-3

TEST SCORES OF C.P.T. STUDENTS
(Original)

Subject Number	Pulse Rate Change	Dexterity Test: Right Hand 1	Right Hand 2	Left Hand 1	Left Hand 2	Together 1	Together 2	Finger Dexterity 1	Finger Dexterity 2	Tweezer Dexterity 1	Tweezer Dexterity 2	Audiometer Loss or Gain for Hearing Speech
1.	6	15	18	17	17	30	31	3-45	3-53	3-55	3-25	3.74
2.	0	14	15	15	14	25	26	4-03	3-37	2-48	2-52	16.85
3.	2	15	18	16	16	31	29	3-45	3-25	3-37	3-00	3.79
4.	0	14	14	14	14	24	23	4-10	3-56	4-07	4-00	7.96
5.	3	15	19	14	17	26	34	4-00	3-53	2-32	2-29	4.36
6.	0	15	18	17	15	25	27	4-14	4-34	2-24	2-33	1.28
7.	1	12	14	13	14	22	22	4-28	4-15	3-00	2-39	8.53
8.	-2	16	16	14	14	27	28	3-50	3-30	3-35	3-24	1.82
9.	-2	13	14	16	15	24	24	3-55	3-32	3-54	3-31	3.42
10.	0	15	15	13	15	25	24	4-05	4-15	3-13	2-52	4.50
11.	0	16	19	12	16	26	28	3-49	3-46	3-41	3-29	6.89
12.	-4	16	15	13	15	23	28	4-00	3-50	3-05	3-15	-0.07
13.	-19	13	16	13	16	24	24	3-25	3-10	4-22	3-00	12.55
14.	0	19	17	17	16	26	26	4-10	3-43	3-27	2-37	7.93
15.	0	14	15	14	14	24	24	4-00	4-01	3-15	3-24	-10.65
16.	0	13	16	17	19	25	29	4-35	3-54	3-27	2-55	8.05
17.	0	14	15	14	14	20	26	4-06	4-09	3-59	3-32	-9.99
18.	-4	15	16	14	14	26	27	3-41	3-41	3-15	3-15	15.03
19.	0	11	14	12	13	24	24	4-19	3-45	4-05	3-25	4.99
20.	0	15	18	17	18	28	26	3-40	3-27	2-57	2-24	6.54
21.	-8	15	13	15	15	24	24	4-02	3-47	3-45	3-33	2.71
22.	6	12	15	14	15	26	28	4-00	4-06	3-25	3-00	1.83
23.	0	16	17	15	17	30	30	4-00	3-51	----	----	4.31
24.	2	16	16	17	15	25	28	3-55	3-14	3-00	3-03	4.63
25.	-2	15	17	13	17	24	24	3-55	3-43	3-42	3-27	0.83
26.	0	12	13	14	13	22	22	4-00	4-02	3-24	3-40	-11.27
27.	0	14	13	12	12	23	24	5-25	4-80	3-50	3-75	6.61
28.	2	13	15	12	13	22	26	4-02	4-10	3-46	3-40	3.94
29.	0	13	17	15	17	24	28	3-57	4-07	3-25	3-15	-1.67
30.	3	14	13	14	15	24	26	3-40	3-55	3-14	2-45	2.77
31.	0	16	18	13	16	26	28	3-47	3-52	2-42	3-05	2.69
32.	-7	16	16	15	16	27	22	3-54	3-34	4-17	3-00	3.10
33.	11	15	14	13	15	22	22	3-42	3-43	3-00	2-42	9.16
34.	0	13	15	14	15	21	26	4-30	4-08	3-00	2-36	4.68
35.	0	13	16	13	13	22	20	4-27	4-05	3-55	3-12	0.57
36.	-6	17	16	14	16	24	28	4-04	4-00	3-30	2-38	5.20
37.	0	13	16	14	16	26	28	3-26	3-12	3-18	2-56	1.81
38.	-6	16	18	15	15	24	28	3-30	3-50	3-32	3-52	1.62
39.	-6	17	16	14	13	26	28	3-40	3-19	3-27	2-53	3.03
40.	-4	16	17	15	17	26	27	3-25	3-25	4-00	3-00	2.20
41.	0	15	16	15	13	21	24	4-04	3-55	2-40	2-48	6.64
42.	1	14	17	15	18	24	32	3-39	3-43	3-13	2-39	-6.74
43.	1	13	15	14	15	28	24	3-47	3-49	4-02	3-07	10.13
44.	0	18	20	16	17	31	32	3-22	3-45	2-32	2-45	-4.83
45.	-2	15	17	15	17	28	28	3-54	3-45	3-12	2-54	-2.64
46.	-15	15	15	14	14	26	26	4-00	3-50	4-25	4-27	7.81
47.	-2	17	16	15	16	24	24	3-20	3-07	2-55	2-25	4.65
48.	-2	13	14	13	13	21	24	4-11	3-55	4-00	4-35	0.29
49.	15	14	17	15	14	23	25	3-34	3-22	4-04	3-00	2.93
50.	0	18	19	16	17	26	24	3-05	3-06	3-06	2-55	7.84

TABLE A-3

VARIOUS FLIGHT RATINGS OF STUDENT PILOTS

Pilot No.	Flight Rating Index*	Chief Instr. Overall Rating	Chief Instr. Rating Scores	CAA Examiner's Rating	Composite Rating	Nav. & Met.	C.A.R.
1.	2.21	Average	5.0	83	4.37	79	83
2.	1.36	Average	5.0	94	7.42	94	89
3.	1.78	Good	7.9	76	5.49	88	70
4.	1.72	Average	5.0	83	5.43	79	79
5.	1.92	Average	5.0	81	4.78	88	87
6.	2.34	Poor	2.1	76	2.39	82	75
7.	1.88	Average	5.0	80	4.76	78	70
8.	2.17	Average	5.0	78	3.93	73	76
9.	1.77	Average	5.0	95	6.66	94	92
10.	1.56	Average	5.0	84	5.87	82	77
11.	1.74	Average	5.0	75	4.50	80	78
12.	1.34	Average	5.0	88	6.78	75	70
13.	1.93	Poor	2.4	84	4.24	91	80
14.	1.86	Poor	0.7	86	4.04	79	70
15.	2.34	Average	5.0	80	3.80	77	71
16.	2.00	Average	5.0	83	4.84	81	77
17.	1.88	Average	5.0	73	3.98	78	70
18.	1.91	Poor	2.6	74	3.68	75	84
19.	2.03	Good	6.6	89	5.88	89	87
20.	1.68	Average	5.0	80	5.18	80	75
21.	1.33	Average	5.0	81	6.01	84	83
22.	1.62	Average	5.0	86	5.97	99	84
23.	2.02	Average	5.0	83	4.80	96	85
24.	2.58	Poor	2.8	80	2.56	85	78
25.	1.44	Average	5.0	87	6.46	70	74
26.	2.49	Poor	2.0	73	1.79	83	70
27.	2.65	Poor	3.1	73	1.47	80	70
28.	1.60	Good	9.3	85	7.34	75	85
29.	1.95	Average	5.0	77	4.32	70	75
30.	1.94	Average	5.0	86	5.30	89	78
31.	2.37	Poor	1.8	77	2.33	83	74
32.	1.92	Average	5.0	80	4.68	89	70
33.	2.15	Good	6.9	93	6.28	85	77
34.	1.52	Average	5.0	96	7.30	89	91
35.	2.49	Average	5.0	80	3.49	80	76
36.	2.08	Average	5.0	90	5.45	84	89
37.	2.11	Average	5.0	84	4.73	94	76
38.	2.34	Good	7.0	82	5.02	85	72
39.	1.92	Poor	3.3	78	3.88	77	84
40.	1.43	Average	5.0	81	5.85	88	72
41.	1.89	Poor	1.4	84	3.59	91	80
42.	1.48	Average	5.0	76	5.35	90	85
43.	2.27	Good	7.2	82	4.90	74	87
44.	1.95	Average	5.0	88	5.51	80	89
45.	2.01	Good	6.7	90	6.17	88	75
46.	1.69	Good	8.2	80	6.33	89	91
47.	1.98	Good	7.5	77	5.08	86	70
48.	2.14	Good	8.5	95	7.09	93	81
49.	1.93	Average	5.0	83	4.99	82	79
50.	1.89	Good	7.4	78	5.15	78	75

* [illegible] a high

PART IV

TESTS OF KINESTHETIC SENSITIVITY IN THE SELECTION OF PILOTS

by

WILLIAM M. LEPLEY

Pennsylvania State College

TESTS OF KINESTHETIC SENSITIVITY IN THE SELECTION OF PILOTS

INTRODUCTION

The primary objective of this investigation was to develop and validate a battery of tests of coordination and balance. It must be noted that the study was in no way concerned with the isolation of the kinesthetic and static functions. It was concerned with devising a test battery that would be of practical value in the prediction of successful progress in flight training.

Since certain other data were available or easily obtained during this experiment, a secondary part of this report includes the description of attempts at the validation of other tests and procedures which are not intended as direct measures of coordination and balance.

THE SUBJECTS

The population involved in this report consisted of the thirty-one male students of The Pennsylvania State College who participated in the 1939-1940 Civilian Pilot Training program.

PROCEDURES

The procedures employed in this study may be classified under the following categories:

A. The Collection of Data from Outside Agencies.

1. The collection and analysis of data from the medical examiner's reports.

2. The analysis of the medical examiner's composite rating.

3. The collection of the ratings made by the local committee on selection, at the time of the formation of the training class.

B. The Administration of Standardized Tests.

4. The Otis Self-Administering Test of Mental Ability, Form A, Higher Examination.

5. The Revised Minnesota Paper Formboard.

6. The Bernreuter Personality Inventory.

7. The Strong Vocational Interest Blank.

C. The Administration of the Battery of Tests of Coordination and Balance.

8. Differential Pressure. This test required the subject to execute and reproduce standard pressures, applied to a pair of spring scales with his hands. For example, the subject would apply two pounds pressure with his left hand and four with right, with his eyes open. He would then close his eyes and try to reproduce the pressures he had already exerted while his eyes were open. Ten trials were given each subject. The test was scored in terms of ounces of error.

9. Reproduction of Postures. This test required the subject to assume a standing, leaning posture with his eyes open. He then attempted to reproduce the posture with his eyes closed. Each subject was given twenty trials. The test was scored in terms of inches of error in the reproduction of the posture with the eyes closed. The errors in reproduction were measured by an appropriate mechanical system for recording movement and position.

10. Finger-to-Finger Coordination. In this test the subject was required to bring the two forefingers together in front of the body at chest height and with the eyes closed. Five trials were given for each hand. Scoring was in terms of millimeters of error.

11. Horizontal Linear Space. The procedure for this test was to have the subject fixate the eighteen inch mark on a horizontally mounted yardstick and then, with the eyes closed, try to move the forefinger along the yardstick from the extremity to the eighteen inch mark. Five trials were given for each hand. Scoring was in terms of inches of error.

12. Walk to Line. The subject was required to fixate a line marked on the floor at a distance of twenty feet. Then, with the eyes closed he tried to walk to that line. Five trials were given. Scoring was in terms of inches of error.

13. Hand Separations. The subject was required to place his hands in front of the body and, with his eyes closed, separate the forefingers a standard distance. He was then required to try to reproduce that distance with the eyes closed. Five trials were given. Scoring was in terms of inches of error.

14. Step to Line. The procedure employed in this test was similar to the procedure used in test No. 12, except that a

single step was required. Five trials were given for the foot preferred in stepping off from the resting position. Scoring was in terms of inches of error.

15. Ataxiameter (eyes open). This test consisted of recording the body-sway of the subject while maintaining an upright posture for one minute. The sway-movements were recorded and summated by a pair of work adders.

16. Ataxiameter (eyes closed). The procedure for this test was the same as that used in test No. 15, except that the record was taken with the eyes closed.

17. Knee Bend. The procedure used in this test was the same as that used in test No. 9, except that the subject was required to stoop by flexing the knees (eyes open) and then try to reproduce the posture with the eyes closed. Five trials were given. Scoring was in terms of inches of error.

18. Foot Span. The subject was required to separate his heels a standard distance and then to try to reproduce the the stance. The task was executed with the eyes closed. Five trials were given. Scoring was in terms of inches of error.

D. Perimetric Examination. A simple perimetric examination was given each subject. Both the right and left eyes were examined. The test was scored in terms of degrees of arc for the external meridians.

E. Flight Ratings. The flight ratings made at the termination of the student's training were collected and analyzed. These ratings were accomplished by means of a graphic rating scale procedure and were made in three different ways, as follows:

1. Each student was rated on nine different aspects of flight performance. These aspects were: taxiing, take-offs, climbs, turns, spins, approaches, glides, slips, and landings.

2. Each student was rated as to his general aptitude for aircraft operation as judged by observation over the whole period of training.

3. The rater was further asked to rate each student with regard to his degree of confidence in the student as a pilot. As before stated, all ratings were made at the close of the flight training, i.e., after the final flight test

had been passed. All ratings were made by the operator of the local field. This operator was also the contractor and the chief instructor. In addition, he was the inspector who administered the final flight tests.

RESULTS

The results of this investigation are presented as briefly as possible in the following summary tables. All coefficients of correlation, unless otherwise described, were computed by the rank difference method and are reported without correction. The total population of 31 cases entered into each calculation.

TABLE 1

INTERCORRELATIONS (RHOS) AMONG THE RATINGS OBTAINED BY THE THREE RATING METHODS (N = 31)

	Gen. Aptitude	Confidence	Mean Rating
General Aptitude	-	.96	.92
Confidence		-	.85
Mean Rating			-

TABLE 2

INTERCORRELATIONS AMONG THE TESTS OF COORDINATION AND BALANCE[1]
(N = 31)

		(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Differential Press. Both Hands (error in oz.)	(8)	--	-.20	-.12	-.02	.04	-.26	-.06	-.03	-.09	-.19	-.08
REPRODUCTION OF POSTURES	(9)		--	.19	-.31	-.03	-.01	-.29	.32	.54	.39	-.09
Finger-to-Finger Coordination	(10)			--	.13	-.12	-.04	-.22	-.27	-.23	.13	-.05
Horizontal Linear Space	(11)				--	.20	-.01	.28	-.08	-.21	-.41	.20
Walk to Line	(12)					--	.15	.02	.12	-.11	-.14	.42
Hand Separation	(13)						--	.02	-.16	.14	.15	-.16
Step to Line	(14)							--	.03	-.06	-.08	.03
ATAXIAMETER. (EYES OPEN)	(15)								--	.61	.19	.18
ATAXIAMETER. (EYES CLOSED)	(16)									--	.29	.13
Knee Bend	(17)										--	-.02
Foot Span	(18)											--

[1] The titles of the supposed tests of balance appear in capitals in this table. Except where otherwise indicated, the scoring is in terms of mean error by linear measure.

TABLE 3

INTERCORRELATIONS BETWEEN THE VARIOUS TESTS OR MEASURES AND THE GENERAL APTITUDE RATINGS (N = 31)[2]

Test or Measure	Rho
A. Data from Outside Agencies:	
1. Medical Examiner's Data.	
Pulse reclining.	.05
Pulse standing.	.06
Reclining B.P. (systolic)	.07
Standing B.P. (systolic)	.13
Schneider Index	.33
2. Medical Examiner's Composite Rating.	.16
3. Local Committee Rating at Time of Selection.	.36
B. Standardized Tests:	
4. Otis Self-Administering Test of Mental Ability.	.11
5. Minnesota Paper Formboard.	.32
6. Bernreuter Personality Inventory.	
B1-N	-.12
B2-S	.09
B4-D	-.02
7. Strong Vocational Interest Blank.	
Group I	.18
Group II	.03
Group V	-.26
Group VIII	-.18
Group IX	.01
Group X	.14

---

2 Because of the high intercorrelations shown in Table 1, only the general aptitude ratings are used. Correlations for the various parts of the balance and coordination tests are shown separately as well as for the tests as wholes. Raw scores were used in the computations involving standardized tests.

Table 3 (continued)

Test or Measure	Rho
C. Tests of Coordination and Balance:	
8. Differential Pressure.	
Right hand	.05
Left hand	-.09
Both hands	-.02
9. Reproduction of Postures.	
Forward movement	.15
Backward movement	.00
Sinistral movement	.02
Dextral movement	.24
Combined movement	.18
10. Finger-to-Finger Coordination.	
Non-preferred hand	.14
Preferred hand	-.28
Combined hands	-.27
A.D. for combined	-.25
11. Horizontal Linear Space.	
Non-preferred hand	-.14
Preferred hand	-.30
Combined hands	-.21
A.D. for combined	.15
12. Walk to Line.	.52
13. Hand Separation.	-.06
A.D.	.08
14. Step to Line.	-.28
A.D.	.00
15. and 16. Ataxiameter.	
Eyes open	-.03
Eyes closed	-.17
Percent difference	-.25
17. Knee Bend.	-.04
18. Foot Span.	-.14
A.D.	.13

TABLE 3 (continued)

Test or Measure	Rho
D. Perimetric Findings	
Left eye	-.17
Right eye	-.13
Both eyes	-.18

TABLE 4

A SAMPLING OF RELIABILITY COEFFICIENTS[3] (N = 31)

	Uncorrected	Corrected
8. Differential pressure. (Right hand - left hand)	.49	.67
9. Reproduction of postures.		
Dextral - Sinistral movement	.38	.55
Forward - Backward movement	.44	.62
10. Finger-to-finger coordination.		
Non-preferred hand	.42	.78
Preferred hand	.77	.94
11. Horizontal linear space.	.51	.84
13. Hand separation.	.59	.88
18. Foot span.	.28	.66

3 These reliability coefficients were computed by the method of average correlations among the separate trials (except where otherwise noted) and corrected by the Spearman-Brown Prophecy Formula.

INTERPRETATIONS AND CONCLUSIONS

Except for certain ambiguities left by the omission of detailed descriptions of the techniques of administration and of scoring, the interpretations of Tables 1 to 4 are quite obvious.

A. A sampling of the reliabilities of the tests of balance and coordination shows them to have fairly high reliabilities for this type of measure.

B. The generally low intercorrelations among the various tests of coordination and balance indicate that these tests are relatively independent measures.

C. The generally low validity coefficients for all tests and procedures suggests that they were inadequate for discriminating between levels of flying aptitude as measured by the rating scale procedures used.

D. The low validity coefficients may be accounted for in several ways:

1. The measures may not have been appropriate.

2. The measures may have been too imperfect to serve the practical objectives considered. However, it should be remembered that some of the measures appear to have fairly acceptable reliabilities.

3. It is very possible that the criterion for aptitude (flight ratings) was inadequate.

4. The homogeneity of the population used (with regard to flying aptitude) was too great to permit significant discrimination by either the criterion or the tests and procedures employed.

5. The sample of student pilots may have been too small.