RESTRICTED Not for Publication

AN ANALYSIS OF PHOTOGRAPHIC RECORDS OF AIRCRAFT PILOT PERFORMANCE

SECTION A: A Study of Criteria of Pilot Proficiency Derived from Motion Photographs of Flight Performance

by

Morris S. Viteles

and

Albert S. Thompson

SECTION B: Supplemental Report - An Analysis of Scores on Aspects of Flight Performance

by

Albert S. Thompson

A report on research conducted at the University of Pennsylvania, Philadelphia, Pennsylvania, on records collected in the Midwest Project at Ohio State University, Columbus, Ohio, by means of a grant-in-aid from the National Research Council Committee on Selection and Training of Aircraft Pilots from funds provided by the Civil Aeronautics Administration.

July 1944

CIVIL AERONAUTICS ADMINISTRATION
Division of Research
Report No. 31
Washington, D. C.

9-2-48

National Research Council

Committee on Selection and Training of Aircraft Pilots

Executive Subcommittee

M. S. Viteles, Chairman

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

National Research Council

1944

LETTER OF TRANSMITTAL

NATIONAL RESEARCH COUNCIL

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

Committee on Selection and Training of Aircraft Pilots

July 21, 1944

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

Dear Dr. Brimhall:

Attached is a report entitled An Analysis of Photographic Records of Aircraft Pilot Performance by Morris S. Viteles and Albert S. Thompson. It 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.

This is one of a series of reports concerned with the problem of obtaining adequate criteria of proficiency through the use of objective recording of flight performance. The report should be reviewed in relation to the material on standard flights and motion photography included in Technical Report No. 15. It is also of interest in relation to the investigation of graphic methods for recording pilot performance discussed in Technical Report No. 25.

Cordially yours,

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

MSV:cv

EDITORIAL FOREWORD

One of the first problems attacked in formulating the research program of the Committee on Selection and Training of Aircraft Pilots involved the evaluation of available criteria of success in flight performance. Early investigations were made of the criteria then in use, including instructor ratings, inspectors' flight test grades, etc. Concurrent with these investigations were studies devoted to the development and evaluation of methods of objective recording of performance during flight designed to provide more adequate and reliable criteria. These studies involved the use of two types of objective records: graphic and photographic.

For the investigation of the former, commercial graphic recorders were available. Research involving their use resulted in an evaluation of the capabilities and deficiencies of commercial recorders[1] and led to the development of a new graphic recorder better designed to serve the research objectives of the Committee.[2]

Research on the photographic method of recording required the construction and development of photographic installations which would record flight and control movement data descriptive of pilot performance. Projects at both the University of Rochester and the University of Pennsylvania initially attacked the problem of the photographic method. While the Rochester project immediately turned towards the development of a concealed photographic unit including an instrument panel and a control movement recorder, the Pennsylvania group, in its first study, directly photographed the plane instrument panel and the actual handling of the controls by the pilot. An exploratory study using this method and yielding quantitative indices of control movements has previously been reported in a study by Viteles and Thompson.[3]

---

[1] Viteles, Morris S. and Backstrom, Oscar, Jr. An analysis of graphic records of pilot performance obtained by means of the R-S Ride Recorder, Part I. Washington, D. C.: Civil Aeronautics Administration Division of Research. Report No. 23. November 1943.

[2] McKay, Walter. The development of the C.A.A.-N.R.C. flight recorder. (Copy in Committee Files)

[3] Viteles, Morris S. and Thompson, A. S. Use of standard flights and motion photography in the analysis of aircraft pilot performance. Washington, D. C.: Civil Aeronautics Administration Division of Research. Report No. 15. May 1943.

This study indicated clearly the desirability of using a control movement recorder, mountable on the plane instrument panel, for obtaining records of pilots' reactions in controlling the plane. The recorder finally developed by the University of Pennsylvania research staff represents a modification of one originally designed by Dr. Brian O'Brien at the University of Rochester in cooperation with Dr. Dean R. Brimhall, Director of Research, Civil Aeronautics Administration. Although the functions of the two recorders are similar, the Pennsylvania control movement recorder differs from the Rochester unit in that the scales are linear and a separate scale is provided for each of the controls. In addition, the Pennsylvania Control Movement Recorder provides for more extended movement of the pointers and easily observed reference lines, facilitating readings of the type desired by the project.[4]

Presented in this report are descriptions of the photographic procedures, methods of analysis, and the results of the application of the methods of recordings and analysis to a group of student pilots acting as subjects in the 1942 Midwest Project. This report may be looked upon as a companion report to "An analysis of graphic records of pilot performance obtained by means of the R-S Ride Recorder", Part I, by Morris S. Viteles and Oscar Backstrom, Jr., which described analogous methods for the analysis of graphic records of performance during flight. Together, they represent contributions of the University of Pennsylvania Project to the Committee research program in the general area of objective recording of flight performance and the development of criteria of flight success.

---

[4]In more recent research, the University of Rochester pattern of a concealed unit has been followed by Dr. R. Y. Walker, with the aid of the Pennsylvania research staff, in adapting the unit for use in the 1943 Midwest-Navy research project.

CONTENTS

CONTENTS (continued)

SUMMARY

This second report of the University of Pennsylvania Project on the use of the photographic method of analysis of pilot performance presents (1) a description of an improved photographic installation for use in photographing flight performance, (2) a description of procedures employed in the analysis of photographic records, and (3) a discussion of measures of flight performance yielded by the analysis, with data as to reliability, interrelationships, performance at two stages of instruction, and other facts bearing upon the usefulness of motion photography in providing criteria of flight performance.

1. The photographic installation provided 16 mm. motion photographs of a special instrument panel mounted in a Piper Cub, the camera field including flight instruments and the Pennsylvania Control Movement Recorder which indicates the position of the rudder, ailerons, and elevator. The films thus yielded data as to the performance of the plane and the performance of the pilot during flight.

2. Student pilots from the 1942 Midwest Project were photographed while performing certain critical maneuvers in Standard Flights developed for use at each of the four stages of C.P.T. primary flight instruction. Since the investigation included a comparison of early and final performance, the records of Flight A and Flight D were selected for analysis. The available records included Flight D performance of 36 student pilots of the 1942 Spring Program, Flight D performance of 33 of the 1942 Summer Program, and Flight A performance of the 34 pilots from both Spring and Summer Programs on whom Flight D records were also obtainable.

3. The analysis of the records took the form of a direct inspection of the films during slow motion projection and yielded qualitative ratings on the performance as a whole and on eight aspects of flight performance - wing control, nose control, directional control, slip-skid tendency, maintenance of altitude, maintenance of air speed, coordination of the controls, and adequacy of execution of the maneuver.

4. The analysis was made by two observers during a rating and re-rating procedure which provided Independent Ratings by each of two observers, Composite Ratings based on a joint observation of discrepant cases, and a final Criterion Rating resulting from the joint comparison of the two sets of Composite Ratings provided by the rating and re-rating procedures. Besides the over-all ratings, flight scores were derived by summing the separate ratings on the eight aspects of flight performance in order to obtain Independent Flight Scores, Composite Flight Scores, and Criterion Flight Scores.

From the results of the investigation, a study was made of the over-all rating and the flight scores in order to determine (1) their reliability, (2) their interrelationship, (3) the predictive value of Flight A performance, (4) the sources of variance affecting the Flight Scores, and (5) possible use of the measures as criterion data.

1. The reliability was measured by determining the extent of agreement between the various sets of measures on the same subjects. It was found that the Composite Ratings were more reliable than the Independent Ratings and that the Criterion Ratings may be considered sufficiently reliable for use as criterion data. Similar results were obtained for the flight scores, the Criterion Flight Scores exhibiting Spearman-Brown reliability coefficients of .87 and .91 on the basis of Flights A and D, respectively, of a group of 33 subjects.

2. Comparisons between corresponding sets of Criterion Ratings and Criterion Flight Scores indicate that the two types of measures of flight performance are highly related and that for a grouping of subjects into three categories both types of criterion data give essentially the same results. The Criterion Ratings have the advantage of representing the combined and inter-acting judgment of two observers while the Criterion Flight Scores possess certain practical advantages.

3. Comparison of Flight A and Flight D performance of a group of 33 subjects revealed that the two performances are significantly related. The prediction of Flight D performance from Flight A performance, as measured by the photographic method, may be represented by an r of approximately .45, assuming that the relationship maintains on additional samples.

4. An analysis of variance of the Flight Scores revealed that differences between the two raters and between two ratings by the same raters were not significant. The major sources of variation lay in differences among pilots and in the fact that Flight A and Flight D performances were significantly but not highly correlated.

5. An evaluation of the Criterion Ratings and Criterion Flight Scores in terms of their reliability, validity, discriminating value, and practicality led to the conclusion that they may be recommended for use as one type of criterion data in aviation research in that they exhibit satisfactory reliability, provide measures of the "skill" aspect of piloting, and differentiate among levels of performance of "successful" student pilots. Practical limitations of the photographic method restrict its use as a source of criterion data to basic research where time and cost are relatively unimportant and accuracy of data on performance is of prime importance.

A supplemental study was made of the Aspect Scores based on the four ratings of each of the eight aspects of performance assigned during the independent rating and re-rating by each of the two observers. From an analysis of the scores on the 31 subjects on whom both Flight A and Flight B records were obtained it was concluded that:

1. The Wing Control Score, Nose Control Score, Directional Control Score, and Slip-Skid Score were more consistently reliable (for the two flights) than were the other scores -- Altitude Score, Airspeed Score, Control Coordination Score, and Execution of Maneuvers Score.

2. The eight aspects of performance, separately rated, are not independent and unitary. A factor analysis of the intercorrelations among the Aspect Scores indicated two factors, tentatively identified as "Coordination of the Controls" and "Longitudinal Control." Since the two factors were correlated, a third influence may be affecting the separate ratings, probably a "general impression" of the flight as a whole.

3. The Aspect Scores most closely associated with the criterion measures of over-all performance (Criterion Ratings and Criterion Flight Scores) are the Wing Control Score, Directional Control Score, Control Coordination Score, and Execution of Maneuvers Score.

4. Aspect Scores on Flight A performance (during Stage A of C.P.T. primary flight instruction) are not highly predictive of Flight B performance (during Stage B of C.P.T. primary flight instruction).

The small number of cases involved and the level of intercorrelation among the Aspect Scores prevent drawing definitive conclusions as to the value of the several Aspect Scores for diagnostic or predictive purposes. On the basis of the present information it seems most practicable to consider the Criterion Ratings and Criterion Flight Scores as possible criterion measures and the Aspect Scores as a source of information concerning the nature and characteristics of these measures of over-all flight performance.

AN ANALYSIS OF PHOTOGRAPHIC RECORDS OF AIRCRAFT PILOT PERFORMANCE

SECTION A
A STUDY OF CRITERIA OF PILOT PROFICIENCY DERIVED FROM MOTION PHOTOGRAPHS

INTRODUCTION

Research at the University of Pennsylvania, under a grant from the National Research Council Committee on Selection and Training of Aircraft Pilots, has been centered on the development and evaluation of objective criteria of success in learning to fly. In this connection, two methods of objective recording of flight performance have been studied:

1. The graphic method, involving the use of commercial flight recorders.

2. The photographic method, involving motion photography of instruments recording plane performance and control movements.

Results of research with graphic recorders have been presented in an earlier report by Viteles and Backstrom.[1] The early work with photographic records has been reported in another study by Viteles and Thompson.[2]

The present report presents (1) a description of an improved photographic installation for use in photographing flight performance, (2) a description of procedures employed in the analysis of photographic records, and (3) a discussion of the criterion measures of flight performance yielded by the analysis, with data as to reliability, interrelationships, performance at two stages of instruction, and other facts bearing upon the usefulness of motion photography in providing objective and acceptable criteria of flight performance.

DESCRIPTION OF THE PHOTOGRAPHIC INSTALLATION

In a previous study on the photographic method referred to in the Introduction,[3] use was made of a Filmo camera Model 141A mounted on a tripod behind the right front seat of a Stinson 105. The camera field

---

[1] Viteles, Morris S. and Backstrom, Oscar, Jr. An analysis of graphic records of pilot performance obtained by means of the R-S Ride Recorder, Part I Washington, D. C.: Civil Aeronautics Administration Division of Research, Report No. 23. November 1943.

[2] Viteles, Morris S. and Thompson, A. S. Use of standard flights and motion photography in the analysis of aircraft pilot performance. Washington, D. C.: Civil Aeronautics Administration Division of Research, Report No. 15. May 1943.

[3] Ibid.

included the instruments on the left side of the standard instrument board (air speed indicator, turn and bank indicator, and tachometer), the control wheel and throttle, and the upper legs, lower arms, and hands of the subject.

Examination of the resulting films gave only partial information as to plane and pilot performance since (1) there was no indication of rudder adjustments, (2) the throttle adjustments had to be inferred indirectly from tachometer readings, and (3) the data on plane performance were limited to those obtained from turn and bank indicator and air speed indicator readings. In addition, the adequacy of the original installation was limited by the fact that a 50' magazine was used, necessitating reloading "in the air" if a flight sufficiently long for criterion purposes was to be photographed.

In the study described in this report, use was made of an improved photographic installation making available fuller data for the analysis of plane performance. A specially designed control movement recorder was also provided which gave an indication of rudder as well as aileron and control movements. The equipment, installed in a C.A.A.-N.R.C. Tandem Piper Cub used specifically for research purposes,[4] included the following:

1. A special instrument panel, mounted over the left half of the standard panel so as to bring the Air Speed Indicator, Rate-of-Climb Indicator, Sensitive Altimeter, Tachometer, Gyro-Horizon, and Turn and Bank Indicator within the camera field.

2. A recorder, designated as the Pennsylvania Control Movement Recorder,[5] providing an indication of position and movements of each of the three.

---

[4]The photographic installation used in the Cub was originally developed for use in a Fairchild and actually installed in a Navy N3-N3 plane. The adaptations on the Cub were made at the Boston Municipal Airport, Boston, Mass., under the supervision of Charles H. Scott of E. A. Wiggins Airways, Inc., in accordance with plans outlined by the University of Pennsylvania Project Staff.

[5]The development of this recorder grew out of an opportunity afforded to the University of Pennsylvania Research Staff to examine a hemispheric recorder designed at the University of Rochester, by Dr. Brian O'Brien. The latter recorder was developed during the course of concurrent independent research on photographic recording conducted under a grant from the Committee on Selection and Training of Aircraft Pilots and with instruments provided, in part, by the Division of Research, Civil Aeronautics Administration. Although the purposes of the two types of recorders are similar, the Pennsylvania Control Movement Recorder differs from the Rochester unit in that the scales are linear and separate scales are provided for each of the controls. In addition, the Pennsylvania Control Movement Recorder provides for more extended movement of the pointers and easily observed reference lines, facilitating readings of the type desired by the project.

each of the three controls[6] (rudder, ailerons, and elevator) by means of three pointers moving along horizontal scales. (Exhibit 1 presents details as to its appearance and construction.)

3. A 16 mm. Bell and Howell Model 70-DA motion picture camera was mounted on a platform located in the upper left corner in the rear of the cockpit, behind and above the head of the rear seat pilot. An attached external magazine containing a 200' reel,[7] provided for approximately 30 minutes of photography at the rate of 8 frames per second. The size and location of the camera necessitated cutting a hole in the roof of the plane and the provision of a fitted cover. Access to the camera for loading of film was obtained by removing this outer cover.

The camera was driven by a 12-volt motor supplied by a 12-volt Exide aircraft battery located in the baggage compartment. A switch controlling the camera was placed so as to be convenient to the "check pilot" in the front seat. The switch also controlled a small spotlight which provided additional illumination of the camera field. The total illumination was sufficient to obtain satisfactory photographs with Eastman Super X film. This spotlight was mounted above the front seat and focused on the special instrument panel

Exhibit 2 shows the special instrument panel and Control Movement Recorder as installed in the research plane. The actual camera field taken by the motion picture camera was slightly larger and included the head and shoulders of the pilot administering the flight and a greater portion of the front seat throttle and control stick. In improvements made after the illustrated photograph was taken, a signal light (indicating that the check pilot had taken over) was mounted above the special instrument panel and the maneuver-number indicator was enlarged and moved to the right of the Control Movement Recorder.

OBJECTIVES OF THE STUDY

Among the various methods available for analysis of the photographic records are the following:

1. A quantitative method involving the examination of detailed readings of the instruments in the camera field in the course of frame by frame projection. From data obtained by this method indices descriptive of total flight performance and of individual maneuvers can be procured.[8]

---

[6]This recorder has subsequently been improved to include the position and movements of the throttle control.

[7]In other installations, such as the Fairchild or Navy N3-N3 plane, a 400' external magazine can be attached, thus providing film for approximately an hour's photography without reloading.

[8]Such as Total Amount of Control Movement, Number of Control Movements, Changes in Direction of Control Movements, etc., as described in Viteles, M.S.

2. A quantitative method involving the recording of observations of specific items of flight performance, such as frequency of slips in entering turns, loss of altitude during 360° steep turns, etc., observed in the course of slow motion projection of a film. Such observations, recorded on check sheets, provide additional quantitative data concerning important characteristics of flight performance.

3. A qualitative analysis of the performance through direct inspection during motion projection of the film from which judgments or estimates may be made on either the flight as a whole or on separate aspects of flight performance.

In this study the "qualitative" method was employed to arrive at judgments in the form of ratings of over-all or "global" pilot proficiency. The primary object of this study was to determine whether these qualitative estimates of over-all flight performance were sufficiently reliable to be used for criterion purposes. The study was therefore designed to provide data on the reliability of such qualitative estimates of performance, arrived at during slow motion projection of photographic records obtained by means of the installation described above. In addition, the experimental design of this study provided an opportunity to examine:

1. The sources of variance affecting the qualitative estimates of over-all flight proficiency.

2. The relationship between the over-all or "global" ratings and flight scores based on the evaluation of specified aspects of the total performance.

3. The relationship between ratings of flight performance at two stages of instruction.

DESCRIPTION OF PROCEDURES IN THE STUDY

A. Photographic Recording of Pilot Performance During Standard Flights:

1. Use of Standard Flights.[9] Student pilots, serving as subjects in the 1942 Midwest Project[10] of the Committee on Selection and Training of Aircraft Pilots, were photographed during standard flights.

---

[9] Standard Flights, developed by the University of Pennsylvania Project, insure controlled observation by requiring each pilot to perform the same maneuvers in the same order, and under essentially the same flight conditions. They provide a "work-sample" test analogous to performance trade tests as used in industry. A complete discussion of the principles underlying the use of Standard Flights may be found in: Standard Check Flight Procedures. Washington, D. C.: Department of Commerce, Civil Aeronautics Administration, Bulletin No. [illegible]

[10] The Midwest Project was set up in 1942 by the National Research Council Committee on Selection and Training of Aircraft Pilots in order to provide a field trial of techniques for the selection of aircraft pilots and for the evaluation of flight performance developed under the supervision of the Committee during the first two years of research. The [illegible] employed in the 1942 Midwest Project [illegible]
[illegible]

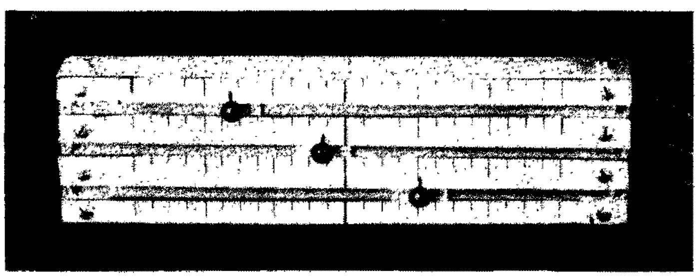
Exhibit 1

DESCRIPTION OF THE PENNSYLVANIA CONTROL MOVEMENT RECORDER

The Pennsylvania Control Movement Recorder is based upon a mechanical cable and pulley system in which three pointers are drawn along parallel linear scales by means of cables attached to moving portions of the airplane controls. The pointers are returned to their original positions during opposite control movements by tension springs located in the recorder. When installed, the cables are adjusted so that the pointers fall on the center scale lines when the control surfaces are streamlined.

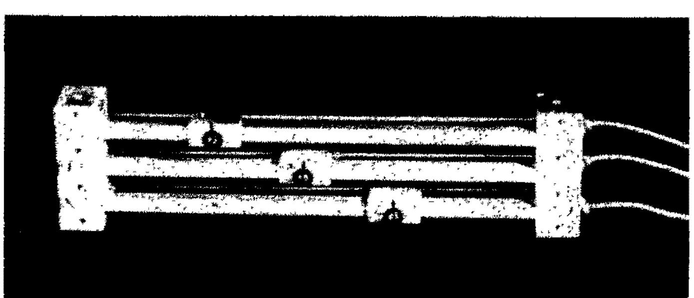
Models of this recorder have been installed in planes of various types, viz. a tandem Piper Cub, a Stinson 105, and a Navy basic training plane, Model N3N3. In the photographed model, shown below, the following specifications apply:

Over-all length:	8"
Over-all height:	$2\frac{1}{4}$"
Over-all depth:	1"
Carriage bars:	$\frac{1}{4}$" square and 8" long
Sliding carriages:	3/4" long and $\frac{1}{2}$" deep
Scales:	8" long, with reference lines $\frac{1}{4}$" apart
Springs:	.014" steel piano wire wound on 1/16 drill rod.



FRONT VIEW

Complete assembly, showing:
Scales, Pointers, and Cables



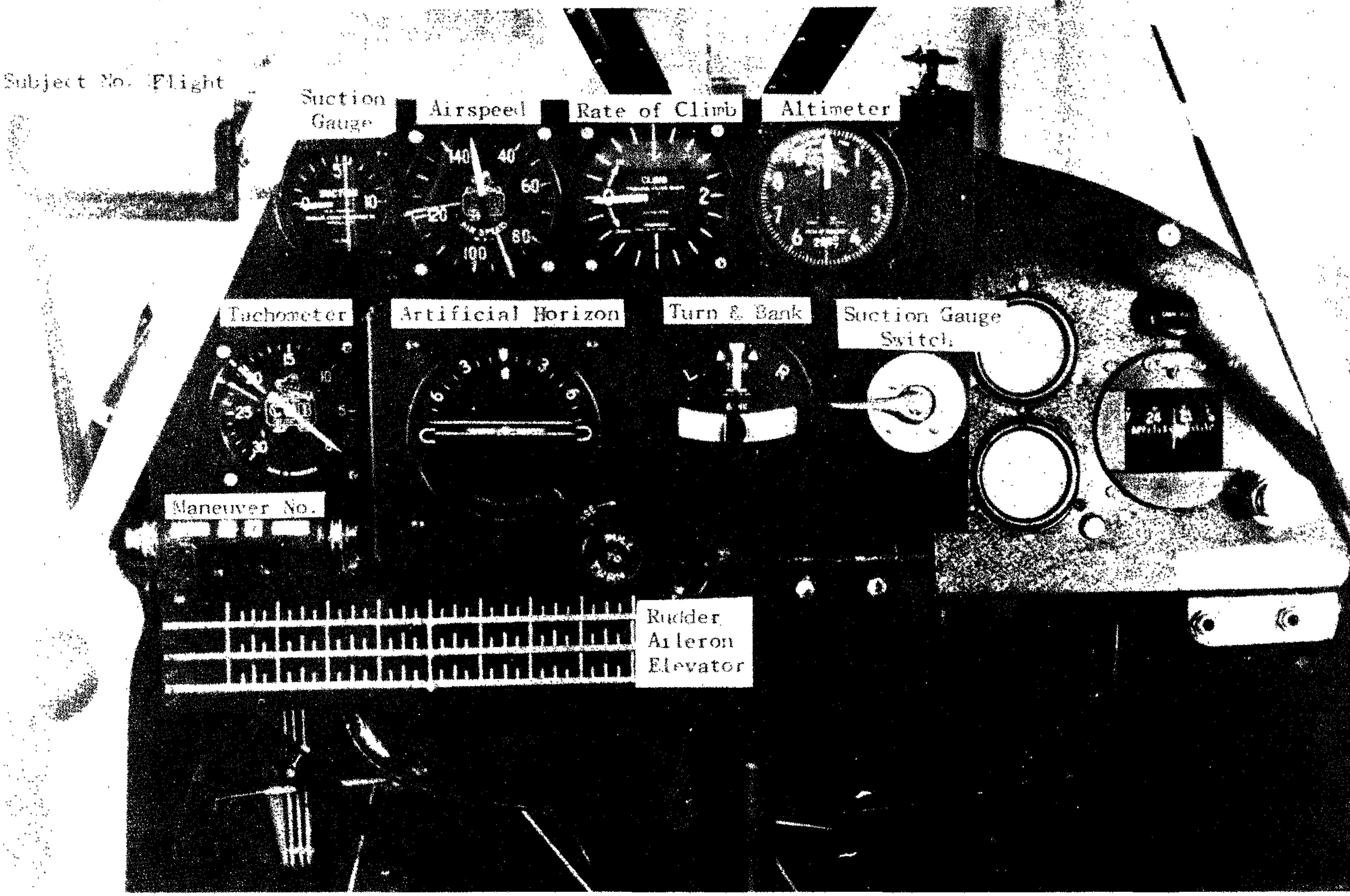
FRONT VIEW

Cover and Scales Removed, showing:
Sliding carriages with pointers attached
Bars on which carriages slide

REAR VIEW

Cover removed, showing:
Rear of carriages and carriage bars
Tension springs & pulleys
Cables leading from carriages to controls

CAMERA FIELD -- PIPER CUB INSTRUMENT PANEL



Actually, a series of four standard flights was used, one for each of the four successive stages of C.P.T. instruction.[11] The flights (Standard Flight A to D, respectively) are of increasing difficulty in that each includes the maneuvers of the preceding flights and requires, in addition, one or more advanced maneuvers appropriate to the stage of instruction. The fourth flight, Standard Flight D, is comparable to the final flight test for the course and includes TAXI, TAKE-OFF, STRAIGHT CLIMB, 90° CLIMBING TURN, STEEP 360° TURNS, NORMAL POWER-OFF STALL, FORWARD SLIP, STRAIGHT GLIDE, 90° LEFT and RIGHT MEDIUM TURNS, STRAIGHT CLIMB, STRAIGHT AND LEVEL FLIGHT, S-TURNS ACROSS A ROAD, RECTANGULAR COURSE, TWO-BANK EIGHT, 180° APPROACH, and PRECISION LANDING. The Standard Flights were spaced in the 35-hour C.P.T. Course as follows:

Hours of Primary Flight Instruction

Standard Flight	Between Minimum	Maximum	Within Stage
A	4 hrs.	6 hrs.	A (1st 8 hrs.)
B	10 hrs.	13 hrs.	B (8 to 13 hrs.)
C	21 hrs.	24 hrs.	C (13 to 24 hrs.)
D	32 hrs.	36 hrs.	D (24 to 36 hrs.)

2. Procedures Used in Photographing Standard Flights. Photographs were taken of all "critical maneuvers"[12] except those definitely oriented with respect to ground reference points, viz., S-Turns across Road, Rectangular Course, Eights, and Approach to Landing. These were omitted since the interpretation and evaluation of changes of plane attitude and control positions depend upon information as to the position and attitude of the plane with respect to specified points on the ground not provided by the photographic record.

Specific instructions for the flight as a whole and for the photography of each selected "critical maneuver" were prepared.[13] The camera was started just before entry into each "critical maneuver"

---

[11] A detailed description of each of the Standard Flights is given in Appendix A, Description of Standard Flights. Details of the procedures used in administering the flights are given in: Walker, R. Y., Lipman, E., and Wantman, M. J. Manual for the administration of the Ohio State Flight Inventory. Washington, D. C.: National Research Council, Dec. 20, 1941 (copy in Committee files).

[12] The "critical maneuvers" are those maneuvers which are being studied, as distinguished from the "transition maneuvers" whose purpose is merely to get the plane into position for the next critical maneuver.

[13] See Appendix B, General Instructions for Photographing Standard Flights.

and was stopped after the pilot recovered into straight and level flight. In the case of Approach to Landing, the camera was started as the plane recovered from the last turn prior to the Final Approach and was stopped after a few seconds of the Landing Run.

B. Photographic Records Available for Analysis. Two groups of student pilots were photographed,[14] one group of 56 subjects from the 1942 Spring C.P.T. Program, and one of 43 subjects from the 1942 Summer Program. The student pilots were those trained on Tandem Piper Cub planes in the Columbus, Ohio area. Due to weather conditions and occasional poor photography, motion photographs were not obtained of all four flights of all subjects.

Since the investigation as a whole included a comparison of initial and final performance, the records of Flight A and Flight D were selected for analysis. The available records included the following:

1. Flight D - 36 student pilots of the 1942 Spring Program.
2. Flight D - 33 student pilots of the 1942 Summer Program.
3. Flight A - 34 student pilots from the Spring and Summer Programs on whom Flight D records were also available.

Photographic records of flight performance taken at the Midwest Project were forwarded to the University of Pennsylvania for analysis by the research staff working under the direction of Dr. Morris S. Viteles. Each film was identified merely by the code number of the pilot being photographed. The research staff at Pennsylvania had no direct contact with the pilots under observation and no knowledge of their performances, except as gained from the photographs.

C. Method of Analysis of the Photographic Records. The analysis took the form of a direct inspection of the films during slow-motion projection,[15] so as to yield the following data on each performance:

1. A rating on a five-point scale (1, good to 5, poor) of each of the following aspects of flight performance:

   a. Wing control, e.g., steadiness of bank or level flight.
   b. Nose control, e.g., maintenance of nose on horizon or angle of glide.
   c. Directional control, i.e., maintenance of direction and rate of turn.
   d. Tendency to slip or skid, i.e., lateral stability.

---

[14] The flights were administered by or under the supervision of Dr. Robert Y. Walker, Project Director of the Midwest Project, in close cooperation with the University of Pennsylvania group.

[15] The films were projected by a Keystone Model A-31 Projector, specially adapted by means of a rheostat and reversible control which permitted slowing down and reversing direction at will. Although the films were taken at half normal speed (8 frames per second) use of this projector provided a slow-motion effect and enabled careful, close observation of the instruments in the camera field.

e. Maintenance of altitude.
f. Maintenance of airspeed.
g. Coordination of controls.
h. Adequacy of execution of maneuvers, e.g., correct degree of bank, complete stall, adequate airspeed, three-point landing, etc.

2. Brief notes on specific items of performance, such as "lost 120' altitude in 360° left, glided at 55 MPH, aileron in stall," etc. In addition, the air conditions were noted as Smooth, Normal, or Rough.

3. An over-all rating of flight performance on a three-point scale: A, best 25% of the group; B, middle 50%; and C, worst 25%.

4. An over-all rating of flight performance on a two-point scale: U, upper half; and L, lower half of the group.

To aid in the observation of the films and in arriving at the ratings, a Manual for Rating Pilot Performance through Direct Inspection of Motion Photographs was prepared. This manual, presented in Appendix C, describes in detail the character of the observations made and the type of ratings obtained. In Exhibit 3 is presented a sample set of notes and ratings made on one subject's performance.

D. Rating Procedure. Each of the three groups of films (i.e., Flight D Spring Program, Flight D Summer Program, and Flight A Spring and Summer Program subjects) was rated by two observers[16] according to the following procedure:

Step 1. Each photographed flight was rated independently by the two observers according to the rating method described above and the resulting three-point and two-point over-all ratings were obtained on each subject. These ratings are designated as Independent Ratings I.

---

[16] A. S. Thompson, University of Pennsylvania and E. S. Ewart, formerly of Purdue University and now at the University of Pennsylvania, were employed as the two observers. Besides a short period of flight instruction for orientation purposes, Thompson, in the course of aviation research since 1940, has had considerable experience in controlled observation of pilots of a wide range of flying experience, and had taken an active part in the development of the standard flights and the methods of photographic recording used in this and other projects.

Ewart, in addition to 30 hours of flying time and private pilot status, was thoroughly familiar with student pilot performance through his work at Purdue University in the preparation of the training aids, "Patter" and "Fundamentals of basic flight maneuvers," published recently by the National Research Council Committee on Selection and Training of Aircraft Pilots.

EXHIBIT 5

Sample Ratings and Notes on Flight Performance

Subject No.:
Group:
Observer:
Rating:

Notes on Performance:

Take-off:
360° Left:

360° Right:
Stall:
Slip:
Glide:
Turns:
Climb:
S & L:
Landing:

Ratings Assigned:

Wing Control:
Nose Control:
Directional Control:
Slip or Skid:
Altitude:
Air speed:
Controls:
Execution:
Rating A,B,C:
Rating D-L:
Air:

In order that the distributions of the ratings by the two observers might be comparable, it was agreed, <u>in advance</u>, that approximately 25% of the sample should be included in the A and C groups, respectively, in the case of the three-point ratings, and that approximately 50% of the sample should be included in the U and L groups, respectively, in the case of the two-point ratings. The actual distributions used were as follows:

	Three-Point Scale			Two-Point Scale	
	A	B	C	U	L
Flight D - Summer Group	7	19	7	17	16
Flight D - Spring Group	9	18	9	18	18
Flight A - Spring and Summer Group	8	16	9	18	15

If, after completing the analysis of all the flights in the group being studied, an observer found that the distribution of his ratings differed from the appropriate distribution (as given above), he then reviewed the records of those subjects whose ratings he considered "doubtful" and forced his final ratings to conform to the required distribution.

<u>Step 2</u>. The two sets of <u>Independent Ratings I</u> were then compared and a <u>Composite Rating I</u> obtained by <u>joint analysis of the films and joint rating</u> of cases in which a discrepancy of rating occurred during the <u>Independent Rating I</u> and by assuming that cases given identical ratings by the two observers in the <u>Independent Rating I</u> would have received a similar rating during joint observation.

This step was taken in order to arrive at ratings which represented the <u>combined judgment of both observers</u>. These <u>Composite Ratings</u> were expected to be more reliable than the <u>Independent Ratings</u> since inconsistencies of the individual observer were brought to light and corrected when the observers worked jointly.[17]

---

[17] This procedure of having two observers make independent judgments and then, in joint session, revise the "discrepant cases" is analogous to the method of "case study of incorrectly predicted cases" discussed in: Horst, Paul (Ed.) <u>The prediction of personal adjustment</u>. New York: Social Science Research Council, 1941, p. 117.

Step 3. The flights were then re-rated independently by the two observers and Independent Ratings II obtained. During this re-rating the identification number of the subject was blocked out of the projection field and the films were presented in a random order. Independent Ratings II were obtained in order to study the consistency of the individual observer and to provide a further check on the comparability of ratings given by different observers.

Step 4. The two sets of Independent Ratings II were then compared, the discrepant cases rated jointly, and a Composite Rating II, prepared. Composite Ratings II were obtained in order to determine the reliability of the previous Composite Ratings I and to provide further data for the assigning of final ratings to the subjects.

Step 5. The cases in which a discrepancy occurred between the rating given in Composite Rating I and that given in Composite Rating II were then jointly reviewed and the final ratings, termed Criterion Ratings were prepared. These final ratings thus represent the combined judgment of two observers on the basis of two sets of independent ratings each and two joint ratings of the discrepant cases.

Step 6. The five-point scale ratings on the eight aspects of flight performance were used to form Flight Scores and Aspect Scores as follows:

1. Independent Flight Score, obtained by summing the eight aspect ratings made by an observer on each subject. The best possible score was 8 and the worst possible was 40.

2. Composite Flight Score, obtained by summing the Independent Flight Scores of the two observers on the same subject. The possible range of this score extended from 16 to 80.

3. Criterion Flight Score, obtained by summing the two Composite Flight Scores on the same subject. The possible range extended from 32 to 160.

4. Aspect Scores, obtained by summing the four ratings on each aspect resulting from the independent rating and re-rating by the two observers separately, i.e., Thompson, I, Ewart I, Thompson II, Ewart II. This score ranged from 4 to 20.

IN SUMMARY: The rating procedure provided two major types of measures of pilot performance:

1. Over-all ratings, composed of the three-point scale and two-point scale ratings of the performance as a whole. These ratings are termed:

a. Independent Ratings, when based on the ratings made by each observer, working independently.
b. Composite Ratings, when resulting from the joint comparison of the two sets of Independent Ratings made by the two observers.
c. Criterion Ratings, when resulting from the joint comparison of the two sets of Composite Ratings provided by the rating and re-rating procedure.

2. Scores, derived from the five-point scale ratings on each of eight aspects of pilot performance, as follows:

a. Independent Flight Scores, obtained by summing the eight aspect ratings assigned by an observer during his independent rating of the performance of a given subject.
b. Composite Flight Score, obtained by summing the Independent Flight Scores given the same subject by the two observers.
c. Criterion Flight Scores, obtained by summing the two Composite Flight Scores of a given subject.
d. Aspect Scores, obtained by summing the four ratings of the same aspect, assigned during the independent rating and re-rating by the two observers, separately.

The above procedure thus resulted in the following sets of over-all ratings and scores on each subject in each group of flights:

Over-all Ratings

1. Independent Rating I, Thompson.
2. Independent Rating I, Ewart.
3. Composite Rating I.
4. Independent Rating II, Thompson.
5. Independent Rating II, Ewart.
6. Composite Rating II.
7. Criterion Rating.

Scores

1. Independent Flight Score I, Thompson.
2. Independent Flight Score I, Ewart.
3. Composite Flight Score I.
4. Independent Flight Score II, Thompson.
5. Independent Flight Score II, Ewart.
6. Composite Flight Score II.
7. Criterion Flight Score.
8. Aspect Scores.
   a. Wing Control Score.
   b. Nose Control Score.
   c. Directional Control Score.
   d. Slip-Skid Score.
   e. Altitude Score.
   f. Airspeed Score.
   g. Control Coordination Score.
   h. Execution of Maneuvers Score.

RELIABILITY OF THE MEASURES OF PERFORMANCE

A. Reliability of the Over-all Ratings. The reliability of the over-all ratings was measured by determining the significance and extent of the agreement between the various sets of ratings on the same subjects.

Since the rating procedure provided (for each subject) two Independent Ratings by each of the two observers and two Composite Ratings, various inter-comparisons were possible. Since three groups of flights were studied, the findings could be compared from sample to sample.[18]

1. Three-Point Scale (A, B, C) Ratings. Comparisons between sets of three-point scale ratings were made by computing Chi-Square and P-values to determine the significance of the associations and Coefficients of Contingency[19] and Percent Rated Identically to measure the extent of agreement.

   The results are presented in Table I, Reliability of the Over-All Ratings: Three-Point Scale (A, B, C) Ratings. Analysis of Table I, page 16, reveals the following:

   a. With the exception of two comparisons in the Flight D Summer Group, a statistically significant association[20] was found between sets of ratings on the same pilots made by:

      (1) The same observer during two independent ratings, e.g., Independent Rating I, Thompson vs. Independent Rating II, Thompson.
      (2) Two observers working independently, e.g., Independent Rating I, Thompson vs. Independent Rating I, Ewart.
      (3) Two observers working jointly, e.g., Composite Rating I vs. Composite Rating II.

   b. The Composite Ratings made by the two raters working jointly were, in general, more reliable than ratings made independently by a single observer. In each flight group, the mC for Composite Rating I vs. Composite Rating II is as high or higher than for any of the Independent Rating comparisons.

---

[18]An analysis of the ratings given to the two Flight D groups has been previously presented to the National Research Council Committee in a preliminary report. See: Viteles, M. S., Thompson, A. S., and Ewart, E. S. An analysis of photographic records of pilot performance: Part I. The reliability of qualitative estimates of flight performance derived from inspection of motion photographs. January 1943. (Copy in Committee files.)

[19]The Contingency Coefficients were corrected for number of categories and for class index according to the procedure described in Guilford, J. P., Psychometric methods. New York: McGraw-Hill Book Co., Inc., 1936, pp. 357-360. In Table I both the uncorrected (uncC) and corrected (mC) coefficients are presented.

[20]Using a P-value of $<.05$ as the criterion. Actually, 11 of the 15 comparisons resulted in P-values less than .01, a value which indicates that the observed association would occur by chance less than one time in 100.

The Percents Rated Identically for Flight D Spring and Flight D Summer Groups were 83% and 76%, respectively.

c. The two observers when rating independently agreed as well with each other as with themselves, as shown by the following comparison:

Percent Rated Identically

Comparison	Flight A		Flight D Spring		Flight D Summer		Average
	I	II	I	II	I	II	
Between two observers	71	68	58	72	67	46	63.6
Between ratings by the same observer	62	62	67	75	67	53	65.1

d. Neither observer was consistently more reliable than the other.

e. The Criterion Ratings may quite possibly be considered as even more reliable than the Composite Ratings since the Criterion Ratings were based on an additional review of those few cases rated differently in Composite Rating I and Composite Rating II. The Criterion Ratings represent final ratings obtained from a careful "sifting" process in which two observers checked each other's independent ratings and in which the joint ratings of the two observers were then checked by a joint re-rating.

2. Two-Point Scale (U-L) Ratings. The reliability with which the raters divided the groups into upper and lower half (U-L Ratings) was determined by computing Chi-Square and P-values to test the significance of the association between sets of ratings, and tetrachoric r's and Percent Rated Identically to measure the extent of agreement.

The results are presented in Table 2, Reliability of the Over-All Ratings: Two-Point Scale (U-L) Ratings. Analysis of Table 2, page 17, reveals the following:

a. With the exception of the intra-rater comparisons of Flight A, a statistically significant association was found between the various pairs of ratings. Twelve of the 15 obtained P-values were less than .01.

b. As contrasted with the three-point scale ratings, the two-point scale Composite Ratings were not found to be consistently more reliable than the Independent Ratings. The tetrachoric r's and Percents Identical of the Composite Rating comparisons are equalled by Independent Rating comparisons in Flight A and Flight D Summer Group and exceeded in the Flight D Spring Group comparisons.

TABLE 1

RELIABILITY OF THE OVER-ALL RATINGS--THREE-POINT SCALE
(A, B, C) RATINGS

Flight	Comparison Made	No. of Cases	χ^2	P	uncC	mC	Percent Identical
A Spring and Summer Group	Composite Rating I vs. Composite Rating II	34	20.64	<.01	.61	.75	68
	Independent Rating I, Thompson vs. Independent Rating I, Ewart	34	19.93	<.01	.61	.74	72
	Independent Rating II, Thompson vs. Independent Rating II, Ewart	34	20.64	<.01	.61	.75	68
	Independent Rating I, Thompson vs. Independent Rating II, Thompson	34	14.27	<.02>.01	.55	.63	62
	Independent Rating I, Ewart vs. Independent Rating II, Ewart	34	17.35	<.01	.59	.71	62
D Spring Group	Composite Rating I vs. Composite Rating II	35	42.22	<.01	.73	.95	83
	Independent Rating I, Thompson vs. Independent Rating I, Ewart	35	11.22	<.05>.02	.49	.54	58
	Independent Rating II, Thompson vs. Independent Rating II, Ewart	36	26.34	<.01	.65	.82	72
	Independent Rating I, Thompson vs. Independent Rating II, Thompson	36	20.00	<.01	.60	.73	67
	Independent Rating I, Ewart vs. Independent Rating II, Ewart	36	32.89	<.01	.69	.88	75
D Summer Group	Composite Rating I vs. Composite Rating II	33	28.97	<.01	.68	.86	76
	Independent Rating I, Thompson vs. Independent Rating I, Ewart	33	19.03	<.01	.61	.74	67
	Independent Rating II, Thompson vs. Independent Rating II, Ewart	33	5.24	<.30>.20	.37	.25	45
	Independent Rating I, Thompson vs. Independent Rating II, Thompson	33	19.78	<.01	.61	.75	67
	Independent Rating I, Ewart vs. Independent Rating II, Ewart	33	9.43	<.10>.05	.47	.49	58

TABLE 2

RELIABILITY OF THE OVER-ALL RATINGS—TWO-POINT SCALE (U-L) RATINGS

Flight	Comparison Made	No. of Cases	X^2	P	r_{tet}	Percent Identical
A Spring and Summer Group	Composite Rating I vs. Composite Rating II	34	9.47	$<.01$	.73	76
	Independent Rating I, Thompson vs. Independent Rating I, Ewart	34	5.70	$<.02$ (.01)	.54	71
	Independent Rating II, Thompson vs. Independent Rating II, Ewart	34	9.47	$<.01$	.73	76
	Independent Rating I, Thompson vs. Independent Rating II, Thompson	34	2.89	$<.10$ (.05)	.43	65
	Independent Rating I, Ewart vs. Independent Rating II, Ewart	34	2.89	$<.10$ (.05)	.43	65
D Spring Group	Composite Rating I vs. Composite Rating II	36	11.12	$<.01$	.77	78
	Independent Rating I, Thompson vs. Independent Rating I, Ewart	36	11.12	$<.01$	.77	78
	Independent Rating II, Thompson vs. Independent Rating II, Ewart	36	16.00	$<.01$	.87	83
	Independent Rating I, Thompson vs. Independent Rating II, Thompson	36	16.00	$<.01$	.87	83
	Independent Rating I, Ewart vs. Independent Rating II, Ewart	36	11.12	$<.01$	.77	78
D Summer Group	Composite Rating I vs. Composite Rating II	33	18.68	$<.01$	.93	88
	Independent Rating I, Thompson vs. Independent Rating I, Ewart	33	13.70	$<.01$	.84	88
	Independent Rating II, Thompson vs. Independent Rating II, Ewart	33	18.68	$<.01$	.93	88
	Independent Rating I, Thompson vs. Independent Rating II, Thompson	33	8.56	$<.01$	.72	76
	Independent Rating I, Ewart vs. Independent Rating II, Ewart	33	13.70	$<.01$	.84	82

c. The two raters when rating independently agreed as well with each other as with themselves, as shown by the following comparison.

Percent Rated Identically

Comparison	Flight A		Flight D Spring		Flight D Summer		Average
	I	II	I	II	I	II	
Between two raters	71	76	78	83	88	89	80.7
Between ratings by the same rater.	65	65	83	78	76	82	74.8

d. Neither of the two raters was consistently more reliable than the other.

e. As shown by the tetrachoric r's and Percent Rated Identically, the two-point scale ratings of Flight A were less reliable, in general, than those of the Flight D groups.

3. Analysis of Composite Ratings and Criterion Ratings. The high relationship between the Composite Ratings I and II could possibly be due to the fact that the ratings of one of the two observers unduly influenced the Composite Rating. To determine whether this actually occurred, comparisons were made between the three-point scale Independent Ratings and the appropriate Composite Rating. These comparisons are presented in Table 3, page 20.

a. Analysis of Table 3 reveals that the Independent Ratings made by Thompson agreed slightly better with the Composite Rating than did those made by Ewart. That the difference was relatively slight is indicated by the following tabulation:

Percent Rated Identically

Comparisons	Flight A		Flight D Spring Group		Flight D Summer Group		Average
	I	II	I	II	I	II	
Independent Rating, Thompson vs. Composite Rating	88	85	78	89	88	76	81.5
Independent Rating, Ewart vs. Composite Rating	76	76	78	83	82	70	77.5

b. A similar type of analysis was made to determine whether the Composite Rating I and Composite Rating II contributed unequally to the final rating (Criterion Rating). The extent of agreement between the Composite Ratings and the Criterion Rating is given in Table 3. Examination of the table reveals that the Composite Ratings II were in higher agreement with the Criterion Ratings than were the Composite Ratings I, as follows:

Percent Rated Identically

Comparisons	Flight A	Flight D Spring Group	Flight D Summer Group	Average
Composite Rating I vs. Criterion Rating	82	89	82	84.3
Composite Rating II vs. Criterion Rating	82	94	94	90.0

The differences, though consistent in direction are relatively small and both Composite Ratings agree highly with the Criterion Ratings. The fact that the Composite Ratings do not agree perfectly with each other or with the Criterion Ratings, however suggests the value of the re-rating procedure which provided a complete and independent re-check before the Criterion Ratings were assigned.

IN SUMMARY: Comparisons between the various sets of over-all ratings provided by the rating procedure revealed that, in general, the Composite Ratings were more reliable than the Independent Ratings and that the three-point scale ratings were somewhat superior to the two-point scale ratings. The Criterion Ratings based on an additional review of the cases rated differently in Composite Rating I and Composite Rating II, may be considered sufficiently reliable for use as criterion data. The results were fairly consistent for the three samples of subjects used.

B. Reliability of the Flight Scores.[21] The reliability of the various Flight Scores was determined from a study of the Flight Scores obtained by 33 of the 34 subjects [22] for whom both Flight A and Flight D records were available. The results are thus based upon a group of 33 flights

---

[21]The treatment of the Flight Scores only is included in the main body of this report since they are the ones used for criterion purposes. An extensive analysis of the Aspect Scores was also made and is described in an attached Supplemental Report.

[22]One of the two raters failed to assign aspect ratings for the Flight D of the omitted case.

TABLE 3

RELATIONSHIPS BETWEEN SETS OF OVER-ALL RATINGS--THREE-POINT SCALE
(A,B,C RATINGS)

Flight	Comparison Made	No. of Cases	χ^2	P	unc. C	Percent Identical
A Spring	Independent Rating I, Thompson vs. Composite Rating I	34	46.74	$<.01$	.76	88
and	Independent Rating I, Ewart vs. Composite Rating I	34	30.03	$<.01$	.68	76
Summer Group	Independent Rating II, Thompson vs. Composite Rating II	34	48.50	$<.01$	.77	83
	Independent Rating II, Ewart vs. Composite Rating II	34	30.08	$<.01$	.68	76
	Composite Rating I vs. Criterion Rating	34	38.25	$<.01$	.73	82
	Composite Rating II vs. Criterion Rating	34	38.25	$<.01$	.73	82
D Spring	Independent Rating I, Thompson vs. Composite Rating I	36	32.39	$<.01$	.69	78
Group	Independent Rating I, Ewart vs. Composite Rating I	36	32.99	$<.01$	.69	78
	Independent Rating II, Thompson vs. Composite Rating II	36	50.22	$<.01$	.76	89
	Independent Rating II, Ewart vs. Composite Rating II	36	41.45	$<.01$	.73	83
	Composite Rating I vs. Criterion Rating	36	50.22	$<.01$	.76	89
	Composite Rating II vs. Criterion Rating	36	61.00	$<.01$	.79	94

TABLE 3 (cont)

Flight	Comparison Made	No. of Cases	χ^2	P	unc. C	Percent Identical
D Summer Group	Independent Rating I, Thompson vs. Composite Rating I	33	44.61	<.01	.76	88
	Independent Rating I, Ewart vs. Composite Rating I	33	34.71	<.01	.72	82
	Independent Rating II, Thompson vs. Composite Rating II	33	27.18	<.01	.67	76
	Independent Rating II, Ewart vs. Composite Rating II	33	19.56	<.01	.61	70
	Composite Rating I vs. Criterion Rating	33	35.64	<.01	.72	82
	Composite Rating II vs. Criterion Rating	33	54.99	<.01	.79	94

TABLE 4

RELIABILITY OF THE FLIGHT SCORES

Comparison Made	Flight A (N = 33) r	Flight A (N = 33) S-Br	Flight D (N = 33) r	Flight D (N = 33) S-Br
Composite Flight Score I vs. Composite Flight Score II	.77		.84	
Independent Flight Score I, Thompson vs. Independent Flight Score I, Ewart	.59		.58	
Independent Flight Score II, Thompson vs. Independent Flight Score II, Ewart	.72		.73	
Independent Flight Score I, Thompson vs. Independent Flight Score II, Thompson	.67		.75	
Independent Flight Score I, Ewart vs. Independent Flight Score II, Ewart	.65		.72	
Criterion Scores (Spearman-Brown) (based on Comp. I, Comp. II)		.97		.91

during Stage A and 33 flights during Stage B, the subjects being the same in both flights.

1. Statistical Treatment of the Flight Scores.

   a. The reliability of the Flight Scores was measured by computing Pearsonian r's among the various sets of scores provided by the rating procedure. The coefficients of correlation are presented in Table 4, page 21.

   b. The statistical significance of the obtained correlation coefficients can easily be determined by applying the "t" test. Since all the correlations were based upon the same N, the value of a coefficient required for significance at the 1% level was read from an appropriate table.[23] In this case, the N is 33 and the minimum coefficient required for significance at the 1% level is .443.

   c. The determination of the significance of differences between coefficients, however, is made difficult due to the fact that the coefficients are based on the same population and are thus intercorrelated. The ordinary critical ratio based upon the standard error of the difference between r's (or between Fisher's z equivalents) may not reveal significant differences since the standard error will be over-estimated.

   d. Since the Criterion Flight Scores are the sum of the two Composite Flight Scores, the reliability of the Criterion Flight Scores can be estimated by the application of the Spearman-Brown Prophecy Formula[24] to the correlation coefficient obtained between Composite Flight Score I and Composite Flight Score II. The latter are essentially equivalent halves of the Criterion Flight Score since they result from an independent re-rating procedure.

   An empirical test of this procedure was made by comparing the correlation actually obtained between Composite Flight Score I and Composite Flight Score II and that which would have been predicted from applying the S-B formula to the average of the four coefficients obtained between pairs of Independent Flight Scores, as given in Table 4 and reproduced below:

---

[23] Lindquist, E.F. Statistical analysis in educational research. New York: Houghton Mifflin Co., 1940. p. 212, Table 13.

[24] Guilford, J. P. Psychometric methods. New York: McGraw-Hill Company, Inc., 1936., p. 419, Formula 196.

Comparison	Flight A r	Flight A Ave.	Flight D r	Flight D Ave.
Independent Flight Score I, Thompson vs. Independent Flight Score I, Ewart	.59		.58	
Independent Flight Score II, Thompson vs. Independent Flight Score II, Ewart	.72		.73	
Independent Flight Score I, Thompson vs. Independent Flight Score II, Thompson	.67		.75	
Independent Flight Score I, Ewart vs. Independent Flight Score II, Ewart	.65	.658	.72	.695
Composite Flight Score I vs. Conposite Flight Score II				
1. As predicted from average		.79		.82
2. As actually obtained		.77		.84

2. Reliability of the Flight Scores (Table 4).

   a. All the coefficients are statistically significant since they are above .443, the minimum value required for significance at the 1% level with an N of 33.

   b. The Composite Flight Scores (based upon the sum of two Independent Flight Scores) show higher reliability coefficients than do the Independent Flight Scores. The Composite I vs. Composite II Flight Scores for Flights A and D are .77 and .84 respectively.

   c. The agreement between raters (Thompson vs. Ewart) is, in general, approximately equal to that obtained when scores of the same rater are compared.

   d. The reliability coefficients for Flights A and D are quite similar in value.

   e. The Spearman-Brown coefficients for the Criterion Flight Scores are .87 and .91 for Flight A and Flight D, respectively. From these estimates, the reliability of the Criterion Flight Scores can be considered as sufficiently high for the purposes of this investigation.

IN SUMMARY: Flight Scores were obtained by summing the five-point scale ratings on eight aspects of flight performance. Statistically significant agreement was found between the various sets of Flight Scores, and the reliability of the Criterion Flight Scores was estimated as approximately .90 on the basis of two flights (Flight A and Flight D) of a group of 33 subjects.

COMPARISON OF CRITERION FLIGHT SCORES AND CRITERION RATINGS

A. Intercorrelation of the Two Measures of Flight Performance. The relationship between the Criterion Flight Scores and the Criterion Ratings was studied in order to determine the extent of agreement between scores obtained by merely summing the sets of aspect ratings made independently by the two observers and the over-all ratings representing the combined and interacting judgments of the two observers in a joint rating procedure.[25]

The comparisons were made in two ways: (1) by computing biserial correlation coefficients between the two sets of measures and (2) by comparing the distributions of Criterion Flight Scores of the sub-groups classified according to Criterion Ratings.

1. Three biserial coefficients were computed for each comparison: first, with the group split into upper 75%, Lower 25% (A and B vs C Ratings); second, with the group split approximately in half (U vs. L Ratings); and third, with the group split into upper 25%, Lower 75% (A vs. B and C Ratings). This was done in order to reveal any tendency for the scores to differentiate unequally at various points in the distribution. The biserial r's obtained in the comparisons are as follows:

	Flight A			Flight D		
	(A+B)vs.(C)	U vs. L	(A)vs. (B+C)	(A+B)vs.(C)	U vs. L	(A)vs. (B+C)
N =	24 - 9	17 - 16	9 - 24	24 - 9	18 - 15	8 - 25
r_{bis}	.88	.95	1.00	.89	1.00	.98

2. When the distributions of Criterion Flight Scores of the groups classified according to the Criterion Ratings are studied the following data are obtained:

---

[25] Since the Criterion Flight Scores were obtained on only 33 cases, selected from both Spring and Summer Groups, Criterion Ratings for Flight D for this composite group had to be obtained. Merely assigning the Criterion Rating given when the subjects were rated within their own groups would have assumed that the Spring and Summer Groups were equivalent samples. An actual cross-comparison was thus made (Thompson and Ewart jointly reviewing the films) and the 33 members of the group assigned into Upper 8 (A rating), Middle 16 (B rating) and Lower 9 (C rating). On the two-point scale, 18 were given U Ratings and 15 were given L Ratings. This re-classification was made before the Flight A ratings were made and before any of the Criterion Flight Scores were computed.

Mean Criterion Flight Scores of Criterion Rating Sub-Groups

Criterion Rating

	A	U	L	C
Flight A	76.2	85.7	110.7	116.3
N =	(9)	(17)	(16)	(9)
Flight D	73.4	92.6	110.3	114.8
N =	(8)	(18)	(15)	(9)

If the two extreme groups (i.e., those given Criterion Ratings A or C) are compared, there is no overlapping at all in the distributions of the Criterion Flight Scores. In fact, the "best" C of Flight A has a Criterion Flight Score 24 points lower (poorer) than that of the "poorest" A. Likewise, a distance of 16 Criterion Flight Score points separates the A and C Criterion Rating sub-groups on Flight D, a value approximately equal to the standard deviation of the entire Criterion Flight Score distribution of Flight D.

The size of the biserial r's and the difference in the distributions of Criterion Flight Scores of the sub-groups warrant the conclusion that the two measures of flight performance are highly related.

B. Evaluation of the Criterion Flight Scores and Criterion Ratings. Since the Criterion Flight Scores and the Criterion Ratings correlate rather highly with each other the question may be raised as to the relative value of these two types of criterion data. A comparison of the two as revealed by this investigation follows:

1. Comparison in Terms of Reliability. A direct comparison of relative reliability is impossible since the reliability of the Criterion Ratings was determined in terms of contingency coefficients and tetrachoric r's in the case of the three-point scale and two-point scale ratings, respectively, while the reliability coeffiecients of the Criterion Flight Scores are in terms of Pearson r's. The absolute values of these statistics are not directly comparable.

   On the basis of a qualitative evaluation of the reliability data on these three sets of criterion data, however, there seems to be little reason for choosing between the three-point scale Criterion Ratings and the Criterion Flight Scores. The reliability of the two-point scale Criterion Ratings fell down somewhat in the case of the Flight A comparisons.

2. Comparison in Terms of Discriminating Value. The discriminating value of the three-point scale and two-point scale Criterion Ratings is limited, of course, by the number of categories in the scale. By the method itself, they do not provide for differentiation within the categories.[26]

---

[26] Whether the scales could be made "finer," i.e., with four, five, or more categories, with no loss of reliability is, of course, a question to be answered only through further experimentation.

The Criterion Flight Scores, on the other hand, provide a "continuous" distribution, the discriminating value of which is determined by the range and reliability of the scores. Taking .90 to be the estimated reliability of the Criterion Flight Scores (See Table 4), it can be shown that the Criterion Flight Scores will accurately place individuals into at least 6 categories.[27]

3. Comparison in Terms of Their Nature. The Criterion Ratings represent the composite judgment of two raters resulting from a "joint" analysis and evaluation of the pilot's performance. The ratings assigned are thus based upon an active "checking" procedure in which the errors of one rater are presumably caught and corrected by the other.[28] The Criterion Flight Scores, on the other hand, are formed merely by an arithmetical summing of aspect ratings made independently by the two raters with no provision for any inter-acting effect of one rater upon the other.

The question is essentially one of "pooled" vs. "composite" group judgment concerning which there has been considerably discussion in the psychological literature. Experiments such as those reported by Gordon[29], [30], Bruce[31], and Eysenck[32] have purported to demonstrate an increase in validity resulting from the averaging (or pooling) of judgments of individual observers to form "group" judgments. Kelley[33], however, pointed out that the increase in correlation between the group orders and the true order closely followed that which would be predicted by the application of the Spearman-Brown Prophecy Formula, and Stroop[34], by means of

---

[27]A formula for determining the number of significant categories from the reliability coefficient is discussed in: Bloom, Benj. S., Test reliability for what? J. educ. Psychol., 1942, 33, 517-526.

[28]In addition, the re-rating procedure as used in this investigation provided a check upon the composite judgments of the two raters, since the Composite Ratings I and II were jointly compared when the Criterion Ratings were assigned.

[29]Gordon, K., Group judgments in the field of lifted weights, J. exp. Psychol., 1924, 7, 398-400.

[30]Gordon, K., Further observations on group judgments of lifted weights, J. Psychol., 1935-36, 1, 105-115.

[31]Bruce, R. S., Group judgments in the fields of lifted weights and visual discrimination, J. Psychol., 1935-36, 1, 117-121.

[32]Eysenck, H. J., The validity of judgments as a function of the number of judges. J. exp. Psychol., 1939, 25, 650-654.

[33]Kelley, T. L., The applicability of the Spearman-Brown formula for the measurement of reliability, J. educ. Psychol., 1925, 16, 300-303.

[34]Stroop, J. R., Is the judgment of the group better than that of the average member of the group, J. exp. Psychol., 1932, 15, 550-562.

a card-sorting experiment, demonstrated that the increase was due to statistical rather than to experimental reasons. Preston[35] further demonstrated that "amalgamation of a sufficiently large number of judgments following Gordon's method might well result in the group judgment being substantially worse than the judgment of the average member of the group ... (as) ... in the case where the median judgment showed a slight negative correlation with the true order of the subjects being judged." Preston concluded that "the increases in question are really increases in reliability."

The results of the present investigation permit a comparison of group judgments of the two types described, i.e., arithmetical amalgamation of independent judgments (Criterion Flight Scores) and composite judgments based on the interaction of observers working jointly (Criterion Ratings). Although there is no external "true" order against which to test the relative validity of the two types of group judgments, the extent to which the Criterion Flight Scores agree with the Criterion Ratings can be determined. As shown above, the Criterion Flight Scores were found to be highly associated with the Criterion Ratings and, for a rough grouping of the subjects into categories, both types of criterion data give essentially the same results.

4. Practical Considerations. The Criterion Flight Scores have certain advantages over the Criterion Ratings with respect to their practical use in research and possible use in the field. The advantages are:

a. Ease of interpretation. Comparisons by means of Criterion Flight Scores are more easily made since the Criterion Ratings are in terms of the particular group being rated. A rating of "A," for example, means that the individual is rated in the "upper 25% of the group being rated." An "A" rating of one group is thus not necessarily comparable to that of another group.[36] A given Flight Score, on the other hand, has no direct reference to the particular group being rated, although it is recognized that the aspect ratings (from which the Flight Score is computed) are undoubtedly influenced by the general level of performance exhibited by the group as a whole.

b. Ease of statistical treatment. Since the Criterion Flight

---

[35] Preston, M.C., Note on the reliability and validity of the group judgment, J. exp. Psychol., 1938, 22, 462-471.

[36] See Footnote 25. Page 24.

Scores form a continuous distribution (in contrast to the large categories provided by the Criterion Ratings), the Criterion Flight Scores can be treated by more extensive statistical methods.

IN SUMMARY: Comparisons between corresponding sets of Criterion Ratings and Criterion Flight Scores indicate that the two types of measures of flight performance are highly related and that for a grouping of subjects into three categories both types of criterion data give essentially the same results. The Criterion Ratings have the advantage of representing the combined and inter-acting judgment of two observers while the Criterion Flight Scores possess certain practical advantages.

COMPARISON OF FLIGHT A AND FLIGHT D

One objective of this investigation was the determination of the extent to which Flight A performance is associated with, or predictive of, performance on Flight D. Inter-comparisons were thus made among the sets of Criterion Ratings and Criterion Flight Scores obtained by the pilots on the two flights.

Comparisons between sets of Criterion Ratings were made by computing X^2 and P-values, between Criterion Ratings and Criterion Flight Scores in terms of biserial r's, and between sets of Criterion Flight Scores in terms of Pearsonian r's.

The results of these comparisons are presented in Table 5. The following relationships were observed:

A. Criterion Ratings Flight A versus Criterion Ratings Flight D.

1. There is a statistically significant association between the three-point scale (A, B, C, Criterion Ratings on Flight A and those on Flight D, as indicated by the P-value of $< .05 > .02$ (See Table 5, Section A).

2. The extent of the agreement is shown by the corrected coefficient of contingency of .49 and by the fact that approximately 60% of the 33 cases were placed in the same category on both flights.

3. The relationship exhibited between the two sets of the two-point scale (U - L) Criterion Ratings is not statistically significant.

B. Flight A Criterion Flight Scores vs. Flight D Criterion Ratings.

1. The highest biserial r between the Criterion Flight Score of Flight A and Criterion Ratings of Flight D is .45. (See Table 5, Section B). This value is not sufficiently high to be statistically significant since for this comparison a biserial r of $\pm$.45 or higher might occur by chance in approximately 7 times in 100,[37] as read

---

[37] With an N of 33 and the proportions .727 and .273, the standard error of a biserial r of .00 is .2329. Peters, C. C. and Van Voorhis, W. R. Statistical procedures and their mathematical bases. New York: McGraw-Hill Book Co. Inc. 1940. p. 365, formula 183.

from a normal probability table. This fact, however, does not preclude the possibility that the two distributions as a whole are significantly related.

2. Inspection of the actual distributions of Criterion Flight Scores of the A, B, and C groups reveals considerable overlapping, even between the two extreme (A and C) groups.

C. Flight A Criterion Flight Scores vs. Flight D Criterion Flight Scores.

1. The relationship between the Criterion Flight Score of both flights is sufficiently high to be statistically significant.

2. The obtained Pearsonian r is only .46, however, indicating that prediction from Flight A to Flight D would be only approximately 10% better than chance if Flight A performance is used as a predictor.[38]

TABLE 5

COMPARISON OF FLIGHT A AND FLIGHT D

A. Criterion Rating Flight A vs. Criterion Rating Flight D

	χ^2	P	$_mC$	Percent Identical
1. Three-Point Scale (A,B,C Ratings)	9.61	$\langle .05 \rangle .02$	.49	59

	χ^2	P	r_{tet}	Percent Identical
2. Two-Point Scale (U-L Ratings)	.92	$\langle .50 \rangle .30$	.23	59

B. Criterion Flight Score Flight A vs. Criterion Rating Flight D (Biserial r's)

		Criterion Rating Flight D		
		(A+B)-(C)	U - L	(A)-(B+C)
Flight A Score Versus:	N =	24 - 9	18 - 15	8 - 25
Biserial r		.45	.41	.38

C. Criterion Flight Score Flight A vs. Criterion Flight Score Flight D .46

Analysis of the above comparisons indicate that, in general, estimates of pilot proficiency made toward the end of Stage A of C.P.T. primary flight training are not highly predictive of similar estimates made toward the end of Stage D, the final stage of C.P.T. primary flight instruction.

[38]Guilford, J. P. Psychometric methods. New York: McGraw-Hill Book Co. Inc., 1936., p. 363, Table 74.

In the various comparisons (Criterion Ratings Flight A versus Criterion Ratings Flight D, Criterion Ratings Flight D versus Criterion Flight Scores Flight A, and Criterion Flight Scores Flight A versus Criterion Flight Scores Flight D) the amount of relationship was approximately 10% greater than that attributable to chance. Whether measured by a Pearsonian r between Criterion Flight Scores, a coefficient of contingency between Three-Point Scale Criterion Ratings,[39] or biserial r's between Criterion Flight Scores and Three-Point Scale Criterion Ratings, the value of the correlation statistic obtained fell in the upper 40's.

The observed degree of relationship between measures of Stage A and Stage D performance certainly cannot be considered sufficiently high to warrant the use of Flight A Criterion Flight Scores or Criterion Ratings as the sole criterion for "washing-out" student pilots at the end of Stage A, or deciding which students are worthy of further training. If the relationship should persist on additional population samples, however, these measures could be used in connection with selection tests, biographical inventories, interviews, etc.[40] To the predictive efficiency of these selection devices (no one of which is highly predictive) can be added that provided by estimates of performance based on a relatively short "try-out" period of instruction.

IN SUMMARY: Comparison of Flight A and Flight D performance of a group of 33 subjects revealed that the two performances are significantly related. The predictive value of Flight A performance, as measured by the photographic method, may be represented by an r of approximately .45 for one sample. Cross validation is needed to establish the exact relationship.

ANALYSIS OF VARIANCE OF THE FLIGHT SCORES

The rating procedure, described on pages 9 to 13, provided eight separate Flight Scores on each of the 33 student pilots in the Flight A - Flight D comparison, two each by the two raters for each of the two flights. The Flight Scores were thus subject to four sources of variation: differences among pilots, differences between raters, differences between first and second rating on the same flight, and differences between flights.

To test the significance of these differences and, in addition, to determine the existence of possible interactions among the sources of variation, an analysis of variance was made of the Flight Scores.[41]

---

[39] The two-point scale (U-L) Criterion Ratings comparison obtained a tetrachoric r of only .23.

[40] The data in this investigation will become a part of the Midwest Project data which include these various types of predictors as well as additional criteria. Inter-comparisons providing further information as to the value of the Criterion Flight Scores and Criterion Ratings will then be possible.

[41] Acknowledgment is made to Dr. Malcolm G. Preston, Department of Psychology, University of Pennsylvania, for his aid in designing the analysis and in the interpretation of the results.

A. Description of the Analysis. The design of the analysis can be determined from an examination of Table 6 which presents the sources of variation tested and the degrees of freedom assigned to each. Since there were 33 pilots and 8 Flight Scores on each (two each by two raters for two flights) the total number of items in the analysis is 264 and the total number of degrees of freedom is 263. They were assigned to the sources of variation as follows:

1. Pilots. The D.F. for Pilots is 32, one less than the number of subjects. The Sum of Squares for this source of variation was determined from a table in which each pilot was assigned the sum of his eight Flight Scores.[42]

2. Flights. The D.F. for Flights is 1, since the pilots were rated on two flights.

3. Raters. The D.F. for Raters is 1, since the Flight Scores from two raters were available.

4. Ratings. The D.F. for Ratings is 1, since each pilot was rated twice for each flight by each rater.

5. Interactions. The D.F.'s for the six possible interactions are the product of the D.F.'s of the two sources of variation involved.

6. Error. The D.F. for Error is 129, the degrees of freedom remaining after the single source and double interaction D.F.'s were subtracted from the total number of degrees of freedom. This remainder is composed of the 97 degrees of freedom attributable to the four possible triple interactions and to the 32 degrees of freedom attributable to the one quadruple interaction. It seemed legitimate to consider these triple and quadruple interactions as the errors of measurement against which to test the significance of the single source and double interaction variance.

B. Results of the Analysis. The results of the analysis are presented in Table 6. Analysis of Variance of Flight Scores, which shows how the total number of degrees of freedom was assigned to the sources of variation.

On the basis of the F test, in which the Source Variance is compared with the Error Variance, the only sources of variation found to be statistically significant were:

1. Differences among Pilots: highly significant.

2. Pilot-Flight Interaction: highly significant.

3. Flight-Rating Interaction: significant between the 5% and 1% levels.

---

[42] Identical with the sum of the Criterion Flight Scores for Flights A and D, respectively, as previously described on Page 12.

TABLE 6

ANALYSIS OF VARIANCE OF FLIGHT SCORES

		Source of Variation	D.F.	Sum of Squares	Mean Square (Variance)	F	F at 5% Level	F at 1% Level
	1.	Pilots	32	3485.13	108.91	12.97	1.55	1.85
	2.	Flights	1	28.02	28.02	3.34	3.92	6.84
	3.	Raters	1	15.52	15.52	1.85	3.92	6.84
	4.	Ratings	1	24.25	24.25	2.89	3.92	6.84
INTERACTIONS	5.	Pilot - Flights	32	1267.98	39.62	4.72	1.55	1.85
	6.	Pilot - Raters	32	211.43	6.61	.79	1.55	1.85
	7.	Pilot - Ratings	32	231.75	7.24	.86	1.55	1.85
	8.	Flights - Raters	1	18.50	18.50	2.20	3.92	6.84
	9.	Flights - Ratings	1	39.36	39.36	4.69	3.92	6.84
	10.	Raters - Ratings	1	31.97	31.97	3.81	3.92	6.84
	11.	Error (Remainder)	129	1084.17	8.40			
		Total	263	6438.13				

C. Interpretation of the Results. The results of the analysis of variance can be used to answer several important questions concerning the rating procedure used in this investigation. The questions are as follows:

1. Are the pilots differentiated by means of the Flight Scores?

   On the basis of the highly significant F of 12.97 obtained for Differences Among Pilots as a source of variation, the answer to this question is clearly "yes." The differences in Flight Scores among pilots cannot be attributed to errors of measurement.

2. Do the other single sources of variation determine the Flight Scores?

   The answer to this question is "No," on the basis of the fact that no significant differences in Flight Scores were obtained between Flights A and D, between the two raters (Thompson and Ewart) or between the first and second ratings. The F ratios for these three comparisons were all smaller than that required for significance at the 5% level. (See items 2, 3, and 4 of Table 6.)

3. Is there a high degree of relationship between the Flight Scores of Flight A and those of Flight D?

Information on this question is provided by the significance of the Pilot-Flight Interaction (Item 5, Table 6). If there were no interaction, it would mean that the score obtained by a given pilot would be independent of which of his two flights was being rated. "Good" pilots would thus obtain "good" scores on both flights and a high positive correlation would be obtained between scores on Flights A and D.

As noted above, the Pilot-Flight Interaction turned out to be significant, indicating that the relationship between the two flights was less than perfect.[43]

4. Are there any significant interactions resulting from the use of two flights, two raters, and two ratings?

The answer to this question is "No," with the possible exception of a Flight-Rating Interaction where the F-ratio is between the 5% and 1% level of significance.

In general, from an analysis of Items 6 to 10 on Table 6, it can be stated that:

(a) how a pilot was rated did not depend upon:

(1) Who was doing the rating (Pilot-Rater Interaction, Item 6).

(2) Whether it was the first or second rating (Pilot-Rating Interaction, Item 7).

(b) the raters exhibited no characteristic difference in their ratings on the two flights (Flight-Rater Interaction, Item 8).

(c) the ratings assigned to Flights A and D, separately, did depend somewhat upon whether it was the first or second rating made, (Flight-Rating Interaction, Item 9.)

(d) the raters did not differ significantly as to any constant tendency to assign the first ratings differently from the second ratings. (Rater-Rating Interaction, Item 10). The F ratio approached rather closely, however, to the value at the 5% level.

---

[43]The actual coefficient of correlation between the Criterion Flight Scores of Flights A and D was found to be .46, as given in Table 5, Section C.

IN SUMMARY: An analysis of variance of the Flight Scores revealed that differences between the two raters and between two ratings by the same raters were not significant. The major sources of variation lay in differences among pilots and in the fact that Flight A and Flight D performances were significantly but not highly correlated.

THE USE OF CRITERION RATINGS AND CRITERION FLIGHT SCORES AS CRITERION DATA

As pointed out in the introduction, this investigation was undertaken to attempt to provide measures of flight proficiency which would be adequate for use as criterion data in the 1942 Midwest Project. The resulting Criterion Ratings and Criterion Flight Scores must therefore be evaluated in the light of standards required for an adequate criterion, namely reliability, validity, discriminating value, and practicality.[44]

A. Evaluation in Terms of Reliability. From the results of the investigation it may be concluded that the measures obtained were sufficiently reliable for use as criterion data. Judged against absolute standards, the reliability of the Criterion Ratings and the Criterion Flight Scores is approximately that represented by an r of .90, the commonly accepted minimum requirement for adequate reliability. Judged against relative standards, the obtained reliability of these two types of measures is superior to that of other criterion measures available in the Midwest Project.

This statement should be tempered, however, by a consideration of the limitations of the present study. The limitations are as follows:

1. The Small Number of Cases Involved. There are only 33 cases of Flight A, 34 of Flight D Summer Group, and 36 of Flight D Spring Group.[45] The fact that the results are fairly consistent from sample to sample, however, increases the confidence with which conclusions can be drawn.

2. The Small Number of Raters Involved. The question of whether the method described in this report would yield reliable measures when used by observers other than the two who took part in the development of the method can be answered only by means of further experimental investigations. There seems no reason to suspect, however, that other observers, adequately indoctrinated with respect to the use of the rating procedure, would not produce ratings equally reliable to those obtained in this investigation.

3. The Use of a Single Sample of Flight Performance. For each subject (at a given stage) only one sample of flight performance was obtained. No information has been obtained, therefore, as to whether the photographed flight performance was representative

---

[44] It is expected that the Midwest Project report will include data concerning the relationship of the photographic criteria to the other criterion measures used in the project and will describe the method used in developing a composite criterion based on photographic criteria, ratings by instructors, flight inventory scores, and total training time.

[45] In the case of the Criterion Flight Scores there were only 33 cases in

of the flight proficiency of the subject being tested. Thus, although the obtained ratings may be a reliable measure of that particular photographed flight, it may not be a reliable measure of the subject's flight proficiency and thereby be an unreliable measure for criterion purposes. In brief, the ratings may be reliable but the single flight unreliable due to pilot variability.

Under ordinary circumstances it is difficult to determine the reliability of a student's performance on a single flight since:

(a) If one "check pilot" rates a student pilot on two successive flights his rating on the second flight may well be influenced by his recall of the student's performance on the first flight, and

(b) if ratings by two "check pilots" who flew with the same student at different times are compared, it is extremely difficult, if at all possible, to determine whether the source of variation between the ratings lies in variation in flight performance by the student, or in the variation between ratings by the two observers.

Although this investigation does not provide the necessary data, the method of photographic recording[46] and the rating procedures used in this study provide excellent means for determining the reliability of a single flight, since independent ratings of successive flights can be obtained from more than one observer and the identity of the pilot can be concealed by blocking out temporarily the identification data in the film.[47]

B. Evaluation in Terms of Validity. The present investigation provides no experimental data as to the validity of the criterion measures provided by the photographic method. For this reason, the justification for their present use as criterion data must come from a consideration of their "logical," or "face" validity.

1. Adequacy of the Photographic Records. As pointed out in the companion report on graphic recording of pilot performance, a "satisfactory performance as a pilot involves a balanced combination of judgment, skill, and emotional stability under the conditions of flight."[48] These three major aspects are stressed

---

[46]It should be noted that, for a study of this problem, similar advantages are possessed by the use of graphic recorders of pilot performance. Viteles, Morris S. and Backstrom, Oscar, Jr. op cit.

[47]Such a study is now in progress as a part of the 1943 Midwest-Navy Training Project in which two successive flights of each subject are being photographed. For this study the reliability not only of the flight as a whole but also of specific characteristics of flight performance will be determined.

[48]Viteles, Morris S. and Backstrom, Oscar, Jr. p. viii, op. cit.

by both a logical analysis of the task of piloting a plane and by a factor analysis of flight instructors' ratings.[49]

The photographic records are ideally suited for the analysis of data descriptive of the "skill" aspect of flying. The camera field[50] provides complete information concerning the attitude and performance of the plane in terms of airspeed, R.P.M., position of nose on horizon, angle of bank, rate of climb, changes in altitude, rate of turn, and extent of slip or skid. In addition, the pilot's adjustments of the three controls (rudder, aileron, and elevator) are recorded. The progress of the plane <u>through the air</u> can be accurately reconstructed by motion projection of the films and the permanent record enables the performance to be observed as many times as desired.

A possible objection to ratings of pilots on the basis of observation of motion photographs of flight performance is that direct evidence of pilot judgment, and of ability to fly the plane in relation to objects on the ground, is not furnished by the method. For example, motion photography does not make it possible to determine directly whether or not a pilot will fly around, or through, a thunderstorm, or whether he stays the correct distance from the pylon when flying the "figure eight" maneuver.

A limited amount of indirect evidence on pilot judgment, however, may be obtained from photographic records. It may be possible to question the judgment as well as the skill of that student pilot who jams the throttle forward abruptly, of one who executes a climbing turn with the plane just above the stalling point, or of the student pilot who makes his approach to a landing in a "grave-yard" or extremely flat glide.

As is evident from the method of analysis of the photographs as described on Page 3, however, the orientation in examining the photographic records was primarily toward performance (i.e., pilot proficiency) as evaluated by observation of control movements and plane performance. No effort was made to evaluate elements of judgment as distinct from those of skill, involved in the correct performance of flight maneuvers. In the evaluation of maneuvers not related to ground patterns, the photographic records would seem to have considerable merit as compared with other graphic methods and with direct observation during flight.

---

[49] Kelly, E. Lowell. <u>The development of a scale for rating pilot competency.</u> Washington, D. C.: Civil Aeronautics Administration Division of Research, Report No. 13. July, 1943.

[50] See Exhibit 2, Page 6.

With respect to the "emotional stability" of the pilot, the photographs obviously provide no information, except in the rare instances in which the pilot "blows up" and makes random movements of the controls. However, it is difficult to measure this aspect of performance by any technique other than subjective judgments based upon direct observation over long periods of time.

2. Nature of the Criterion Measures. The method used in evaluating the recorded performance involved essentially a comparison of how the maneuvers in the standard flight were performed against how they should have been performed. This comparison resulted in ratings representing the judgment of the observer (or rater) as to how closely the observed performance met the standards of performance set up by the observer.

In this rating situation, the "ideal" performance was determined by (a) principles of aerodynamics, e.g., turns should be made so that the plane neither slips nor skids, direction should be maintained in straight maneuvers, wings should remain level in level maneuvers, altitude should be maintained during level turns, etc.; and (b) certain specifications, e.g., medium turn with a bank of between 30 and 45 degrees, release of back pressure soon after break in Normal Power-off Stall, etc.[51]

Two precautions were taken so that both of the two observers (or raters) in this investigation would develop and maintain the same relative "standards" against which to rate the observed performance:

(a) A Manual of Rating Procedure[52] was prepared which described the rating procedure and listed (for each maneuver) the "critical" observations to be made.

(b) The rating procedure provided (in the case of the overall ratings) a cross-checking procedure by means of which all cases of discrepancies in rating between the two observers were reviewed jointly and a "combined" rating made.

---

[51]As noted in the report on graphic records, "records of airplane attitude during flight can be meaningful as criterion data in evaluating pilot performance only if maneuvers to be performed are fully set forth in rigid prescriptions to the pilot (or reported by the pilot to the experimenter). Criterion data then are essentially represented by the deviations of the pilot's performance from these prescriptions." For further discussion of the use of records as a source of criterion data see Viteles and Backstrom. Pages 1 and 2, op. cit.

[52]Presented in Appendix C.

The rating procedure, however, did not involve the use of an explicit method of evaluation, in the sense that a predetermined system of differential weighting of observed deviations from the "ideal" performance was rigorously applied. The evaluations, in the form of the assigned ratings, were subjective and based upon implicit standards which were developed by the two observers involved, and which may or may not agree with those of other observers.

The ratings are thus "general impressions," designed to take into account the dynamic aspects of the performance as a "gestalt." That these general impressions were subject to the varying influences of "successive contrast," personal biasses and prejudices, arbitrary weighting of specific faults, etc. is undeniable. A more "objective" evaluation, i. e., in terms of numerical weights assigned to specific items of performance, would eliminate some of these weaknesses. It seemed desirable, however, at least for this investigation, to utilize an over-all rating procedure, based upon careful observation of accurate recording of the performance being rated and upon a cross-checking technique in which two observers jointly re-rated the performance of cases rated differently during their independent rating.

C. Evaluation in Terms of Discriminating Value. The usefulness of a criterion depends somewhat upon the degree to which it is able to differentiate levels of success, or performance. For example, a commonly used criterion is that based upon success or failure in a training program, the Pass-Fail criterion. It has limited value, however, except when large groups are involved and when the proportion of "failures" is relatively large. It has limited application in research involving C.P.T. students where the washout rate is negligible. In addition it does not provide levels, or grades, of success, merely seperating the "sheep" from the "goats."

The Criterion Ratings, as provided by this study, permit only as many levels as the rating scale provided, namely, three for the A, B, C rating and two for the Upper-Lower rating. Combined, they permit a division of the group into quartiles (A, Upper B, Lower B, and C) but the results do not warrant dividing the "B's" into sub-groups. They have been shown, however, to differentiate among levels of performance of "successful" student pilots.

Since the Criterion Flight Scores are continuous, their discriminating efficiency is determined by their range and reliability. As pointed out previously,[53] the Criterion Flight Scores can be said to place individuals accurately into at least 6 categories.

For practical purposes, and in the light of our present knowledge of pilot proficiency, measures of performance which reliably identify the

[53] See Page 26.

"cream of the crop," the "average" group, and the "border-line cases" of a total population of "passing" student pilots are of sufficient value (if validity can be logically assumed) to warrant their use as criterion data in studies of predictors, training methods, etc.

D. Evaluation in Terms of Practicality. The photographic method as a means of obtaining criteria is admittedly one with practical limitations which restricts its use to research rather than field use. Its practical limitations are as follows:

1. It requires a plane with specially installed photographic equipment.

2. It is expensive, the film and equipment being costly.

3. The records are not immediately available, requiring time for printing of the films.

4. The length of the flight is limited by the film supply.

5. The time required by the rating procedure is rather long.

On the other hand, when time and cost are relatively unimportant in comparison with accuracy and adequacy of data and when basic research is the primary aim, the photographic method (and the resulting criterion data) have definite advantages:

1. The flight performance is permanently recorded in a form readily interpretable, i.e., in terms of standard units such as airspeed, rate of climb in feet per minute, excursions of the ball in a standard ball-bank instrument, etc.

2. Opportunity is provided for more than one observer to analyze the same performance and for more than one viewing by the same person without the contamination resulting from knowledge of the identity of the subject.

IN SUMMARY: The Criterion Ratings and Criterion Flight Scores yielded by the photographic method may be recommended for use as one type of criterion data in aviation research in that they exhibit satisfactory reliability, provide measures of the "skill" aspect of piloting, and differentiate among levels of performance of "successful" student pilots. Practical limitations of the photographic method restrict its use as a source of criterion data to basic research where time and cost are relatively unimportant and accuracy of data is the primary aim.

FURTHER DEVELOPMENTS IN THE USE OF THE PHOTOGRAPHIC METHOD

As pointed out in the Introduction, inspection during motion projection of the film can also be directed toward the observation and recording of specific items of performance such as "slips during entry," "loss of altitude during steep turns," "angle and maintence of bank in turns," etc.

This procedure has the basic advantage that the data obtained are primarily descriptive rather than evaluative in nature and questions of personal bias of the observer, shifting standards, etc. do not arise. The performance as a whole could then be "scored" if an explicit and experimentally determined system of differential weighting of specific items or "patterns" of items were prepared.

Such a procedure (involving the preparation and use of check sheets, analogous to the Ohio State Flight Inventory) is now being used in connection with the Midwest-Navy Training Project. From this current project (involving a larger number of subjects) will be obtained data which will provide:

1. Further information as to the reliability of the photographic method of analysis of flight performance.

2. Comparison of items of flight performance as revealed by the photographic records and as checked on C.A.A.-A.C.A. Form 342Z by inspectors on the basis of direct observation during the photographed flight.

3. Information concerning the reliability of the single flight as a work-sample of flight proficiency.

4. Comparison of groups trained with and without use of flight inventories as training aids.

ADDITIONAL APPLICATIONS OF THE PHOTOGRAPHIC METHOD

The photographic method is particularly applicable in research problems in which permanent recording and detailed analysis of flight performance are required. Besides the examples given above, other problems suggest themselves which could profitably be attacked through the application of the photographic method. Some of the possible applications are as follows:

1. Determination of the extent of dependence upon flight instruments, by comparisons of flight performance with and without vision of instruments.

2. Study of the differences in performance among experienced and inexperienced pilots, especially instructors and inspectors, leading to greater standardization of flight instruction and inspection.

3. Determination of the amount of "inspectoritis," by comparison of student performance with and without an inspector in the plane.

4. Comparison of photographic and graphic records of flight performance.

5. Comparison of student performance at various stages of the training program.

In addition, the photographic method has practical application as a field (as well as research) technique, especially in connection with flight instruction. Possible applications in this area are:

1. Development of an instructional film as a visual training aid. Ideal performance of maneuvers could be filmed and made available for study on the ground. Common errors and their effect on flight performance could be visually presented.

2. Use of a photographic plane as a field instrument for analysis of student pilots having difficulty with a given maneuver. Photographs taken during flight could be used in a manner similar to the Monday morning analysis by a coach of Saturday's football game.

3. Use of photographic records in doubtful cases of washout. The films could serve as objective evidence for consideration by review boards.

APPENDIX A

DESCRIPTION OF STANDARD FLIGHTS USED IN THE 1942 MIDWEST PROJECT

Prepared by

Morris S. Viteles
and
Albert S. Thompson

University of Pennsylvania
Philadelphia, Pennsylvania

STANDARD FLIGHT A

(Revised May 15, 1942, in accordance with changes in C.P.T.P. described in C.A.A. General Release No. 4, "Supplementary Instructions for Civilian Pilot Training Flight Courses, Spring Session.")

This flight includes the basic maneuvers taught in Stage A of the C.P.T. program and describes the flight to be made by the student pilots while being observed and photographed.

The "critical maneuvers," representing those maneuvers which are being studied, are designated by numbers and are capitalized. The intervening maneuvers are "transition maneuvers," representing those maneuvers whose purpose is to get the plane into position for the next critical maneuver.

Order and Description of Maneuvers

1. TAXI: to take-off line plus pivot turn for observation of approaching aircraft and turn back into wind.

2. TAKE-OFF: into wind or as near as appropriate runway allows.

3. STRAIGHT CLIMB: level off in accordance with local requirements as to altitude and distance from field boundary.

   Transition Maneuver A: depart from airport in accordance with local traffic rules and proceed to practice area. While flying to practice area make certain that the plane is properly trimmed for straight and level flight.

4. 90° LEFT CLIMBING TURN & RECOVERY: enter from and recover in straight climb, beginning at signal from observer, and recovering when 90° from direction at entry.

   Transition Maneuver B: proceed to correct location for succeeding maneuver attaining 1000' altitude and heading into wind.

5. STRAIGHT GLIDE: into wind with optimum gliding angle, beginning at 1000' and recovering at signal from observer.

   Transition Maneuver C: obtain altitude of 700' and head into wind.

6. 90° MEDIUM RIGHT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting into wind and ending in cross-wind direction.

   Transition Maneuver D: short straight and level flight at 700', heading in cross-wind direction.

7. 90° MEDIUM LEFT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting cross-wind and ending into wind.

   Transition Maneuver E: short straight and level flight, at 700' altitude, heading into wind.

8. 90° MEDIUM LEFT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting into wind and ending in cross-wind direction.

Transition Maneuver F: short straight and level flight at 700', heading cross-wind.

9. 90° MEDIUM RIGHT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting cross-wind and ending into wind.

Transition Maneuver G: short straight and level flight at 700' heading into wind.

10. STRAIGHT CLIMB: into wind with optimum climbing angle, starting at 700' altitude and leveling off at signal from observer.

Transition Maneuver H: obtaining altitude of 1000', turn down wind.

11. STRAIGHT AND LEVEL FLIGHT: down wind at cruising speed at altitude of 1000', obtaining straight and level flight as soon as possible and maintaining straight and level flight until observer gives signal for next maneuver.

Transition Maneuver I: upon instruction from observer, proceed to cross-roads suitable for S-turns and rectangular course, turning into wind en route.

12. S-TURNS ACROSS ROAD: begin into wind and make first 180° turn to right followed by a 180° left turn. End maneuver heading into wind.

Transition Maneuver J: upon instruction from observer, proceed to correct location for entering rectangular course.

13. RECTANGULAR COURSE: begin into wind, parallel to one side of the rectangular course, and make four turns to left, ending maneuver immediately after fourth turn and when heading in same direction as at entry.

Transition Maneuver K: upon instruction from observer, return to airport according to local flight rules.

14. 180° APPROACH TO LANDING: from altitude specified by local flight rules. Close throttle opposite intended landing spot and make two separate 90° turns.

15. FINAL APPROACH AND LANDING: into wind, beginning final approach at least 1000' from airport boundary.

16. LANDING RUN: at least 100' without use of brakes.

17. PIVOT TURN AND TAXI: to desired position.

STANDARD FLIGHT B

(Revised May 15, 1942, in accordance with changes in C.P.T.P. described in CAA. General Release No. 4, "Supplementary Instructions for Civilian Pilot Training Flight Courses, Spring Session.")

This flight is designed for use in Stage B and is composed of the maneuvers included in Flight A with the addition of Power-off Stall and Elementary Eight.

The "Critical maneuvers," representing those maneuvers which are being studied, are designated by numbers and are capitalized. The intervening maneuvers are "transition maneuvers," representing those maneuvers whose purpose is to get the plane into position for the next critical maneuver.

<u>Order and Description of Maneuvers</u>

1. TAXI: to take-off line plus pivot turn for observation of approaching aircraft and turn back into wind.

2. TAKE-OFF: into wind or as near as appropriate runway allows.

3. STRAIGHT CLIMB: Level off in accordance with local requirements as to altitude and distance from field boundary.

   Transition Maneuver A: depart from airport in accordance with local traffic rules and proceed to practice area. While flying to practice area make certain that the plane is properly trimmed for straight and level flight.

4. 90° LEFT CLIMBING TURN & RECOVERY: enter from and recover in straight climb, beginning at signal from observer, and recovering when 90° from direction at entry.

   Transition Maneuver B: proceed to correct location for succeeding maneuver, attaining 2000' altitude and heading into wind.

5. NORMAL POWER-OFF STALL AND RECOVERY: without use of ailerons.

   Transition Maneuver C: reduce altitude to 1000' and head into wind, in appropriate location for the succeeding maneuver.

6. STRAIGHT GLIDE: into wind with optimum gliding angle, beginning at 1000' and recovering at signal from observer.

   Transition Maneuver D: obtain altitude of 700' and head into wind.

7. 90° MEDIUM RIGHT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting into wind and ending in cross-wind direction.

   Transition Maneuver E: short straight and level flight at 700' heading in cross-wind direction.

Standard Flight B (Revised 5/15/42)

8. 90° MEDIUM LEFT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting cross-wind and ending into wind.

Transition Maneuver F: short straight and level flight, at 700' altitude, heading into wind.

9. 90° MEDIUM LEFT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting into wind and ending in cross-wind direction.

Transition Maneuver G: short straight and level flight at 700', heading cross-wind.

10. 90° MEDIUM RIGHT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting cross-wind and ending into wind.

Transition Maneuver H: short straight and level flight at 700', heading into wind.

11. STRAIGHT CLIMB: into wind with optimum climbing angle, starting at 700' altitude and leveling off at signal from observer.

Transition Maneuver I: obtaining altitude of 1000', turn down wind.

12. STRAIGHT AND LEVEL FLIGHT: down wind at cruising speed at altitude of 1000', obtaining straight and level flight as soon as possible, and maintaining straight and level flight until observer gives signal for next maneuver.

Transition Maneuver J: upon instruction from observer, proceed to cross-roads suitable for S-turns and rectangular course, turning into wind en route.

13. S-TURNS ACROSS ROAD: begin into wind and make first 180° turn to right, followed by a 180° left turn. End maneuver heading into wind.

Transition Maneuver K: upon instruction from observer, proceed to correct location for entering rectangular course.

14. RECTANGULAR COURSE: begin into wind, parallel to one side of the rectangular course, and make four turns to left, ending maneuver immediately after fourth turn and when heading in same direction as at entry.

Transition Maneuver L: after completing rectangular course make a Half-Eight, beginning cross-wind. Proceed to correct position for entry to elementary eight.

Standard Flight B (Revised 5/15/42)

15. ELEMENTARY EIGHT: make one complete eight, beginning with right turn and passing through intersection in level flight.

Transition Maneuver M: upon instruction from observer, return to airport. Proceed to position appropriate for closing throttle for 180° side approach.

16. 180° SIDE APPROACH: close throttle opposite landing spot.

17. FINAL APPROACH AND LANDING: into wind, beginning Final Approach at least 1000' from airport boundary.

18. LANDING RUN: at least 100' without use of brakes.

19. TURN AND TAXI: to desired position.

STANDARD FLIGHT C

(Revised May 15, 1942, in accordance with changes in C.P.T.P. described in C.A.A. General Release No. 4, "Supplementary Instructions for Civilian Pilot Training Flight Courses, Spring Session.")

This flight is designed for use in Stage C and is composed of the maneuvers included in Flights A and B with the addition of 360° Left and Right Steep Turns.

The "critical maneuvers," representing those maneuvers which are being studied, are designated by numbers and are capitalized. The intervening maneuvers are "transition maneuvers," representing those maneuvers whose purpose is to get the plane into position for the next critical maneuver.

<u>Order and Description of Maneuvers</u>

1. TAXI: to take-off line plus pivot turn for observation of approaching aircraft and turn back into wind.

2. TAKE-OFF: into wind or as near as appropriate runway allows:

3. STRAIGHT CLIMB: level off in accordance with local requirements as to altitude and distance from field boundary.

   Transition Maneuver A: depart from airport in accordance with local traffic rules and proceed to practice area. While flying to practice area make certain that the plane is properly trimmed for straight and level flight.

4. 90° LEFT CLIMBING TURN & RECOVERY: enter from and recover in straight climb beginning at signal from observer, and recovering when 90° from direction at entry.

   Transition Maneuver B: proceed to correct location for succeeding maneuver attaining 2000' altitude and heading into wind.

5. STEEP 360° LEFT TURN: at 2000' beginning into wind and recovering completely.

   Transition Maneuver C: short straight and level flight after recovery.

6. STEEP 360° RIGHT TURN: beginning into wind and recovering completely.

   Transition Maneuver D: short straight and level flight after recovery.

7. NORMAL POWER-OFF STALL AND RECOVERY: without use of ailerons.

   Transition Maneuver E: reduce altitude to 1000' and head into wind.

8. STRAIGHT GLIDE: into wind with optimum gliding angle, beginning at 1000' and recovering at signal from observer.

Transition Maneuver F: obtain altitude of 700' and head into wind.

9. 90° MEDIUM RIGHT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting into wind and ending in cross-wind direction.

Transition Maneuver G: short straight level flight at 700', heading into cross-wind direction.

10. 90° MEDIUM LEFT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting cross-wind and ending into wind.

Transition Maneuver H: short straight and level flight, at 700' altitude, heading into wind.

11. 90° MEDIUM LEFT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting into wind and ending in cross-wind direction.

Transition Maneuver I: short straight and level flight at 700', heading cross-wind.

12. 90° MEDIUM RIGHT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting cross-wind and ending into wind.

Transition Maneuver J: short straight and level flight at 700', heading into wind.

13. STRAIGHT CLIMB: into wind with optimum climbing angle, starting at 700' altitude and leveling off at signal from observer.

Transition Maneuver K: obtaining altitude of 1000', turn down wind.

14. STRAIGHT AND LEVEL FLIGHT: down wind at cruising speed at altitude of 1000', obtaining straight and level flight as soon as possible and maintaining straight and level flight until observer gives signal for next maneuver.

Transition Maneuver L: Upon instruction from observer, proceed to cross-roads suitable for S-turns and rectangular course, turning into wind en route.

15. S-TURNS ACROSS ROAD: begin into wind and make first 180° turn to right followed by a 180° left turn. End maneuver heading into wind.

Transition Maneuver M: Upon instruction from observer, proceed to correct location for entering rectangular course.

Standard Flight C (Revised 5/15/42)

16. RECTANGULAR COURSE: begin into wind, parallel to one side of the rectangular course, and make four turns to left, ending maneuver immediately after fourth turn and when heading in same direction as at entry.

Transition Maneuver N: after completing rectangular course, make a Half-Eight beginning cross-wind, followed by an Elementary Eight. Upon completion of Elementary Eight, proceed to position for entry to Two-Bank Eight.

17. TWO-BANK EIGHT: make one complete Two-Bank Eight with first turn to left.

Transition Maneuver O: upon instruction from observer, return to airport. Proceed to position appropriate for closing throttle for 180° side approach.

18. 180° SIDE APPROACH: close throttle opposite landing spot.

19. FINAL APPROACH AND LANDING: into wind, beginning Final Approach at least 1000' from airport boundary.

20. LANDING RUN: at least 100' without use of brakes.

21. TURN AND TAXI: to desired position.

STANDARD FLIGHT D

(Revised May 15, 1942, in accordance with changes in C.P.T.P. described in C.A.A. General Release No. 4, "Supplementary Instructions for Civilian Pilot Training Flight Courses, Spring Session.")

This flight is designed for use in Stage D and is composed of the maneuvers included in Flights A, B, and C with the addition of Forward Slip.

The "critical maneuvers," representing those maneuvers which are being studied, are designated by numbers and are capitalized. The intervening maneuvers are "transition maneuvers," representing those maneuvers whose purpose is to get the plane into position for the next critical maneuver.

Order and Description of Maneuvers

1. TAXI: to take-off line plus pivot turn for observation of approaching aircraft and turn back into wind.

2. TAKE-OFF: into wind or as near as appropriate runway allows.

3. STRAIGHT CLIMB: level off in accordance with local requirements as to altitude and distance from field boundary.

   Transition Maneuver A: depart from airport in accordance with local traffic rules and proceed to practice area. While flying to practice area make certain that the plane is properly trimmed for straight and level flight.

4. 90° LEFT CLIMBING TURN & RECOVERY: enter from and recover in straight climb, beginning at signal from observer and recovering when 90° from direction at entry.

   Transition Maneuver B: proceed to correct location for succeeding maneuver attaining 2000' altitude and heading into wind.

5. STEEP 360° LEFT TURN: at 2000' beginning into wind and recovering completely.

   Transition Maneuver C: short straight and level flight after recovery.

6. STEEP 360° RIGHT TURN: beginning into wind and recovering completely.

   Transition Maneuver D: short straight and level flight after recovery.

7. NORMAL POWER-OFF STALL AND RECOVERY: without use of ailerons.

   Transition Maneuver E: straight and level flight at whatever altitude obtained during recovery from stall.

Standard Flight D (Revised 5/15/42)

8. FORWARD SLIP: into wind with 30° left bank, recovering at signal from observer.

Transition Maneuver F: reduce altitude to 1000' and head into wind in appropriate position for succeeding maneuvers.

9. STRAIGHT GLIDE: into wind with optimum gliding angle, beginning at 1000' and recovering at signal from observer.

Transition Maneuver G: short straight and level flight at 700', heading into wind.

10. 90° MEDIUM RIGHT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting into wind and ending in cross-wind direction.

Transition Maneuver H: short straight and level flight at 700', heading into cross-wind direction.

11. 90° MEDIUM LEFT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting cross-wind and ending into wind.

Transition Maneuver I: short straight and level flight, at 700' altitude, heading into wind.

12. 90° MEDIUM LEFT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting into wind and ending in cross-wind direction.

Transition Maneuver J: short straight and level flight at 700' and with 45° heading cross-wind.

13. 90° MEDIUM RIGHT TURN AND RECOVERY: at altitude of 700' and with 45° bank, starting cross-wind and ending into wind.

Transition Maneuver K: short straight and level flight at 700', heading wind.

14. STRAIGHT CLIMB: into wind with optimum climbing angle, starting at 700', altitude and leveling off at signal from observer.

Transition Maneuver L: maintaining altitude of 1000' turn down wind.

15. STRAIGHT AND LEVEL FLIGHT: down wind at cruising speed at altitude of 1000', obtaining straight and level flight as soon as possible and maintaining straight and level flight until observer gives signal for next maneuver.

Transition Maneuver M: upon instruction from observer, proceed to cross-roads suitable for S-turns and rectangular course, turning into wind en route.

APPENDIX B

GENERAL INSTRUCTIONS FOR PHOTOGRAPHING STANDARD FLIGHTS

Prepared by

Morris S. Viteles
and
Albert S. Thompson

University of Pennsylvania
Philadelphia, Pennsylvania

GENERAL INSTRUCTIONS FOR PHOTOGRAPHING STANDARDS FLIGHTS

Field work in connection with N.R.C. research projects at the University of Pennsylvania and at Boston has shown that motion pictures of sufficient clarity to permit micromotion analysis can be taken of student performance during standard flights. Satisfactory photographs can be obtained if the operator carefully carries out the following suggestions based upon practical experience to date:

A. Care of equipment:

1 - Keep camera oiled according to standard instructions. The Bell & Howell 70 camera should be oiled after each 600' of film.
2 - Keep the lens dry and clean.
3 - Check the camera mount for indications of vibration.
4 - Keep a good supply of film on hand. 200' reels are difficult to obtain quickly.
5 - Make periodic checks of the control recorder to see that the zero positions are correct and that the cable system is running smoothly and freely.
6 - Set up definite provisions for re-charging battery at necessary intervals.
7 - Keep extra supplies on hand, especially bulbs for spotlight.

B. Before each flight:

1 - Check amount of unexposed film in camera.
2 - Read light meter on instrument panel and control recorder. In doing so, have plane facing into the wind since this position best represents the conditions of illumination occurring during most of the maneuvers.
3 - Set lens diaphragm in accordance with obtained light reading and camera speed.
4 - Check camera speed (8 frames per second).
5 - Run camera momentarily (just before starting flight) as final check.

C. During flight:

1 - Follow carefully the accompanying detailed procedures for photographing each of the standard flights. Make sure that the correct maneuver numbers appear for each maneuver. Note that a given maneuver has the same number in each of the four flights. This provides for ease in identification of maneuvers but necessitates the omission of some of the numbers [illegible]. For example, since Flight A does not include steep turns, stalls or slips, the maneuver numbers 5, 6, 7, and 8 will not appear in Flight A photographs.
2 - Be careful to include all of the entry and all of the recovery in the photographs. The tendency to stop the camera too soon must be guarded against. Otherwise, [illegible] of turns such as the Medium Turns or the Steep Turns, [illegible] for the pilot to "roll" from one turn to the next, [illegible] the control movements which occur during an entry from straight and level flight. Ideally, each maneuver should begin and end [illegible] level flight and with the controls around neutral.

3 - Jot down (on the back of the flight inventory) any irregularities which occur, such as interference with flight path, unusual air conditions during a specific maneuver, etc. Any observations which will aid in the interpretation of the film, however minute, should be included.

4 - Keep out of the camera field at all times, except in cases of emergency. Determine experimentally how far to the right you must sit in order to stay out of the field and then limit your body movements accordingly. Remember that a portion of your flying gear may "stick out" into the field, thus covering up some important flight instrument.

D. Instruction for specific maneuvers

1 - In Straight Glide, Straight Climb, Forward Slip, and Straight and Level Flight the specifications call for recovery after a signal from observer. The procedure for these maneuvers should be as follows.

a. At signal to subject for beginning of maneuver, start camera and stopwatch.
b. After 30 seconds, signal subject to recover from maneuver.
c. Leave camera running until recovery is complete.

2 - In the 90° Medium Turn and 720° Steep Turns, if the subject fails to recover when 30 degrees beyond the required recovery direction, the observer should signal the subject to recover. The camera should be left running until recovery and on the flight inventory the error of recovery should be designated as 30°+.
Essentially this means that a maximum tolerance of 30° is allowed in both 90° and 720° turns.
This procedure will provide for complete photography of turns and recovery while reducing the possibility of long sections of film due to error in recovering after the required number of degrees of turn.

3 - During certain maneuvers, a signal light (mounted on the instrument panel within the camera field and activated by a switch on the front cock-pit stick) should be turned on momentarily so as to identify the location of significant portions of the maneuvers, as follows:

a. Take-off - light on when wheels leave ground.
b. Stall - light on at moment of stall.
c. Power Turns - light on at moment of hitting slip-stream.
d. Landing - light on when wheels touch ground for first time.
e. Whenever observer takes over controls in case of emergency or need for correction.

E. Photographing of comparison flight data.

The observer should run a short test flight at the beginning of each day (or repeated during the day, if necessary) in order to obtain flight instrument data against which to compare the student flights of that series. This flight will include short series of each type of: (1) optimum climb, (2) optimum glide, (3) straight and level at cruising speed, (4) straight and level at maximum speed.

These flights should be identified by number and each of the subject flights of the following series should be identified by that number as well as the student's number and flight letter.

PROCEDURE FOR PHOTOGRAPHING STANDARD FLIGHT B

Place name card of subject in card holder. Draw a heavy black circle around the letter designating the flight to be taken. Expose the number 2 in the maneuver number indicator. Then photograph maneuvers as follows:

Man. No.	Name of Maneuver	
1	TAXI -	not photographed.
2	TAKE-OFF -	start camera at beginning of Take-off run. Leave camera running as plane leaves ground.
3	STRAIGHT CLIMB -	Stop camera when 100' altitude is reached. Change maneuver number to 4, (omitting number 3).
4	90° LEFT CLIMBING TURN -	start camera just before signal to subject for turn. Stop camera after complete recovery. Change maneuver number to 7, (omitting numbers 5 and 6).
7	POWER-OFF STALL & RECOVERY -	start camera just before signal to subject for stall. Stop camera after complete recovery. Change number to 9, (omitting number 8).
9	STRAIGHT GLIDE -	start camera just before signal to subject for glide. Stop camera after complete recovery. Change number to 10.
10	90° MEDIUM RIGHT TURN -	start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 11
11	90° MEDIUM LEFT TURN -	start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 12.
12	90° MEDIUM LEFT TURN -	start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 13
13	90° MEDIUM RIGHT TURN -	start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 14.
14	STRAIGHT CLIMB -	start camera just before signal to subject for climb. Stop camera after complete recovery. Change number to 15.
15	STRAIGHT & LEVEL FLIGHT -	start camera just before signal to subject. After 30 seconds of Straight & Level Flight stop camera. Change maneuver number to 16.
	S-TURNS ACROSS ROAD -	not photographed.
	RECTANGULAR COURSE -	not photographed.
	Elementary Eight -	not photographed.
	180° APPROACH -	not photographed.
16	FINAL APPROACH & LANDING -	start camera as pilot heads into wind from the final turn. Stop camera after several seconds of Landing Run.
	LANDING RUN -	partially photographed as provided in 16 above.
	TURN & TAXI -	not photographed.

PROCEDURE FOR PHOTOGRAPHING STANDARD FLIGHT C

Place name card of subject in card holder. Draw a heavy black circle around the letter designating the flight to be taken. Expose the number 2 in the maneuver number indicator. Then photograph maneuvers as follows:

Man. No.	Name of Maneuver
1	TAXI - not photographed.
2	TAKE-OFF - start camera as beginning of Take-off. Leave camera running as plane leaves ground.
3	STRAIGHT CLIMB - stop camera when 100' altitude is reached. Change maneuver number to 4, (omitting number 3).
4	90° LEFT CLIMBING TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change maneuver number to 5.
5	360° STEEP LEFT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 6.
6	360° STEEP RIGHT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 7.
7	POWER-OFF STALL & RECOVERY - start camera just before signal to subject for stall. Stop camera after complete recovery. Change number to 9, (omitting number 8).
9	STRAIGHT GLIDE - start camera just before signal to subject for glide. Stop camera after complete recovery. Change number to 10.
10	90° MEDIUM RIGHT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 11.
11	90° MEDIUM LEFT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 12.
12	90° MEDIUM LEFT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 13.
13	90° MEDIUM RIGHT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 14.
14	STRAIGHT CLIMB - start camera just before signal to subject for climb. Stop camera after complete recovery. Change number to 15.
15	STRAIGHT & LEVEL FLIGHT - start camera just before signal to subject. After 30 seconds of Straight & Level Flight stop camera. Change maneuver number to 16.
	S-TURNS ACROSS ROAD - not photographed.
	RECTANGULAR COURSE - not photographed.
	TWO-BANK EIGHT - not photographed.
	180° APPROACH - not photographed.
16	FINAL APPROACH & LANDING - start camera as pilot heads into wind from the final turn. Stop camera after several seconds of Landing Run.
	LANDING RUN - partially photographed as provided in 16 above.
	TURN & TAXI - not photographed.

PROCEDURE FOR PHOTOGRAPHING GRADED FLIGHT

Place name card of subject in card holder. Draw a heavy black circle around the letter designating the flight to be taken. Expose the number 2 in the maneuver number indicator. Then photograph maneuvers as follows:

Man. No.	Name of Maneuver
1	TAXI - not photographed.
2	TAKE-OFF - start camera at beginning of take-off run. Leave camera running as plane leaves ground.
3	STRAIGHT CLIMB - stop camera when 100' altitude is reached. Change maneuver number to 4, (omitting number 3).
4	90° LEFT CLIMBING TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change maneuver number to 5.
5	360° STEEP LEFT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 6.
6	360° STEEP RIGHT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 7.
7	POWER-OFF STALL & RECOVERY - start camera just before signal to subject for stall. Stop camera after complete recovery. Change number to 8.
8	FORWARD SLIP - start camera just before signal to subject for slip. Stop camera after complete recovery. Change number to 9.
9	STRAIGHT GLIDE - start camera just before signal to subject for glide. Stop camera after complete recovery. Change number to 10.
10	90° MEDIUM RIGHT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 11.
11	90° MEDIUM LEFT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 12.
12	90° MEDIUM LEFT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 13.
13	90° MEDIUM RIGHT TURN - start camera just before signal to subject for turn. Stop camera after complete recovery. Change number to 14.
14	STRAIGHT CLIMB - start camera just before signal to subject for climb. Stop camera after complete recovery. Change number to 15.
15	STRAIGHT & LEVEL FLIGHT - start camera just before signal to subject. After 30 seconds of Straight and Level Flight stop camera. Change maneuver number to 16.
	S-TURNS ACROSS ROAD - not photographed.
	RECTANGULAR COURSE - not photographed.
	TWO-BANK EIGHT - not photographed.
	180° APPROACH - not photographed.
16	FINAL APPROACH & LANDING - start camera as pilot heads into wind from the final turn. Stop camera after several seconds of Landing Run.
	LANDING RUN - partially photographed as provided in 16 above.
	TURN & TAXI - not photographed.

APPENDIX C

MANUAL FOR RATING PILOT PERFORMANCE THROUGH DIRECT INSPECTION OF MOTION PHOTOGRAPHS

Prepared by

Albert S. Thompson
Edwin S. Smart
Morris S. Viteles

University of Pennsylvania
Philadelphia, Pennsylvania

MANUAL FOR RATING PILOT PERFORMANCE THROUGH DIRECT INSPECTION OF MOTION PHOTOGRAPHS

In rating by inspection the performance of individuals from the data provided by the photographic records, it is required that the photographs of the entire flight be run through three times (forward, backward, and forward) at moderately fast speed. During this time observations and ratings on a five-point scale should be made on the following aspects of flight performance: (1, high, to 5, low):

1. Wing Control, e.g., steadiness of bank.
2. Nose Control, e.g., maintenance of nose on horizon.
3. Directional Control, i.e., maintenance of direction or rate of turn.
4. Tendency to slip or skid, i.e., lateral stability.
5. Maintenance of altitude.
6. Maintenance of air speed.
7. Coordination of controls.
8. Adequacy of execution of maneuvers, e.g., <u>correct</u> degree of bank, <u>complete</u> stall, <u>adequate</u> air speed, three-point tendency, etc.

The aspects can be defined as follows:

1. Wing Control – refers to the plane's lateral stability during maneuvers, e.g., whether the plane's wings are consistently level during straight and level flight, whether or not the degree of bank varies markedly during a turn, etc. Cues for rating this maneuver come primarily from the artificial horizon.

2. Nose Control – refers to the student's ability to keep the nose in the proper relationship to the horizon, e.g., steadily above the horizon during a climb, steadily on the horizon during turns, etc. This variable can also be best rated by reference to the artificial horizon.

3. Directional Control – refers to the plane's directional stability, e.g., whether or not the plane turns during straight and level flight, straight climb, straight glide, and the maintenance of rate of turn during turns, etc. This variable can best be rated by reference to the Rate of Turn Indicator.

4. Tendency to slip or skid. The ball bank instrument is the best indicator of this tendency. It is useful, however, sometimes to refer to the coordination of movement of the Rate of Turn Indicator and the banking movement as indicated by the artificial horizon. Usually the rate of turn has a slight lag. If it is observed that the Rate of Turn Indicator moves at the same time as does the horizon, one can be sure that the student has too great a rate of turn for his bank and is skidding.

5. Maintenance of altitude. This variable can best be rated from (1) noting changes in the sensitive altimeter, and (2) noting movements in the rate of climb indicator, although this latter instrument has considerable lag.

6. Maintenance of proper air speed. This variable obviously can best be rated through reference to the air speed indicator.

7. Control adjustment. This refers to the coordination exhibited in movement of the controls. Tendencies to lead with rudder or aileron, to hold rudder during a turn, to carry top rudder during a turn, whether or not rudder alone is used to maintain lateral stability during a stall, etc., should be looked for. Specific details to be noted in the various maneuvers will be discussed later.

8. Adequacy. This refers to the method in which the student performed the various maneuvers, e.g., did he have the correct amount of bank in turns; did he stall the plane completely during the stall maneuvers, etc.

After inspecting the flight of a given student, the performance should be rated on a three-point scale: A, excellent; B, average; and C, poor. It would be expected that approximately 20 percent of the students would fall in the A and C categories, and about 60 percent in the B category, although no hard and fast rule in this connection can be stated. One point should be noted: If uncertain as to whether or not to rate a student in the extreme categories, that is, if he is on the borderline, place him in the extreme category. The ratings themselves can be made either at the end of first, second, or third viewing of a given flight. If made before the end, subsequent observation may cause some ratings to be changed. In addition, the performance should be rated on a two-point scale: U, above average, and L, below average. The air conditions should also be rated as Smooth, Normal, or Rough. The ratings should be recorded on the Data Sheet.

No hard and fast rule can be presented regarding methods of viewing the films. It is suggested that on the first viewing the observer look particularly at the turn and bank indicator, artificial horizon, and control indicators. Then variations in air speed, altitude, etc., can be noted while running the film backward, and specific observations can be confirmed while running through forward again. Other systems are possible, however. For instance, the observer may find it expedient to obtain just a general picture of the performance while running it through the first time, and then confirm certain tentative hypotheses regarding the student's performance on subsequent viewings. The method used is relatively unimportant. It is important, however, that all of the salient features of the flight be noted. For this reason a list of observations pertaining to each of the eight flight variables has been prepared to direct the observer's attention and to insure complete observation of the salient features of the performance. The list is found at the end of the manual.

One further caution should be added. The fact that the air conditions cannot be held constant for all flights should be realized, and factors such as rough air may cause variations in performance which are not the fault of the student. One cue to the presence of rough air is that under such conditions the rate of climb indicator will fluctuate frequently, sometimes almost continually, but usually to only a small degree. Sharp movements of the air speed indicator pointer are also suggestive of bumpy air. Rapid gain or loss of altitude for no apparent reason indicates the presence of thermals. Reference should also be made to any notes on unusual air conditions recorded by the administrator of the flight. At the time of the rating, notes should be kept of any unusual air conditions obtaining at the time of the flight, as determined from the photographs or from the flight notes. Notes should also be made of specific items of performance, such as "lost 150' in 360° left, glided at 65 MPH, aileron in stall," etc.

COMMENTS ON OBSERVATIONS TO BE MADE DURING DIRECT INSPECTION OF FILMS

1. Wing Control: Artificial Horizon

    (a) Take-off: Disregard because of non-functioning of artificial horizon during first maneuver.

    (b) 90° Left Climbing Turn
        1. Ability to hold shallow banks is important
        2. Note over recovery of banks

    (c) 360° Steep Turns
        1. Smoothness of entry
        2. Steadiness in maintaining bank
        3. Smoothness of recovery
        4. Accuracy in returning to level - rather important, but disregard if slight and corrected quickly
        5. Right bank harder to hold than left bank

    (d) Stall
        1. Maintaining wings level during pull-up
        2. Slight drop immediately at break permissible
        3. Rapid returning of wings to level during recovery desired
        4. Note wings at bottom of dive and at end of recovery

    (e) Slip
        1. Smoothness of entry
        2. Maintenance of bank during slip, returning to level during the slip especially [illegible]
        3. Accuracy in recovery [illegible] seems to occur rather frequently in students

    (f) Straight Glide & Climb and Straight & Level Flight
        1. Maintenance of level [illegible]

    (g) Medium Turns
        1. Smoothness of entry
        2. Maintenance of bank
        3. Accuracy in recovery
        4. Rights not much harder than left
        5. Slight variations less important than in 360°

    (h) Landing
        1. Variations during gliding approach not so important as during leveling off and just before and after landing
        2. Erratic reading just after moment of landing probably [illegible] [illegible]. Confirm by noting [illegible]

2. Nose Control: Artificial [illegible]

(a) Take-off: Disregard because of non-functioning of artificial horizon during first maneuver.

(b) 90° Left Climbing Turn, Straight Glide and Straight Climb
1. Smoothness of entry
2. Maintenance of angle
3. Accuracy of recovery especially important

(c) 360° Steep Turns
1. Maintenance of nose on horizon throughout turn
2. Slight variations permissible unless high or low air speed results. Also note whether corrected. The steeper the bank, the more variation is permissible
3. Note diving in recovery. Check by noting air speed and rate of climb

(d) Stall
1. Nose held constant during climb rather than steadily increasing until break
2. Abrupt pull-up undesirable
3. Accuracy of final recovery; especially bad if nose high on recovery

(e) Slip
1. Maintenance of angle of glide during entry and recovery
2. Nose should not be very much below horizon
   Check by air speed; if higher than 60, then too steep a gliding angle

(f) Medium Turns
1. Should be on horizon throughout
2. Slight variations more important than in 360's
   Especially bad in entry and recovery, unless corrected

(g) Straight and Level Flight
1. Nose steady on horizon
2. Note corrections for maintenance of altitude, especially if there is a suggestion of thermals

3. Directional Control: Turn Indicator

(a) Take-off
1. Note particularly tendency to change direction immediately after leaving ground and at point of take-off
2. Tendency to turn left particularly bad

(b) Climbing Turn
1. Note if turn is maintained or varied
2. Particularly poor if rate of turn returns to zero

(c) 360° Steep Turns

1. Smoothness of entry and recovery - regular or in steps
2. Steadiness <u>during</u> turn
3. Note tendency to over-recover, directionally; note whether immediately corrected

(d) Stall

1. Note directional control particularly in pull-up
2. Variations permissible during dive, but not during recovery from dive

(e) Slip

1. During slip no turn allowed. Indicator will turn right during entry, left during recovery in left bank slip
2. Note tendency to over-recovery directionally

(f) Medium Turns: Same as 360° turns, except less variation allowable, particularly as regards over-recovery

(g) Straight and Level, Straight Climb & Straight Glide

1. No variation allowable, except as may be due to rough air
2. Note tendency to turn right when engine cut, left when RPM increased.

(h) Landing

1. Variations during approach permissible, due to "jockeying" around into landing position. After leveling off reached, no turn permissible
2. Note indicator just before wheels touch ground
3. Note further tendency to turn <u>immediately</u> after landing
4. Slight variations during landing run expected

4. Tendency to Slip or Skid: Ball Bank Indicator

(a) Take-off and Straight Climb

1. Any consistent or continuous tendency bad
2. Note that rough air may cause slight variations of ball; note over-correction.

(b) 90° Left Climbing Turn

1. Tendency to slip or skid bad, since extremely slight bank
2. Skidding particularly bad in climbing turn

(c) 360° Steep Turns

1. Slight slipping tendencies in entry not too bad; somewhat worse in recovery
2. Occasional slips or skids during turn may be due to rough air; watch for immediate correction
3. Prolonged tendency in either direction bad, particularly skidding
4. More variation of ball allowed than in Medium 90° turns

(d) Stall

1. Variations at break and during first part of dive permissible

(e) Slip

1. Note smoothness of entry and recovery
2. Note particularly tendency to skid after recovery
3. Steadiness rather than position of ball is important
4. Check position of ball against air speed and rate of climb

(f) Glide, Climb, and Straight and Level

1. Only variations due to rough air permissible
2. Note consistent tendencies

(g) Medium 90° Turns

1. In general, same comments as regarding 360's except less variation allowable

(h) Landings

1. Note tendency to slip and skid other than that due to rough air, which is often encountered close to ground
2. Note tendency particularly during leveling-off process

General: In notes describe extent of slip or skid in three degrees

1. Slight: more than half of ball showing within lubber lines
2. Moderate: half or less of ball showing between lubber lines
3. Extreme: ball more than half way over to end of tube. Lower limit of extreme is when ball is out of the lubber lines

5. Maintenance of Altitude: Altimeter

(a) No observation necessary in Take-off, Straight Climb, or Straight Glide

(b) 360° Steep Turns

1. Observe altitude at beginning and end, noting loss or gain; if rate of climb or nose indicates much variation, look for changes of altitude <u>within</u> the turn
2. Correction for loss or gain is better than failure to do so
3. Variation of 50' or less not serious

(c) Medium Turns: No loss or gain expected

(d) Stall: Note whether altitude is <u>lost during pull-up</u>

(e) Straight and Level Flight

1. No change expected
2. Corrected change better than non-corrected

(f) Landing: Merely check at beginning of film to see whether the photographs begin early or late in approach

Note: Variations in altitude may be due to thermals. A check can be made by comparing nose position and rate of climb readings, taking into account the lag in the Rate of Climb Indicator

6. Air speed: Air speed Indicator

(a) Take-off
1. Note air speed at point of take-off--should be higher than 40 MPH
2. Note also indications of leveling off after take-off; usually shown by air speed building up to 60 MPH or more soon after climb begun

(b) Climbing Turn
1. Indications of too shallow climb indicated by too great air speed, i.e., higher than in normal climb
2. Latter not so serious as too low air speed, e.g., 40 or 45 MPH

(c) 360° Steep Turns
1. Air speed should be at cruising or slightly below; too high or low MPH indicates nose high or low
2. Watch particularly for diving during turn (entry or recovery included)
3. More variation is permitted than in 90° Medium Turns

(d) Stall
1. Note if air speed drops smoothly
2. Note if excessive air speed built up in dive; should remain less than 75 MPH
3. Should not level off with air speed still low (45 MPH)

(e) Slip
1. Air speed during slip is underestimated, due to fact that nose is at an angle to relative wind. During slip, air speed should not be higher than normal for glide
2. Note particularly A.S. at recovery from slip; if high this indicates that the purpose of the slip has been defeated

(f) 90° Medium Turns
1. Air speed should remain at cruising, or only slightly lower
2. Note variations

(g) Glide
1. Air speed should not be over 60 or less than 50
2. Too low air speed probably worse than too high, but latter also bad
3. Note variations

(h) Climb
1. Note particularly too low air speed (under 50 MPH)
2. Note variations

(i) Straight and Level
1. Should be 70-75 MPH
2. Note variations

(j) Landing
1. Air speed during approach should be normal gliding speed
2. At leveling-off should be less than normal, i.e., under 50 MPH
3. Note air speed at moment when wheels touch ground

General: Rough air indicated by slight and consistently abrupt variations in air speed

7. Control Adjustment: Control Recorder

The emphasis is on coordination. Coordination of the controls means that correct pressures are applied to one or more controls simultaneously or in sequence, in such a manner that the plane does exactly what the pilot wants it to do. Thus, if all instruments indicate that the execution of a given maneuver is satisfactory, the chances are that control movements were coordinated. If lack of coordination is indicated, the specific factors involved should be noted. In general, the three points to be noted under control adjustment are:

(1) Temporal coordination, e.g., were rudder and aileron applied simultaneously, was the throttle coordinated with elevator control in entering the climb, etc.

(2) Over-controlling, e.g., was it necessary to "correct a correction" for a bump by applying controls first in one direction and then in the other, or to apply opposite pressures to correct for over-recovery from a turn

(3) Specific control habits, e.g., did the pilot hold right or left rudder, exhibit tendency to hold top tudder during turns, etc.

Specific points, within the general categories above, to be noted in connection with given maneuvers are as follows:

(a) Take-off
1. Note rudder action during take-off run. Considerable activity is usually required, but note tendency to "fan" or to "walk" the rudder. Little aileron activity expected
2. Look for suggestions of raising tail too soon or too late and for suggestions of "pulling" plane off ground

(b) Climbing Turn
1. Note consistency with which pressures held

(c) All turns
1. Note points outlined above, particularly leading with aileron or rudder.

2. Note tendency to hold rudder during turn
3. If evident that altitude gained or lost, note coordination of back pressure with other controls, e.g., point at which back pressure first applied
4. Note smoothness with which controls released (Rudder and Ailerons after entry; Rudder, Ailerons and Elevators in recovery)
5. Note coordination in recovering from bumps

(d) Stall

1. Note smoothness with which back pressure applied, and whether stick full back at break
2. Note whether was used; especially bad around break
3. Note whether rudder correction was appropriate

(e) Slip

1. Note whether controls were applied smoothly, and pressure held
2. Note if leads with Aileron or Rudder; latter worse
3. Noté coordination on recovery

(f) Climb and Glide and Straight and Level: In all these note whether corrections were made using coordinated movements

(g) Climb and Glide specifically

1. Note relationships of elevator movement to throttle
2. Note rudder coordination with throttle, slight right rudder as RPM increased, slight left movement as motor cut to idling in glide

(h) Landing

1. Note whether low wing, if any was raised by using rudder alone or with aileron
2. Note tendency to fan or walk rudder during Landing Run

3. Adequacy of execution

(a) Take-off

1. Adequate air speed at moment of take-off
2. Full RPM

(b) 90° Left Climbing Turn

1. Bank about 15°
2. Note whether actual climb occurred

(c) 360°

1. Bank of 45-50° expected

(d) Medium Turns

1. Bank of 30° expected

(e) Stall

1. Normal stall asked for--including nose to landing position, air speed 40 MPH or less at stall, stick forward and throttle on after break
2. Note especially if plane incompletely stalled

(f) Slip

1. Left bank expected, but either OK
2. Bank of 15° expected
3. Disorganized recovery bad

(g) Climb and Glide

1. Change of altitude expected unless hampered by thermals
2. Low rate of climb or descent indicates failure to achieve purpose of maneuver
3. Angle of glide not too shallow or too steep

(h) Straight and Level Flight

1. Note correction for changes of altitude and direction
2. Maintenance of altitude and direction good

(i) Landing

1. Note whether bounce landing, or heavy landing occurred
2. Note whether correction was made for poor landing and whether student made the correction unaided by observer
3. Note whether plane was under control just previous to moment of landing and immediately after.

DATA SHEET DIRECT OBSERVATION ANALYSIS PROJECT: RATING:

Student #																					
Wing Control																					
Nose Control																					
Directional Control																					
Slip or Skid																					
Altitude																					
Air Speed																					
Controls																					
Execution																					
Rating A B C																					
Rating U L																					
Air Cond.																					

DIRECT OBSERVATION ANALYSIS
Comments on Flight Performance

Take-off

Cl. Turn	Glide
Turn R	Turn L
Turn R	Turn L
Climb	St.& L.

Landing

Remarks:

Take-off

Cl. Turn	Glide
Turn R	Turn L
Turn R	Turn L
Climb	St.& L.

Landing

Remarks:

Take-off

Cl. Turn	Glide
Turn R	Turn L
Turn R	Turn L
Climb	St.& I.

Landing

Remarks:

AN ANALYSIS OF PHOTOGRAPHIC RECORDS
OF AIRCRAFT PILOT PERFORMANCE

SECTION B

SUPPLEMENTAL REPORT

AN ANALYSIS OF SCORES ON ASPECTS OF FLIGHT PERFORMANCE

by

Albert S. Thompson

University of Pennsylvania
Philadelphia, Pennsylvania

ACKNOWLEDGMENTS

The data for this Supplemental Report were obtained in a study undertaken by Albert S. Thompson for doctoral thesis purposes at the University of Pennsylvania. Acknowledgment is made to Dr. Morris S. Viteles, under whose direction the investigation was carried out, and to the National Research Council Committee on Selection and Training of Aircraft Pilots for permission to use the data for this purpose.

INTRODUCTION

The major interest in the University of Pennsylvania Project was the development of measures of over-all flight performance which could be used as criterion data. For this reason, the discussion in the main body of the report has been limited to an analysis of the Criterion Ratings and the Criterion Flight Scores, the measures recommended for criterion use.

The rating procedure, however, also provided Aspect Scores,[1] obtained by summing the four ratings on each aspect of flight performance assigned during the independent rating and re-rating by each of the two observers, i.e., Independent Rating I, Thompson, Independent Rating I, Ewart, Independent Rating II, Thompson, and Independent Rating II, Ewart. There were eight Aspect Scores in all -- Wing Control Score, Nose Control Score, Directional Control Score, Slip-Skid Score, Altitude Score, Airspeed Score, Control Coordination Score, and Execution of Maneuvers Score.

A separate analysis of these Aspect Scores has been made not only for the sake of obtaining information concerning the aspect scores themselves, but also because they are essentially "part-scores" making up the Criterion Flight Scores and additional information concerning the Flight Scores should thus result from such a study.

This Supplemental Report, therefore, describes an analysis of the Aspect Scores and presents results pertaining to:

a. the reliability of each of the Aspect Scores.

b. the inter-correlations among the Aspect Scores, including a factor analysis of the inter-correlation matrices of Flight A and Flight D respectively.

c. the relationship between the Aspect Scores and the criterion measures, i.e., the Criterion Ratings and the Criterion Flight Scores.

d. the relationship between the Flight A Aspect Scores and measures of Flight D performance.

As in the case of the Criterion Flight Scores, the results are based upon the scores of the 33 subjects on whom both Flight A and Flight D photographic records were obtained.[2]

ANALYSIS OF THE ASPECT SCORES

A. Reliability of the Aspect Scores. The reliability of the Aspect Scores was measured by computing Pearsonian r's (for each aspect) between the

---

[1] See description of the rating procedure in the main report, Page 12.

[2] See Page 19 of the main report.

TABLE A

RELIABILITY OF THE ASPECT SCORES

Comparison Made	Flight A (N = 33) r	Flight A (N = 33) S-Br	Flight D (N = 33) r	Flight D (N = 33) S-Br
Wing Control Score				
Composite I vs. Composite II	.73		.77	
Spearman-Brown		.85		.87
Nose Control Score				
Composite I vs. Composite II	.82		.76	
Spearman-Brown		.90		.86
Directional Control Score				
Composite I vs. Composite II	.76		.82	
Spearman-Brown		.86		.90
Slip-Skid Score				
Composite I vs. Composite II	.79		.72	
Spearman-Brown		.88		.84
Altitude Score				
Composite I vs. Composite II	.69		.31	
Spearman-Brown		.81		.47
Airspeed Score				
Composite I vs. Composite II	.72		.58	
Spearman-Brown		.84		.74
Control Coordination Score				
Composite I vs. Composite II	.54		.61	
Spearman-Brown		.70		.76
Execution of Maneuvers Score				
Composite I vs. Composite II	.55		.66	
Spearman-Brown		.71		.80

NOTE: With an N of 33, coefficients above .443 are statistically significant at the 1% level. See Table 13, Page 212, in: Lindquist, E. F., "Statistical Analysis in Educational Research," New York. Houghton Mifflin Co., 1940.

Composite I Aspect Score (Independent I, Thompson plus Independent I, Ewart) and the Composite II Aspect Score (Independent II, Thompson plus Independent II, Ewart). Since this is essentially a split-half technique, the reliability of the total Aspect Score was estimated by use of the Spearman-Brown Prophecy Formula.

Table A presents the correlation coefficients and the Spearman-Brown estimates for each Aspect Score for each of the two flights. Analysis of this table reveals the following:

1. With the exception of the Altitude Score, Flight D, the Composite I vs. Composite II coefficients for the single Aspect Scores are statistically significant, being greater than .443.

2. Considering both Flight A and Flight D, the Aspect Scores for Wing Control, Nose Control, Directional Control and Slip-Skid are, in general, more reliable than those for Altitude, Airspeed, Control Coordination, and Execution of Maneuvers.

   In Flight A, the difference between the highest coefficient (.82 for Nose Control) and the lowest (.54 for Control Coordination) almost attains the 1% level of significance. Using the "z" test, the Critical Ratio is 2.14, indicating that the positive difference in favor of the Nose Control Score would occur by chance in approximately 2 cases in 100.

   The difference between the lowest and highest reliability coefficients for Flight D (.31 for Altitude Score and .82 for Directional Control, respectively) is highly significant. The exceptionally low values for the Altitude Score may be due somewhat to the narrow range of scores obtained, the number of step intervals being only 7. Correcting for broad categories,[3] however, raises it to only .33.

3. The four Aspect Scores mentioned above (Wing Control, Nose Control, Directional Control, and Slip-Skid) have Spearman-Brown reliability coefficients almost as high as the Criterion Flight Scores. The consistency of these particular Aspect Scores may be due to the fact that they were based upon ratings made from the more easily read and the more dependable flight instruments in the camera field. Both the Gyro-Horizon (upon which the Wing Control and Nose Control ratings were based) and the Turn Indicator (upon which the Directional Control ratings were based) are "gyro" instruments in contrast to the Airspeed Indicator, Rate of Climb Indicator, and Altimeter which are "air pressure" instruments.

---

[3] By means of Formula 214, Page 397 in: Peters and Van Voorhis, op. cit.

TABLE B

INTER-CORRELATIONS AMONG ASPECT SCORES

A. Flight A (N = 32)	Aspect Scores						
	2	3	4	5	6	7	8
1. Wing Control Score	.52	.66	.49	.40	.47	.62	.60
2. Nose Control Score		.44	.28	.78	.73	.48	.48
3. Directional Control Score			.59	.41	.44	.82	.50
4. Slip-Skid Score				.29	.37	.68	.54
5. Altitude Score					.54	.41	.42
6. Airspeed Score						.52	.46
7. Control Coordination Score							.63
8. Execution of Maneuvers Score							-

B. Flight D (N = 33)	2	3	4	5	6	7	8
1. Wing Control Score	73	.69	.64	.62	.62	.62	.80
2. Nose Control Score		.71	.55	.62	.62	.81	.76
3. Directional Control Score			.55	.57	.66	.70	.73
4. Slip-Skid Score				.48	.28	.65	.62
5. Altitude Score					.70	.59	.56
6. Airspeed Score						.55	.66
7. Control Coordination Score							.75
8. Execution of Maneuvers Score							-

B. Inter-correlation and Factor Analysis of the Aspect Scores.[4] A study of the inter-correlations among the Aspect Scores was made in order to determine whether the several Aspect Scores were unitary and independent and, if not, to identify the underlying factors involved.

Table B presents the inter-correlations among the Aspect Scores for Flight A and D, respectively. Inspection of the table reveals that, although the inter-correlations are sufficiently high (especially in the case of Flight D) to suggest that each of the eight Aspect Scores is not measuring an independent aspect of flight performance, there was probably something other than mere "over-all" general impression determining the separate ratings.[5] It was therefore decided to make a factor analysis of the inter-correlation matrices for Flight A and Flight D, respectively, realizing that the small number of cases involved (N = 33) would prevent any definitive conclusions from being drawn. The relatively high reliability of the Aspect Scores, however, and the intention to obtain merely a tentative identification of the factors involved seemed to warrant the application of factor analysis techniques to the available data.

The matrices of inter-correlations among the Aspect Scores for Flight A and Flight D separately were factor analyzed by the Thurston Centroid Method. In the case of each matrix, two factors were extracted, the extraction being terminated at that point since all of the second factor residuals were within one standard error of a correlation coefficient of zero (.17 in the case of an N of 33).

Since in each matrix only two factors were involved, the configuration can be represented graphically on a plane surface. The plottings of the centroid loading for the Flight A matrix, and for the Flight D matrix are presented in Figures 1 and 2, respectively, shown on page 100.

Inspection of Figure 1 indicates that in the Flight A analysis, the variables fall into two distinct clusters. Aspect Scores 2, 5, and 6 (Nose Control Score, Altitude Score and Airspeed Score) form one cluster, and the remaining Aspect Scores constitute the second. The pattern in the case of Flight D, while not so clear, is similar. Aspect Scores 5 and 6 (Altitude Score and Airspeed Score) can be considered to represent a cluster, with Aspect Score 3 (Directional Control Score) lying between the two clusters.

---

[4]Special acknowledgment is made to Edwin S. Ewart for his aid in the conduct of the factor analysis.

[5]It should be recalled that the aspect ratings were made concurrently with the over-all ratings. (See Page 6 of the main report.) In fact, the aspect ratings were originally intended to be merely "short-hand notes" to aid in arriving at the over-all rating and their use as data for computing Aspect Scores was not contemplated at the time the Flight D ratings were made. The use of the aspect ratings in order to compute scores was suggested by Dr. Jack W. Dunlap of the National Research Council Committee on Selection and Training of Aircraft Pilots.

TABLE C

FACTOR ANALYSES

Flight A

Aspect Score	Centroid Factor Loadings I	Centroid Factor Loadings II	Orthogonal Rotation Factor Loadings I	Orthogonal Rotation Factor Loadings II	Oblique Rotation Factor Loadings I	Oblique Rotation Factor Loadings II	Communality h^2	Uniqueness $1 - h^2$	Specificity $r_{11} - h^2$	Reliability* r_{11}
1	.74	.17	.75	.12	.52	.12	.58	.42	.27	.85
2	.76	-.53	.51	.78	.08	.78	.86	.14	.04	.90
3	.79	.30	[illegible]	.02	.65	.02	.74	.26	.12	.86
4	.66	.37	.75	-.10	.65	-.10	.57	.43	.31	.88
5	.68	-.44	.47	.66	.04	.66	.66	.34	.15	.81
6	.72	-.34	.54	.59	.07	.59	.63	.37	.21	.84
7	.84	.33	.90	.01	.71	.01	.80	.20	-.10	.70
8	.72	.13	.72	.15	.67	.15	.54	.46	.17	.71

Flight D

Aspect Score	Centroid Factor Loadings I	Centroid Factor Loadings II	Orthogonal Rotation Factor Loadings I	Orthogonal Rotation Factor Loadings II	Oblique Rotation Factor Loadings I	Oblique Rotation Factor Loadings II	Communality h^2	Uniqueness $1 - h^2$	Specificity $r_{11} - h^2$	Reliability* r_{11}
1	.85	.09	[illegible]	.28	.47	.28	.73	.27	.14	.87
2	.87	.12	.85	[illegible]	[illegible]	.36	.77	.23	.09	.86
3	.83	-.04	.78	.29	.34	.29	.69	.31	.21	.90
4	.67	.31	.73	-.09	.58	-.09	.55	.45	.29	.84
5	.75	-.28	.63	.50	.09	.50	.64	.36	-.17	.47
6	.75	-.46	.57	.67	-.07	.67	.77	.23	-.05	.74
7	.85	.19	.87	.08	.56	.08	.76	.24	.00	.76
8	.87	.11	.86	.16	.49	.16	.77	.23	.05	.80

*Estimated by Spearman-Brown Prophecy Formula (see Table A).

In the case of both matrices, the reference axes were rotated orthogonally so that the number of zero loadings was maximized (see Table C). In the Flight A analysis Factor I appears with significant loadings on all Aspect Scores, but with relatively lower loadings on Aspect Scores 2, 5, and 6. Factor II appears with low or zero loadings on all Aspect Scores except 2, 5, and 6. From the Flight D analysis a somewhat similar factor configuration is evident after orthogonal rotation. Factor I contains significant loadings on all Aspect Scores, and Factor II appears with low or zero loadings on all except Aspect Scores 5 and 6 (Altitude Score and Airspeed Score).

Factor II, characterized by zero loadings on all Aspect Scores except Altitude Score, Airspeed Score, and Nose Control Score (in the Flight A analysis only) can be tentatively identified as "Longitudinal Control" since these three are affected by changes in the longitudinal attitude of the plane. Factor I, although exhibiting significant loadings on all Aspect Scores, is most heavily weighted with those not directly associated with longitudinal control. The cluster of Wing Control Score, Directional Control Score, Slip-Skid Score, and Control Coordination Score (appearing rather clearly in the Flight A analysis) suggests a possible identification of this factor as "Coordination."

The two factors, however, are correlated, as is evident from inspection of Figures 1 and 2, in which it is clearly seen that the clusters of Aspect Scores are not at right angles. The correlation between the two factors is approximately .60 in the case of the Flight A analysis, and .70 in the Flight D analysis, as estimated from the angle formed by vectors passing through the clusters. The angle is 52° in the Flight A analysis and 45° in the Flight D analysis. That the two factors, although correlated, appear to be "unitary" is indicated by the fact that after oblique rotation of the reference vectors, "oblique simple structure" is clearly obtained from the Flight A matrix and is closely approximated in the case of the Flight D matrix. The loadings after oblique rotation are also presented in Table C.

This correlation suggests that the "general impression" of over-all flight performance, which determined the over-all ratings, may have influenced the Aspect Scores as well, since the latter are based upon the aspect ratings made concurrently with the over-all ratings.[6] On the basis of this hypothesis it may be conjectured that the Aspect Scores are determined by three influences:

1. The over-all impression of the flight performance as a whole.

2. Longitudinal control as exhibited by variations in altitude, airspeed, and position of nose on the horizon.

3. Coordination of the controls, as exhibited by the pilot's control of the wings, direction, lateral balance, and the plane as a whole during the execution of the maneuvers in the flight.

---

[6] The factor analysis, however, does not yield a "general impression factor," since measures of the over-all performance were not included in the inter-correlation matrix.

TABLE [illegible]

CORRELATION BETWEEN ASPECT SCORES AND CRITERION FLIGHT SCORES

Aspect Score	Flight A Criterion Score	Flight A Criterion Score Minus Aspect Score*	Flight B Criterion Score	Flight B Criterion Score Minus Aspect Score*
1. Wing Control Score	[illegible]	.63	.85	.81
2. Nose Control Score	.74	.65	[illegible]	.84
3. Directional Control Score	[illegible]	.74	[illegible]	.80
4. Slip-Skid Score	[illegible]	[illegible]	.71	.61
5. Altitude Score	.59	[illegible]	[illegible]	.69
6. Airspeed Score	[illegible]	.65	.75	.68
7. Control Coordination Score	[illegible]	.78	[illegible]	.82
8. Execution of Maneuvers Score	.77	.70	[illegible]	.85

*The Criterion Score Minus Aspect Score refers to the sum of the seven aspect scores other than the aspect score being correlated.

C. Relationship between Aspect Scores and the Criterion Measures.

1. Correlation with Criterion Flight Scores. The relationship between the Aspect Scores and the total score of which they were a part was determined by correlating the Aspect Scores with the Criterion Flight Scores. The relevant data are presented in Table D, Correlation between Aspect Scores and Criterion Flight Scores.

The correlations between Aspect Scores and the Criterion Flight Scores are, of course, spuriously high since the Criterion Flight Score is the sum of the eight Aspect Scores and thus includes the Aspect Score being correlated with it.[7] In addition, the amount of intercorrelation among the Aspect Scores prevents these correlations from being interpreted as measures of the unique contribution of the Aspect Scores to the Criterion Flight Score.

That the differences among Aspect Scores in amount of correlation between the Aspect Scores and the Criterion Flight Scores are due not merely to the differences in means and standard deviations of the distributions of Aspect Scores is evidenced from the fact that when the Aspect Scores are ranked according to the obtained correlation with the Criterion Flight Score, the rank order is practically the same as that obtained when they are ranked according to the obtained correlation with the Criterion Flight Score Minus the Aspect Score. In the latter case, the mean and standard deviation of the distribution of the specific Aspect Score has no effect on the total score with which it is being correlated.

An examination of Table D, however, reveals that, considering both Flight A and Flight D, the single Aspect Scores which best predict the Criterion Flight Scores are Control Coordination, Execution of Maneuvers, and Directional Control. The least predictive are Slip-Skid and Altitude.[8]

---

[7] Table D also presents the correlations between each of the Aspect Scores and the sum of the other seven Aspect Scores. These coefficients were computed by means of Formula 136, Page 217, in:
Peters, C. C. and Van Voorhis, W. R. Statistical Procedures and Their Mathematical Bases. New York: McGraw-Hill Co., Inc., 1940.

[8] It was considered unnecessary, because of the small N and the number of variables involved, to determine their relative contribution to the total score by means of a multiple correlation technique.

TABLE E

COMPARISON OF ASPECT SCORES WITH CRITERION RATINGS

Biserial r Between Scores and Over-All Ratings When
Grouped as Follows: (A+B) versus (C)
(U) versus (L)
(A) versus (B+C)

Score	Flight A (A+B)-(C)	Flight A U - L	Flight A (A)-(B+C)	Flight D (A+B)-(C)	Flight D U - L	Flight D (A)-(B+C)
N =	24 - 9	17-16	9 - 24	24 - 9	18-15	8 - 25
1. Wing Control Score	.70	.78	.87	.84	.77	.77
2. Nose Control Score	.51	.66	.71	.79	.89	.77
3. Directional Control Score	.72	.64	.82	.63	.82	.85
4. Slip-Skid Score	.78	.75	.67	.63	.65	.83
5. Altitude Score	.49	.53	.71	.70	.75	.75
6. Airspeed Score	.35	.78	.75	.69	.74	.60
7. Control Coordination Score	.84	.80	.86	.64	1.00	.99
8. Execution of Maneuvers Score	.81	.80	.86	.83	.95	.85

[illegible] Criterion Ratings. As in the case of the [illegible] between the Criterion Flight Scores and the Criterion Ratings biserial correlation coefficients were computed with the group split into categories according to the Criterion Ratings.

From an examination of Table E, in which the results are presented, it may be seen that, in general, the Control Coordination Score and Execution of Maneuver Score exhibit the highest relationship with the Criterion Ratings while the Altitude Score and Airspeed Score exhibit the lowest relationship.[a]

3. It is interesting to note that the correlations between the [illegible] Criterion Ratings [illegible] higher in Flight D [illegible]

[illegible] tendency: (1) [illegible] progress through the [illegible] (within themselves) [illegible] performance and thus are [illegible] the investi- [illegible] interest in the [illegible] for computing Aspect [illegible] before the [illegible] have been more [illegible] ratings independ- [illegible] ratings were [illegible]

4. [illegible] Flight D Performance. [illegible] value of Flight A perform- [illegible] the [illegible] Aspect Scores and [illegible] (2) [illegible] Criterion Flight Scores, and [illegible]

---

[a] [illegible] of this report

[b] [illegible] Coordination Score and Execution of Maneuvers Score were also found to be [illegible] predictive of the Criterion Flight Scores and to have high [illegible] on Factor I in the factor analysis, while the Altitude Score and Airspeed Score [illegible] high [illegible] on Factor II.

[c] It was also noted in the factor analysis that the inter-correlation among the Aspect Scores tended to be higher for Flight D than for Flight A.

TABLE F

COMPARISON OF FLIGHT A ASPECT SCORES AND FLIGHT D PERFORMANCE

A. Scores on Flight A vs. Criterion Ratings Flight D (Biserial r's)

Scores on Flight A	Criterion Ratings Flight D		
	(A+B)-(C)	U - L	(A)-(B+C)
N =	24 - 9	18 - 15	8 - 25
1. Wing Control Score	.45	.53	.65
2. Nose Control Score	.20	.23	.03
3. Directional Control Score	.62	.57	.47
4. Slip-Skid Score	.43	.28	.40
5. Altitude Score	.01	.19	.07
6. Airspeed Score	.02	-.12	.16
7. Control Coordination Score	.56	.48	.40
8. Execution of Maneuvers Score	.30	.31	.29
9. Criterion Flight Score	.45	.41	.38

TABLE F (Cont'd)

COMPARISON OF FLIGHT A AND FLIGHT D

B. Scores on Flight A vs. Criterion Flight Scores on Flight D (Pearson r's)

	Flight A		Flight D	r*
1.	Wing Control	vs.	Criterion Flight Score	.59
2.	Nose Control Score	vs.	Criterion Flight Score	.17
3.	Directional Control Score	vs.	Criterion Flight Score	.46
4.	Slip-Skid Score	vs.	Criterion Flight Score	.45
5.	Altitude Score	vs.	Criterion Flight Score	.14
6.	Airspeed Score	vs.	Criterion Flight Score	.07
7.	Control Coordination Score	vs.	Criterion Flight Score	.48
8.	Execution of Maneuvers Score	vs.	Criterion Flight Score	.47
9.	Criterion Flight Score	vs.	Criterion Flight Score	.46

C. Aspect Scores Flight A vs. Aspect Scores Flight D (Pearson r's)

	Flight A		Flight D	r
1.	Wing Control Score	vs.	Wing Control Score	.50
2.	Nose Control Score	vs.	Nose Control Score	.20
3.	Directional Control Score	vs.	Directional Control Score	.37
4.	Slip-Skid Score	vs.	Slip-Skid Score	.31
5.	Altitude Score	vs.	Altitude Score	.01
6.	Airspeed Score	vs.	Airspeed Score	.07
7.	Control Coordination Score	vs.	Control Coordination Score	.45
8.	Execution of Maneuvers Score	vs.	Execution of Maneuvers Score	.53

*With N = 33, coefficients .443 or above are significant at the 1% level.

(3) Flight D Aspect Scores. The resulting coefficients are presented in Table F.

1. Correlation with Flight D Criterion Ratings (Table F, Section A). The only biserial r's approaching a satisfactory degree of significance are those between the Wing Control Score of Flight A and the (A) vs. (B+C) grouping according to Flight D Criterion Ratings (+.66), and between the Directional Control Score in Flight A and the (A+B) vs. (C) grouping in Flight D (+.62). These two values suggest that those pilots exhibiting good wing control in Flight A tended to be given high over-all ratings in Flight D and those exhibiting pour directional control in Flight A tended to be rated low for Flight D performance as a whole.

2. Correlation with Flight D Criterion Flight Scores (Table F, Section B). Five of the eight Aspect Scores of Flight A exhibit a statistically significant relationship with the Criterion Flight Scores of Flight D. As in the Aspect Score -- Criterion Rating comparison, the highest relationship is exhibited by the Wing Control Score.

3. Correlation with Flight D Aspect Scores (Table F, Section C). Three of the correlations between Flight A and Flight D scores on the individual Aspect Scores are statistically significant, viz., Wing Control Score, Control Coordination Score, and Execution of Maneuvers Score.

   Even the significant r's, however, are not sufficiently high to suggest that any particular aspects of flight performance are unusually consistent from initial to final stage of primary C.P.T. Flight Instruction.

SUMMARY AND CONCLUSIONS

During a rating and re-rating by two independent observers of photographic records of flight performance taken during Stage A and Stage D of a group of 33 student pilots, ratings (on a five-point scale) of eight aspects of performance were assigned by each of the two observers for each flight for each subject. An analysis of Aspect Scores, computed from these ratings, revealed the following:

1. The Wing Control Score, Nose Control Score, Directional Control Score, and Slip-Skid Score were more consistently reliable (for the two flights) than were the other scores — Altitude Score, Airspeed Score, Control Coordination Score, and Execution of Maneuvers Score.

2. The eight aspects of performance, separately rated, are not independent and unitary. A factor analysis of the inter-correlations among the Aspect Scores indicated two factors, tentatively identified as "Coordination of the Controls" and "Longitudinal Control." Since the two factors were correlated, a third influence may be affecting the separate ratings, probably a "general impression" of the flight as a whole.

3. The Aspect Scores most closely associated with the criterion measures of over-all performance (Criterion Ratings and Criterion Flight Scores) are the Wing Control Score, Directional Control Score, Control Coordination Score, and Execution of Maneuvers Score. These four also have their highest factor loadings for the factor identified as "Coordination of the Controls."

4. Aspect Scores on Flight A performance (during Stage A of C.P.T. primary flight instruction) are not highly predictive of Flight D performance (during Stage D of C.P.T. primary flight instruction). In various comparisons, however, statistically significant prediction was obtained most consistently in the case of Wing Control Score, Directional Control Score, Control Coordination Score, and Execution of Maneuvers Score.

The limitations of the study (due to the small number of cases involved) do not permit any definitive conclusions as to the value of the separate Aspect Scores for diagnostic or predictive purposes. On the basis of the present information it would seem most practicable to use the Criterion Ratings and Criterion Flight Scores (described in the main report) as the criterion measures and to consider the Aspect Scores as a source of information concerning the nature and characteristics of these criterion measures of over-all flight performance.

Figure 1
Plot of Flight A Factor Loadings

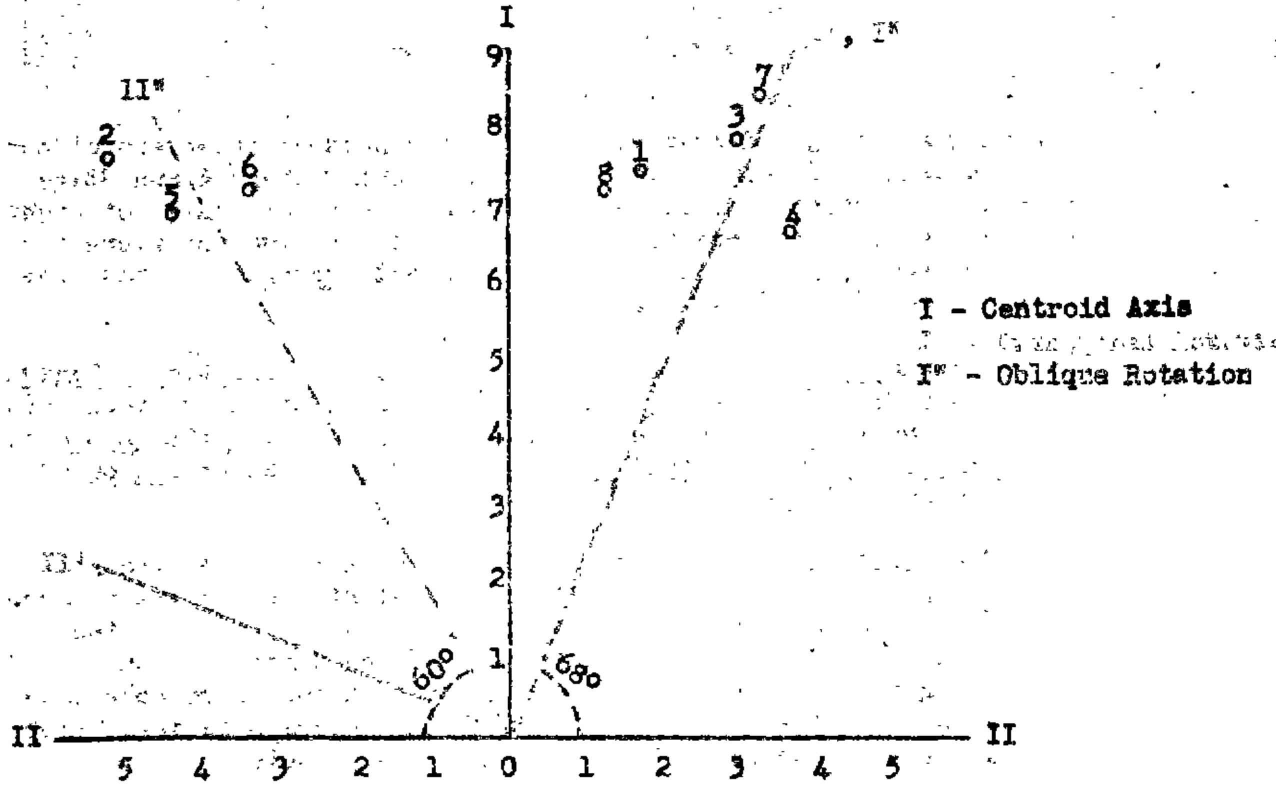


Figure 2
Plot of Flight D Factor Loadings