HISTORY AND DEVELOPMENT OF THE OHIO STATE FLIGHT INVENTORY

PART I: EARLY VERSIONS AND BASIC RESEARCH

by

Harold A. Edgerton
Robert Y. Walker

A report of studies conducted at the Ohio State University, Columbus, Ohio, by means of grant-in-aid from the National Research Council Committee on Selection and Training of Aircraft Pilots from funds provided by the Civil Aeronautics Administration.

July [illegible]

CIVIL AERONAUTICS ADMINISTRATION
Division of Research
Report No. 47
Washington, [illegible]

National Research Council

Committee on Selection and Training of Aircraft Pilots

Executive Subcommittee

M. S. Viteles, Chairman

E. C. Andrus	H. M. Johnson
C. W. Bray	W. E. Kellum
D. R. Brimhall	W. R. Miles
L. A. Carmichael	P. J. Rulon
J. W. Dunlap	G. R. Wendt
J. C. Flanagan	

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

July 10, 1945

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

Dear Dr. Brimhall:

Attached is a report entitled History and Development of the Ohio State Flight Inventory, Part I: Early Versions and Basic Research, by Harold A. Edgerton and Robert Y. Walker. 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.

The present report describes the early research which led to the formulation and preparation of a comprehensive and standardized check list of items descriptive of a pilot's performance during flight. This check list, when used in conjunction with standard flights, has provided a valuable source of criterion data in aviation research.

In general, the principle of standardized observation in recording specific items of flight performance has found wide application in pilot evaluation. Copies of Ohio State Flight Inventory sheets and Manual were requested in 1943 by the Army Air Forces for tryout on an experimental basis. CAA Form ACA-342Z, issued by the Civil Aeronautics Administration in 1943, embodies principles for observing and recording flight performance earlier used in the Ohio State Flight Inventory.

This report does not describe the present form of the Ohio State Flight Inventory. Part II, to be issued separately, will trace the development of the Inventory from early versions to its present form.

Cordially yours,

MSViteles

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

MSV:rm

rec 10-3-45

CONTENTS

EDITORIAL FOREWORD

Prior to the initiation of research by the Committee on Selection and Training of Aircraft Pilots, evaluation of pilot performance was limited largely to the assignment of ratings on individual maneuvers and grades on over-all flight performance without detailed or controlled reference to specific aspects of the performance. Beginning in 1939, research at Ohio State University was directed towards the development of a rating technique, including a standardized procedure for recording and scoring observations on specific items of pilot performance.

This research has led to the preparation of the Ohio State Flight Inventory, a comprehensive and standardized check list of items descriptive of a pilot's performance during flight. Check sheets are available for each maneuver taught in the CAA elementary course. Items are grouped, whenever possible, according to the portion of the maneuver being observed, as for example, entry, turn proper, and recovery in turns. When used in conjunction with standard flights, the Inventory provides for standardized observation and recording of the details of pilot performance during each maneuver.

The present form of the Ohio State Flight Inventory is the outcome of a series of revisions based upon continuous try-out in Committee research projects and in the field since 1940. This report constitutes Part I of a complete report on the Ohio State Flight Inventory, describing early versions and basic research upon which the later versions were developed. Part II, to be issued separately, will describe intermediate forms and the current version of the Ohio State Flight Inventory and the outcomes from the use of the Inventory in Committee research projects.

Combined with standard flights, the Ohio State Flight Inventory has served as a valuable source of criterion data in Committee research. It has become familiar to many CAA flight instructors and inspectors, both through its use in field research and through the course on training methods given at the Institutes held at the University of Minnesota and at Ohio State University in 1943.

Acknowledgment is due to A. S. Thompson, of the Editorial Staff of the Committee on Selection and Training of Aircraft Pilots, for considerable assistance in revising and editing the materials of this report.

The development of the Ohio State Flight Inventory represents an attempt to devise a method of evaluating pilot performance which would provide a satisfactory measure of flight competence and also serve as a diagnostic device for the analysis of flight proficiency. Its development involved three basic steps: (1) construction of descriptive items of flight performance representative of various levels of pilot proficiency; (2) preparation of standardized check lists of items for individual maneuvers; and (3) modification and refinement based on field trial and use in research projects.

The present form of the Ohio State Flight Inventory is the outcome of a series of versions, each designed to improve the Inventory on the basis of findings obtained during experimental use in the field. Experimentation with a preliminary form which was essentially a rating scale with descriptive items at various points on the scale revealed this approach to be unsatisfactory. Beginning with the 1940 version, therefore, the Inventory took the form of a check list of items descriptive of various grades of pilot proficiency. Check lists were prepared for each maneuver in the standard flights being used in pilot evaluation. The items for each maneuver were grouped according to function or portions of the maneuver such as Entry, Turn, and Recovery in Turns. In addition, general items referring to coordination, tension, judgment, etc., were placed in a separate check sheet. Check lists were prepared for all maneuvers except Taxiing in which the rating scale method was retained.

Since the items checked by the observer on the basis of observation during test flights were merely descriptive in nature, the evaluation of the performance was a function of the scoring method. Each item for each maneuver was assigned a scale value ranging from 1, excellent, to 5, very poor, based on judgments of experienced pilots. Three scores were then obtained for each maneuver: (1) the mean scale value of the items checked for that maneuver (mean score); (2) a weighted mean score, taking into account the importance as well as the scale value of the item; and (3) a variability score based on the variance about the mean maneuver score of the item values checked for that maneuver.

As part of the 1940 Ohio State University project the performance of 30 student pilots was observed during two successive flights each. Inventory data on these 30 paired sets of observations were analysed as follows:

1. The frequency with which each item was checked by the 6 instructors involved was determined. This tabulation showed that many items were checked infrequently and that the majority of the items checked were those descriptive of superior flight performance.

2. Low correlations were obtained between the number of hours of flight training at the time of the flight and the three types of scores obtained from the Inventory.

3. Correlations between scores for the first and second flights were found to be generally low. These correlations cannot be considered

in reliability coefficients, however, since both student pilot variability and observer variability were present in the two flights.

4. Intercorrelations among the three types of scores were found to be positive and high particularly between the mean scores and weighted mean scores.

5. The problem of devising an over-all flight score was attacked in two ways: (1) first, by weighting maneuvers according to their importance as judged by a group of 74 flight instructors and (2) by obtaining regression weights based on a factor analysis of maneuver intercorrelations. Because of the lack of consistency among the maneuver importance ratings by flight instructors, this approach to weights for composite scores was not considered satisfactory. The factor analysis of the maneuver intercorrelations suggested the presence of two factors. Regression coefficients and gross score weights to be used in obtaining an estimate of the first factor score were obtained.

The 1941 version of the Ohio State Flight Inventory was essentially a simplification of the previous version, obtained by weeding out the items found to be infrequently checked. Maneuver sheets were developed for all the maneuvers in the CPT primary course and data were obtained on 143 flights made by CPT students at three flying schools. Analysis of the inventory data from these flights was made as follows:

1. The data were examined, maneuver by maneuver, to determine the variance ascribable to instructor, observer, school, wind velocity, flight number, and familiarity with plane. Of the 31 maneuver values of epsilon-squareds, 26 were greater than the 1% value for observer, 22 for instructor, 13 for flight number, 12 for school, and only 4 for wind velocity.

   Similar study of a restricted sample consisting of the 69 observations made by the two best trained observers suggested that trained observers would give more homogeneous results.

2. A factor analysis of maneuver intercorrelations was made based on the restricted sample of 69 observations and 15 of the 31 maneuvers. In general, the intercorrelations were low and positive. Three factors were identified: Factor A, with high loadings for Steep 8's and Shallow 8's; Factor B, with fairly high loadings in Climbs, Left Climbing Turns, and Left Power Turns; and Factor C, with fairly high loadings in Climbs, Left Climbing Turns, and Left Power Turns; and Factor C, with fairly high loadings in Straight and Level Flight, Take-offs, 180° Approaches, and Landings. Regression weights were obtained for each of the three factors.

In general, experience with the 1941 version confirmed its value as a useful instrument in measuring flight performance and indicated the advisability of continued experimentation, especially with respect to further simplification of the form, to improvements in the methods of scoring, and to the development of training methods in the proper use of the Inventory.

[illegible] FLIGHT INVENTORY

PART I: EARLY INVESTIGATIONS AND BASIC RESEARCH

INTRODUCTION

One of the major objectives of the research program of the National Research Council Committee on Selection and Training of Aircraft Pilots since its inception in 1939 has been the development of measures of success in the task of piloting a plane. The need for a sound, practical method of evaluating pilot performance was felt both by the research investigator and by the professional pilot in the field. The research investigator needs an adequate "criterion" by which he can test the efficiency of selection techniques, training procedures, the use of new instruments, etc. The flight instructor has to decide whether a student's proficiency has reached the level where he can solo, i.e., pilot a plane without risking his own life. The flight inspector has to decide whether a pilot has reached a level of proficiency which will warrant a license permitting him to risk a passenger's life. A good measure of pilot proficiency is thus of considerable practical as well as theoretical importance.

In 1939, when the investigations described in this report were initiated, two methods of evaluation of pilot performance were in use in the newly developed Civilian Pilot Training Program of the Civil Aeronautics Administration. One, designed for use by flight instructors, required ratings in the Student Pilot Log Book for each maneuver practiced by the student during each instruction flight. The other, designed for use by the CAA flight inspectors at the time of the final flight test for licensing purposes, required grades on each maneuver and on the flight as a whole. The log book ratings were on a five-point scale: "1" for excellent; "2" for above average performance; "3" for average performance; "4" for below average performance; and "5" for poor performance. The flight test grades were on a "per cent" system with 70 as passing.

Early studies soon revealed that the ratings and grades as obtained in the ordinary field situation possessed many limitations, if they were to be used as criterion data in research. As might be expected, instructors differed as to what was meant by the terms "excellent," "average," "poor," etc. Average, for example, was variously interpreted by instructors as: (1) the average of their present group of students; (2) the average performance of all the students they had trained; (3) the average performance of the student considering his amount of training; (4) the average performance required to pass the flight test for private license. In addition, "fixed ideas" appeared, such as that no student could possibly get better than a "3" rating until he had at least 30 hours in the air. As shown in the following distribution of ratings for landings at the end of 5, 10, 15, 20, and 25 hours, respectively, for the CPT student pilots of one flight instructor, the rating method no longer differentiated one performance from another by the end of 15 hours of training.

Ratings	Number of Ratings at the end of: 5 hrs.	10 hrs.	15 hrs.	25 hrs.
1. (Excellent)				
2. (Above Average)				
3. (Average)	1	6	9	7**
4. (Below Average)	6*	4*		
5. (Poor)	3			

* One individual receiving ratings of "4" at 5 and at 10 hours was dropped before completing 15 hours of flying.

** Two had not, at the time of record, completed 25 hours of flying.

The limitations of the rating and grading systems in field use in the early period of the CPT program were clearly revealed in a study by H. M. Johnson and Mary L. Boots,[1] based on the training records of student pilots trained at twelve CPT centers in the spring of 1939. Low correlations were obtained between instructor's ratings and inspector's grades. Inconsistencies were found between instructor's comments and the assigned rating, such as: "Still not satisfied with Landings" -- rating of 3; "Landing improved" -- rating of 3; "Landings O.K. now" -- rating of 3.

Such findings led to the conclusion that the systems of grading used in the field would not yield adequate criterion data for use in research. The inadequacies of the systems seemed to be one possible explanation of the frequent occurrence of low correlations between predictors and such criteria as instructor's final grades, composite scores from log book ratings, final flight test grades, etc.[2]

OBJECTIVES OF THE STUDY

The major objective of the Ohio State University Project, begun in 1940, was to develop a method of evaluating pilot performance in which the above

---

[1]Johnson, H. M., and Boots, M. L. Analysis of ratings in the preliminary phase of the C.A.A. training program. Washington, D. C.: C.A.A. Division of Research, Report No. 21, October 1943.

[2]Kelly, E. L., and Johnson, G. P. Analysis of test data on 1938-39 C.A.A. students in Studies of Predictors of Achievement in Learning to Fly. Washington, D. C.: C.A.A. Division of Research, Report No. 27, March 1944.

limitations would be minimized and which would serve at least three important purposes:

1. As a measure of flight competence, to be used as evidence of progress in training and as a criterion for validation of selection procedures and for evaluation of instructional methods.

2. As a device for training instructors to observe flight performance more adequately.

3. As an instructional device, to enable instructors to make a good periodic inventory of flight performance, to be used in directing training and practice toward those skills in which the student pilot was most deficient.

In line with this objective the Ohio State Flight Inventory was devised. This inventory, essentially a standardized form for recording flight performance, was developed with the following principles as guides:

1. The items should be descriptive rather than evaluative. The observer should be required merely to <u>record what the plane and pilot are doing</u>, rather than to <u>judge how well</u> the pilot is flying.

2. The scoring method should be objective. The measure of proficiency should not be dependent upon subjective standards of performance which were vaguely defined, subject to individual variation, etc.

3. It should be discriminating throughout the training period. The inventory should reveal differences among pilots at the early as well as the later stages of training and should yield information as to the individual's progress.

4. It should be applicable for use in the ordinary field situation. It should not be time-consuming, should yield information directly usable, should require little in the way of special training in its use, etc.

5. It should yield information of value for instructional as well as criterion purposes. In particular, it should give not only an over-all measure of proficiency, but also diagnostic information as to the student pilot's specific strengths and weaknesses.

DEVELOPMENT OF THE OHIO STATE FLIGHT INVENTORY

The development of the inventory involved three basic steps: (1) the construction of a check list of items descriptive of various levels of pilot performance; (2) organization of the items on the basis of a standardized flight observation situation; and (3) modification and refinement based on field trial and use in research projects.

The present form of the inventory is the outcome of a series of versions, each designed to improve the inventory with respect to one or another of the principles listed above. This report, presented in two parts, shows how the inventory developed from its preliminary form in 1940 to the present version, describes the changes incorporated into each version, and presents the research findings upon which the changes were based.

Part I of this report is a history of the OSFI as originally developed by the Ohio State University Project. Part II, to be issued separately, will describe later versions and their use in various projects in the research program of the Committee on Selection and Training of Aircraft Pilots.

In the succeeding sections of Part I, each of the early versions will be described, methods of scoring indicated, and research data and conclusions based on experimental tryout will be presented.

A. Preliminary Form (1940)

1. Description. Items descriptive of flight performance were obtained from current CAA flight instructor's manuals, from interviews and discussions with flight instructors, and through the experience of R. Y. Walker in learning to fly. Descriptions of all grades of performance were obtained by asking such questions as, "What did the student do or fail to do which made you give him a '4' on Climbs?", "When you judge how well a student does a Gliding turn, what is involved?" Statements descriptive of performance on a given maneuver were grouped together to form maneuver sheets.

Maneuver rating sheets were constructed for the following maneuvers: Taxiing, Take-offs, Climbs, Climbing turns, Straight and Level flight, Medium turns, Glides, Gliding turns, Landings. For field use these were assembled with a description and a diagram of the proposed flight pattern.[3]

The general form of each maneuver rating sheet was as follows. The items describing grade 1 (excellent flying) were placed at the top of the sheet. Next came the items describing grade 2 (above-average performance). In the middle of the sheet came the grade 3 (average performance) items. These were followed in order by grade 4 and by grade 5 items.

As can be seen in Exhibit 1,[4] the rating sheet for the maneuver "Landing," a "grade 1 performance" was "perfect" while a "grade 5 performance" was characterized by items such as "leveling off too high," "improper rudder control," "swinging on the ground," etc.

---

[3]The standard flight situation used with this preliminary form was Standard Flight A, described in: Viteles, M. S., and Thompson, A. S. The use of standard flights and motion photography in the analysis of aircraft pilot performance. Washington, D. C.: C.A.A. Division of Research, Report No. 15, May 1943.

[4]Since the early versions of the OSFI are not in use at present, only sample maneuvers will be presented as exhibits. Complete copies of each of the versions are on file in the office of the National Research Council Committee on Selection and Training of Aircraft Pilots and are available for inspection. The present version will be presented in its entirety in Part II of the report.

Sample Page From Preliminary Form (1940)

1+		
1	Perfect	
1-		
2+		
2	No errors in approach or landing 3-point landing but excessive speed No coaching necessary	
2-		
3+		
3	Any <u>one</u> of the errors listed in V to a marked degree or Slight degree of error in 3 of the headings listed in V Coaching permitted	
3-		
4+		
4	Same as V but inaccuracies to a smaller degree or Three of the errors listed in V	
4-		
5+	Failure to estimate approach correctly. Overshoots or undershoots Too steep a glide	
5	Leveling off too high Leveling off too low Failure to keep wings level	
5-	Improper rudder control Failure to keep tail down. Wheel landing Failure to maintain constant glide Swinging on the ground	

LANDINGS INTO WIND

[illegible] experimental tryout of the preliminary form was made at the [illegible] Service Flying School during the Spring of 1940. Each student flew the Standard Flight twice with his own instructor and then twice with another instructor. Each instructor made separate ratings for the two successive flights, but on the same form. Since the group was small no statistical treatment of the data was made, but inspection of the results revealed the following limitations in this preliminary form:

a. The ratings were not sufficiently discriminating. The range of ratings for any one maneuver was very limited and the agreement in the ratings of a single instructor for the two successive flights was better than many instructors believed two successive flights by a student pilot should provide.

b. The ratings were still being made on an evaluative rather than a descriptive basis. The results suggested that the instructors were selecting a value such as 2, 3, or 4 for the particular maneuver and then checking those items in the scale which were consistent with this judgment.

B. <u>1940 Version.</u>

1. <u>Description of the 1940 Version.</u> On the basis of experience with the preliminary form, three types of changes were made:

a. The items were classified on the basis of <u>operation</u> rather than on their rating value. As shown in the check sheet for "Landing" (see Exhibit 2), instead of having all the items descriptive of grade 1 flying grouped together, all the items in grade 2 grouped together, etc., items were grouped according to <u>function</u>. Those items referring to "leveling-off" were put in one group, those referring to "stalling" in another group, those having reference to "throttle control" in another, etc. The observer, flying with a student, thus was requested merely to check those items <u>descriptive</u> of the flight performance of the student. The evaluation of the performance became a matter of scoring the Inventory.

b. An attempt was made to eliminate such qualifying adjectives as "good," "fair," and "poor," by using more objective descriptions of the attitude of the airplane.

c. Another improvement was the segregation of items referring to general coordination, tension, and relaxation onto a separate sheet. This separated the description of the performance of the plane from inferences regarding the coordination and tension of the student pilot.

The varying sets of conditions which affect performance in <u>Taxiing</u> made it necessary to retain a <u>value</u> scale for this maneuver. A change in wind conditions, for example, may change an item in Taxiing from good to poor or vice versa. Attention was called to various factors which need to be considered (see Exhibit 3), but for each factor the rating was in terms of excellent, above average, average, below average, or very poor.

EXHIBIT 2

Sample Page from 1940 Version

Lat. Balance On Glide		Lat. Balance During Roll		Direction to Wind On Glide		Direction to Wind During Roll	
()	()	()	()	()	()	()	()
()	()	()	()	()	()	()	()
()	()	()	()	()	()	()	()
()	()	()	()	()	()	()	()
()	()	()	()	()	()	()	()
()	()	()	()	()	()	()	()
1st Flight	2d Flight	1st Flight	2d Flight	1st Flight	2d Flight	1st Flight	2d Flight

LATERAL BALANCE & DIRECTION

Constant
Not more than 5° r. or l. deviation
Not more than 10° r. or l. deviation
More than 10° r. or l. deviation
Continuous fluctuation of less than 10°
Continuous fluctuation of more than 10°

LEVEL OFF

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

STALLING

()()Slow, continuous action in stalling plane for landing
()()Periodic application of controls in stalling plane
()()Jerky in stalling plane
()()Has stick all the way back just as plane touches ground
()()Has stick all the way back just before plane touches ground
()()Slow on getting stick back for landing
()()Fails to hold stick all the way back
()()Fails to get stick back for landing
()()Holds stick back until plane stops
()()Does not hold stick back till plane stops

LANDING

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

1st 2d
Flight

EXHIBIT 2 (Contd.)

PRECISION

()()Lands on spot
()()Lands on spot or within 100 ft. beyond
()()Lands on spot or within 200 ft. beyond
()()Overshoots spot landing more than 500 ft.
()()Undershoots spot landing less than 50 ft.
()()Overshoots spot by more than 300 ft.
()()Undershoots spot by more than 50 ft.
()()Overshoots field
()()Undershoots field

THROTTLE

()()Keeps hand on throttle
()()Fails to keep hand on throttle

1st 2d
Flight

Into wind__________
Cross wind__________
Down wind__________
Gusty air__________
Forced landing________

Wind velocity________

Landing glides entered under glides and so noted__________

Final Approach

LANDINGS

EXHIBIT 3

Sample Page from 1940 Version

	Excellent				Above Average				Average				Below Average				Very poor			
	Into Wind	Cross Wind	Down Wind	Gusty Air	Into Wind	Cross Wind	Down Wind	Gusty Air	Into Wind	Cross Wind	Down Wind	Gusty Air	Into Wind	Cross Wind	Down Wind	Gusty Air	Into Wind	Cross Wind	Down Wind	Gusty Air
Traffic check																				
Field "																				
Throttle Control) Speed																				
) Coord.																				
Elevator control																				
Rudder "																				
Aileron "																				

Add Judgment

TAXIING

Rating forms of this type were developed for all maneuvers contained in Standard Flights A and C-1.[5] In addition, rating forms for Stalls, Slips, and Spins were constructed.[6]

2. Method of Scoring the 1940 Version. Each item for each maneuver was assigned a scale value of excellence of performance ranging from 1, excellent, to 5, very poor, based, as in the case of the preliminary form, on discussions with flight instructors, information from the CAA flight instructor's manual, and the experience of one of the authors in flying. The assignment of these values was subjective and somewhat arbitrary.

Three methods of scoring the check sheets were developed:

a. Method 1: For each maneuver, the score was the mean scale value of the items checked for that maneuver. This method gave equal importance to each item.

b. Method 2: The second type of score developed for each maneuver was called a "weighted mean score" designed to take into account the importance as well as the scale value of the items. For purposes of this method of scoring, the importance of an item was considered as directly related to its scale value, thereby giving greater weight to those items representing poor performance.[7] This was done by giving items of value 1 a weight of 1, items of value 2 a weight of 2, items of value 3 a weight of 3, and so on, and dividing the total by the sum of the unweighted values. For example, if for one maneuver five items were checked, one with value 1, one with value 2, one with value 3, one with value 4, and one with value 5, the usual mean value (Method 1) would be 3.0. Using the importance weights, the weighted mean value would be 3.7.

c. Method 3: A "variability score" was obtained for each maneuver by computing the variance about the mean maneuver score of the item values checked for the maneuver.

---

[5] Viteles, M. S., and Thompson, A. S. Op. cit.

[6] These, with the maneuvers included in Standard Flights A and C-1, were all of those taught in the primary course leading to the private license.

[7] Discussion of the item scale values with flight instructors revealed that one poor item of flight performance may have considerable influence in lowering the instructor's over-all rating. For example, in landing, a student may come in with proper glide angle, proper air speed, stick all the way back, wings level, direction satisfactory, but may level off ten feet too high. This one item -- leveling-off ten feet too high -- would perhaps result in an over-all rating of "4" in spite of the fact that in all other respects the landing was done correctly. From such evidence it seemed that each item in the standard flight check list should be given not only a value indicating the level of performance, but also an importance weight as well.

3. <u>Collection of Data</u>. As part of the 1940 Ohio State University Project, the performance of thirty student pilots being trained at three CPT operations in the Columbus, Ohio, area was observed during two successive flights each. Standard Flight A was used and a 1940 version of the Ohio State Flight Inventory filled out for each flight (Exhibit 2). The student's own instructor served as observer for one flight; another instructor for the second.

The procedure yielded thirty paired sets of inventory ratings, representing the performance of thirty students from three schools and the observations of three pairs of instructors. Each student flew the same plane during both flights.

4. <u>Frequency of Item Checks</u>. Appendix 1 shows the number of times each of the six instructors checked each item for each maneuver. It is apparent that the majority of the items checked are those descriptive of superior flight performance.[8]

Many items were not checked at all. As shown in Table 1, for example, for Turns in Climbs 23 of the 51 possible items did not receive a check mark for any of the 60 flights. This may have been due to the fact that, at the time the flights were made, most of the students were well beyond the Stage A maneuvers comprising the standard flight pattern used in the study.

5. <u>Correlation Between Total Flight Time and OSFI Scores</u>. The correlations between number of flight hours and the mean score for each maneuver, the weighted mean score for each maneuver, and the variability score for each maneuver are shown in Table 2. All of these correlations are low. It may be noted that the correlations with mean scores and with weighted mean scores are negative with the exception of Taxiing (.25 and .09) and coordination (.01 and .06).[9] That the correlations were so low may be due to the fact that most of the students were well beyond the stage for which Standard Flight A was designed. Another factor which may account for the low relationship between amount of flight training and inventory scores is that the items do not measure "judgment," or "anticipation." Flight instructors consider these factors of particular importance in the evaluation of students in the later stages of training.

6. <u>Agreement Between First and Second Rater</u>. In Table 3 are shown the correlations between the ratings for each maneuver by the first rater and those by the second rater, i.e., first versus second flight. The correlations for mean scores range from -.13 for Medium turns to .56 for Take-offs while those for the <u>weighted</u> scores range from -.07 for Medium turns to .50 for Take-offs.

The first rater-second rater coefficients in Table 3 should not be interpreted as reliability coefficients since the pairs of observations were made on separate performances and by different observers. The correlation between the

---

[8]The only exception to this is in Taxiing (a maneuver in which the instructor did the <u>valuing</u> rather than only description) in which most of the ratings were "average."

[9]A negative correlation with number of flight hours indicates that those with a greater number of hours (i.e., farther along in the flight training program) exhibited <u>better</u> flight performance as measured by the inventory scores.

TABLE 1

TURNS IN CLIMBS

Entry	Instructor A	B	C	D	E	F	Total
Simultaneous application of controls	4	3	11	14	6	5	43
Successive application of controls	1	2	3		3	6	15
Uses aileron, no rudder, on entry					1		1
Uses rudder, no aileron, on entry					1		1
Keeps nose up during entry							0
Nose wanders during entry							0
Fails to keep nose up during entry							0

Bank	A	B	C	D	E	F	Total
Accurate angle of bank	3	1	11	14	6	2	37
Bank varies not more than + or - 5°	2	2	3		2	4	13
Bank varies not more than + or - 10°		1				3	4
Bank varies more than + or - 10°		1				2	3
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0
No slipping					1		1
Slips							0
Banks more than 30°							0
Banks less than 10°							0

Turn	A	B	C	D	E	F	Total
Accurate rate of turn	5	2	11	14	6	3	41
Varies rate of turn			2		5	6	13
Fluctuation in rate of turn						1	1
Rides same rudder during turn							0
Rides opposite rudder during turn							0
No skidding		2					2
Skids							0

Climb	A	B	C	D	E	F	Total
Holds accurate rate of climb in turn	3	1	10	13	6	3	36
Climbs at opt. angle of climb + or - 5°	1	4	4	1	2	4	16
Climbs at opt. angle of climb + or - 10°	1				3	4	8
Varies angle of climb up to + or - 5° from optimum							0
Varies angle of climb up to + or - 10° from optimum							0
Climbs too steeply for bank and power							0
Holds off bank with ailerons and uses rudder, little turn in climb							0
Cross-controls, slip counteracted by climb, no gain in altitude							0
Stalls plane							0

TABLE 1 (Contd.)

TURNS IN CLIMBS

Recovery	Instructor						
	A	B	C	D	E	F	Total
Simultaneous application of controls	4	5	10	14	9	1	43
Successive application of controls	1		3			9	13
Uses aileron, no rudder, on recovery			1				1
Uses rudder, no aileron, on recovery							0
Keeps nose up during recovery							0
Nose wanders during recovery					1		1
Fails to keep nose up during recovery							0

Precision	A	B	C	D	E	F	Total
Recovers from turn "on course"	4	1	11	11	7	1	35
Recovers from turn 5° r. or l. of "on course"		2	2	2	1	7	14
Recovers from turn not more than 10° r. or l. of "on course"		2			1	2	5
Recovers from turn more than 10° r. or l. of "on course"	1				1	1	3
R. and l. turns equal in precision							0
R. and l. turns not equal in precision							0

Throttle	A	B	C	D	E	F	Total
Throttle control or reduction in climbing angle for turn		2	2		4	5	13
Coordinates throttle with attitude of plane	3	2	12	14	6	3	40
Slow on throttle adjustment to attitude of plane	1	1				1	3
Abrupt on throttle control							0
Disregards throttle	1				3	4	8

[illegible]

CORRELATIONS OF TOTAL FLIGHT TIME WITH MEAN SCORES, WEIGHTED MEAN SCORES, AND VARIABILITY FOR EACH MANEUVER
(N = 60)

<u>Maneuver</u>	Mean <u>Score</u>	Weighted <u>Mean Score</u>	<u>Variability</u>
Taxiing	.25	.09	.16
Take-offs	-.22	-.20	.20
Climbs	-.22	-.03	.26
Climbing turns	-.19	-.03	.31
Straight & Level	-.22	-.13	-.01
Medium turns	-.09	-.17	.16
Glides	-.09	-.06	.03
Gliding turns	-.00	-.03	.29
Landings	-.06	-.09	.01
Coordination	.01	.06	-.01

TABLE 3

CORRELATION BETWEEN RATINGS OF FIRST AND SECOND RATER FOR MANEUVERS IN STANDARD FLIGHT A
(N = 30)

<u>Maneuver</u>	Mean <u>Score</u>	Weighted <u>Mean Score</u>
Taxiing	.03	.14
Take-offs	.56	.50
Climbs	.17	.22
Climbing turns	.12	.17
Straight and level	.13	.28
Medium turns	-.13	-.07
Glides	.24	.16
Gliding turns	.06	.27
Landings	.55	.35
Coordination	.53	.41

ratings of paired observers is thus affected not only by differences between raters but also by differences in the actual performance during the two flights.[10]

There seems to be some relationship between the complexity of the maneuver and the magnitude of the first rater-second rater coefficient. Examination of Appendix 1 shows that for those maneuvers having low coefficients the observers checked a greater proportion of those items characterising high quality flying, and few of the items representing defects in flying. The items for Take-offs, Landings, and Coordination show a much wider range of items checked as well as higher first rater-second rater coefficients.

It seemed of interest to investigate further the correlations between first and second rater for Medium turns, the only maneuver for which a negative correlation was obtained. Table 4 presents the distribution, means, and standard deviations (computed from ungrouped data) of the mean scores for Medium turns for each of the six instructors, A to F, inclusive.

TABLE 4

DISTRIBUTION OF MEAN SCORES FOR MEDIUM TURNS
ASSIGNED BY THE SIX FLIGHT INSTRUCTORS

	Instructor					
Score	A	B	C	D	E	F
3.2 - 3.3						2
3.0 - 3.1						1
2.8 - 2.9			1			2
2.6 - 2.7						
2.4 - 2.5						
2.2 - 2.3						1
2.0 - 2.1		1	2		1	1
1.8 - 1.9		1	2			
1.6 - 1.7		2		2	1	1
1.4 - 1.5	2	1	3		3	1
1.2 - 1.3	1		1	4	3	2
1.0 - 1.1	2		5	8	3	
Mean	1.24	1.72	1.52	1.16	1.37	2.29
Standard Deviation	.22	.19	.60	.24	.29	.76

[10] Editor's Note. An attempt to isolate and measure separately the variability of the student's performance from flight to flight is currently being made through the use of motion photographs of the flight performance. See: Wapner, S., Festinger, L., and Odbert, H. S. Consistency of student pilot performance as observed in photographic records. Progress Report, January 15, 1945. (Copy in Committee files.)

When the ratings of all raters were corrected to the same mean, the correlation between the mean scores for the two flights became .05, and when the ratings of Raters E and F were omitted,[11] the correlation became .51. When the ratings of all six raters were equated for variability as well as for mean, the correlation became .20.

7. Comparison of Scoring Methods. The intercorrelations among the three types of scores are presented for each maneuver in Table 5. As might be expected, the mean scores and the weighted mean scores exhibit high correlation (.88 to .98), although not identical results.

With the exception of Taxiing, the correlations of the variability scores with the mean scores and with the weighted mean scores are fairly high and positive, due perhaps partly to the fact that there were limits to the score values assigned to the individual items. If the mean score falls at one extreme or the other of the total possible range of values, there can be little variability. If the mean score falls somewhere near the middle of the distribution, it may or may not have considerable variability. Since practically all the mean scores fall in only the bottom half of the total possible range, a positive correlation between variability and maneuver scores would be expected.

TABLE 5

INTERCORRELATIONS AMONG CSFI SCORES
(N = 60)

Maneuver	Mean Scores vs. Variability Scores	Weighted Mean Scores vs. Variability Scores	Mean Scores vs. Weighted Mean Scores
Taxiing	-.20	-.03	.98
Take-offs	.89	.74	.86
Climbs	.92	.94	.96
Climbing turns	.79	.82	.94
Straight & Level	.70	.40	.90
Medium turns	.74	.61	.96
Glides	.73	.54	.91
Gliding turns	.83	.52	.88
Landings	.92	.74	.90
Coordination	.90	.70	.88

[11] Observers E and F exhibited the greatest disagreement. This may have been due partly to the fact that one observer made the flights in the same type plane as he used for instruction purposes while the other was observing in a different type of plane. Another source of discrepancy may be the fact that some of the students of Observers E and F did not receive the same amount of indoctrination as to the purpose of the check flights as did the subjects in the group as a whole. In the case of two of the student pilots there was definite evidence of "inspectoritis." In one of these cases the student's own instructor had no opportunity to introduce him to the strange instructor. The strange instructor merely called to the student, "Come on, let's go," with no explanation as to the purpose. These two cases of Instructors E and F stand out in practically all first rater-second rater comparisons as extreme deviates from the general trend.

The low negative correlation of variability with mean scores for Taxiing is additional evidence of the advisability of separating functions of description and evaluation. As described on page 15 the "taxiing" scale was in the form of ratings rather than descriptive items.

8. Derivation of Over-all Scores. From the point of view of validation of aptitude tests or other selection devices, a single over-all criterion measure is desirable. In the case of the OSFI, the problem was to combine the scores for the several maneuvers into an over-all score. The following methods of obtaining a total score for the flight as a whole were considered:

a. Summing the maneuver scores to obtain a total flight score.

b. Deriving aggregate scores on aspects of flight performance. For example, the observations for entry into turn could be treated together (regardless of the maneuver in which they occurred) to obtain an aggregate score for "entry into turn."

c. Weighting each of the maneuver scores proportionally to its importance as judged by flight instructors or other flight experts.

d. Weighting each of the maneuver scores on the basis of its relationship to the other scores, as determined through factor analysis. One might assume that the first factor (if sufficiently general) would be the best single summary of information contained in the data and that whatever is common to the various maneuver scores is the over-all criterion being sought.

The first two methods although the simplest, were discarded since they provided no adequate control of the weights.

The third method was investigated experimentally. It was assumed that flight instructors would be the most competent to judge the importance of the several maneuvers relative to flight competence. A preliminary ranking of 30 maneuvers in the CPT primary flight course was made by five flight instructors in the Columbus, Ohio, area. The average intercorrelation of their rankings was .31. Rankings were then obtained from 74 CAA rerated flight instructors from various parts of the country. Correlations of the first 18 of these rankings with the mean rank (based on the 74 instructors) ranged from -.47 to .66 with a mean value of .27. Intercorrelations among the 74 rankings were not computed.

Table 6 shows the various maneuvers in the order of their mean rank importance based on the rankings of all 74 instructors. The maneuver Stalls obtained the highest mean rank (7.96) and thus was given a composite rank of 1. Turns in Climbs and Glides was ranked as the second most important maneuver with a mean

TABLE 6

IMPORTANCE OF MANEUVERS LISTED IN LOG BOOK, BASED ON RANKINGS BY 74 FLIGHT INSTRUCTORS RERATED FOR C.P.T.P.

Maneuver	Composite Rank	Mean Rank	S.D. of Rankings
Stalls	1	7.96	5.9*
Turns in Cl. and Gl.	2	8.14	5.9*
Climbs and glides	3	9.36	7.4
Straight and level	4	10.68	8.4
Medium turns	5	10.76	7.5
Forced Land. - take-off	6	11.68	7.4
Steep turns	7	11.86	6.1*
Coordination exercises	8	11.91	9.7
Spins	9	12.91	7.5
Landings into wind	10	12.99	6.9
Take-offs into wind	11	13.27	7.7
30° eights	12	14.18	7.5
Forced land. 90° to wind	13	14.26	7.9
70° power turns	14	15.05	6.8*
70° eights	15	15.18	8.5
Forced land. - 180° to wind	16	15.30	8.0
Taxiing down wind	17	15.74	8.5
Landings cross wind	18	16.27	8.0
Taxiing cross wind	19	17.04	8.3
Taxiing gusty air	20	17.20	8.3
180° side approach	21	17.47	6.0*
Taxiing into wind	22	17.64	9.1
Slips	23.5	17.68	7.1
Rectangular course	23.5	17.68	8.0
Take-offs cross wind	25	18.07	7.6
720° spiral approach	26	19.85	7.5
360° overhead approach	27	19.86	5.2*
Power approaches	28	24.43	5.6*
Power landings	29	25.43	5.2*
Dragging areas	30	25.97	4.7*
Non-maneuver items			
Aptitude		10.49	12.0
Confidence maneuvers		19.32	8.8

*Significantly different from chance value of 8.65

rank of 8.14.[12]

The standard deviation about the mean rank for each maneuver indicates considerable variability of the ranks assigned for each maneuver. These standard deviations range from 4.7 to 9.7. If the rankings of maneuvers were made on a purely chance basis, the standard deviation would be 8.65. The nine standard deviations marked in Table 6 with an asterisk are the only ones which differ significantly from the chance value of 8.65. In these nine cases, the chances are less than one in a hundred that such a deviation could be produced by chance variation. For the remaining 21 maneuvers, the standard deviation aobut the mean rankings could be considered a chance variation in the universe where the true standard deviation is 8.65. Because of the lack of consistency among the ratings of importance, and because other attacks upon the problem seemed more promising, this approach to weights for composite scores was not used.

Two "non-maneuver" items were included but not considered in the mean ranks. Aptitude obtained approximately the same rank as straight and level flight, but had the largest standard deviation. It was ranked either very important or quite unimportant by most raters. Confidence maneuvers is not a single maneuver, since various maneuvers are used to give the student pilot confidence in himself and the plane.

In the fourth method of obtaining an over-all score, the intercorrelations among the maneuver scores were considered. Table 7 shows the intercorrelations among the maneuvers scored by using the mean value of the items checked for each maneuver and Table 8 the intercorrelations using the weighted mean scores.[13] These correlations are based on 60 sets of ratings (two sets of ratings for each of the 30 students) obtained as described on page 10.

Examination of Table 7 reveals that of the six coefficients greater than .60, three are due to the intercorrelation among Climbing turns, Straight & Level Flight, and Climbs, indicating the possibility of a factor common to these three performances. The other three correlations above .60 are correlations of Gliding turns with Climbing turns, with Medium turns, and with Glides.

---

[12] Editor's Note. It is interesting to compare these rankings in terms of "judged importance" with a later study on the association between ratings on maneuvers and success or failure in flight training. In the latter study, Takeoff into wind, Climbs and Glides, Straight & Level, Gliding approach, Medium turns, and S's and 8's were found to exhibit consistently high relationship with a pass-fail criterion. Stalls and spins were among the group of maneuvers showing consistently low relationship. See: Viteles, M. S., Franzen, R., and Rogers, R. The association between ratings on specific maneuvers and success or failure in flight training of RAF cadets. Washington, D. C.: C.A.A. Division of Research, Report No. 37, October 1944.

[13] These maneuver intercorrelations furnish additional evidence that the first first rater-second rater coefficients in Table 3 cannot be considered reliability coefficients. For example, the correlation between Climbing turns and Climbs for the weighted mean scores (Table 8) is .62. The first rater-second rater coefficients for Climbs is .22 and for Climbing turns is .17 (Table 3). It would be impossible to obtain a correlation of .62 between two variables whose reliabilities were .22 and .17, respectively, except by chance.

TABLE 7

MANEUVER INTERCORRELATIONS (Mean Scores)
(N = 60)

Maneuver	Take-offs	Climbs	Cl. Turns	Str. & Level	Med. Turns	Glides	Gl. Turns	Land.	Coord.
Taxiing	-.219	-.249	-.224	-.197	-.201	-.183	-.198	-.174	-.010
Take-offs	----	.416	.543	.476	.380	.266	.567	.296	.310
Climbs		----	.689	.634	.386	.331	.508	.127	.286
Cl. Turns			----	.678	.593	.503	.671	.128	.210
St. & Level				----	.397	.372	.493	.317	.254
Med. Turns					----	.465	.659	.326	.318
Glides						----	.634	.281	.232
Gl. Turns							----	.278	.230
Landings								----	.431
Coord.									----

It may be noted that Taxiing correlates negatively with each of the other variables. Because of the nature of the ratings for Taxiing, such correlations do not necessarily indicate that competence in Taxiing is inversely related to competence in other maneuvers. The negative correlation with other maneuvers may have been due to the fact that Taxiing is the only maneuver for which the observer rated the performance of the student as superior, above average, average, below average, or poor, as against checking descriptive items of performance for the other maneuvers.

Table 8 presents the correlations based on the weighted mean scores. These correlations, in general, are somewhat similar to those of Table 7, but tend to be lower. Only two of the coefficients are above .60, and three between .50 and .60. The fact that they are lower does not necessarily reflect a real difference in the relationship of competence in the several maneuvers. It may be due to the fact that the weighting method made those scores which were not consistent with the general trend (for example, Raters E and F) even more extreme deviates.

9. Factor Analysis of Maneuver Intercorrelations. To get a better picture of the grouping of maneuvers, a factor analysis was made of the weighted mean score intercorrelations in Table 8. The centroid factor loadings without rotation, are shown in Table 9. The first factor loadings are the more significant. Of the first factor residuals, only one was greater than two times the standard error of a correlation coefficient of zero where N = 60. Except for Taxiing, all first factor loadings are positive and of somewhat the same magnitude. The largest first factor loading is .732 for Climbing turns. The second largest is for Gliding turns, .673. While these loadings are but little larger than those for other maneuvers, it is interesting to note that these maneuvers are considered by many experienced instructors to reveal better than other simple maneuvers the pilot's recognition of the attitude of the plane and his ability to coordinate.[14]

Inspection of the factor plot indicates that the maneuvers fell into two groups. Group 1, characterized by positive loadings in both factors, indluded Take-off's, Medium turns, Glides, Gliding turns, Landings, and Coordination. Group 2, with positive first factor loadings and negative second factor loadings included Climbs, Climbing turns, and Straight & Level Flight.

It is difficult to interpret such a classification of maneuvers. With additional observations it might be feasible to carry such an analysis to more definite conclusions.[15]

---

[14]Since centroid first factor loadings are often considerably changed with rotation, new factor loadings based on a 14° rotation were computed. Gliding turns received the highest Factor I loading .686, followed by Medium turns .638, Landings .621, and Climbing turns .611.

[15]Editor's Note. This early study was purely exploratory and no conclusions could be drawn. Additional information on the interrelationships among maneuvers is found in: Johnson, H. M., and Boots, M. L. Op. cit.

A related study, involving a factor analysis of ratings on "aspects of flight performance," such as wing control, nose control, maintenance of altitude, etc., suggested the presence of three factors: (1) the over-all impression of the flight performance as a whole; (2) longitudinal control; and (3) coordination of the controls. See: Viteles, M. S., and Thompson, A. S. An analysis of photographic records of aircraft pilot performance. Washington, D.C.: C.A.A. Division of Research, Report No. 31, July 1944, pp. 89-91.

TABLE 8

MANEUVER INTERCORRELATIONS (Weighted Scores)
(N = 60)

Maneuver	Take-offs	Climbs	Cl. Turns	Str. & Level	Med. Turns	Glides	Gl. Turns	Land.	Coord.
Taxiing	-.157	-.289	-.231	-.146	-.215	-.232	-.216	-.270	.019
Take-offs		.126	.312	.255	.170	.063	.304	.303	.246
Climbs			.616	.500	.260	.192	.396	.113	.183
Cl. turns				.610	.517	.369	.524	.091	[illegible]
St. & Level					.308	.352	.298	.354	.205
Med. turns						.409	.494	.370	.197
Glides							.440	.361	.228
Gl. turns								.301	.219
Landings									.342
Coord.									----

TABLE [illegible]

FACTOR LOADINGS FOR WEIGHTED SCORES
(Centroid Factor Loadings)

Maneuver	Factor I	Factor II	h^2
Taxiing	-.367	-.054	.138
Take-offs	.405	.124	.176
Climbs	.596	-.429	.704
Climbing turns	.732	-.410	.702
Straight & Level Flight	.659	-.328	.542
Medium turns	.626	.128	.408
Glides	.559	.180	.345
Gliding turns	.673	.138	.472
Landings	.521	.476	.498
Coordination	.380	.161	.170

On the assumption that whatever is common to all the weighted maneuver scores is the best estimate of competence in flying practice, regression coefficients and the gross score weights to be used in obtaining an estimate of the first factor score were obtained (see Table 10). These scores represent Factor I as obtained by the centroid factor method without rotation. Use of these weights produced a composite score which correlated .92 with Factor I.

10. Summary of the 1940 Version of the OSFI. The 1940 version of the OSFI represented a successful attempt to obtain a standardized form for recording flight performance in descriptive rather than evaluative terms.

For each maneuver with the exception of Taxiing, the inventory provided a list of items descriptive of pilot performance with the observer required merely to check whether or not the pilot exhibited those items during a standard flight. The inventory had value for diagnostic as well as criterion purposes, since inspection of the item checks revealed the specific strengths and faults in the individual's performance.

The scoring method was objective in the sense of yielding a score based on the average scale value of the checked items. The scale values themselves were determined subjectively, since based upon instructor opinion, but they represented a "consensus" of opinion rather than the judgment of one person. Of particular value was the fact that, once determined, the standards of performance represented by the scale values did not fluctuate as in the case of the usual rating method.

Research with the 1940 version revealed certain practical limitations in its present form. The inventory was found to be too long, and to contain too many items for convenient use in the field. Satisfactory filling in of this form required careful training and a high level of motivation in the observer.[16] Many of the items were deadwood in the sense of being checked very infrequently.

On the basis of inventory data on two flights each by 30 student pilots, correlations between the first and second raters were found to be generally low. These correlations cannot be considered as measures of the reliability of the inventory data, however, since both student pilot variability and observer variability were present in the two flights.

TABLE 10

NET REGRESSION COEFFICIENTS AND GROSS SCORE WEIGHTS FOR THE VARIOUS MANEUVERS

Maneuver	Net Regression Coefficients	Gross Score Weights
Taxiing	-.0805	-2.98
Take-offs	.1167	1.80
Climbs	.1643	1.66
Climbing turns	.1643	1.94
Straight & Level Flight	.2176	3.92
Medium turns	.1851	2.62
Glides	.1490	2.24
Gliding turns	.2019	2.58
Landings	.0780	1.03
Coordination	.0989	1.50

The problem of devising an over-all score was attacked: first by weighting maneuvers according to their importance as judged by a group of 74 flight instructors, and second, by obtaining regression weights based on a factor analysis of maneuver intercorrelations. Neither method was considered satisfactory.

C. The 1941 Version of the OSFI.

1. Description. The 1941 version was essentially a simplification of the previous version, obtained by weeding out the items found to be infrequently checked. The 1941 maneuver sheet for Landing, as shown in Exhibit 4, contained only 27 items, in contrast with the 46 items for Landing in the 1940 version. In addition the format was improved by grouping mutually exclusive items and the actual size of the sheet was reduced from 8½" by 14" to 8½" by 5½".

---

[16] Editor's Note. Later experience at Tulane University indicated that the 1940 version of the inventory was not suited for use by the ordinary flight instructor. An account of the difficulties experienced in a trial situation is given in: Viteles, M. S., and Backstrom, O. An analysis of graphic records of pilot performance obtained by means of the R-S Ride Recorder. Washington, D. C.: C.A.A. Division of Research, Report No. 23, November 1943, p. 115.

The method was retained of classifying items under functions such as "throttle" or "planning," and under steps in the maneuver such as "level off," "stall," etc. In turns, for example, the items were grouped under entry, turn, and recovery. As in the 1940 version, each item was assigned a scale value of 1 to 5 on the basis of consultation with flight instructors and a weighted mean score obtained for each maneuver.

Maneuver sheets were developed for all the maneuvers in the CPT primary course, including Taxiing, Take-offs, Climbs, Climbing turns, Spins, Power turns, Stalls, Roll from left gliding turn to right gliding turn, Slips, Gliding turns, 8's, Straight & Level, Glides, 180° Approach, Landings, Forced landings, and for coordination, tension and judgment.

2. Collection of Data. During the Spring and Summer of 1941, 143 flight observations were made, using a standard flight which included all the maneuvers in the primary CPT course. These were arranged in such sequence that the entire flight could be made in approximately one hour. The observations were made on CPT students at three flying shcools.[17]

Three flights were made with each student, using the same flight pattern for each flight. Flight 1 was made with the chief instructor of the shcool, just prior to the last hour or so of "polishing off" for the private license. Flights 2 and 3, made within a week after the student obtained his private license, were flown with Walker and Ceschiat. They were so scheduled that the student made Flight 2 with one observer and, immediately upon landing, changed to the plane of the other observer and repeated the flight. Sometimes Ceschiat, and sometimes Walker,[18] was the observer for Flight 2.

3. Study of Sources of Variance. Since there were five different observers, and flights were made under varying conditions following training in different schools and by different instructors, the data were first examined, maneuver by maneuver, for possible sources of variance which might be ascribed to factors other than those which could be thought of as a function of individual differences among the students. The major sources of variance which could be isolated in this study were as follows:

a. Instructor. Due to differences in instructional methods, personality, and other factors, different instructors obtain different results in terms of flight competence of students.

b. Observer. Different observers emphasize different factors, note different items of performance, react differently to the same student. They may obtain different observations of the same student.

---

[17]The Miller Flying School at Port Columbus, Columbus, Ohio; the Midwest School of Aviation at Norton Field, Columbus, Ohio; and the Ann Arbor Flying Service, Ann Arbor, Michigan.

[18]R. Y. Walker, as research assistant in the Ohio State University Project had obtained a flight instructor's license. Angelo Ceschiat, a flight instructor at the Miller Flying School, Port Columbus, Columbus, Ohio, was hired to participate in this research project

EXHIBIT 4

Sample Page from 1941 Version

<u>Deviation in Direction & Lateral Balance</u>

During Glide:	Direction_______	degrees R or L
	Lat. Bal._______	degrees R or L
During Roll:	Direction_______	degrees R or L
	Lat. Bal._______	degrees R or L

<u>Throttle</u>
Does) or (does not keep hand on throttle

<u>Planning</u>
Does) or (does not plan glide path
Does) or (does not consider other traffic
Recognizes) or (fails to recognize drift
Corrects) or (fails to correct for drift

<u>Level off</u>
Levels off_______feet too high

<u>Stall</u>
Abrupt), (jerky), or (continuous action on elevators
Complete stall) or (wheel landing
Stalls near the ground), (pancakes), or (more than 3 feet above ground

<u>Poor Landing</u>
Recognizes) or (fails to recognize) poor landing and (corrects) or (fails to correct) by adequate use of throttle

(Landing begins at level-off for landing)

LANDING

c. School. The school in which the student was trained may make a difference. Differences in equipment, general attitude, discipline, quality of "hangar flying," etc., may be operative.

d. Wind Velocity. Wind velocity at the time of the check flight may make a difference, even within the narrow range of velocities under which the check flights were made. Wind velocity might be particularly important in those maneuvers made relative to points or patterns on the ground, such as Pylon 8's, Rectangular course, Landings, and Take-offs.

e. Flight Number. In a series of check flights, there may be difficulties associated with the order of the check flight. Nervousness may increase or decrease from first to second flight.

f. Familiarity with Plane. Previous experience with a plane would yield information as to its particular flight characteristics and thus might influence flight performance. In this study Flights 2 and 3 were made in different planes: one in which the student pilot had had the bulk of his training (designated as the "N" plane); and one brought to the field by the research investigators (designated as the "F" plane). All planes used were J-3 Cubs, and all the students had been training in J-3 Cubs. Each student made two flights in an "N" plane (Flight 1 and either Flight 2 or Flight 3) and one in "F" plane.

In addition to using the customary analysis of variance, it was decided to use ε^2, a correlation function developed by T. L. Kelley.[19] Table 11 shows the values of ε^2 for each maneuver in the standard flight for each of the following bases of classification: Flight Number, Instructor, Observer, School, and Wind Velocity. Of the 31 maneuver values of ε^2, 26 were greater than the 1% value for Observer, 22 for Instructor, 13 for Flight Number, 12 for School, and only 4 for Wind Velocity.

Experience had shown that inadequately trained observers were a source of considerable error. If the above significant variances ascribed to Flight Number, Instructor, Observer, etc., markedly decreased when the observations were made with better trained and supervised observers, the Flight Inventory technique would be shown to be more valuable, and the conditions of its use better described. In the light of such a situation, it was thought advisable to investigate further the problem of homogeneity of observations.

Since Walker and Ceschiat were the best trained observers and since they made all of the observations for Flights 2 and 3 and none for Flight 1, it

---

19Epsilon squared is derived from the index of correlation. It yields a value which can be used both for estimation and for description. It can be interpreted in the same way as the analysis of variance, and also as a correlation coefficient indicating the degree of relationship. See: Kelley, T. L. An unbiased correlation measure. Proc. Nat. Acad. Sci. Vol. 21, pp. 554-559. The same information is also shown in Table 12 in terms of the F ratio (ratio of mean square between classes to mean square within classes).

[illegible]

[illegible]	[illegible]	[illegible]structor	Observer	School	Wind Velocity
[illegible]	[illegible]	.162	.325	.099	.038
[illegible]ke offs	[illegible]	.067	.130	.011	-
[illegible]	[illegible]	.122	.171	.003	.044
[illegible] in Climbs - L	[illegible]	.173	.134	.077	.011
[illegible] in Climbs - R	[illegible]	.184	.109	.064	.022
[illegible] in Climbs - Both	[illegible]	.162	.121	.096	.027
[illegible] - L	[illegible]	.078	.047	-	-
[illegible] - R	-	.233	.091	.094	-
[illegible] - Both	[illegible]	.149	.118	.055	.000
[illegible] - L	[illegible]	.147	.073	.035	.024
[illegible] - R	[illegible]	.099	.193	.004	.027
[illegible] - Both	[illegible]	.140	.216	.015	.045
[illegible]	-	.112	.068	.002	-
[illegible]	[illegible]	.105	.281	.007	-
[illegible]	[illegible]	.190	.071	.050	.009
[illegible] - [illegible]	-	.060	.001	.000	.009
[illegible] - [illegible]	-	.119	.022	.053	-
[illegible] - Both	[illegible]	.079	.018	.034	.006
[illegible] in Glides - [illegible]	-	.147	.270	.083	.021
[illegible] in Glides - [illegible]	-	[illegible]	.399	.064	-
[illegible] in Glides - [illegible]	-	.171	.413	.092	.013
Steep 8's	[illegible]	.084	.329	.018	.073
[illegible]rced Landings	.009	.064	.055	.010	.030
[illegible]raight and Level	.054	.156	.316	.005	.061
[illegible]allow 8's	[illegible]	.138	.414	.023	.060
[illegible]ides	[illegible]	.108	.123	-	.043
[illegible]ndings	.054	.034	.087	.002	.065
[illegible] Approach	.049	.118	.192	.044	.013
[illegible]rdination	.052	.160	.345	.072	.045
[illegible]dgment	.052	.053	.158	.002	.015
[illegible]anning	.022	.075	.104	.004	-
[illegible] Level	[illegible]	.055	.036	.026	.032
[illegible] Level	[illegible]	.086	.060	.047	.054

[illegible]anks [illegible] negative values, [illegible] where the mean square for error is [illegible]eater than the [illegible] square between classes.

TABLE 12

VALUE OF F FOR DIFFERENT CLASSIFICATIONS
(N = [illegible])

Maneuver	Flight Number	Instructor	Observer	School	Wind Velocity
Taxiing	11.35	24.11	23.22	8.76	2.87
Take-offs	1.97	9.73	6.29	1.77	-
Climbs	8.85	17.71	3.34	1.19	3.20
Turns in Climbs - L	1.94	25.08	6.51	6.95	1.52
Turns in Climbs - R	-	23.14	5.36	5.89	2.07
Turns in Climbs - Both	2.24	24.22	5.88	8.52	2.31
Spins - L	1.66	11.16	2.76	-	-
Spins - R	-	37.50	4.55	8.34	
Spins - Both	1.85	12.10	5.73	5.10	1.60
Power Turns - L.	6.52	11.65	3.81	3.55	2.37
Power Turns - R	11.94	24.19	9.47	1.32	2.30
Power Turns - Both	11.38	20.56	10.78	2.06	2.75
Stall-off	-	16.09	4.63	1.10	-
Stall-on	3.38	15.10	14.90	1.48	-
Rolls	-	37.22	3.70	7.15	1.41
Slips - L	-	3.60	1.02	.98	1.42
Slips - R	-	17.18	1.81	4.93	-
Slips - Both	1.07	10.19	1.65	3.51	1.27
Turns in Glides - L	-	21.67	14.16	7.39	2.04
Turns in Glides - R	-	24.94	24.60	5.90	
Turns in Glides - Both	-	25.72	25.92	8.23	1.63
Steep 8's	15.81	11.96	18.42	2.33	4.73
Forced Landings	1.21	9.30	3.06	1.72	2.45
Straight and Level	5.03	13.26	17.41	1.39	4.19
Shallow 8's	8.91	20.26	26.10	2.70	4.72
Glides	5.63	15.54	5.99	-	3.13
Landings	5.05	5.27	4.38	1.14	4.30
180° Approach	4.65	17.04	9.43	4.26	1.53
Coordination	4.88	23.88	19.70	6.39	3.25
Judgment	4.93	7.65	7.64	1.14	1.72
Planning	2.58	10.77	5.13	1.28	-

Blanks are negligible values, i.e., cases where the mean square for error is greater than the mean square between classes.

was thought that perhaps their inventory results would be more homogeneous. Further study was therefore made of a restricted sample consisting of the 69 observations made by these two observers.

Tables 13 and 14 present the values of ϵ^2 and F for the 69 observations made by these two observers. The results, in general, showed a decrease in variance due to the sources studied and suggested greater homogeneity of ratings when trained observers are used. The absolute values of ϵ^2 in Table 13, for example, were numerically smaller than those in Table 11 in the case of 18 of the 31 maneuvers for Flight Number, in 28 of the 31 maneuvers for Observer, in 20 of the 31 maneuvers for School, and in 19 of the 31 maneuvers for Wind Velocity.

4. Factor Analysis of Maneuver Intercorrelations. As in the case of the 1940 Revision (see page 17) intercorrelations among maneuver scores were obtained so as to investigate consistency in performance from maneuver, the existence of "factors" in flight performance, etc. Since the analysis described above had revealed the restricted sample of 69 observations by Walker and Ceschiat to be more homogeneous the intercorrelations were based on these data. In addition only 15 of the 31 maneuvers were used.[20]

Table 15 presents the intercorrelations. In general, these correlations are low and positive. Only 8 of the correlation coefficients are as high as .30, the value required for significance at the 1% level with an N of 69. The correlation of Take-offs and Climbs is .344; left and right Climbing turns correlated .423; right Power turns correlated .316 with Landings; Shallow 8's correlated .675 with Steep 8's, .314 with 180° Approach, and .362 with Straight & Level Flight. Landings and Straight & Level Flight correlated .302.[21]

The correlation matrix, Table 15, was factored by the Thurstone Centroid method. Table 17 presents both the centroid and the rotated factor loadings for each maneuver. The rotated factors are of particular interest. Factor A has high loadings for Steep 8's and Shallow 8's. Factor B is characterized by fairly high loadings in Climbs, left Climbing turns, and left Power turns. Factor C, well defined, is characterized by fairly high loadings in such maneuvers as Straight & Level Flight, Take-offs, 180° Approaches and Landings.

On the basis of the rotated factors, it was thought advisable to construct factor scores for each of the 69 cases so that an item analysis of the inventory could be made for each factor.

Two procedures were used to set up factor scores. One utilized weighted values of scores of all maneuvers for each factor; the second only selected maneuvers for each factor. The first procedure yielded a multiple correlation of .366 for Factor A, a multiple correlation of .794 for Factor B, and a multiple correlation of .841 for Factor C.

---

[20]Selected for inclusion in the succeeding (1942) version of the OSFI.

[21]For comparison, the intercorrelations for all 143 flights for the same 15 maneuvers are shown in Table 16. In this Table 17 coefficients are as high as .30, as compared with 8 in Table 16. This difference may be due to the greater heterogeneity of the larger sample.

TABLE 33

VALUES OF E^2 FOR DIFFERENT CLASSIFICATIONS
(Observers Walker and Geschist. N = 69)

Maneuver	Flight Number	Plane	Observer	School	Wind Velocity
Taxiing	-	.028	.007	.047	-
Take-offs	.025	.058	.020	.047	.010
Climbs	-	.000	.058	.204	-
Turns in Cl - L	-	.009	-	.009	-
Turns in Cl - R	-	-	-	-	-
Turns in Cl - Both	-	.020	-	.023	-
Spins - L	.015	.335	.831	.018	.361
Spins - R	-	.014	.047	-	.191
Spins - Both	-	.029	.113	-	-
Power Turns - L	.084	.077	-	.103	.051
Power Turns - R	.056	.043	.053	.003	-
Power Turns - Both	.054	-	.051	.061	-
Stall-off	-	.048	-	-	-
Stall-on	-	-	.361	-	-
Rolls	-	.037	.055	.023	.002
Slips - L	-	-	-	.002	-
Slips - R	.005	-	.016	.030	-
Slips - Both	-	-	.002	.042	.031
Turns in Glides L	-	.016	.028	.022	.011
Turns in Glides R	-	-	.059	.095	-
Turns in Glides Both	-	.019	.060	.064	.046
Steep 8's	-	.021	.017	-	.015
Forced Landings	.008	.018	.011	.111	-
Straight and Level	-	.023	.384	-	.007
Shallow 8's	-	.030	.141	.001	.016
Glides	-	-	-	-	-
Landings	-	-	.034	.018	.006
180° Approach	-	-	-	-	-
Coordination	.016	-	.031	-	-
Judgment	.017	-	.069	-	-
Planning	.007	-	.052	-	-
1% Values of E^2	.082	.082	.082	.105	.120

Blanks are negative values, i.e., cases where the mean square for error is greater than the mean square between classes.

VALUES OF F FOR DIFFERENT CLASSIFICATIONS
([illegible] N = 69)

Maneuver	Flight Number	Plane	Observer	School	Wind Velocity
Taxiing	-	2.93	1.45	2.66	-
Take-offs	2.74	5.16	2.35	2.67	1.22
Climbs	-	1.01	5.20	9.73	-
Turns in Cl - L	-	1.60	-	1.30	-
Turns in Cl - R	-	-	-	-	-
Turns in Cl - Both	-	2.41	-	1.81	-
Spins - L	2.01	35.30	337.73	1.61	13.81
Spins - R	-	1.96	4.39	-	1.44
Spins - Both	-	3.04	9.62	-	-
Power Turns - L	7.23	6.70	-	4.91	2.21
Power Turns - R	5.05	3.87	4.82	1.11	-
Power Turns - Both	4.36	-	4.65	3.22	-
Stall-off	-	4.46	-	-	-
Stall-on	-	-	39.31	-	-
Rolls	-	3.61	4.92	1.79	1.05
Slips - L	-	-	-	1.06	-
Slips - R	1.31	-	2.11	2.05	-
Slips - Both	-	-	1.17	2.47	1.73
Turns in Glides L	-	2.12	2.97	1.77	1.26
Turns in Glides R	-	-	5.25	4.56	-
Turns in Glides Both	-	2.30	5.32	3.32	2.10
Steep 8's	-	2.46	2.14	-	1.35
Forced Landings	1.55	2.27	1.74	5.26	-
Straight and Level	-	2.32	43.36	-	1.16
Shallow 8's	-	3.11	12.13	1.03	1.37
Glides	-	-	-	-	-
Landings	-	-	3.38	1.61	1.14
180° Approach	-	-	-	-	-
Coordination	2.11	-	3.21	-	-
Judgment	2.17	-	6.05	-	-
Planning	1.46	-	4.73	-	-

Blanks are negligible values, i.e. cases where the mean square for error exceeds the mean square between classes.

TABLE 15

INTERCORRELATION OF SELECTED MANEUVERS
(Flights 2 and 3: N = 69)

	Take-off	Climbs	Turns in Cl-L	Turns in Cl-R	Power Turns R	Stall-off	Slips L	Turns in Gl-L	Steep 8's	Str. and Level	Shallow 8's	Glides	Landings	180° Approach
Taxiing	.175	-.015	.219	.010	-.031	.023	-.082	-.033	.096	.156	-.038	.142	.055	.124
Take-offs		.344	.176	.283	.088	.194	-.121	.157	.019	.323	-.067	.174	.137	.013
Climbs			.271	.185	.022	-.150	-.170	-.096	.037	.138	-.145	.059	.072	.122
T in Cl-L				.423	.027	.095	.048	.214	-.031	.023	.114	-.013	.149	.136
T in Cl-R					.240	.223	.115	.251	.155	.213	.140	.247	.180	-.001
Power Turns R						.113	[illegible]	.116	.101	.110	.087	[illegible]	[illegible]	[illegible]
Stall-offs							[illegible]	.092	[illegible]	[illegible]	[illegible]	[illegible]	.108	.092
Slips-L								.076	.026	-.066	-.176	.153	.038	.126
T in Gl-L									.122	-.134	.032	.193	.127	.188
Steep 8's										-.124	.675	.219	.098	.196
Str. and L											-.362	.145	.302	.088
Shallow 8's												.092	.182	.314
Glides													.235	.184
Landings														.126
180° Appr.														

TABLE 16

INTERCORRELATIONS OF SELECTED MANEUVERS OF FLIGHT X
(All Flights; N = 143)

	Take-off	Climbs	Turns in Cl-L	Turns in Cl-R	Power turns R	Stall-off	Slips L	Turns in Gl-L	Steep 8's	Str. and Level	Shallow 8's	Glides	Landings	[illegible]
Taxiing	.102	.081	.035	-.221	.075	-.141	-.079	-.057	.146	.145	.024	-.046	.142	[illegible]
Take-offs		.454	.121	.152	.088	.165	.025	-.069	-.026	.265	.091	.189	.182	[illegible]
Climbs			.306	.249	.083	.052	.174	.058	.182	.377	.194	.380	.297	[illegible]
Turns in Cl-L				.546	.192	.132	.189	.336	.176	.119	.382	.142	.190	-[illegible]
Turns in Cl-R					.116	.214	.129	.367	.231	.150	.337	.232	.043	[illegible]
Pow. Turns-R						.184	.150	.169	.341	.293	.303	.074	.320	[illegible]
Stall-off							.145	.170	.106	.213	.136	-.028	.159	[illegible]
Slips - L								.102	.065	.135	.185	.231	.178	.149
T in Gl-L									.358	.061	.368	.280	.097	-[illegible]
Steep 8's										.212	.621	.164	.159	[illegible]
S and L											.096	.315	.347	[illegible]
Shallow 8's												.216	.224	[illegible]
Glides													.289	[illegible]
Landings														[illegible]
180° Appr.														

TABLE 17

FACTOR MATRICES* OF FLIGHT DATA BY MANEUVER
FOR SUMMER, 1941
(N = 69)

Maneuver	Centroid Matrix A	B	C	Rotated Matrix A	B	C	h^2
Taxiing	159	116	230	-014	-043	300	.092
Take-offs	426	330	052	-054	315	437	.293
Climbs	149	284	-397	-162	481	054	.261
Turns in Climbs L	402	266	-271	-012	522	183	.306
Power turns L	390	089	-429	132	571	000	.343
Power turns R	354	-031	085	214	123	269	.133
Stall-off	243	-153	244	259	-100	255	.142
Slips L	144	-032	-115	103	-017	155	.035
Turns in Glide L	334	133	040	064	204	292	.131
Steep 8's	490	-574	-075	746	129	033	.575
Straight & Level	239	334	247	-156	065	448	.229
Shallow 8's	413	-659	-290	777	218	-191	.688
Glides	340	089	116	106	134	329	.137
Landings	497	044	036	227	266	358	.250
180° Approach	407	-029	188	241	077	372	.202

*Decimal points omitted.

On the basis of the net regression coefficients obtained in computing the multiple correlations reported above, a smaller number of maneuvers was selected to be used in the actual construction of the factor scores. Those maneuvers having the highest regression weights, regardless of sign, were selected and each factor predicted by its selected group of maneuvers.

Right Power turns, Steep 8's, and Shallow 8's had the highest regression weights for Factor A. These three maneuvers, properly reweighted, correlated .843 with the theoretical Factor A as compared with .866 when all the maneuvers were used. Climbs, Climbing turns, left Power turns, and Shallow 8's correlated .766 with Factor B as compared with .794 for all 15 maneuvers. For Factor C, six maneuvers were used: Take-offs, Climbs, Steep 8's, Shallow 8's, Landings, and 180° Approaches. These correlated .789 with the theoretical Factor C as compared with .841 for all 15 maneuvers. The actual factor scores computed on the basis of the selected maneuvers, showed a high degree of independence r_{AB} = .16, r_{AC} = .02, r_{BC} = .06).

As in the factor analysis of the 1940 version,[22] the factors were difficult of interpretation and had little practical value. An item analysis of the 1941 version was unproductive in that it did not yield results consistent with any *a priori* meanings of the rotated factors. In addition, the inventory as a field instrument was required to provide separate measures for each maneuver. Factor scores were therefore not recommended for use in the field.

[22] See Page 22.

5. Summary of the 1941 Version of the OSFI. The chief advantages of the 1941 version of the OSFI lay in the reduction in number of items, in further refinement of the items, and in a more convenient format for field use.

Research with the 1941 version gave further evidence of the need for careful training in the proper use of the inventory. Research data showed that, with well trained observers, flight inventory scores were little affected by sources of variance such as differences in Schools, Observers, Flight Number, Wind Velocity, or Familiarity with Plane. A factor analysis of maneuver score intercorrelations yielded three factors which, however, had little practical value for field scoring purposes. An item analysis of the inventory, designed to find items associated with each of the three factors, gave little useful information.

In general, the experience with the 1941 version confirmed its value as a useful instrument for measuring flight performance and indicated the advisability of continued experimentation especially with respect to further simplification of the form and to improvement in the method of scoring.

CONCLUSION OF PART I

With the 1941 version of the Ohio State Flight Inventory, the Ohio State University project drew to a formal close, having devised and put to experimental test an inventory which showed considerable promise as a practical tool for the measurement of flight performance from direct observation during flight.

Additional improvements, however, resulted from subsequent use of the inventory in related projects in the research program of the Committee on Selection and Training of Aircraft Pilots. Part II of the report on the Ohio State Flight Inventory, to be issued separately, will describe these changes and the construction of the present version and its current applications in field research.

APPENDIX 1

FREQUENCY DISTRIBUTIONS FOR EACH ITEM OF THE 1940 VERSION

1. TAXIING

FREQUENCY DISTRIBUTIONS FOR EACH ITEM OF [illegible] VERSION

Items	Excellent				Above Average				Average				Below Average				Very Poor			
	Into wind	Cross wind	Down wind	Gusty air	Into wind	Cross wind	Down wind	Gusty air	Into wind	Cross wind	Down wind	Gusty air	Into wind	Cross wind	Down wind	Gusty air	Into wind	Cross wind	Down wind	Gusty air
Traffic Check					7	3	6	1	31	13	18		2							
Field					4	2	5	1	23	20	9	1		1						
Throttle) Speed					6	1	6		26	19	17	2	1	2	2					
Control) Coord.					6	1	6		24	20	21	2								
Elevator Control			1		5		3		26	17	30		1							
Rudder Control					5	1	3		30	22	27									
Aileron Control					6		3		33	23	25			1						

Note: These are total figures for all ratings and not itemized as to instructors.

2. TAKE-OFFS

Lateral Balance During Roll	A	B	C	D	E	F	Total
Constant	5	2	15	13	5	4	44
Not more than 5° r. or l. deviation		2			4	6	12
Not more than 10° r. or l. deviation					1	1	2
Not more than 15° r. or l. deviation							0
More than 10° r. or l. deviation							0
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

Lateral Balance on Take-Off	A	B	C	D	E	F	Total
Constant	3	3	12	13	7	3	41
Not more than 5° r. or l. deviation	2	1	0	3	6	1	13
Not more than 10° r. or l. deviation				1	4		5
Not more than 15° r. or l. deviation							0
More than 10° r. or l. deviation							0
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

Direction During Roll	A	B	C	D	E	F	Total
Constant	3	3	3	11	12	5	37
Not more than 5° r. or l. deviation	2	1	2	2	6	6	19
Not more than 10° r. or l. deviation			1				1
Not more than 15° r. or l. deviation					1		1
More than 10° r. or l. deviation							0
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

Direction on Take-Off	A	B	C	D	E	F	Total
Constant	2	3	8	13	4	3	33
Not more than 5° r. or l. deviation	3	1	5	1	6	7	23
Not more than 10° r. or l. deviation					1		1
Not more than 15° r. or l. deviation						1	1
More than 10° r. or l. deviation							0
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

2. TAKE-OFFS (Contd.)

Throttle	A	B	C	D	E	F	Total
Smooth and deliberate throttle control	5	5	13	13	6	4	46
Full open throttle after starting to roll		1		1	2	4	8
Full open throttle to start roll					3	3	6
Abrupt on throttle control					2		2
Eases throttle to cruising speed on completion of take-off					1		1
Adjusts throttle to cruising speed too soon or too late on completion of take-off						1	1
Uses only part throttle in take-off attempt						1	1

Tail	A	B	C	D	E	F	Total
Eases tail to optimum angle	5	3	14	13	5	8	48
Abrupt on raising tail				1	4	3	8
Fails to lift tail							0
Tail held high, prolonged run		1			1		2
Tail low, tends to stall on take-off		1			1	1	3

Speed	A	B	C	D	E	F	Total
Levels off after take-off to gain speed	5	4	13	10	6	5	43
Fails to level off to gain speed		1	1	4	2	3	11
After gaining speed assumes optimum angle of climb	2				3	2	7

Climb	A	B	C	D	E	F	Total
Assumes optimum angle of climb	3	2	13	13	3	1	35
Assumes + or − 5° opt. angle of climb	1	3		2	8	6	22
Assumes + or − 10° opt. angle of climb	1					2	3
Climbs more than 10° steeper than optimum							0

Traffic	A	B	C	D	E	F	Total
Infrequent traffic check			8		6		14
Does not check traffic							0
Uses all available field	4	4	6	13	5	11	43
Does not use all available field		1		1	1		3

3. CLIMBS

Lateral Balance

	Instructor						
Lateral Balance	A	B	C	D	E	F	Total
Constant	4	1	13	14	5	3	40
Not more than 5° r. or l. deviation		2			5	8	15
Not more than 10° r. or l. deviation						1	1
More than 10° r. or l. deviation						1	1
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

Direction

Direction	A	B	C	D	E	F	Total
Constant	2		8	13	4	1	28
Not more than 5° r. or l. deviation		1	2	1	5	7	16
Not more than 10° r. or l. deviation						2	2
More than 10° r. or l. deviation						1	1
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

Angle of Climb

Angle of Climb	A	B	C	D	E	F	Total
Sets plane in normal climb	2	2	13	13	6	2	38
Climbs at optimum angle + or - 5°	2	2	1	1	4	4	14
Climbs at optimum angle or -10°	1					3	4
Climbs at optimum angle + or - 10°		1			1	1	3
Varies angle of climb up to + or -5° from optimum							0
Stalls ship						1	1
Mushes on climb							0
Varies angle of climb up to + or -10° from optimum							0

Throttle Control

Throttle Control	A	B	C	D	E	F	Total
Smooth and deliberate throttle control	3	4	12	2	3	3	27
Abrupt on throttle control				2			2
Disregards throttle	2				1	3	6
Coordinates throttle with attitude of plane		2	3	5	6		21
Slow on throttle adjustment to attitude of plane		1			1	5	7

TURNS IN CLIMBS

Entry	Instructor A	B	C	D	E	F	Total
Simultaneous application of controls	4	3	11	14	6	5	43
Successive application of controls	1	2	3		3	6	15
Uses aileron, no rudder, on entry					1		1
Uses rudder, no aileron, on entry					1		1
Keeps nose up during entry							0
Nose wanders during entry							0
Fails to keep nose up during entry							0

Bank	A	B	C	D	E	F	Total
Accurate angle of bank	3	1	11	14	6	2	37
Bank varies not more than + or - 5°	2	2	3		2	4	13
Bank varies not more than + or - 10°		1				3	4
Bank varies more than + or - 10°		1				2	3
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0
No slipping					1		1
Slips							0
Banks more than 30°							0
Banks less than 10°							0

Turn	A	B	C	D	E	F	Total
Accurate rate of turn	5	2	11	14	6	3	41
Varies rate of turn			2		5	6	13
Fluctuation in rate of turn						1	1
Rides same rudder during turn							0
Rides opposite rudder during turn							0
No skidding		2					2
Skids							0

Climb	A	B	C	D	E	F	Total
Holds accurate rate of climb in turn	3	1	10	13	6	3	36
Climbs at opt. angle of climb + or - 5°	1	4	4	1	2	4	16
Climbs at opt. angle of climb + or - 10°	1				3	4	8
Varies angle of climb up to + or - 5° from optimum							0
Varies angle of climb up to + or - 10° from optimum							0
Climbs too steeply for bank and power							0
Holds off bank with ailerons and uses rudder, little turn in climb							0
Cross-controls, slip counteracted by climb, no gain in altitude							0
Stalls plane							0

4. TURNS IN CLIMB (Contd.)

Recovery	A	B	C	D	E	F	Total
Simultaneous application of controls	4	5	10	14	9	1	43
Successive application of controls	1		5			9	[illegible]
Uses aileron, no rudder, on recovery			1				1
Uses rudder, no aileron, on recovery							0
Keeps nose up during recovery							0
Nose wanders during recovery					1		1
Fails to keep nose up during recovery							0

Precision	A	B	C	D	E	F	Total
Recovers from turn "on course"	4	1	11	11	7	1	35
Recovers from turn 5° r. or l. of "on course"		2	2	2	1	7	14
Recovers from turn not more than 10° r. or l. of "on course"		2			1	2	5
Recovery from turn more than 10° r. or of "no course"	1				1	1	3
R. and l. turns equal in precision							0
R. and l. turns not equal in precision							0

Throttle	A	B	C	D	E	F	Total
Throttle control or reduction in climb angle for turn		2	2		4	5	13
Coordinates throttle with attitude of plane	3	2	12	14	6	3	40
Slow on throttle adjustment to attitude of plane	1	1				1	3
Abrupt on throttle control							0
Disregards throttle	3				3	4	[illegible]

5. STRAIGHT AND LEVEL

Instructor

Lateral Balance	A	B	C	D	E	F	Total
Constant	4	3	13	13	7		40
5° deviation		3	1	1	4	10	19
Not more than 10° deviation	1						1
More than 10° deviation							0
Continuous fluctuation or deviation, less than 10°							0
Continuous fluctuation or deviation, greater than 10°							0

Direction	A	B	C	D	E	F	Total
Constant	2	3	12	14	7	2	40
5° deviation	1		3		4	9	17
Not more than 10° deviation							0
More than 10° deviation							0
Continuous fluctuation or deviation, less than 10°							0
Continuous fluctuation or deviation, greater than 10°							0

Altitude	A	B	C	D	E	F	Total
Constant	1	1	13	11	7		33
5° deviation	2	2	1	3	4	9	21
Not more than 10° deviation	1					2	3
More than 10° deviation							0
Continuous fluctuation or deviation, less than 10°							0
Continuous fluctuation or deviation, greater than 10°							0

Drift	A	B	C	D	E	F	Total
Immediate recognition and correction for drift	2	2	9	6	4		23
Recognizes and corrects for drift	2	1	4	6	4	8	25
Recognizes but fails to correct for drift	1		1	2	2	1	7
Fails to recognize or correct for drift					1		1

5. [illegible] TURNS

Instructor

Entry	A	B	C	D	E	F	Total
Simultaneous application of controls	4	3	9	13	9	5	43
Successive application of controls	1	2	1	1	1	4	10
Uses aileron, no rudder, on entry			2			1	3
Uses rudder, no aileron, on entry							0

Bank	A	B	C	D	E	F	Total
Accurate angle of bank	4		7	11	6	1	29
Bank varies not more than + or - 5°	1	5	5	2	3	2	18
Bank varies not more than + or - 10°			1		1	4	6
Bank varies more than + or - 10°						3	3
Contiguous fluctuation of less than 10° in bank					1		1
Continuous fluctuation of more than 10° in bank							0
No slipping					1		1
Slips						2	2
Banks more than 45°			1				1
Banks less than 10°							0

Turn	A	B	C	D	E	F	Total
Constant rate of turn	5	4	12	14	9	2	46
Varies rate of turn			2		2	7	11
Fluctuation in rate of turn						1	1
Rides same rudder during turn							0
Rides opposite rudder during turn			1				1
No skidding					3		3
Skids						1	1

Altitude	A	B	C	D	E	F	Total
No loss or gain of altitude	2	1	10	9	4	2	28
Changes altitude not more than 50 ft.	3	2	4	4	6	2	21
Changes altitude not more than 100 ft.		1		1		6	8
Changes altitude over 100 ft.						2	2

Recovery	A	B	C	D	E	F	Total
Simultaneous application of controls	5	5	10	14	10	7	51
Successive application of controls			3		1	4	8
Uses aileron, no rudder, on recovery			1				1
Uses rudder, no aileron, on recovery							0

5. MEDIUM TURNS (Contd.)

Precision	A	B	C	D	E	F	Total
	Instructor						
Recovers from turn "on course"	4		12	13	8		37
Recovers from turn 5° r. or l. of "on course"	1	2	2	1	1	7	14
Recovers from turn not more than 10° r. or l. of "on course"		3			2	3	8
Recovers from turn more than 10° r. or l. of "on course"						1	1
R. and l. turns equal in precision							0
R. and l. turns not equal in precision							0

7. GLIDES

Lateral Balance	A	B	C	D	E	F	Total
	Instructor						
Constant	2	1	10	14	8		35
Not more than 5° r. or l. deviation	2	4	3		2	10	21
Not more than 10° r. or l. deviation							0
More than 10° r. or l. deviation							0
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

Direction	A	B	C	D	E	F	Total
Constant	4	1	[illegible]	14	8		35
Not more than 5° r. or l. deviation	1		3		1	8	13
Not more than 10° r. or l. deviation						1	1
More than 10° r. or l. deviation							0
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

Glides	A	B	C	D	E	F	Total
Coasts to dissipate extra speed then eases nose down to normal glide			[illegible]		4	5	11
Eases plane in glide	4	5	11	12	6	2	40
Sets plane in glide					2	1	3
Glides at normal angle + or - 5°			4				[illegible]
Glides at normal angle or - 10°						2	[illegible]
Glides at normal angle + or - 10°							0
Dives on glides							0
Mushes on glides							0
Stalls							0

Throttle	A	B	C	D	E	F	Total
Smooth and deliberate throttle control	5	5	14	11	5	4	44
Abrupt on throttle control				1	6	1	8
Clears engine on glides and maintains proper glide angle				1		1	2
Clears engine on glides			6	2		1	9
Fails to clear engine on glides			1	4	1	3	9
Disregards throttle						2	2

Instructor

Entry	A	B	C	D	E	F	Total
Simultaneous application of controls	6	2	7	13	10	3	41
Successive application of controls			3			6	9
Uses aileron, no rudder, on entry			2				2
Uses rudder, no aileron, on entry			1				1
Nose wanders during entry						1	1

Bank	A	B	C	D	E	F	Total
Accurate angle of bank	3	2	10	14	5	2	36
Bank varies not more than + or - 5°	1	3	5		3	4	16
Bank varies not more than + or - 10°	1				2	3	6
Bank varies more than + or - 10°						1	1
Continuous fluctuation of less than 10° in bank							0
Continuous fluctuation of more than 10° in bank							0
No slipping					2		2
Slips							0
Banks more than 50°							0
Banks less than 10°							0

Turn	A	B	C	D	E	F	Total
Constant rate of turn	4	4	10	14	6	4	42
Varies rate of turn			3		5	3	11
Fluctuation in rate of turn						2	2
Rides same rudder during turn (mild)							0
Rides same rudder during turn (strong)							0
Rides opposite rudder during turn (mild)							0
Rides opposite rudder during turn (strong)							0
No skidding		1			1	1	3
Skids						1	1

Glide	A	B	C	D	E	F	Total
Glides at normal angle	5	2	12	14	7	4	44
Increases glide angle for turns		1			4	1	6
Holds constant rate of glide in turn		1			1	3	5
Fails to increase glide angle for turn							0
Glides at normal angle + or - 5°		1			1	1	3
Glides at normal angle + or - 10°							0
Glides too shallow, mushes or dives							0
Glides too steep							0
Varies angle of glide up to + or - 5° from optimum						2	2
Varies angle of glide up to + or - 10° from optimum							0

8. TURNS IN GLIDES (Contd.)

Recovery	Instructor A	B	C	D	E	F	Total
Simultaneous application of controls	4	5	9	14	11	5	48
Successive application of controls	1		4			3	8
Uses aileron, no rudder, on recovery			1				1
Uses rudder, no aileron, on recovery							0
Nose does not "bob" on recovery						1	1
Nose 'bobs" up on recovery						1	1

Throttle	A	B	C	D	E	F	Total
Clears engine on glides and maintains proper glide angle		1	11				12
Clears engine on glides			1	4	4		9
Fails to clear engine			2	5	7	6	20
Coordinates throttle with attitude of plane	1	2		3			6
Slow on throttle adjustment to attitude of plane		2				2	4
Abrupt on throttle control				1			1
Disregards throttle	3					3	6

Precision	A	B	C	D	E	F	Total
Recovers from turn "on course"	5		13	14	8		40
Recovers from turn 5° r. or l. of "on course"		3			1	5	9
Recovers from turn not more than 10° r. or l. of "on course"		1			1	2	4
Recovers from turn more than 10° r. or l. of "on course"		1					1
R. and l. turns equal in precision							0
R. and l. turns not equal in precision							0

[illegible]. LANDING (Final Approach)

Lateral Balance During Glide	A	B	C	D	E	F	Total
	Instructor						
Constant	3	3	12	13	5	1	37
Not more than 5° r. or l. deviation	2	2	2	1	4	10	21
Not more than 10° r. or l. deviation					2		2
More than 10° r. or l. deviation							0
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

Lateral Balance During Roll	A	B	C	D	E	F	Total
Constant	5	4	11	14	6	2	42
Not more than 5° r. or l. deviation		1	3		3	8	15
Not more than 10° r. or l. deviation						1	1
More than 10° r. or l. deviation							0
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

Direction During Glide	A	B	C	D	E	F	Total
Constant	5	5	11	13	6	4	44
Not more than 5° r. or l. deviation			3	1	5	7	16
Not more than 10° r. or l. deviation						1	1
More than 10° r. or l. deviation							0
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

To Wind During Roll	A	B	C	D	E	F	Total
Constant	5	5	10	14	6	4	44
Not more than 5° r. or l. deviation			4		5	6	15
Not more than 10° r. or l. deviation						1	1
More than 10° r. or l. deviation							0
Continuous fluctuation of less than 10°							0
Continuous fluctuation of more than 10°							0

9. LANDINGS (Final Approach) (Contd.)

Level off	A	B	C	D	E	F	Total
Starts to level off proper distance from the ground	4	5	14	12	7	7	49
Starts to level off 5 ft. above proper distance					4	3	7
Starts to level off 5 ft. below proper distance	1						1
Levels off more than 5 ft. above proper distance							0
Levels off on the ground				2		1	3
Pushes stick forward to resume glide within 10 ft. of the ground							0

Stalling	A	B	C	D	E	F	Total
Slow, continuous action in stalling plane for landing	5	4	11	13	5	4	42
Periodic application of controls in stalling plane			3		3	1	7
Jerky in stalling plane				1			1
Has stick all the way back just as plane touches ground	1	1			1	2	5
Has stick all the way back just before plane touches ground				1		2	3
Slow on getting stick back for landing					1		1
Fails to hold stick all the way back							0
Fails to get stick back for landing				1		1	2
Holds stick back until plane stops							0
Does not hold stick back till plane stops							0

Landing	A	B	C	D	E	F	Total
Three-point landing	5	5	7	8	4	9	38
Stalls plane 3 ft. off ground, "pancakes"					3	1	4
Stalls plane 5 ft. off the ground							0
Stalls plane 10 ft. off the ground				1			1
Wheel landing			5	5	4		14
No drifting or skidding on landing			1				1
Slight drift or skid on landing			1			2	3
Drifts or skids on landing							0
Recognizes bad landing			2				2
Pulls out of bad landing and makes recovery			2	1			3
Fails to recognize the need of pulling out of a bad landing							0
Fails to pull out of bad landing				1			1
Recognizes the need but fails to pull out of bad landing							0

9. LANDINGS (Final Approach) (Contd.)

Precision	Instructor A	B	C	D	E	F	Total
Lands on spot		1			2	2	5
Lands on spot or within 100 ft. beyond		2					2
Lands on spot or within 200 ft. beyond							0
Overshoots spot landing more than 500 ft.		1					1
Undershoots spot landing less than 50 ft.		1					1
Overshoots spot by more than 300 ft.					1		1
Undershoots spot by more than 50 ft.						1	1
Overshoots field				1		2	3
Undershoots field				1	1	2	4

Throttle	A	B	C	D	E	F	Total
Keeps hand on throttle	5	5	13	14	10	5	52
Fails to keep hand on throttle			1			5	6

10. COORDINATI[illegible] AND T[illegible]SION

Instructor

Coordination	A	B	C	D	E	F	Total
Coordinates controls by pressures, "Feels"	1	3	3	1	7		15
Smooth and coordinated	4	2	6	10	2	4	28
Coordinated, not so smooth			4	3	2	6	15
No coordination						1	1
Mechanical			2				2
Over-controls							0
Jerky							0
Cross controls							0

Tension	A	B	C	D	E	F	Total
Relaxed, and flies plane without bein[illegible] aware of handling controls "witho[illegible] thinking of movements"	3	3	8		3		17
Relaxed, and observes conditions outside of plane	2	1	3	14	1	7	28
Tense, observes conditions outside o[illegible] plane		1	3		2	1	7
Relaxed, centers attention on plane						1	1
Tense					2	1	3
Tense, centers all attention on plan[illegible]					1		1
Leans against banks							0
Freezes on controls							0

Judgement	A	B	C	D	E	F	Total
Blocks wheels on starting engine					2		2
Fails to block wheels on starting en[illegible]							0
Uses brakes on starting engine	[illegible]	2	13	14	7	1	42
Fails to use brakes on starting engi[illegible]			1		1		2
Careful on checking throttle and swi[illegible]	[illegible]	1	1	3	2	9	[illegible]
Careless on checking throttle and switch							0
Cautious around propeller							0
Careless around propeller							0
Asks for help when taxiing in stro[illegible] wind or gusty air							0
Does not ask for help when taxiing i[illegible] strong wind or gusty air					2		2
Cautious of other traffic when taxi[illegible]	[illegible]	3	14	13	9	4	46
Careless of other traffic when taxi[illegible]							0
Blind to other traffic when taxiing							0
Prompt observance of field traffic r[illegible]	[illegible]	4	12	13	11	8	[illegible]
Fails to observe field traffic rul[illegible]			1				1

[illegible]

Judgment (continued)	A	B	C	D	E	F	Total
Uses all available field for take-off	2	3	14	13	9	9	50
Uses two-thirds of available field for take-off	3	1		1	1	1	7
Uses less than two-thirds of available field for take-off		1			1		2
Observes other traffic when in the air	5	3	8	12	4	3	35
Looks for other traffic when in the air		2	6	2	8	5	23
Fails to look for other traffic when in the air					2	1	3
Lands in first third of field	5	3	12	9	5	5	39
Lands in middle of field			1	2	4	1	8
Lands in last third of field					2	2	4

11. STALLS-ON AND STALLS-OFF

Entry

Pulls nose up slowly and steadily
Pulls nose up suddenly
Gets nose too high
Fails to stall plane, mushes

Stall

Holds wings level
Fails to hold wings level
Uses rudder only to hold direction and level
Uses aileron in attempt to hold wings level
Holds direction
Loses direction
Recognizes approach to stall point
Fails to recognize approach to stall point, completely stalls plane

Recovery

Recovers smoothly
Dives in recovery
Brings plane up slowly and smoothly to level flight
Brings plane up suddenly
Brings plane up into a second stall
On power-off stall opens throttle when plane returns to level flight
On power-off stall fails to open throttle when plane returns to level flight

Note: Frequency tabulations for this maneuver were not obtained since it was not included in the standard flight used in the collection of data on the 1940 Version.

12. SLIPS (Forward)

Entry

Holds constant flight path while lowering wing
Plane turns from flight path less than 10°
Plane turns from flight path more than 10°
Coordinates rudder with aileron and elevator
Holds nose down but slightly higher than in glide
Does not gain speed
Gains speed
Does not stall or mush
Tends to stall plane

Slips

Holds direction constant
Yaws less than 10°
Yaws more than 10°
Direction fluctuates
Holds constant speed
Gains speed
Gains excessive speed

Recovery

Holds direction constant
Changes direction less than 10°
Changes direction more than 10°
Maintains constant speed
Drops nose and gains speed
Drops nose and gains excessive speed
Pulls nose high and stalls

Note: Frequency tabulations for this maneuver were not obtained since it was not included in the standard flight used in the collection of data on the 1940 Version.

13. SPINS

Entry

Dissipates speed before entering stall
Fails to dissipate speed before entering stall
Stalls plane slowly and steadily
Stalls plane suddenly
Fails to stall plane
Gets nose too high
Recognizes stalling point
Fails to recognize stalling point
Applies rudder before stalling plane
Applies rudder after stalling plane
Applies rudder at stalling point
Holds ailerons neutral
Fails to hold ailerons neutral
Gets stick all the way back
Blasts engine to enter spin

Spin

Holds full rudder
Fails to hold full rudder
Holds stick all the way back
Fails to hold stick all the way back
Holds ailerons neutral
Fails to hold ailerons neutral

Recovery

Neutralizes rudder
Fails to neutralize rudder
Uses opposite rudder
Eases stick to neutral
Puts plane into dive
Recovers at optimum speed
Recovers at 10 M.P.H. above optimum speed
Recovers with excessive speed
Recovers "on course"
Recovers + or - 10° "on course"
Recovers + or - 20° "on course"
Recovers + or - 90° "on course"
Recovers more than + or - 90° of "on course"

Note: Frequency tabulations for this maneuver were not obtained since it was not included in the standard flight used in the collection of data on the 1940 Version.