A PRELIMINARY STUDY OF CERTAIN PREDICTORS OF SUCCESS IN CIVILIAN PILOT TRAINING

CINIL ASSOCIATION

bу

E. Lowell Kelly and E. Ewart

A Report on research conducted at Purdue University, Lafayette, Indiana, 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.

December 1942

CIVIL AERONAUTICS ADMINISTRATION

Division of Research

Report No. 7

Washington, D. C.

NATIONAL RESEARCH COUNCIL

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

Committee on Selection and Training of Aircraft Pilots

December 30, 1942.

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

Dear Dr. Brimhall:

The study on training aids reported in C.A.A. Division of Research Report No. 6 provided an opportunity to investigate the relationship between a number of "predictors" and achievement in learning to fly, as well as the interrelations among the measures or "criteria" of achievement in learning to fly.

Results of this study are presented in the attached report, by L. Lowell Kelly and B. Ewart, entitled A Preliminary Study of Certain Predictors of Success in Civilian Pilot Training. It is the recommendation of the Committee on Selection and Training of Aircraft Pilots that this report be included in the series of technical reports published by the Division of Research, Civil Aeronautics Administration

In making this recommendation, however, the Committee desires to point to the fact that the study, based upon a small and somewhat selected group of student pilots taking training under somewhat specialized conditions, must be viewed as tentative in character. The relationships discussed in this report are being submitted to more extensive analysis in the unlegrated Boston and Midwest studies which will be discussed in later reports. Nevertheless, the present study reveals interesting preliminary findings, such as, for example, the lack of relation between instructors' ratings and flight inspectors' ratings; the value of the instructor's ratings at the end of the period of dual instruction as an index of later performance in training, etc. However, even such interesting items, involving a small number of inspectors, instructors, and student pilots, must be treated with caution until evidence from other studies is available.

Very truly yours,

Morris S. Viteles. Chairman. Committee on Selection and Training of Aircraft Pilots

National Research Council

115V:k

A PRELIMINARY STUDY OF CERTAIN PREDICTORS OF SUCCESS IN CIVILIAN PILOT TRAINING*

Summary

The relative predictive values of An Inventory of Personal Data for Prospective Pilots, of an interview involving the use of both a Personal History Inventory and an Aviation Interview Rating Scale, and of a Flight Work Sample were investigated, using a group of 40 primary Civilian Pilot Training students as subjects. The criteria were instructors' ratings at solo and at completion of course, check pilot's ratings, and C.A.A. inspectors' ratings.

The pre-training 20-minute interview was conducted by three persons. The ratings of all three interviewers correlated positively with student success, but the median rating of the three had more predictive value than that of any of the three singly.

Scores on the <u>Inventory of Personal Data for Prospective Pilots</u> (referred to as the <u>Biographical Inventory</u>) predicted student success as well as the median Interview rating. Because of the high intercorrelation between the Biographical Inventory scores and the Interview ratings, a combination of them did not improve the prediction. Combining instructor's rating at solo with either the Biographical Inventory score or the Interview rating greatly increased the accuracy of the prediction of final success.

C.A.A. inspectors' flight test grades were found to be unrelated to any of the other criteria and were not predicted by either the Biographical Inventory or the Interview.

^{*}This study was made during the spring of 1942 at Purdue University with the cooperation of the local Civilian Pilot Training staff. It was financed by a grant from the National Research Council Committee on Selection and Training of Aircraft Pilots from funds provided by the Civil Aeronautics Administration.

The subjects and measures are the same as those referred to in the report by:

Kelly, E. L. & Ewart, E. The Effectiveness of <u>Patter</u> and of <u>Fundamentals of Basic Flight Maneuvers</u> as Training Aids. Washington, D. C.: Civil Aeronautics Administration, Division of Research, Report No. 6, December 1942.

THE PROBLEM

The purpose of this study was to investigate the relative value of three measures in predicting the success of primary Civilian Pilot Training students.

Two of the measures were obtained before the beginning of the course:

- (1) A score on <u>An Inventory of Personal Data for Prospective Pilots</u> (popularly known as the <u>Biographical Inventory</u> and hereinafter referred to as the <u>B.I.</u>).
- (2) The median rating of a three-man board after a 20-minute interview involving the use of a <u>Personal History Inventory</u>² and an <u>Aviation Interview Rating Scale</u>³

The third measure was the rating by the instructor at the time the student was ready to solo (after at least eight hours of training). This measure may be entitled the <u>Flight Work Sample</u>.

The B.I. contains 150 questions of a biographical nature. Part A, headed "Biographical Data," has questions of a multiple choice type. Part B, headed "How I Feel and Act," contains 45 "Yes-No" questions. Sixty minutes of testing time were allowed for the B.I.

Prior to the interview the applicant filled in a <u>Personal History Inventory</u> (P.H. <u>Inventory</u>) containing 100 "Yes-No" questions on personal background. (See Exhibit B.) A maximum of 35 minutes was allowed for completing this form. Each interviewer was supplied with a copy of the

An Inventory of Personal Data for Prospective Pilots, Form 2c. Washington, D. C.: Civil Aeronautics Administration, 1941.

Sample items are presented as Exhibit A. Early stages in the construction of the Biographical Inventory are described in:

Kelly, E. L. The Relationship of Background and Personality Factors to Pilot Competency. (Purdue Research Foundation, Project No. 3: Progress Report, September, 1940.) Washington, D. C.: National Research Council Committee on Selection and Training of Aircraft Pilots, 1940.

²Sample items are presented as Exhibit B. See: P-H Inventory, Form XP.I., 1942.

³For a description of both the <u>Personal History Inventory</u> and the <u>Av-iation Interview Rating Scale</u> see:

Kogan, L. S., Tiedemann, D. V., Wantman, M. J., & Dunlap, J. W. Directions for the Use of the Aviation Interview Rating Scales. Washington, D. C.: National Research Council Committee on Selection and Training of Aircraft Pilots, January 29, 1942.

. While I r

EXHIBIT A

his should ge in

Sample items from

An Inventory of Personal Data for Prospective Pilots. (B.I.)

EXHIBIT B
Sample items from

Personal History (P-H) Inventory

candidate's answers on these questions and in addition, with an <u>Aviation</u> <u>Interview Rating Scale</u> on which the interviewer entered a rating on each of eight areas presumably related to flying success, as listed below, and an over-all rating on fitness for flight training.

- A. Academic background as related to flying.
- B. Family and socio-economic background as related to flying.
- C. General social adjustment as related to flying. (Exhibit C.)
- D. Desire to fly.
- E. Hobbies, diversions, and outside activities as related to flying.
- F. Athletic activities and coordination.
- G. Personality as related to flying.
- H. Appearance, mannerisms, and physique as related to flying.

As indicated above, three interviewers participated in the 20-minute interview. For the purposes of the present study, the median over-all rating of the three interviewers on "Fitness for flight training" (Exhibit D) is considered as a predictive measure.

The total time spent by each applicant in filling out the <u>P.H. Inventory</u> and in being interviewed is about the same as the time spent on the Biographical Inventory. The total staff time and expense are obviously much greater for the interview.

The controlled C.P.T. flight course is divided into four stages. The first stage ends when the student solos, usually between the eighth and the tenth hours. It is a fixed rule of the course that a student must have eight hours of dual instruction before soloing, even though he may seem qualified earlier. The predictor designated as Flight Work Sample was a rating made on Item 14 of A Scale for Rating Pilot Competency. This item requires the instructor to answer the question: "In your opinion, considering skill, emotional stability, judgment, etc., how good an 'all-around pilot' is he likely to become?" The judgment is made by the instructor on the basis of his close observation of the student pilot during the first eight to ten hours of flight instruction.

PROCEDURE

<u>Intercorrelations</u>. Intercorrelations were computed between the following variables:

- 1. Score on Section A, Biographical Inventory.
- 2. Score on Section B, Biographical Inventory.

4This scale, frequently referred to as the <u>Purdue Rating Scale</u>, was prepared by E. L. Kelly and is printed by the Purdue Research Foundation. See:

Kelly, E. L. Development of a Scale for Rating Pilot Competence. (Purdue Research Foundation, Project No. 2: Progress Report, September 1, 1940.) Washington, D. C.: National Research Council Committee on Selection and Training of Aircraft Pilots, 1940.

		gan var sen mensen sist greek var had sing med tomograph engine segiment var sammen a mensely of mediene frequency limits, and	an commission men
140			
	WHIR MAR		r jir piloti. Si Milato o Angli Sirato
67 846			
	page 100		a valg Taman
#4.			
			100
gardy March			
Andre		TNE NOT BE A CONTROL OF THE STATE OF THE STA	
		Control to the control of the contro	

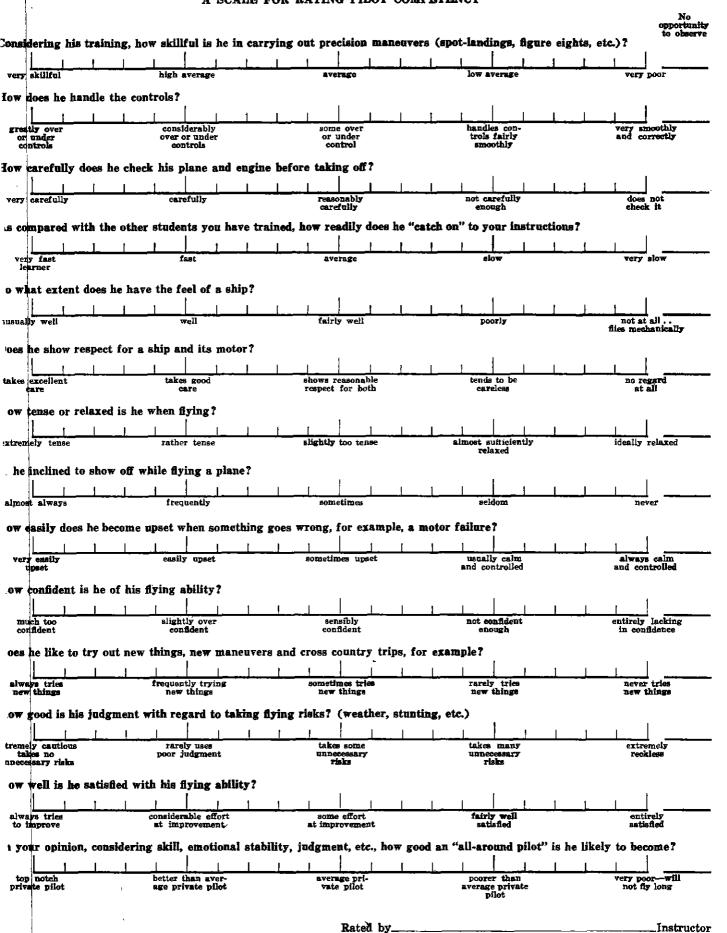
THE REPORT OF THE PARTY OF THE

BELLETING C

Sample rating sheet

	Chart for "over-all" rating from

A SCALE FOR RATING PILOT COMPETENCY



- 3. Total score, Biographical Inventory.
- 4. Median rating on Interview.
- 5. Instructor's rating of pilot competency at actual time of solo, as recorded on Item 14 of the Purdue Scale for Rating Pilot Competency. (Flight Work Sample.)
- 6. Instructor's final rating of pilot competency at completion of the course. This rating was also made on Item 14 of the Purdue Scale for Rating Pilot Competency.
- 7. Check pilot's rating of pilot competency at completion of the course. This rating was also made on Item 14 of the Purdue Scale for Rating Pilot Competency. In each case the check pilot giving this final over-all rating was one who had rated the student pilot at the end of each stage of the course. However, steps were taken to withold from the check pilot the identity of the instructor by whom the student pilot had been trained.
- 8. C.A.A. Flight Inspector's grade on flight test made in terms of a percentage scale, 70% being the lowest passing grade.

Variables 1 to 4 are predictive variables. Variables 6, 7, and 8 are "criterion variables." Variable 5 can be considered both as a "predictive" and as a "criterion" variable.

In most cases two correlations were computed between each set of variables: (1) the correlation in which N included all cases available—the range of N under these circumstances was 36 to 40, and these coefficients are denoted \underline{r} ; (2) the correlations in which all students having previous flight training were excluded from the population—the range of N in these cases was from 30 to 33, and these coefficients are denoted \underline{r} .

Since one of the instructors had also served as an interviewer, it was thought advisable to determine the correlation between the Interview and the three chief criterion variables—instructor's rating at solo (5), instructor's final rating (6), and check pilot's rating (7)—when the students of this instructor were eliminated from the population. These coefficients are denoted \underline{r} '. The range of N in these cases was 26 to 30.

Multiple correlations and correlations with composite criteria. A number of multiple correlations was computed between two or more of the predictive variables, and the criterion variables. Furthermore, in order to determine the best combination of predictors and criteria, several of the criterion variables were combined into a composite criterion, and the maximum average correlation between the composite criterion and the weighted independent variables was determined. The procedure was as

⁵This step was taken in conducting the investigation of the effectiveof <u>Patter</u> and of <u>Fundamentals of Basic Flight Maneuvers</u> as training aids,
as described in C.A.A. Division of Research Report No. 6. (Kelly and
Ewart, op. cit.)

follows: using Horst's approximation formula for combining criteria so that the greatest possible variance is obtained, the weights for the criterion measures were determined; then Kurtz' formula was applied to determine the maximum average correlation between the composite criterion and the weighted independent variables. Due to the extremely small number of cases, it was not deemed wise to resort to more elaborate statistical techniques in combining the criteria.

RESULTS

The correlations are summarized in Tables I and II. Since N is small, these correlations should be viewed as merely suggestive of relationships. However, five points appear significant.

- 1. Section A of the Biographical Inventory has a higher correlation with instructor's rating at solo (5), instructor's final rating (6), and check pilot's final rating (7), than has either the score on Section B, or the total score. Section B actually correlates negatively (but not significantly) with ratings at solo and with final ratings of the instructors (see Table I). It appears that for the C.P.T. population, Section A of the Biographical Inventory is carrying the weight of the prediction of the specific criteria available in this project, while Section B detracts slightly from the validity of the instrument, at least if combined with Section A in an additive manner. Further investigation of the validity of Section B, using additional criteria and different populations, is suggested.
- 2. The value of the Flight Work Sample (i.e., ratings of students by instructor at the end of the dual period, or after eight hours of instruction) as a predictive variable should be emphasized since the correlation between ratings by instructors at time of solo and check pilot's final ratings (criterion) is .63, when persons having previous flight training are eliminated from the population. The correlation between ratings at solo and final ratings by instructors is even higher, r' being .74.

Furthermore, the value of combining the Flight Work Sample score with other predictive variables should be noted. The multiple correlation between check pilot's ratings (criterion) and the combination of ratings at solo and the score on Section A, Biographical Inventory is .69; between this criterion variable and the combination of ratings at solo and median Interview ratings, the multiple correlation is .64. When a composite criterion is predicted from a combination of independent variables (using Kurtz' technique) the advantage of using the Flight Work Sample as a predictive variable is even more pronounced (see Table II, correlations IV, V, VI, and VII).

⁶Horst, P. Obtaining a composite measure from a number of different measures of the same attribute. <u>Psychometrika</u>, 1936, 1, 53-60.

⁷Kurtz, A. K. The simultaneous prediction of any number of criteria by the use of a unique set of weights. <u>Psychometrika</u>, 1937, 2, 95-101.

SIT is to be noted that the weights assigned to items on this test were derived from a Navy population in a study to be reported elsewhere in this series of technical reports.

20 LE 1

CHEACHAR FOLDS IN CHICAGO AND CHEER OF RANKS AND TARGETAL

Revee of N = 35-40

(r' denotes correlation when persons having previous flight instruction are excluded from the correlation table. Range of N = 30-33.)

(\underline{r}^{*}) denotes correlation when students of Instructor B, who also served as an interviewer, are dropped from the correlation chart. Hange of N = 26-30.)

INTERCORRELATIONS OF CRITERIA

Final Ratings (instructor) againsts	T-P CA MATE	1 4
Ratings at solo (instructor)	.753	.740
Check pilot's ratings	.344	.371
Inspector's ratings	s Interior	
Check pilot's ratings z ratings at solo	.481	.630
Check pilot's ratings x inspector's ratings	.040	

CORRELATIONS HETVELN BIOGRAPHICAL INVENTORY AND INTERVIEW

Interview against:	2	
Section A, Biographical Inventory	*594	
Section B, Biographical Inventory	~,040	
Total score, Biographical Inventory	. 386	

CORRELATIONS AMONG BIOCHAPHICAL INVESTIGAT, INTERVEL, AND CATELIA

	B.I. Section A		E.T Secti				AND SHOW IN A RESIDEN		
; !	I	<u>r</u> †		7. 1	17	y t	**************************************	P ?	E * P
Criteria Ratings at Solo (5)	.335	. 294	065	034	-188	.203	,1.64	# EAR	.279
Final Ratings(6)	.342	. 336		004	122	.091	-430	.423	.311
Check Pilot's	, 502	.46 9	.006	.001	,249		.426	.370	.459

TABLE II

MULTIPLE CORRELATIONS AND CORRELATIONS WITH COMPOSITE CRITERIA

Ţ	redicting Variables	Criterion Variables	Correlations based on:		
			Cases	Those with no previous instruction (N=30-33)	
I.	B.I., Section A Ratings at solo	Check pilot's ratings	.601	.693	
I.	Median Interview ratings Ratings at solo	Check pilot's ratings	.551	.640	
ı.	B.I., Section A Median Interview ratings	Check pilot's ratings	.526		
	B.I., Section A Ratings at solo	Check pilot's ratings Instructor's final ratings		.710 *	
	Median Interview ratings Ratings at solo	Check pilot's ratings Instructor's final ratings		.681 *	
I.	B.I., Section A Median Interview ratings	Check pilot's ratings Instructor's final ratings		•454*	
I.	B.I., Section A Median Interview ratings	Check pilot's ratings Instructor's final ratings Ratings at solo		•43 9*	
				· · · · · · · · · · · · · · · · · · ·	

^{*}Calculated by the formula for \bar{r} (wy_i) (x₁ ÷ bx₂) presented by Kurtz (op. cit.).

One question should be raised, however. The multiple correlations obtained on this sample, when the ratings at solo are used as a predictive variable, are high largely because these ratings at solo have a high correlation with check pilot's ratings, and relatively low correlations with both Interview and Biographical Inventory. Furthermore, it should be noted that the high correlations between the composite criterion (final ratings and check pilot's ratings) and the combination of ratings at solo and either Interview or Biographical Inventory scores are influenced by the fact that the check pilot's ratings at the end of the course correlate more highly with the instructor's ratings at solo than with the instructor's final ratings. If these facts result merely from peculiarities in the sample, these multiple correlations are spuriously high. Due to the sime of N, none of these differences between correlations is statistically significant. These relationships should be investigated using a larger sample.

3. Comparisons of the correlations between the criteria and the Interview and between the criteria and Section A of the Biographical Inventory indicate that there is little choice between them so far as prediction is concerned (see Table I).

While again it must be remembered that N is extremely small, it seems significant that when the students of the instructor who also served as an interviewer are eliminated from the sample, the correlations between Interview and criteria, and Section A of the Biographical Inventory and criteria, are of about the same order. Furthermore, it should be noted that the intercorrelation between Section A and the Interview is so high as to render combination of these two variables of little value for predicting the criteria (see Table I).

To summarize, in comparing the predictive value of the Interview with the Biographical Inventory, three points should be noted: (a) the correlations between Section A of the Biographical Inventory and the criteria are of about the same order when we eliminate the spurious effect introduced by the interviewer who also was a flight instructor; (b) regardless of the criterion of success in flight training (single or a weighted composite criterion) a combination of the Flight Work Sample (rating at solo) with either the Interview or the Biographical Inventory yields correlations of about the same magnitude; (c) the correlation between Interview and Section A of the Biographical Inventory is so high as to render the combination of Interview and Biographical Inventory for prediction purposes impractical.

Superficially, it appears that the Interview correlates more highly with instructor's ratings than does Section A of the Biographical Inventory, while Section A correlates more highly with ratings of the check pilot than does the Interview. However, the fact that one of the flight instructors also served as an interviewer, and thus rated ten of the students both on Interview and on their flying competency, renders the correlations between Interview and instructor's ratings spuriously high. When the students of this instructor are eliminated from the sample, the correlations between Interview and instructor's ratings are decreased, while the correlation between Interview and the check pilot's rating is increased slightly (see r'', Table I). Possible bias on the part of the instructor in his final and solo ratings is suggested.

Therefore, from the statistical evidence obtained from this sample, there seems to be little to choose between the two. However, due to the advantage which the Biographical Inventory holds in ease of administration, lower cost, and time required, it appears to be more applicable to selection in the field. 10

- 4. Correlations between C.A.A. inspectors' grades on the flight test at the end of the student's training, and the ratings of the students by their instructor and by the check pilot at the end of the course. are extremely low. Ratings by three inspectors are included in these results. Each student was rated by only one inspector. Inspectors' grades correlate only .12 with final ratings by instructors, and only .04 with the final ratings by the check pilot. The correlation between inspectors' grades and Section A of the Biographical Inventory is .12. The range of inspectors' grades is 70 to 80. Of the 36 students checked by one of the three inspectors, 21 received grades of 80, 11 received grades of 75, and 4 received grades of 70. Such a distribution is obviously not well suited to the calculation of a Pearson r. When 2x2 tables are prepared relating C.A.A. flight test ratings with three other variables, the relationships are as near chance as could be obtained. (The Chi-squares between C.A.A. inspectors' grades and check pilot's ratings, final ratings by instructor, and grade on Section A of the Siographical Inventory are .124, .302, and .020 respectively. These figures are such as would be obtained by chance in more than 80 per cent of samples drawn from a uniform population.) Such facts raise serious questions concerning the usefulness of the flight test as now applied.
- 5. On the basis of this sample, the fact that a student has had previous flight training apparently attenuates the prediction. It seems probable that the student's previous training introduces certain extraneous factors into the instructor's judgment of that student's flying competency. Also, it should be noted that previous flying experience was taken into consideration in the interviewers' rating of the student. The correlation between the check pilot's rating and ratings by instructors at time of solo is increased markedly when students having previously soloed are eliminated from the sample. This increase in correlation might be anticipated in view of the fact that previous flight training would most influence an instructor's judgment during the first eight hours of instruction.

However, the possibility should be recognized that this apparent attenuation may merely result from change in the sample when the seven

¹⁰ The point might be raised that the correlations between predictors and criteria are spuriously high, due to the fact that two methods of training were used on the students, as described in an earlier report by the same authors (op. cit.). Such differences in training might be expected to increase the range of scores on the criteria. However, the results outlined in the earlier report indicate that as far as the criteria included in this project are concerned, the two methods of training resulted in no significant differences in criterion scores.

students having previous training are dropped. Further investigation, using a large population, might yield evidence on this point.

CONCLUSIONS

As mentioned previously, since the number of cases is small, these conclusions can be considered only tentative. However, on the basis of the analysis of data from this sample, the following points seem significant:

- 1. Section A of the Biographical Inventory when scored with the Navy weights appears to be carrying the weight of the prediction of the available criteria. Section B detracts from the validity of the instrument, if the scores on Sections A and B are combined in an additive manner.
- 2. The most efficient predictor of flying competency at the end of the course (in terms of our criteria) is the Flight Work Sample or ratings of instructor at time of solo. The largest multiple correlations with individual criteria are obtained when the Flight Work Sample is one of the independent variables. When two independent variables are combined to predict a composite criterion, the superior advantage of using the Flight Work Sample as one of the independent or predicting variables is even more pronounced.
- 3. In terms of their efficiency as predicting variables, there is little choice between Section A of the Biographical Inventory and the median rating on the Interview. The Biographical Inventory, however, because of obvious advantages in ease of administration, is recommended for practical use in selection in preference to an Interview.
- 4. C.A.A. inspectors' grades on the final flight test show low correlations with instructors' ratings of students at the end of the course, and with ratings by a check pilot at the completion of the course. The data raise doubts as to the value of the C.A.A. Flight Test which now requires so much time and energy on the part of inspectors. Further checks as to the validity and reliability of these Flight Tests are needed.
- 5. The fact that some students in the population have had previous flight training apparently tends to make prediction from early measures less accurate, although the value of the observation is questionable by reason of the very small number of cases (N \mp 7) involved. To the extent that the trend exists, students with any other training should be considered as separate groups in the study of predictive measures.

Editor's note. It should be noted that the conclusions drawn from this exploratory research are being checked by an examination of results obtained in two extensive studies known as the "Boston Project" and the "Midwest Project," conducted under the direction of the Committee on Selection and Training of Aircraft Pilots during 1941-42. Findings of this analysis will be presented in a later technical report.