AFHRL-TR-71-18

AIR FORCE

HUMAN RESOURCES

AD732616

DEVELOPMENT OF AUTOMATED GAT-1 PERFORMANCE MEASURES

By

John W. Hill
Stanford Research Institute

Ronald A. Goebel

FLYING TRAINING DIVISION
Williams Air Force Base, Arizona

May 1971

Approved for public release; distribution unlimited.

LABORATORY

AFPS SA

AIR FORCE SYSTEMS COMMAND

BROOKS AIR FORCE BASE, TEXAS

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Stanford Research Institute 333 Ravenswood Avenue Menlo Park, California 94025	2a. REPORT SECURITY CLASSIFICATION 2b. GROUP

3. REPORT TITLE

DEVELOPMENT OF AUTOMATED GAT-1 PERFORMANCE MEASURES

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Final Report (April 1970 to May 1971)

5. AUTHOR(S) (First name, middle initial, last name)

John W. Hill
Ronald A. Goebel

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
May 1971	33	6

8a. CONTRACT OR GRANT NO. F41609-70-C-0041 b. PROJECT NO. 1123 c. Task No. 112301 d. Work Unit No. 112301004	9a. ORIGINATOR'S REPORT NUMBER(S) AFHRL-TR-71-18 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) SRI Project 8733

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Flying Training Division Williams Air Force Base, Arizona 85224

13. ABSTRACT

This report describes a systematic search for flight parameters that correlate with pilot proficiency. The system for making psychophysical measurements of the parameters consists of a Link General Aviation Trainer (GAT-1) connected to a small on-line digital computer (LINC-8). The computer can simultaneously monitor eight flight variables and input pseudo-random command signals (rough air) to three GAT-1 flight variables. By using this small computer in conjunction with a second computer program for further processing, 266 flight parameters were measured for each of the 30 subjects (*S*s) run through the experiment series. The parameters were means, standard deviations, correlations between variables, and compensatory tracking gains and phase shifts.

The experiment series consisted of four tasks of increasing difficulty: a holding task, a holding task with power changes, a five-part flight profile, and an ILS landing approach. First, an analysis of variance on each of the 266 variables was used to select the most important ones. Second, these selected variables were entered in a multivariate discriminant analysis to determine which contributed most to differences in pilot experience. Although between 10 and 15 variables sufficed for perfect separation of the *S*s into the three experience groups from which they were chosen, 27 variables significantly contributed to the separation. A single criterion variable, a linear weighted sum of these 27 flight parameters, is suggested as a measure of pilot proficiency.

DD FORM 1473 1 NOV 65

Unclassified
Security Classification

NOTICE

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
GAT-1 performance measures performance measures flight monitoring compensatory tracking tracking						

Unclassified

Security Classification

AFHRL-TR-71-18

May 1971

DEVELOPMENT OF AUTOMATED GAT-1 PERFORMANCE MEASURES

By

John W. Hill
Stanford Research Institute

Ronald A. Goebel

Approved for public release; distribution unlimited.

FLYING TRAINING DIVISION
AIR FORCE HUMAN RESOURCES LABORATORY
AIR FORCE SYSTEMS COMMAND
Williams Air Force Base, Arizona

ABSTRACT

This report describes a systematic search for flight parameters that correlate with pilot proficiency. The system for making psychophysical measurements of the parameters consists of a Link General Aviation Trainer (GAT-1) connected to a small on-line digital computer (LINC-8). The computer can simultaneously monitor eight flight variables and input pseudo-random command signals (rough air) to three GAT-1 flight variables. By using this small computer in conjunction with a second computer program for further processing, 266 flight parameters were measured for each of the 30 subjects (Ss) run through the experiment series. The parameters were means, standard deviations, correlations between variables, and compensatory tracking gains and phase shifts.

The experiment series consisted of four tasks of increasing difficulty: a holding task, a holding task with power changes, a five-part flight profile, and an ILS landing approach. First, an analysis of variance on each of the 266 variables was used to select the most important ones. Second, these selected variables were entered in a multivariate discriminant analysis to determine which contributed most to differences in pilot experience. Although between 10 and 15 variables sufficed for perfect separation of the Ss into the three experience groups from which they were chosen, 27 variables significantly contributed to the separation. A single criterion variable, a linear weighted sum of these 27 flight parameters, is suggested as a measure of pilot proficiency.

FOREWORD

This study was designed by the Performance Measurement Branch of the Flying Training Division of the Air Force Human Resources Laboratory (AFHRL) and represents a portion of the work being done under Project 1123, Flying Training Development, under the direction of Dr William V. Hagin. This effort was documented under Task 112301, Development of Performance Measurement Techniques for Air Force Flying Training, Dr Ronald A. Goebel, Task Scientist. Special thanks are due Dr Lonnie D. Valentine of the Personnel Research Division of AFHRL for his valuable guidance in the preparation of the contract. Thanks are also due to David R. Baum, A1C, USAF, of AFHRL/FT for his technical assistance throughout the course of the study.

The research was carried out by Dr John W. Hill, Bioinformation Systems Group of the Stanford Research Institute, under contract number F41609-70-C-0041. Dr Ronald A. Goebel, Performance Measurement Branch of AFHRL/FT, served as technical monitor. This effort covered the period of time between April 1970 to January 1971; the final report was submitted in May 1971.

Prominent among the supporting personnel at SRI were Mr Kenneth W. Gardiner who gave valuable assistance in designing and personally test flying the four flight tasks used in the study and also helped in gathering subjects; Mrs Elizabeth Primrose, who searched far and wide to obtain pilot subjects and ran them through the test series; and Mr A. F. Ferrera, who kept the GAT-1 flying. Thanks also go to the test pilots whose curiosity and interest brought them back for repeated tests when equipment failure nullified their data.

This technical report has been reviewed and is approved.

GEORGE K. PATTERSON, Colonel, USAF
Commander

SUMMARY

Hill, J. W., and Goebel, R. A., DEVELOPMENT OF AUTOMATED GAT-1 PERFORMANCE MEASURES. AFHRL-TR-71-18. Williams AFB, AZ: Flying Training Division, Air Force Human Resources Laboratory, May 1971.

Problem

Pilot proficiency measurement has always posed one of the major problems for research workers due to the fact that it tends to be subjective, to lack sensitivity, and to be unreliable. Although attempts have been made to make scales more sensitive and more objective, the pay-off has been minimal due to operational constraints. The objective of this effort was to develop an automated data recording and reduction technique for obtaining flight parameter correlates of pilot proficiency.

Approach

Three different experience level groups of general aviation pilots were given four flight tasks in a Link General Aviation Trainer (GAT-1). The trainer was interfaced with a Linc 8 computer, which monitored eight different flight parameters. Subsequent processing of these data yielded 266 potential indices of proficiency.

Results

An analysis of variance on each of these 266 dependent variables was performed to identify a subset of the most important parameters. These variables were then analyzed collectively by a multiple discriminant analysis which showed that 27 measures of this subset significantly contributed to perfect separation of the three groups.

Conclusions

This research has shown that automated pilot performance measurement can be utilized successfully in objectively reflecting pilot proficiency, which supports other earlier research in this area. Automated techniques such as this one are advantageous for both research and operations in that they are objective, sensitive and entail a minimum of effort to use. Efforts should continue in this area to apply the same basic system to other ground trainers and aircraft in the Air Force flying training inventory.

This summary was prepared by R. A. Goebel, Performance Measurement Branch, Flying Training Division, Air Force Human Resources Laboratory.

BLANK PAGE

CONTENTS

TABLES

ILLUSTRATIONS

BLANK PAGE

DEVELOPMENT OF AUTOMATED GAT-1 PERFORMANCE MEASURES

I INTRODUCTION

Previously, pilot proficiency has been measured subjectively, and the resulting unreliability and lack of precision have made comparisons between pilot training programs a major problem. Consequently, various automated measurements have been sought by previous investigators. Early work by Krendel and McRuer (1960) showed no differences in single-axis compensatory tracking ability with pilot proficiency: The ultimate skill levels were achieved after a few practice sessions for all the subjects (Ss) involved. More recently, Garber and Krasnitskii (1968) found a small group of psychophysical variables that correlate with progress on initial flight training. Connelly, Schuler, and Knoop (1969) derived an extensive set of models for making a study of automated pilot proficiency measurements, but they have not yet used their models on actual pilot performance data.

In this study a small on-line computer was used to monitor flight performance in a Link GAT-1 trainer. The measurements were based on variables that were to be held constant over a test run (changing variables, such as altitude in a climb, could not be used). Means and standard deviations were obtained for each such variable, and correlation coefficients were obtained for each pair of such variables. It was felt that in sufficiently complex simulated test flights Ss' differing abilities to share attention and hold a number of variables constant could be measured.

Since compensatory tracking affords stable measurements of human performance (especially in the crossover region, see McRuer et al, 1965) and can provide a number of models of pilot performance, a tracking task was included in the tests. It was hoped that two-dimensional tracking in the GAT-1, with its dependencies among several flight variables, would provide better discrimination of pilots on the basis of their experience than the single-variable tracking using a joystick.

II METHOD

Apparatus

To carry out this experiment an existing interface between a Link General Aviation Trainer (GAT-1) and a LINC-8 computer system (described by Hill, Gardiner, and Bliss, 1969) was enlarged to permit continuous monitoring of eight flight variables (airspeed, altitude, climb, roll, pitch, heading, glideslope, and localizer) and to allow the computer to supply signals to three GAT-1 variables (roll, pitch, and heading) for tracking tasks. Computer processing of the data to provide usable output was divided into two stages because of the limited memory size and computational power of the LINC-8 computer. The LINC-8 was used to accumulate simple sums, cross products, and Fourier sums and to generate the sum of sines command signals on line. The time scale for these accumulations was based on a 15-Hz clock*. The variables were sampled, and sums, products, and command signals were updated every 1/15 sec. The sums output by the LINC-8 on paper tape at the end of the test runs were further processed by a FORTRAN program on a commercial time-sharing computer to calculate means, standard deviations, correlation coefficients, and gains and the phase shifts for the tracking tasks.

Subjects

Thirty Ss (three groups of ten) were selected on the basis of their flying experience and run through the test series. Subjects in the beginner (B) group had had little or no flying experience (less than ten hours); Ss in the intermediate (I) group had had moderate experience (25 to 50 hours); and Ss in the advanced (A) group had had extensive experience (more than 100 hours) and in addition had logged 20 or more hours in the previous six months. The majority (approximately 80 percent) of the Ss had had 1-1/2 hours of experience in the GAT-1 trainer from a previous flight test. Many Ss, especially the advanced pilots, volunteered for the tests; others were obtained by offering them $5.00 for participating.

---

*A stable clock was obtained by dividing the 60-Hz line frequency by four.

Procedure

Four simulated flight tests were chosen for the experiment. Although the four tests had been thought to be of increasing difficulty, Test III, which included a synchronized turn and descent, was found to be the most difficult to perform by most Ss. In each of the monitored runs all the unchanging flight variables were monitored, while the remaining variables were not. For example, in a descending turn, altitude and heading were not monitored, but airspeed, rate of descent, roll angle, and pitch angle were monitored. In the first two tests a sum of eight sine waves of different frequencies was added to the pitch axis, a second similar sum was added to the roll axis of the GAT-1, and the S was asked to fly level. This is a classical two-dimensional compensatory tracking task. The ability of a S to cancel out the different frequencies in these two pseudo-random rough air signals is measured by his gain and phase shift at each individual command signal frequency.

Since 266 flight variables were measured for each S over the test series, many having the same name but measured in different tests, a numbering system is used here to designate the variables. The numbers for each variable are given in figures (1 through 8) as numbers placed to the right of the computer printout in the same relative positions as the values of the variables themselves.

1. Pilot Warm-up

Before the flight tests each S was asked to perform the following series of maneuvers to get used to the GAT-1 and its flight characteristics:

(1) A take-off and climb to 1000-ft heading due east

(2) A 180° left standard level turn in 1 minute

(3) A 180° right standard level turn in 1 minute

(4) A 500-ft climb at 500 ft/minute

(5) A power change to 1800 rpm and back to 2400 rpm

(6) A 500-ft descent at 500 ft/minute while keeping airspeed at 100 mph

(7) A level flight while switching on half flaps and then full flaps and maintaining altitude at 1000 ft.

TRACKING ANALYSIS 1/11 11:11

SUBJECT= PILØT #1

CØMMENT: TRK1

PRAM	MEAN	SD	UNITS	VARIABLE NUMBERS	
SPEED	105	2	MPH	1	7
ALT.	978	25	FEET	2	8
CLIMB	20	108	F/MIN	3	9
RØLL	-.2	1.8	DEG	4	10
PITCH	3.6	2.0	DEG	5	11
HEAD	.0	1.9	DEG	6	12

CØRRELATIØN CØEFFICIENTS, R(I,J)

					Variable numbers
.398					13
.591	.232				14 15
.262	.085	.140			16 17 18
-.212	.153	-.249	.062		19 20 21 22
.252	.077	.133	.995	.064	23 24 25 26 27

TRACKING GAIN & PØWER

28, 29, 30 not shown

FREQ	MAG.	PHASE		
.002	374.366	4.23	31	39
.010	131.734	96.86	32	40
.029	37.977	129.15	33	41
.057	14.001	152.72	34	42
.125	5.242	187.06	35	43
.191	3.844	198.49	36	44
.307	2.336	214.93	37	45
.467	1.914	-115.64	38	46

TRACKING GAIN & PØWER

47, 48, 49 not shown

FREQ	MAG.	PHASE		
.004	87.843	180.78	50	58
.014	32.653	206.66	51	59
.027	22.979	244.70	52	60
.053	3.467	-77.01	53	61
.115	1.016	-68.56	54	62
.158	.140	-140.66	55	63
.283	.688	126.43	56	64
.535	1.039	141.40	57	65

FIGURE 1 FLIGHT PARAMETERS FOR THE FIRST TRACKING TASK. Variables 28, 29, and 30 are correlated error power, correlated response power, and command power from roll tracking. Variables 47, 48, and 49 are the same three powers from pitch tracking.

TRACKING ANALYSIS 1/11 11:12

SUBJECT= PILØT #1

CØMMENT: TRK2

PRAM	MEAN	SD	UNITS	VARIABLE NUMBERS
SPEED	101	9	MPH	66 72
ALT.	987	25	FEET	67 73
CLIMB	16	118	F/MIN	68 74
RØLL	-1.8	2.2	DEG	69 75
PITCH	4.2	2.2	DEG	70 76
HEAD	-1.6	2.2	DEG	71 77

CØRRELATIØN CØEFFICIENTS, R(I,J)

					VARIABLE NUMBERS
.131					78
.130	.248				79 80
.101	.218	.172			81 82 83
-.562	.113	-.299	.080		84 85 86 87
.099	.216	.169	.997	.084	88 89 90 91 92

TRACKING GAIN & PØWER — 93, 94, 95 not shown

FREQ	MAG.	PHASE	VARIABLE NUMBERS
.002	877.138	2.96	96 104
.010	40.039	215.61	97 105
.029	29.042	144.60	98 106
.057	22.791	143.02	99 107
.125	4.376	179.62	100 108
.191	3.165	200.06	101 109
.307	2.490	214.32	102 110
.467	1.946	235.02	103 111

TRACKING GAIN & PØWER — 112, 113, 114 not shown

FREQ	MAG.	PHASE	VARIABLE NUMBERS
.004	23.887	70.65	115 123
.014	17.242	-91.53	116 124
.027	12.454	-16.89	117 125
.053	2.328	-23.00	118 126
.115	1.258	-11.74	119 127
.158	.377	11.18	120 128
.283	.819	131.95	121 129
.535	1.175	142.16	122 130

FIGURE 2 PARAMETERS FROM THE SECOND TRACKING TASK. Variables 93, 94, and 95 are correlated error power, correlated response power, and command power, respectively, from roll tracking. Variables 112, 113, and 114 are the same three powers from pitch tracking.

TRACKING ANALYSIS 1/11 11:13

SUBJECT= PILØT #1

CØMMENT: CLIMB

PRAM	MEAN	SD	UNITS	VARIABLE NUMBERS	
SPEED	95	6	MPH	131	136
CLIMB	402	105	F/MIN	132	137
RØLL	-1.3	2.1	DEG	133	138
PITCH	8.2	1.9	DEG	134	139
HEAD	-1.1	2.1	DEG	135	140

CØRRELATIØN CØEFFICIENTS, R(I,J)

.555				141			
-.707	-.227			142	143		
-.483	-.515	.459		144	145	146	
-.703	-.226	1.001	.447	147	148	149	150

FIGURE 3 PARAMETERS FOR THE 1000-FOOT CLIMB

TRACKING ANALYSIS 1/11 11:14

SUBJECT= PILØT #1

CØMMENT: RT TURN

PRAM	MEAN	SD	UNITS	VARIABLE NUMBERS	
SPEED	109	4	MPH	151	156
ALT.	1444	41	FEET	152	157
CLIMB	-15	160	F/MIN	153	158
RØLL	6.9	26.7	DEG	154	159
PITCH	2.9	2.1	DEG	155	160

CØRRELATIØN CØEFFICIENTS, R(I,J)

.614				161			
.571	.174			162	163		
.456	.039	.337		164	165	166	
-.239	.253	-.639	-.280	167	168	169	170

FIGURE 4 PARAMETERS FOR THE RIGHT, LEVEL 360° TURN

TRACKING ANALYSIS 1/11 11:15

SUBJECT= PILØT #1

CØMMENT: PØWER CHA

PRAM	MEAN	SD	UNITS	VARIABLE NUMBERS	
SPEED	79	5	MPH	171	177
ALT.	1433	38	FEET	172	178
CLIMB	8	118	F/MIN	173	179
RØLL	3.1	1.7	DEG	174	180
PITCH	7.3	1.5	DEG	175	181
HEAD	3.4	1.7	DEG	176	182

CØRRELATIØN CØEFFICIENTS, R(I,J)

-.185					183				
.012	.174				184	185			
.386	.041	-.454			186	187	188		
-.248	.117	-.750	.258		189	190	191	192	
.375	.043	-.450	.995	.262	193	194	195	196	197

FIGURE 5 PARAMETERS FROM THE 80-mph SLOW FLIGHT

TRACKING ANALYSIS 1/11 11:16

SUBJECT= PILØT #1

CØMMENT: DES TURN

PRAM	MEAN	SD	UNITS	VARIABLE NUMBERS	
SPEED	80	4	MPH	198	202
CLIMB	-405	96	F/MIN	199	203
RØLL	-16.3	20.0	DEG	200	204
PITCH	2.4	1.9	DEG	201	205

CØRRELATIØN CØEFFICIENTS, R(I,J)

.341			206		
-.278	-.147		207	208	
-.045	-.580	-.333	209	210	211

FIGURE 6 PARAMETERS FROM THE DESCENDING 360° TURN

TRACKING ANALYSIS 1/11 11:16

SUBJECT= PILØT #1

CØMMENT: DESCENT

PRAM	MEAN	SD	UNITS	VARIABLE NUMBERS	
SPEED	74	3	MPH	212	217
CLIMB	-195	68	F/MIN	213	218
RØLL	-3.7	4.5	DEG	214	219
PITCH	.2	2.4	DEG	215	220
HEAD	-3.5	4.6	DEG	216	221

CØRRELATIØN CØEFFICIENTS, R(I,J)

.681				222			
.611	.435			223	224		
-.773	-.620	-.486		225	226	227	
.610	.440	.998	-.486	228	229	230	231

FIGURE 7 PARAMETERS FOR THE 500-FOOT DESCENT

TRACKING ANALYSIS 1/11 11:17

SUBJECT= PILØT #1

CØMMENT: LAND

PRAM	MEAN	SD	UNITS	VARIABLE NUMBERS	
SPEED	97	7	MPH	232	239
CLIMB	-268	147	F/MIN	233	240
RØLL	-11.0	3.7	DEG	234	241
PITCH	2.8	2.2	DEG	235	242
HEAD	-10.9	3.8	DEG	236	243
SLØPE	-1.1	2.0	DEG	237	244
L/R	-.0	.4	DEG	238	245

CØRRELATIØN CØEFFICIENTS, R(I,J)

.707						246					
.192	.133					247	248				
-.656	-.670	-.060				249	250	251			
.190	.130	.991	-.057			252	253	254	255		
.063	-.282	.109	-.174	.111		256	257	258	259	260	
.258	.111	.203	-.134	.182	.115	261	262	263	264	265	266

FIGURE 8 PARAMETERS FOR THE ILS LANDING APPROACH

This warm-up procedure took from 15 to 30 minutes, depending on the S. More warming-up time was spent with the B group than the I or A groups. If a S did not reasonably complete one of the warm-up maneuvers, he was given a second opportunity before going on to the next.

2. Test I--Altitude and Heading Holding

In this test the S was requested to hold altitude at 1000 ft and heading at 90° (due east) while maintaining level flight with a power setting of 2400 rpm. He was told that simulated rough air would make the trainer pitch and roll slowly and that he was to keep as level as possible during the test run while maintaining altitude and heading.

A long warm-up* time (34.2 sec) was provided in this test, to allow the S to adapt to the command signals which started abruptly with the onset of the experiment. During the remainder of the test (545 sec) the S's performance was continuously monitored. The six flight variables monitored in this test are listed in the computer printout shown in Figure 1. The frequencies of the roll and pitch command signals in hertz are also listed in Figure 1.

3. Test II--Altitude and Heading Holding with Power Changes

Test II was identical to Test I, except that the S was additionally required to perform the series of power changes in synchronism with the trainer's clock as listed in Table 1. As in Test I, the S first set up the trainer with proper altitude, heading, and power, then he

Table 1

SEQUENCE OF POWER CHANGES

Time after Start (min)	Power (rpm)
1	2000
3	2600
5	2000
7	2600
9	2400

*The warm-up time was a time before the S's performance was monitored during which the LINC-8 generated the sum-of-sines command signals.

set the minute hand of the aircraft clock about 1 or 2 minutes before an hour and continued to hold. When the time reached 0 minutes and 0 seconds, the S called "mark" to the experimenter (E), who started the computer. The variables measured in Test II are shown in Figure 2.

4. Test III--Flight Profile

This test consisted of five short (2- to 3-minute) maneuvers performed in quick succession. The sequence of maneuvers is in Table 2. The S set up the clock as in Test II and called "mark" at the beginning of the first maneuver (time 0) and at the beginning of the other maneuvers, as indicated in Table 2. When E heard "mark," he started the computer monitor. The S was monitored for only the first 75 percent (approximately) of each maneuver, to keep the mean flight parameters from being biased if the S completed the maneuver too quickly and to allow the E time for entering commands at the teletype to set up the next maneuver.

The flight variables monitored and the output in each of the five maneuvers are given in Figures 3 through 7. No command signals were generated by the computer for the flight profile, but the rough air (internally generated by the GAT-1) was kept at its maximum value for the duration of the test.

5. Test IV--ILS Landing Approach

In this final test the S's task was to climb to 2300 ft, with power at 2400 rpm, and heading at 288°, and fly toward the simulated airport at Seaport Beach. In addition to the instrument heading, a visual airport-reference was used consisting of a white piece of paper pinned to the cloth drape surrounding the trainer at a compass heading of 288°. The position plotter was used to position the trainer 12 miles (approximately) from the airport, directly on the approach beam. As the S flew into the glide slope beam and the glide-slope needle crossed zero, he was instructed to call "mark" to E and to begin descending, while keeping the glide-slope and localizer needles crossed and his airspeed at 100 mph.

Since many of the S's had never flown an ILS approach, even many of the Ss in the A group, a warm-up was given before the monitored test. The warm-up consisted of placing an S on the localizer beam before he started climbing to 2300 ft and letting him follow the beam while climbing, and then again while descending several hundred feet on the beam after

Table 2

SEQUENCE OF EVENTS IN THE FLIGHT PROFILE

Maneuver (Monitoring duration in Seconds)	GAT-1 Time (Min)	Subtasks
Pretest		Hold altitude at 500'. Hold heading at 90° Hold power at 2400 rpm.
Climb (95)	0	Call "mark." Change power to 2600 rpm. Climb at 500'/min to 1500'.
		Level off at 1500'. Attain airspeed of 110 mph.
Right Turn (95)	2-1/2	Call "mark." Start right standard 360° turn. Hold altitude of 1500'. Hold airspeed of 110 mph.
Slow flight (145)	5	Call "mark." Reduce speed to 80 mph. Hold altitude at 1500'. Hold heading at 90°.
Descending Turn (85)	8	Call "mark." Switch on half flaps. Lower power to 1200 rpm. Start left standard 360° turn. Descend at 500'/min to 500'.
Descent (85)	10	Call "mark." Switch on full flaps. Raise power to 2500 rpm. Descend at 250'/min to ground. Hold heading at 90°.

the glide slope had been intercepted. During the warm-up E gave S verbal help to keep him on the beam and suggested the strategy for returning if he drifted off. After this warm-up, the test began, and the variables given in Figure 8 were monitored.

III RESULTS

Each of the 266 flight variables was given a one-way analysis of variance to determine its relative importance in separating the three groups. The distribution of the F values obtained from these analyses is shown in Table 3. The 40 variables that were statistically significant at the 5-percent level are shown in Figure 9 together with the significance level achieved.

Table 3

NUMBER OF VARIABLES WITH AN F VALUE GREATER THAN THE LIMIT SHOWN AND LESS THAN THE NEXT HIGHER LIMIT SHOWN

F Limits	Significance	Variables
0	.50	138
1.0		33
1.5		20
2.0		18
2.5	.10	12
3.0		5
3.35	.05	15
4.24	.025	17
5.49	.01	2
6.49	.005	3
7.81	.002	2
8.81	.001	1
Total		266

The distribution of F values shown in Table 3 was used to select the variables for input to the multivariate discrimination program. The set of variables with F values larger than a given F criterion consists of two groups: those representing a real difference between the three experience groups and those representing chance statistical fluctuations in the data. As the criterion is raised, fewer variables are included,

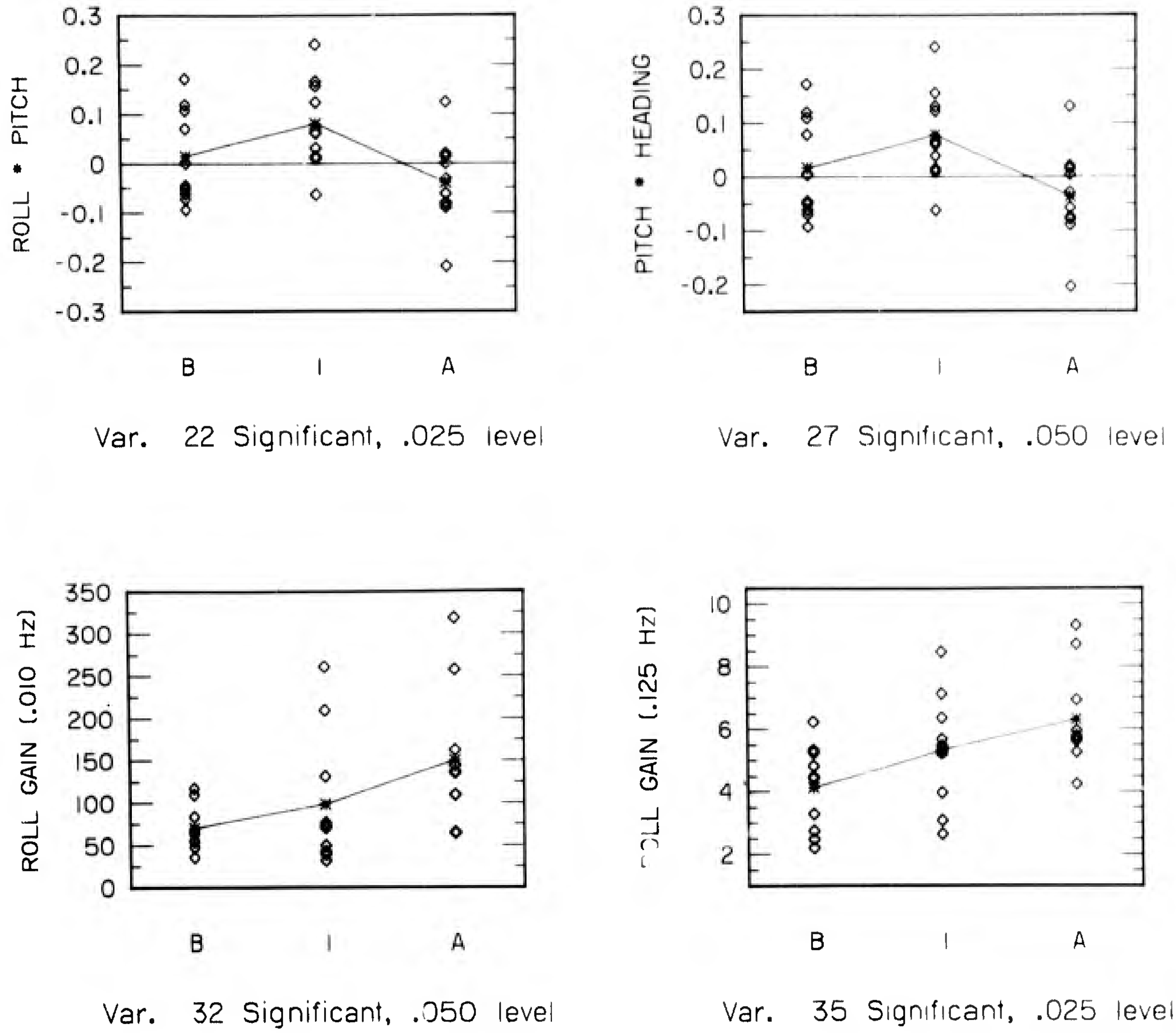
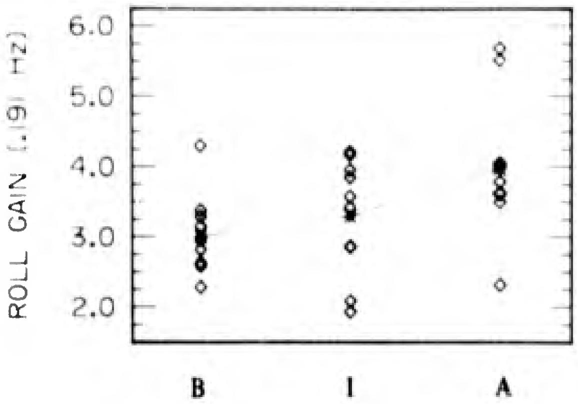
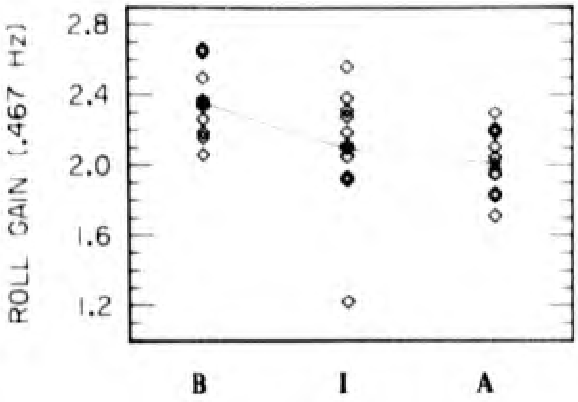


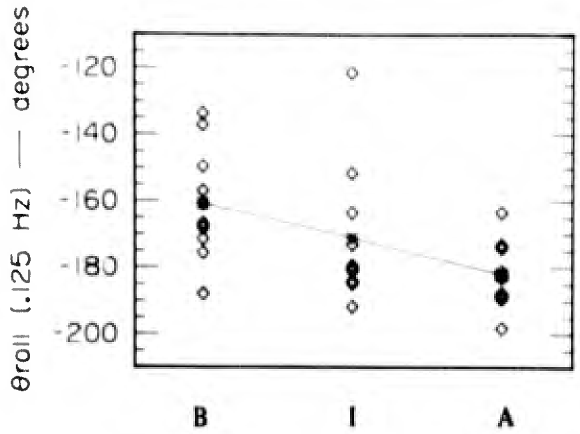
FIGURE 9 PLOTS OF THE STATISTICALLY SIGNIFICANT VARIABLES. The diamond-shaped data points represent the values measured for each S, and the asterisks connected by lines are the mean values obtained for each group (B = Beginner, I = Intermediate, and A = Advanced).



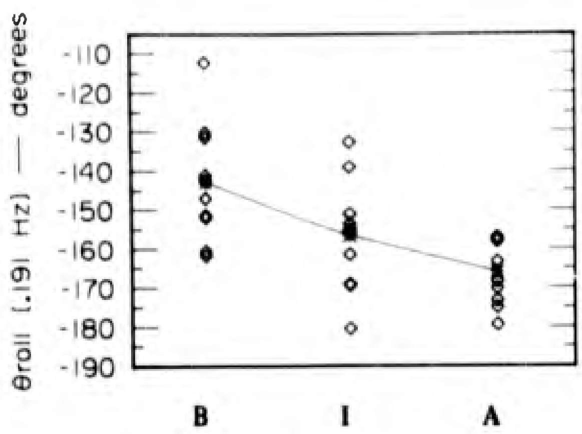
Var. 36 Significant, .050 level



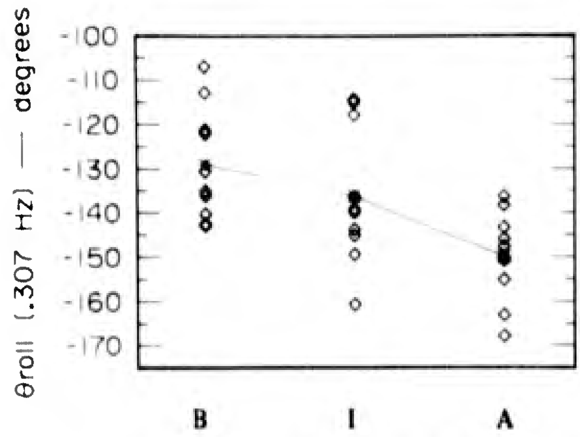
Var. 38 Significant, .025 level



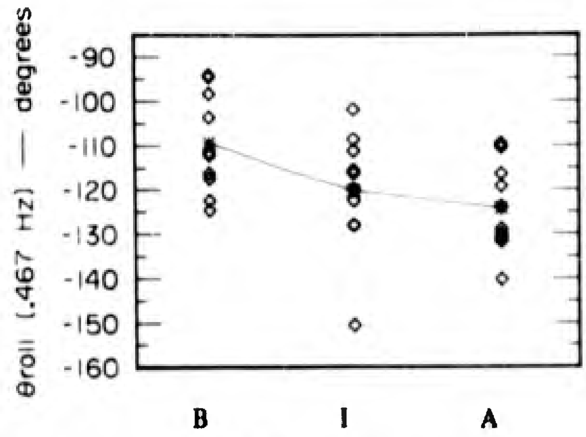
Var. 43 Significant, .050 level



Var. 44 Significant, .002 level



Var. 45 Significant, .005 level



Var. 46 Significant, .025 level

FIGURE 9 PLOTS OF THE STATISTICALLY SIGNIFICANT VARIABLES. The diamond-shaped data points represent the values measured for each S, and the asterisks connected by lines are the mean values obtained for each group (B = Beginner, I = Intermediate. and A = Advanced). (Continued)

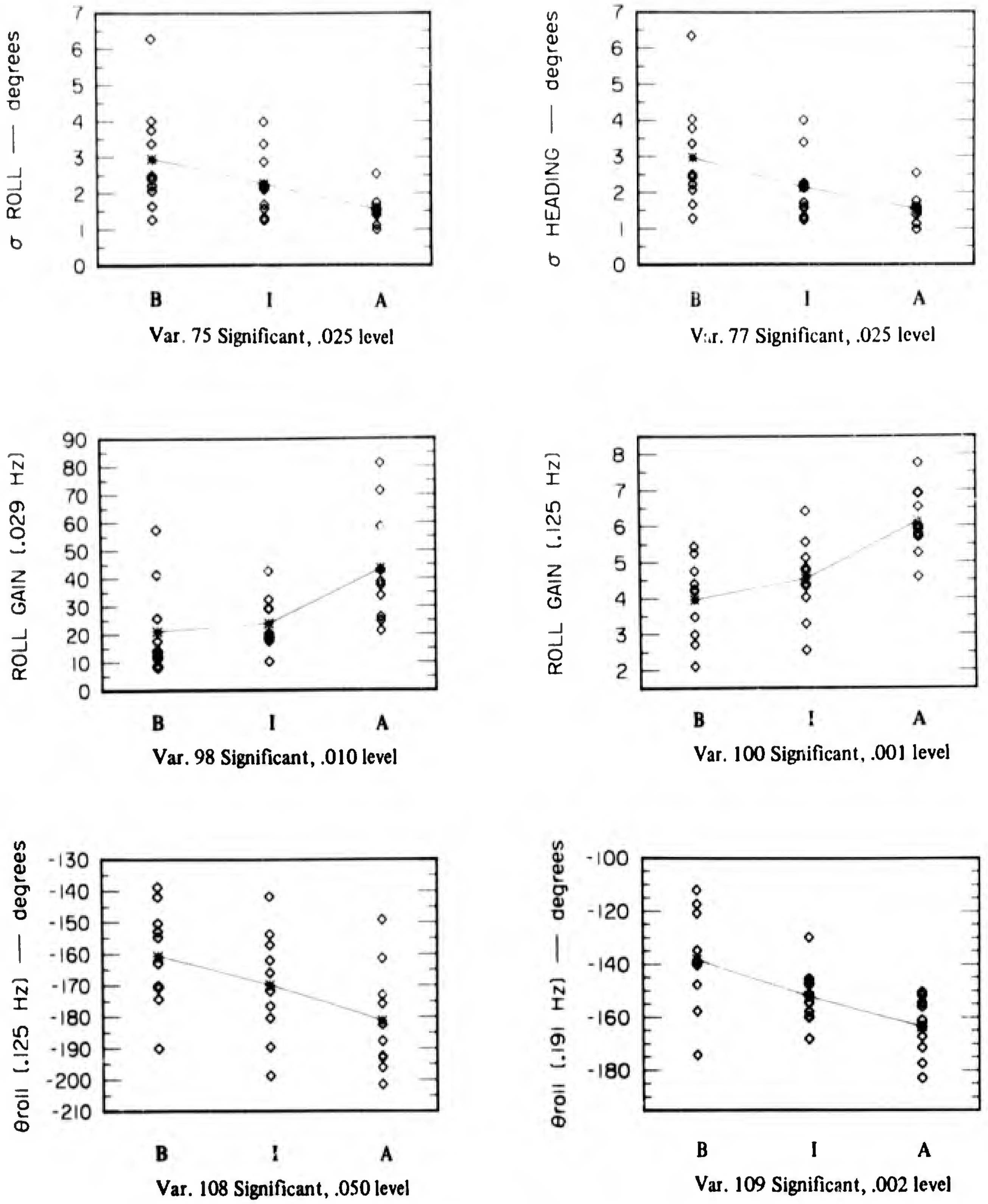
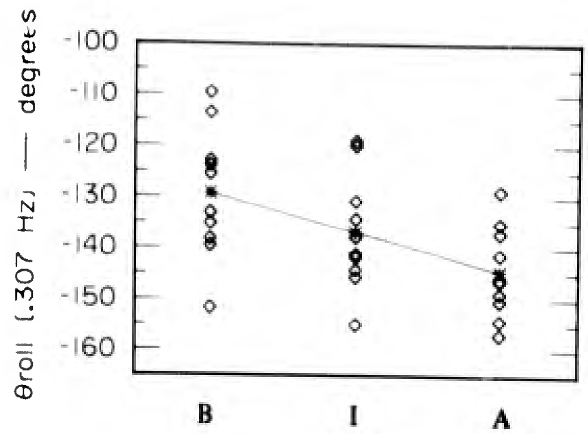
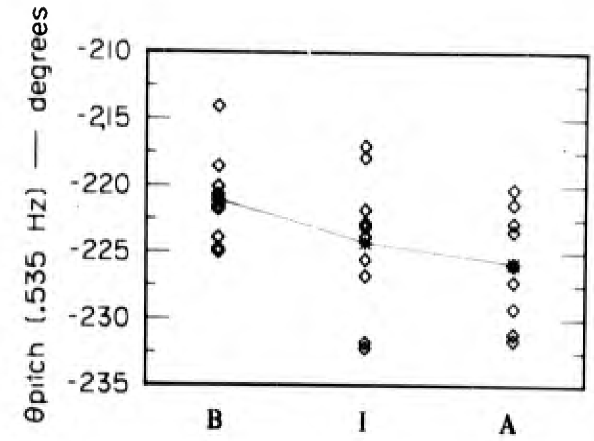


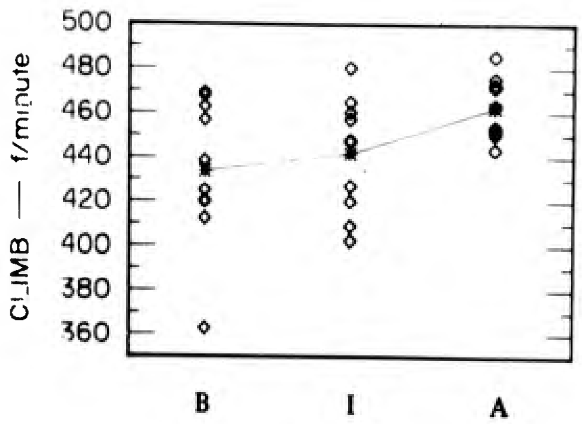
FIGURE 9 PLOTS OF THE STATISTICALLY SIGNIFICANT VARIABLES. The diamond-shaped data points represent the values measured for each S, and the asterisks connected by lines are the mean values obtained for each group (B = Beginner, I = Intermediate, and A = Advanced). (Continued)



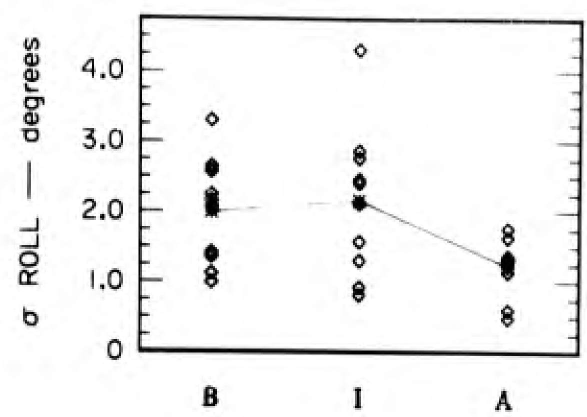
Var. 110 Significant, .025 level



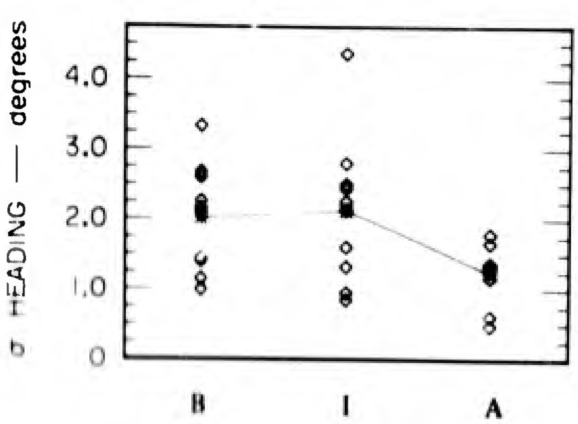
Var. 130 Significant, .050 level



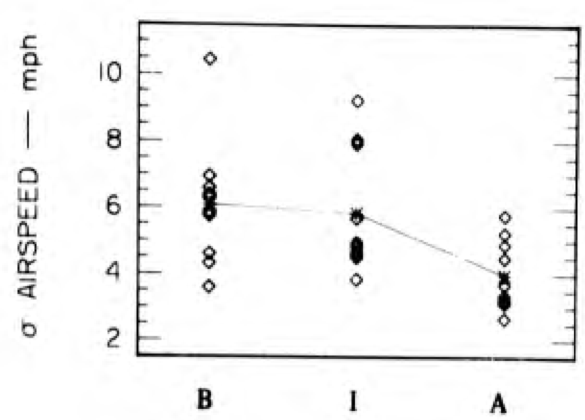
Var. 132 Significant, .050 level



Var. 138 Significant, .050 level

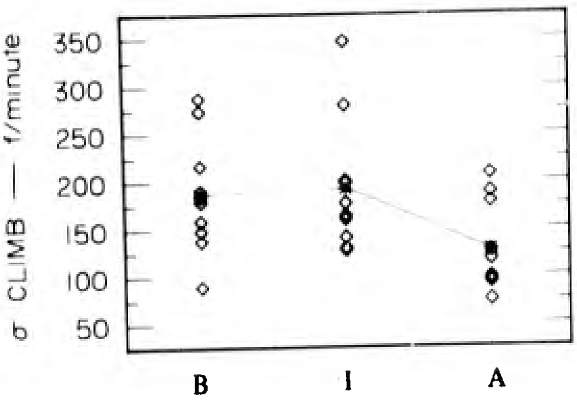


Var. 140 Significant, .050 level

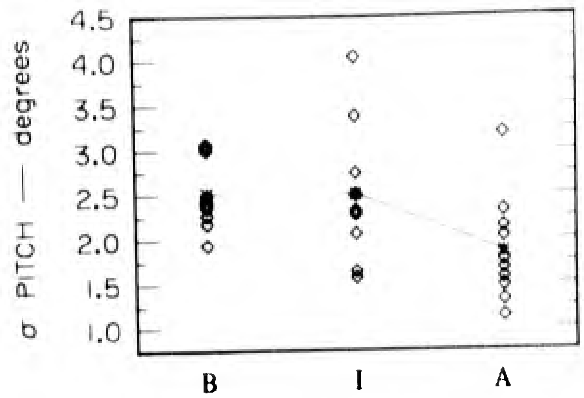


Var. 156 Significant, .025 level

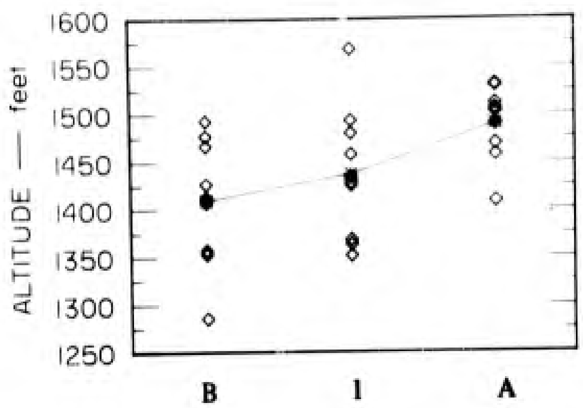
FIGURE 8 PLOTS OF THE STATISTICALLY SIGNIFICANT VARIABLES. The diamond-shaped data points represent the values measured for each S, and the asterisks connected by lines are the mean values obtained for each group (B = Beginner, I = Intermediate, and A = Advanced). (Continued)



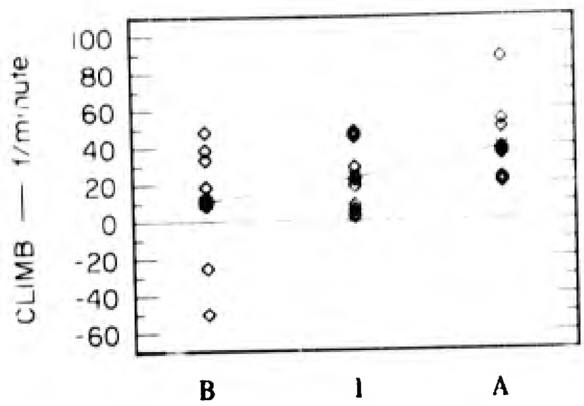
Var. 158 Significant, .050 level



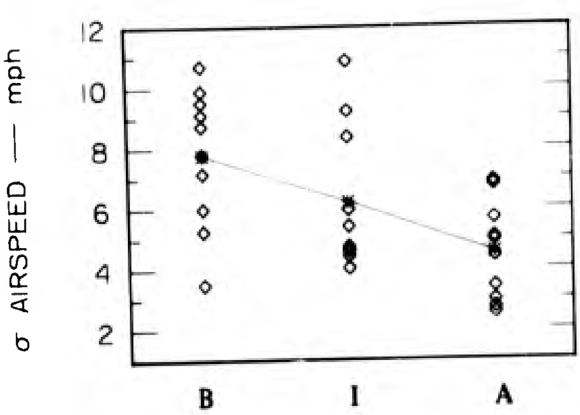
Var. 160 Significant, .050 level



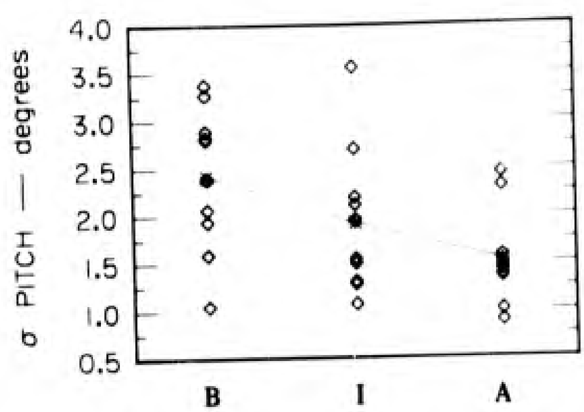
Var. 172 Significant, .025 level



Var. 173 Significant, .050 level

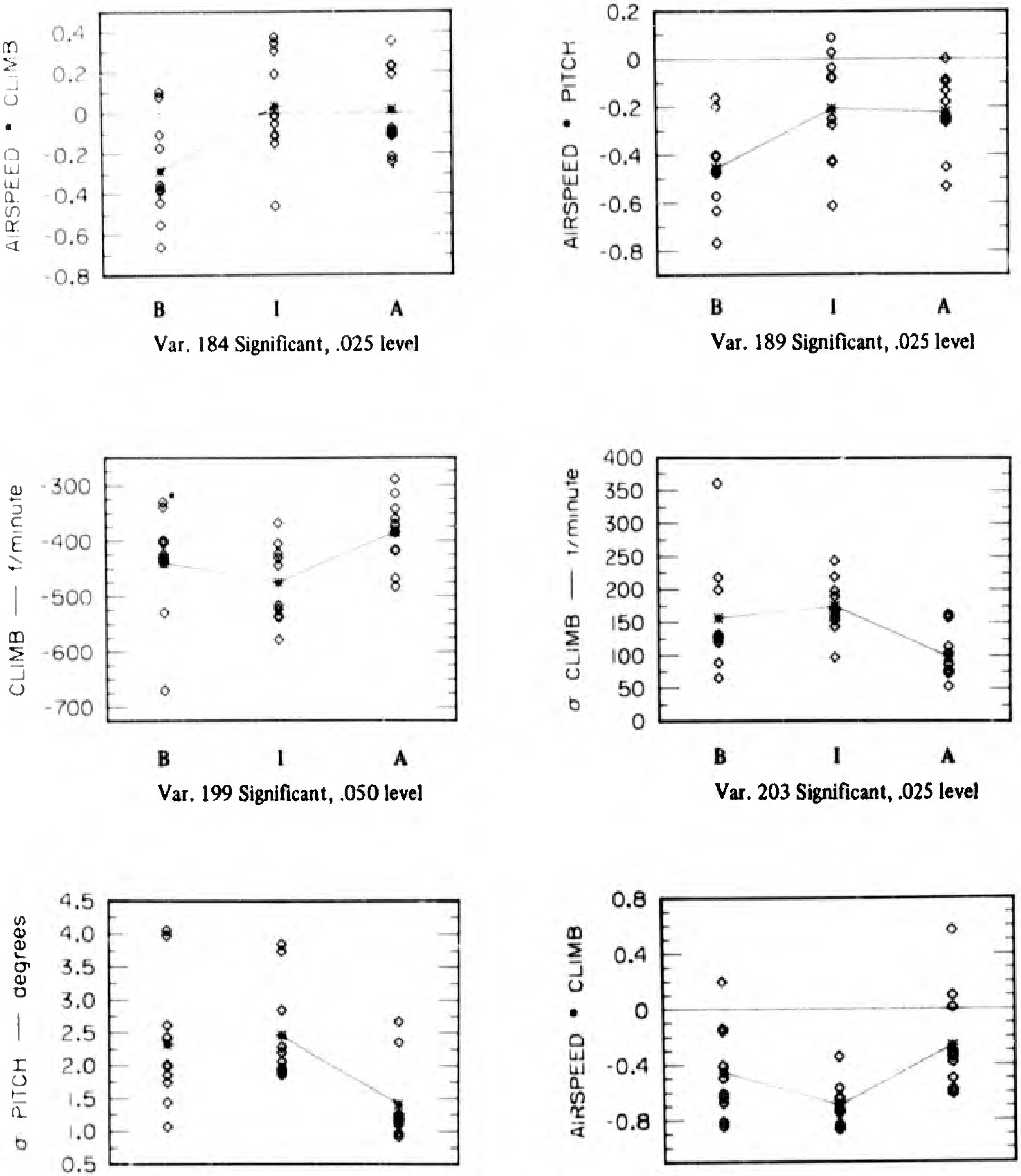


Var. 177 Significant, .010 level



Var. 181 Significant, .025 level

FIGURE 9 PLOTS OF THE STATISTICALLY SIGNIFICANT VARIABLES. The diamond-shaped data points represent the values measured for each S, and the asterisks connected by lines are the mean values obtained for each group (B = Beginner, I = Intermediate, and A = Advanced). (Continued)



Var. 184 Significant, .025 level

Var. 189 Significant, .025 level

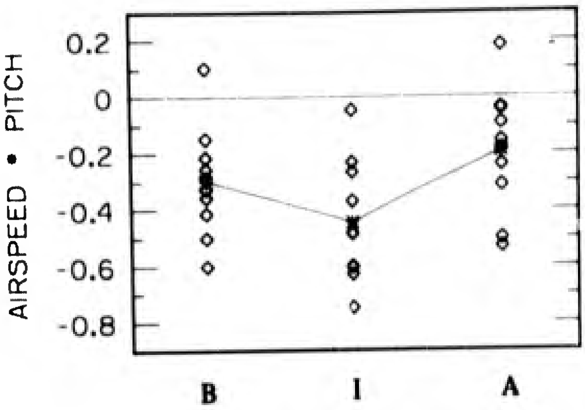
Var. 199 Significant, .050 level

Var. 203 Significant, .025 level

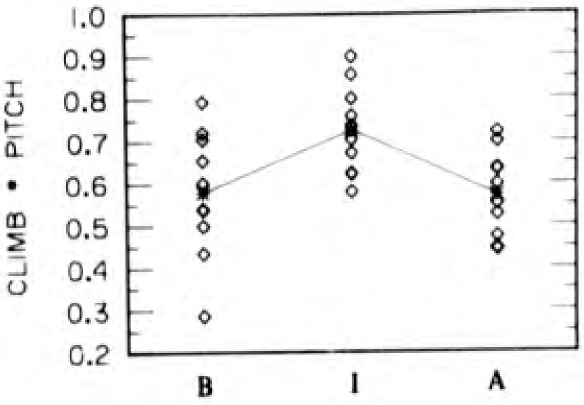
Var. 205 Significant, .025 level

Var. 206 Significant, .025 level

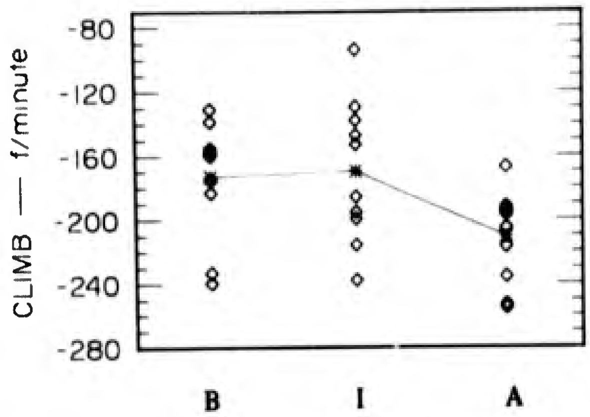
FIGURE 9 PLOTS OF THE STATISTICALLY SIGNIFICANT VARIABLES. The diamond-shaped data points represent the values measured for each S, and the asterisks connected by lines are the mean values obtained for each group (B = Beginner, I = Intermediate, and A = Advanced). (Continued)



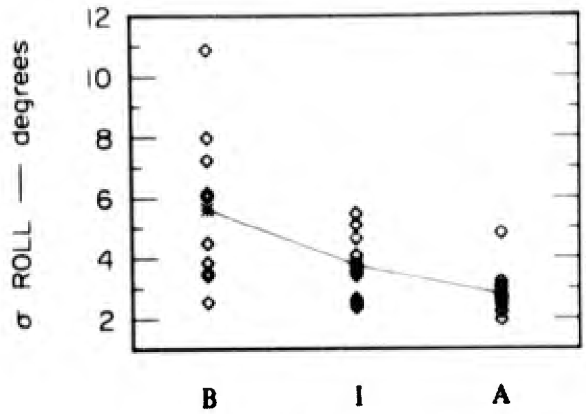
Var. 209 Significant, .050 level



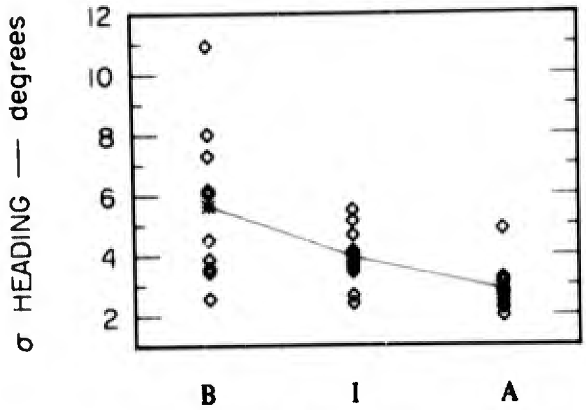
Var. 210 Significant, .025 level



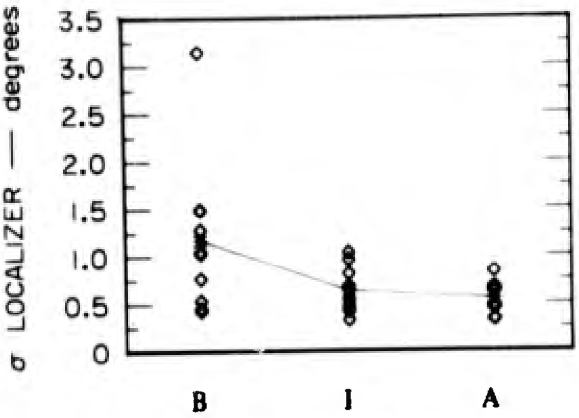
Var. 213 Significant, .050 level



Var. 241 Significant, .005 level



Var. 243 Significant, .005 level



Var. 245 Significant, .025 level

FIGURE 9 PLOTS OF THE STATISTICALLY SIGNIFICANT VARIABLES. The diamond-shaped data points represent the values measured for each S, and the asterisks connected by lines are the mean values obtained for each group (B = Beginner, I = Intermediate, and A = Advanced). (Concluded) but they are more likely to show real, repeatable differences. As the criterion is lowered, more variables are included, and a larger portion of them are variables significant by chance, showing differences between groups that will not be repeated in subsequent tests.

As a compromise between loosing important variables and gaining variables significant only by chance, an F criterion of 2.5 was chosen. This value is achieved by 10 percent of the F's computed on samples drawn from the same population. A smaller F criterion would begin to include more false positives than positives and a larger F criterion would eliminate too many discriminating variables.

The 57 variables for the 30 Ss were entered into the stepwise multivariate discriminant analysis program, BMD07M,* for the final selection of the most important flight variables. The group means and grand means for these variables are given in Table 4 as obtained from this program. The program went through 27 iterations before it stopped adding the variable with the highest F value at each step and recomputing the F values of the remaining unselected variables on the basis of the remaining pooled variance. In this way 27 of the 57 flight variables were selected by the program as being significant in discriminating between the three groups of pilots before the F level of the remaining variables became lower than that required for further computation. These final variables are given in Table 5, with their rank of importance.

After selecting the variables that best discriminate between the three groups, the BMD07M program performs a coordinate rotation in 27-space to find new orthogonal variables (canonical variables or factors) that account for most of the dispersion (or variance) in the data. For the analysis described, the first canonical variable accounted for 99.4 percent of the total dispersion and completely separated the three groups of pilots. Other canonical variables are obviously not important.

The value of the canonical variable, y, for each pilot in the test series is given by

$$y = \sum_{j=1}^{r} u_j (x_j - \bar{x}_j) \quad ,$$

---

*A description of BMD07M, including a description of the step-by-step calculations performed, is given in Dixon (1967).

Table 4

GROUP MEANS AND GRAND MEANS OF THE 57 SELECTED FLIGHT VARIABLES

VARIABLE	GROUP B	I	A	GRAND
	2.54623	2.27871	1.53095	2.11863
10	2.23927	1.87304	1.63540	1.91590
11	2.56576	2.17797	1.53010	2.09127
12	0.01516	0.08027	-0.03974	0.01956
22	0.01633	0.07706	-0.03708	0.01877
27	69.85413	98.50087	150.04590	106.13362
32	9.63837	14.94364	17.84645	14.14282
34	4.13256	5.33747	6.31923	5.26309
35	3.00001	3.29179	2.96922	3.42034
36	2.35219	2.09285	2.00957	2.15153
38	-160.75146	-171.08241	-182.10034	-171.31133
43	-143.04228	-156.78362	-166.06708	-155.29765
44	-129.04201	-136.18913	-149.94662	-138.39258
45	-109.45081	-120.23413	-124.22620	-117.97037
46	4.39890	4.19608	3.99781	4.19760
70	2.94824	2.24893	1.53017	2.24244
75	2.60037	2.27751	2.25681	2.37823
76	2.96704	2.13739	1.53466	2.21303
77	21.05275	23.79065	43.50586	29.44975
98	3.95644	4.52719	6.13003	4.87188
100	3.10814	3.44017	3.87676	3.47502
101	-160.66289	-169.81158	-181.32009	-170.59830
108	-138.31097	-152.20261	-163.73010	-151.41444
109	-129.27719	-136.91345	-144.47656	-136.88905
110	50.09497	26.44533	22.49861	33.01297
116	-221.12643	-224.25851	-225.87091	-223.75182
130	433.52490	441.59985	462.22485	445.78320
132	1.98837	2.15357	1.22291	1.78828
138	2.00794	2.12129	1.24231	1.79051
140	-1.97500	19.00000	51.62500	22.88333
153	6.07864	5.83223	3.96751	5.29279
156	54.78842	49.61769	35.73451	46.71355
157	186.62119	192.16701	126.79636	168.52812
158	2.51859	2.50618	1.85784	2.29420
160	-0.42905	-0.41167	-0.30733	-0.40268
167	1409.25537	1436.81348	1489.07910	1445.04932
172	10.67500	22.92499	38.72499	24.10832
173	7.76723	6.21369	4.51055	6.16383
177	3.88833	2.94373	1.33141	2.72116
180	2.41710	1.91738	1.52141	1.95196
181	-0.28477	0.03262	0.01576	-0.07880
184	-0.45458	-0.20661	-0.22519	-0.29546
199	-439.37500	-476.09985	-385.00000	-433.49146
200	-10.69159	-13.61504	-5.15651	-11.15438
201	1.43479	1.16146	2.15676	1.58434
203	156.23405	174.02374	56.98511	142.41418
205	2.31822	2.46031	1.40291	2.06048
206	-0.45720	-0.69671	-0.26603	-0.47331
209	-0.29859	-0.44596	-0.19242	-0.31232
210	0.57779	0.72550	0.57133	0.62487
213	-173.12500	-169.72499	-211.39999	-184.75000
217	4.15295	3.64201	7.43291	5.07596
225	-0.45556	-0.50013	-0.23476	-0.39692
239	7.03551	5.59606	4.40942	5.68033
241	5.60559	3.74911	2.80310	4.05260
243	5.64246	3.93252	2.82877	4.13459
245	1.16360	0.62990	0.54360	0.77903

Table 5

COEFFICIENTS FOR THE FIRST CANONICAL VARIABLE

Variable	Rank	Coefficient
22	6	-466.65
32	17	0.0088
34	11	0.81
35	9	-60.60
36	16	24.18
43	10	2.98
44	15	-7.73
75	24	11.25
76	22	-67.67
98	13	1.74
100	1	83.69
109	5	-1.37
110	23	2.42
116	2	-1.39
130	12	1.36
132	3	1.29
138	8	13.01
160	20	97.90
167	4	38.82
172	18	0.52
173	27	-0.041
200	21	-6.37
201	26	3.00
206	19	-69.60
213	14	-0.57
225	25	20.16
245	7	-72.99

where r is the number of discriminating variables chosen (27); u_j are the weighting coefficients listed in Table 5; x_j are the values of the 27 variables (listed in Table 5) obtained for the given pilot; and $\bar{x}_j$ are the grand means for each of the discriminating variables included in Table 4. The average values of y for the B, I, and A groups are -176.4, -2.7, and 179.2, respectively.

Since the separation of the pilots by the 27 most important variables was so complete, the BMD07M program was run again but stopped, after 5, 10, 15, 20, and 25 variables had been selected, in order to determine how the separability between the groups increases with an increasing number of variables. If only one or two variables could be used to separate the three groups, there would be no use making more than a few measurements. The graphical output supplied by the computer program for 10, 20, and the final 27 most important variables is shown in Figures 10, 11, and 12. These figures show that within-groups dispersion decreases slowly as the number of discriminating variables become larger: There does not seem to be an abrupt change after accepting any additional variable. With 10 variables only one pilot is classified in the wrong group, but with 15 or more variables all are correctly classified.

The results of this multivariate analysis suggest that as many of the 27 variables as possible should be used to make the actual discriminations between pilots. They also suggest that several (at least ten) variables must be used to adequately measure pilot ability.

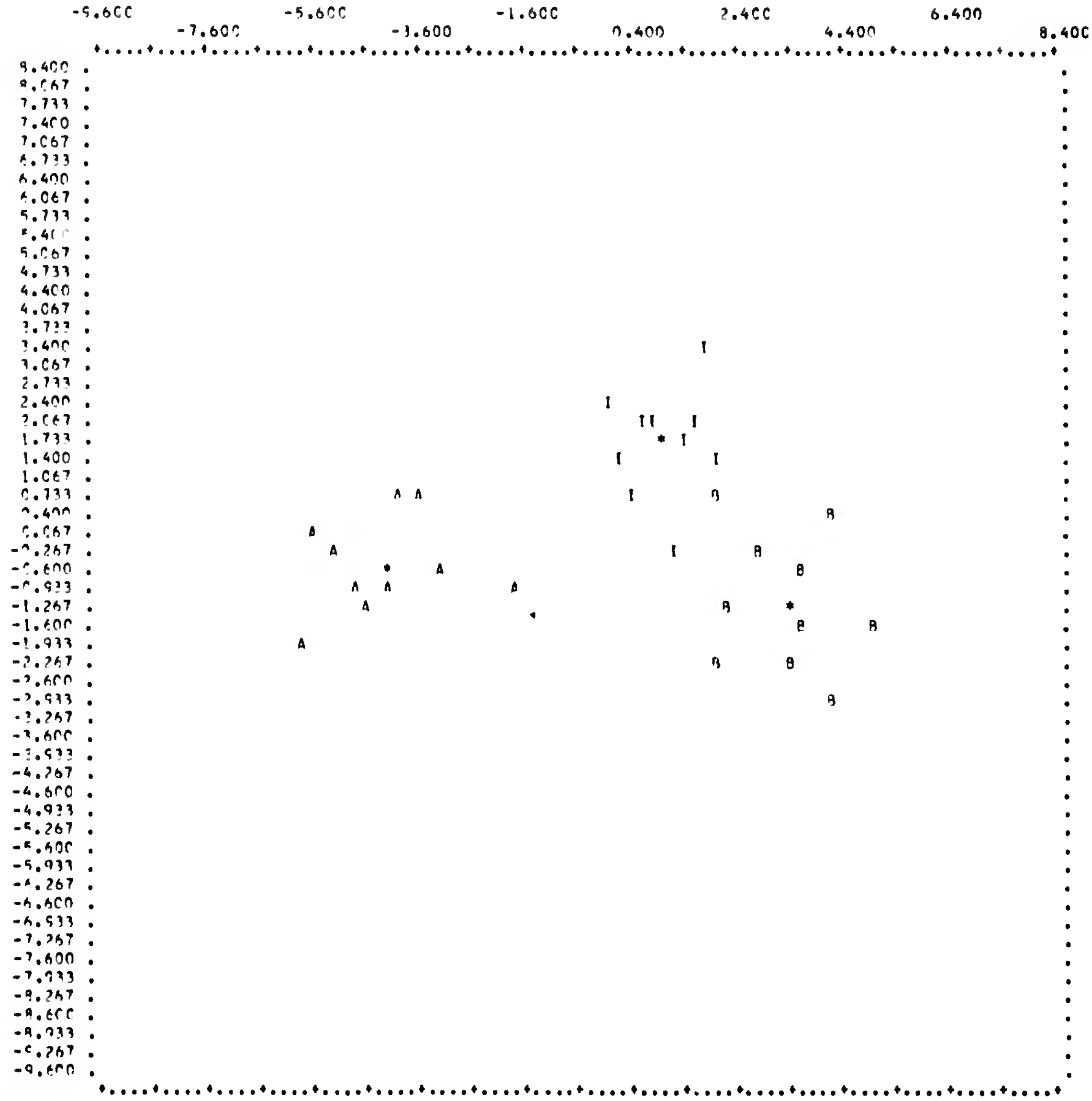


FIGURE 10 DISPERSION OF THE DATA ON THE BASIS OF THE TEN MOST IMPORTANT VARIABLES. The abcissa and ordinate are the first and second canonical variables, respectively. Each S is plotted with the symbol for his group. Group means are plotted with the * symbol.

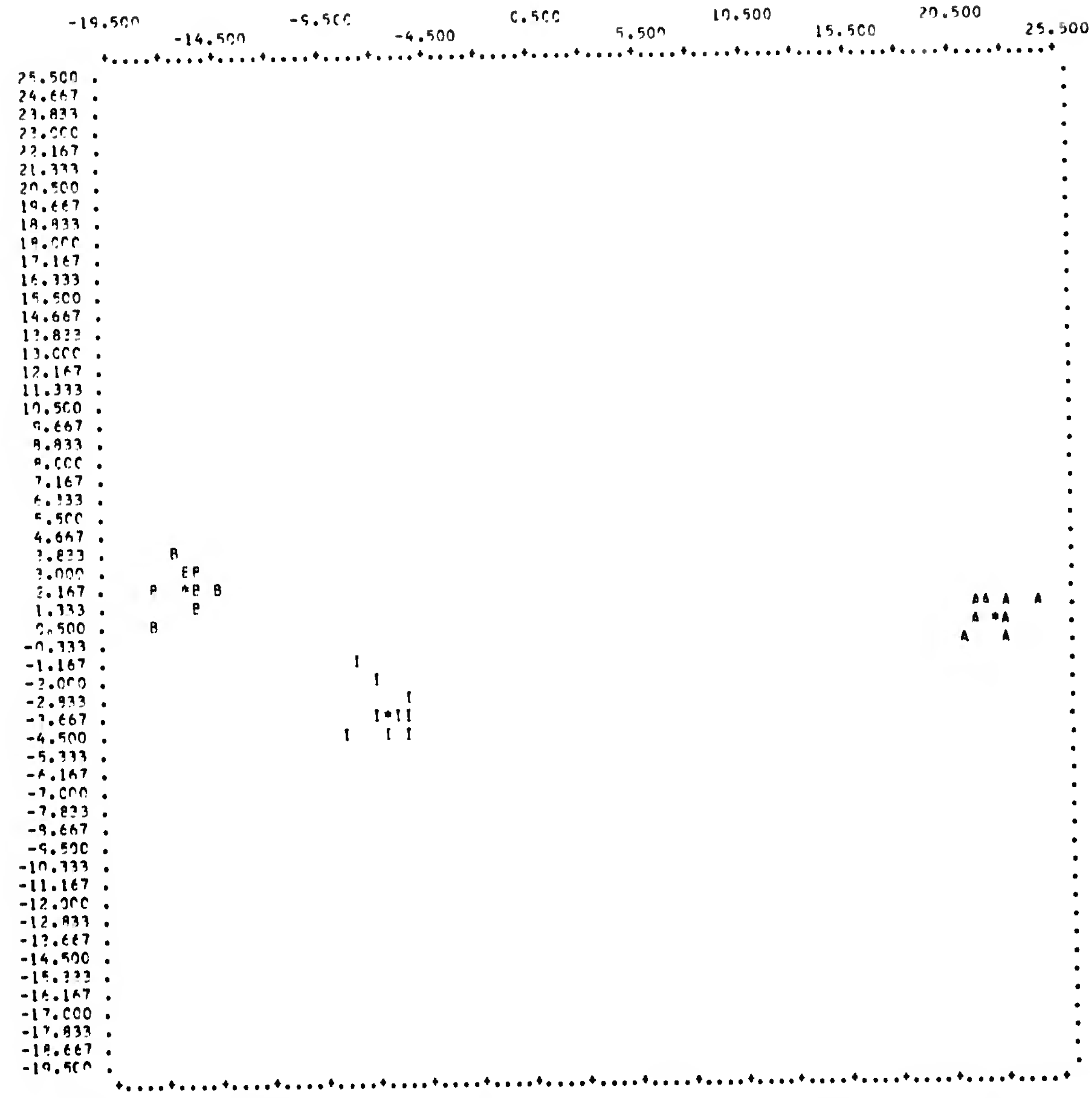


FIGURE 11 DISPERSION OF THE DATA ON THE BASIS OF THE 20 MOST IMPORTANT VARIABLES. The plot is explained in the caption of Figure 10.

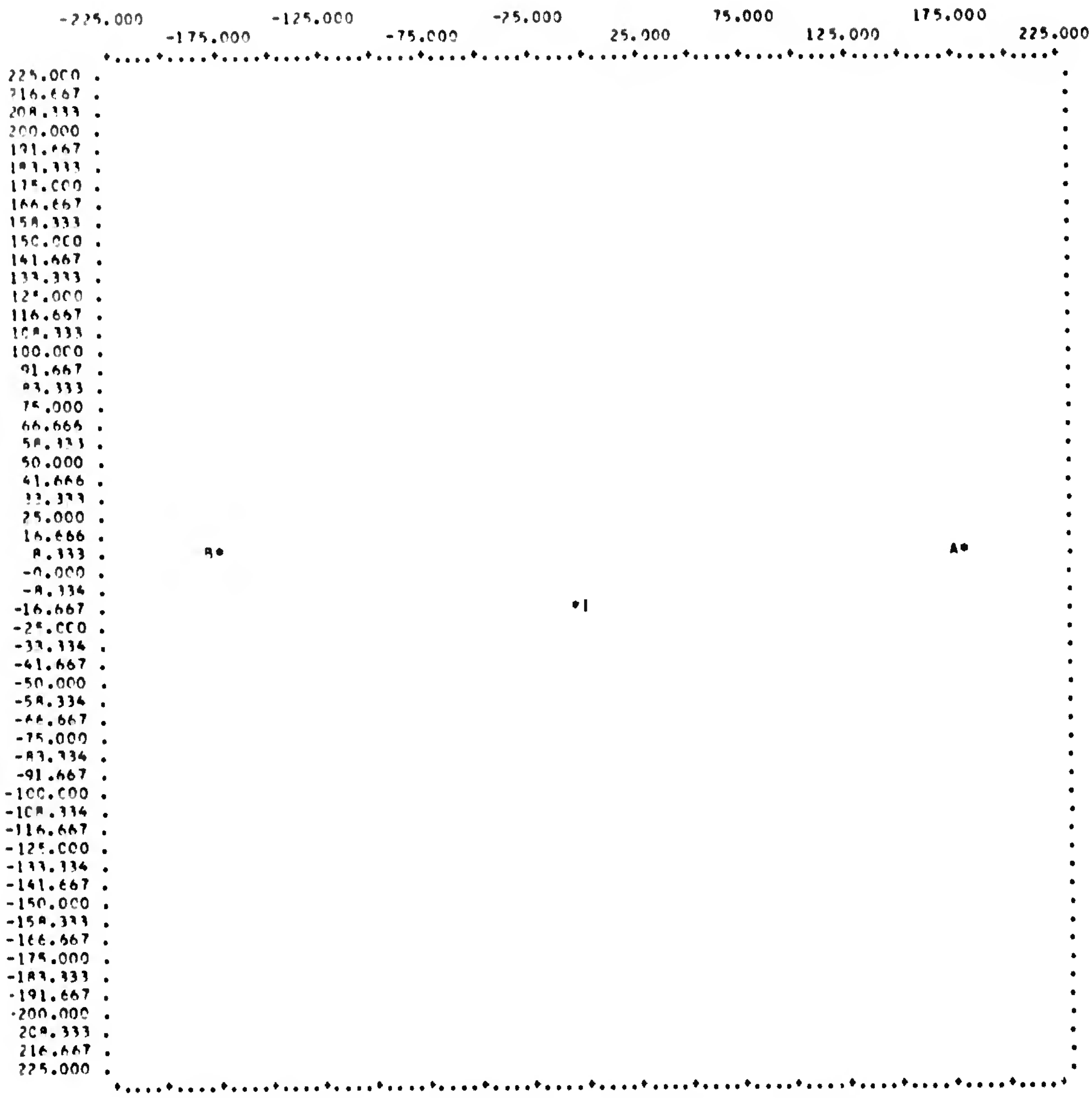


FIGURE 12 DISPERSION OF THE DATA ON THE BASIS OF ALL 27 IMPORTANT VARIABLES. The plot is explained in the caption of Figure 10.

IV DISCUSSION

To determine which of the four flight tests and which of the five generic types of variables measured in the tests contributed most to pilot separation, the summary shown in Table 6 was made. Of the four

Table 6

BREAKDOWN OF THE IMPORTANT VARIABLES

Source of Variables	Total Number of Variables	Number Significant (5% test)	Number Selected by BMD07M
Test I	65	10	7
Test II	65	8	8
Test III	101	21	11
Test IV	35	3	1
Means	44	5	6
Deviations	44	13	5
Correlations	114	6	4
Gain	32	7	6
Phases	32	8	6
Roll Tracking	32	13	10
Pitch Tracking	32	2	2

tests shown in Table 6, the flight profile (Test III) contributed the highest percentage of significant variables (21 percent). At the other extreme, the ILS landing (Test IV) contributed only 9 percent of the significant variables and only one of the variables selected by BMD07M. Even though Test IV produced only one important variable, the localizer standard deviation, this variable was the seventh most important in separating the three groups of Ss and should not be deleted from the analysis. Since all of the tests contributed variables for pilot selection, it appears that a group of tests requiring several different skills, types of coordination, tasks, etc., should be used to separate the pilots, instead of only one test.

Of the five types of measurements made in the test (means, standard deviations, correlations, tracking gains, and tracking phase shifts), the standard deviations produced the highest proportion (30 percent) of significant differences between Ss. Although more of the deviations (30 percent) than means (11 percent) were significant at the 5-percent level, slightly more means than deviations were chosen by BMD07M, suggesting that both measurements are important for separating pilots.

The gains and phases measured in the tracking tasks showed the next-best ability to separate Ss, with 22 percent and 25 percent, respectively, of the variables being significant. A breakdown of the tracking variables into those obtained in roll and pitch tracking, also given in Table 6, shows that almost all the significant tracking variables were obtained from roll tracking. In fact, roll tracking was the most efficient task for generating significant variables (41 percent were significant at the 5-percent level) and for separating the groups (ten tracking variables were chosen by BMD07M). Pitch tracking, on the other hand, was barely above chance, with only 6 percent of the variables significant at the 5-percent level and only two variables chosen by BMD07M.

The correlation coefficients were the least productive in showing significant differences between pilots. Since only 5 percent of them were significant at the 5-percent level, the correlations should probably be eliminated from the tests. However, a few of the correlations were chosen by the BMD07M program do differentiate between Ss. Two of the variables (No. 22, roll correlated with pitch and No. 167, airspeed correlated with pitch) ranked sixth and fourth, respectively, in separating the three S groups and should not be eliminated without more study.

V SYSTEM MODIFICATION

An instrumentation type tape recorder could replace the on-line computer in monitoring flight variables. The variables from the GAT-1 could be played back at a later time through a high-speed data link to a remote computer to obtain final performance measurements. An advantage in tape recording the signals is that the test runs could be played back at slower speed, for transmission over slow-speed data links, or played back at faster speeds, compressing hours of data runs into minutes if direct connections to a high-speed computer were available. Since only six of the eight variables measured in this experiment played a role in separating the three S groups (heading and glide slope were variables never selected by BMD07M), only six need be recorded. In addition, since the localizer signal was used only when the altitude signal was changing, these two signals are mutually exclusive and could share a channel on the recorder, reducing the number of simultaneously used channels to five.

The use of a tape recorder for the tracking tasks involves some additional steps however. The roll and pitch command signals must be prerecorded on two analog channels, and synchronization (or timing) marks must be recorded on another. The roll and pitch command signals must be provided by the tape recorder simultaneously as the other variables are being recorded. The sync channel is prerecorded at the same time as the command signals and may be a steady train of 15-Hz pulses corresponding to the incremental phase changes in the sum-of-sines command signal. The remote computer would use the sync channel to determine when to sample the other channels and would thus keep synchronized to the phase changes in the command signal. The synchronization and command signal channels would be required only for the tracking runs.

In summary, a tape recorder for repeating the measurements useful in discriminating between pilot groups should have eight channels, five for recording flight variables, two for playing back the prerecorded roll command signal, and one for the prerecorded synchronization signal.* In

---

*A seven-channel recorder could be used in this application if a voice channel were available for recording the sync track.

addition, the seven analog channels (five flight variables and two command signals) should have characteristics better than the following:

(1) Signal-to-noise ratio of 40 dB (100 to 1)

(2) Drift of 1 percent per hour

(3) Accuracy of 1 percent

(4) Bandwidth of 0 to 16 Hz (0.5 dB).

This set of characteristics limits the accuracy in the measurement of means and standard deviations made over a 1-hour period to 1 percent. Since most of the important variables changed by 10 percent or more, a 1-percent random perturbation of these measurements due to the tape recorder will add an acceptably small variance to the results.

Some variables are more sensitive to tape recorder fluctuations than others, however. Variable 172, the average altitude, shows an extreme range of only 80 ft out of an average of 1500 ft. This is only a 5-percent change from Group B to Group A. A 1-percent additional variation in these numbers is still small, but approaches the point where measurement differences will be limited by the tape system. If better performance is desired for monitoring tasks other than those reported here, all of the first three recorder characteristics listed above must improve simultaneously for the system performance to improve. For example, if 0.5-percent differences are to be detected, signal-to-noise ratio must go to 46 dB, and both drift and accuracy must be 0.5 percent or less.

VI RECOMMENDATIONS FOR FURTHER WORK

Since the results of Section III may depend on the two-stage selection process (analysis of variance followed by multivariate discriminant analysis), other means of treating the data are recommended before a final set of discriminating variables is fixed. Another two-step selection processs is that of subjecting all the means from the tests to one discriminant analysis and similarly all the standard deviations, correlations, and tracking parameters to other separate analyses. Not only would these four analyses tell whether each of the four families of measurements could be individually used to discriminate among the pilots, but they would also provide four sets of discriminating variables for a single, final discriminant analysis. This final analysis would double check the analysis described in Section III. These individual analyses would also indicate how well performance monitoring systems that measure only a single family of parameters could discriminate between Ss.

A second means of checking the validity of the discriminating functions is to divide the subject pool randomly into two sets of 15 each (five in each experience group), obtain the discriminant function for the 27 previously chosen flight variables by using each set of 15 Ss, and then determine how well each set is separated by the discriminant function of the other set. Although this approach is simple, not requiring any additional data, it may suffer from noisy discriminant functions based on so few Ss. However, if half of the Ss can be successfully used to classify the other half, and vice versa, this speaks very well for the set of flight variables chosen. A method less subject to noise is that of running 15 new Ss through the same test (five in each group) and seeing whether they can be separated by using the data of the original 30 Ss.

Since the tracking results furnished the largest portion of variables significant in separating the three groups of Ss (12 variables out of 27), the role of tracking tasks in predicting pilot performance should be studied further. The success of tracking as a performance measure may depend strongly on several parameters which have not been varied. These are:

(1) Tracking dimensionability or number of simultaneously varied flight variables (only two-dimensional tracking was used in this study),

(2) The particular flight variables that are tracked (only pitch and roll used in this study),

(3) The amplitude of the command signals, and

(4) The frequencies in the command signals.

Additional analyses of tracking performance not carried out in this study may also yield important performance measurements. One such on-line measurement is the isolation of the command frequencies in other control channels from a given response channel as measured in multidimensional tracking. The amount of this cross-talk between variables may depend on pilot experience. Another analysis of tracking performance is the off-line fitting of a model to the gain and phase characteristics of each pilot. (A series of such models of increasingly better accuracy is given in McRuer et al., 1965.) The resulting model parameters (time delay, time constants, and gain constants) may depend on pilot experience. The off-line Fourier analysis of the S's error signals over a wide range of frequencies (not just the frequencies in the command signals) may additionally allow quantitative separation of S's with jerky control and smooth control.

Finally, one of the limitations in this study was the inability to meaningfully monitor a flight variable with a changing reference. In a climb or turn the changing variable (altitude or heading) must be compared with a changing reference in order to calculate the mean and standard deviation of the error. A change in the on-line program to monitor such changing references would allow both additional tasks to be included in the original battery of flight tests (such as a 360° standard turn synchronized with the clock), ad additional flight variables to be obtained from the original battery of tests.

REFERENCES

Connelly, E. M., A. R. Schuler, and P. A. Knoop, "Study of Adaptive Mathematical Models for Deriving Automated Pilot Performance Measurement Techniques," Technical Report AFHRL-TR-69-7, Vol. I, Air Force Human Resources Laboratory, Wright Patterson Air Force Base, Ohio (October 1969).

Dixon, W. J., "Biomedical Computer Programs," University of California Press, Los Angeles, California (1969).

Garber, E. I., and V. S. Krasnitskii, "Use of Multidimensional Regression Analysis for Predicting Flight Training Quality," A70-38964, Technical Information Service, American Institute of Aeronautics and Astronautics, Inc., 740 Third Avenue, New York, N.Y. 10017 (1968).

Hill, J. W., K. W. Gardiner, and J. C. Bliss, "Design Study of a Tactile Cuing System for Pilot Training," Technical Report AFHRL-TR-69-12, Air Force Human Resources Laboratory, Wright-Patterson Air Force Base, Ohio (August 1969).

Krendel, E. S., and D. T. McRuer, "A Servomechanism Approach to Skill Development," J. Franklin Institute, Vol. 269, No. 1, pp. 24-42 (1960).

McRuer, D., D. Graham, E. Krendel, and W. Resner, Jr., "Human Pilot Dynamics in Compensatory Systems--Theory, Models, and Experiments with Controlled Elements and Forcing Function Variations," Technical Report 115-1, AFFDL-TR-65-15, Systems Technology, Inc., Hawthorne, California (1965).