

MEASURING WORKLOAD DIFFERENCES BETWEEN SHORT-TERM MEMORY
AND LONG-TERM MEMORY SCENARIOS IN A SIMULATED FLIGHT ENVIRONMENT

Scott L. Berg and Thomas B. Sheridan
Man-Machine Systems Laboratory
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract

Four highly experienced Air Force pilots each flew four simulated flight scenarios. Two scenarios required a great deal of aircraft maneuvering. The other two scenarios involved less maneuvering, but required remembering a number of items. All scenarios were designed to be equally challenging. Pilot's Subjective Ratings for Activity-level, Complexity, Difficulty, Stress, and Workload were higher for the maneuvering scenarios than the memory scenarios. At a moderate workload level, keeping the pilots active resulted in better aircraft control. When required to monitor and remember items, aircraft control tended to decrease. Pilots tended to weigh information about the spatial positioning and performance of their aircraft more heavily than other items.

Research Supported by NASA Ames Research Center

I. Introduction

Deregulation is having a profound impact on the airline industry. It has brought increased competition, cut-throat "fare wars", and demands by management for greater employee productivity. This new economic environment has intensified the pressure to cut cockpit crews from three to two persons.

In addition, the nation's airways are becoming more crowded every day, absorbing an ever greater mix of aircraft types, sizes, and performance characteristics. Thus, the need for pilots to spend more time looking outside the cockpit is of major concern.

These conflicting demands for less "in cockpit" workload while simultaneously cutting the cockpit workforce by 33%, have accelerated the push to automate and computerize today's aircraft. New display technologies and microprocessors have led to the widespread use of programmable calculators and a growing number of computer-monitored, computer-flown, and computer-display-dominated flight decks.

This technology has relieved a great deal of the pilot's physical labor in aircraft configured with the latest equipment. However, this equipment has generated its own set of concerns and problems:

- (1) At what point and to what extent does boredom or the lack of "something to do" impact performance?
- (2) Given that pilots must plan, program, and monitor "automatic" equipment, when do these mental tasks begin to overwhelm a crewmember?
- (3) How can this mental workload be measured?
- (4) Can one determine how close a given crew member is operating to his or her "breaking point"?

This research examines these questions using a fixed-base simulator. This interim report looks at the problem of measuring mental workload by subdividing it into short-term mental operations and long-term mental functions such as information "storage" and "retrieval".

II. Objective

This research examines whether certain objective and subjective measures can distinguish between two types of simulated flight scenarios: (1) a scenario emphasizing short-term memory tasks; (2) a scenario with a large number of long-term memory tasks.

III. Simulator Configuration

The flow of information to and from various elements of this experiment is represented schematically on Figure A-1. The PDP-11 Computer acts upon an aircraft dynamics simulation program (four engine Lockheed Jetstar) and presents information on the present flight condition to a Megatek CRT display. (For an in-depth description of the simulation program and aircraft dynamics, see Mental Workload in Supervisory Control of Automated Aircraft, by Tanaka, Buharali, and Sheridan, 1983).

The Megatek display (Figure A-2) simulates an aircraft cockpit display. The upper part of this CRT display gives a simplified "out the window" perspective of an airport and three runways. Below this is a set of instruments in the familiar "T" pattern. An Airspeed Indicator, Attitude Deviation Indicator (ADI) with Glide Slope Deviation Indicator (GSI), and Altimeter comprise the top row. A Horizontal Situation Indicator (HSI) with the selected course (CRS) and distance (DME) to a selected navigation aid, is directly beneath the ADI. A Vertical Velocity Indicator (VVI) is to the right of the HSI. Landing Gear Position (Up, Down), Flap Position (Up, Down), Thrust Setting, Stability Augmentation Selection (On, Off), Navigation Radio Selection (Off, VOR, ILS, channel number), Lateral Autopilot Selection (Off, Manual Heading, VOR Course, Localizer Course), and the Longitudinal Autopilot Selection (Off, Altitude Hold, Speed Hold, Altitude/Speed Hold, Glide Slope/Speed Hold) are also presented.

The subject interprets the displayed flight information and manipulates the controls on the Control Box (Figure A-3) to make the "aircraft" respond in a desired fashion. The Control Box contains an aircraft-type control-stick or joy-stick, a throttle, and a number of other controls. On the top-rear of the box are eight Radio Toggles. To the left of the Throttle are the Course Set Knob and the Flaps and Landing Gear Selector. To the right of the joy-stick is a longitudinal Trim Control. The front panel has six controls: Heading Set Knob; VOR/ILS Selector; Lateral Autopilot Selector; Longitudinal Autopilot Selector; Radio-Navigation Channel Selector; and Stability Augmentation Selector.

Electrical signals convey information on control positions from the Control Box to the Computer. The Computer then uses these inputs to update the flight condition, aircraft dynamics, and display.

The Experimenter (XPRMNT) interacts with the Computer via a separate Video Display Terminal (VDT). After experimental runs are completed, the experimenter can get an output of data stored by the Computer, on a Line Printer.

IV. Data

Every ten seconds, the computer stores aircraft x, y, and z positions. In addition, it stores every control box manipulation along with the magnitude and time of the event. This data yielded Ground Track information. By correlating the aircraft's x, y position with time and the chosen scenario, altitude error was derived.

Since part of each flight consisted of maintaining certain magnetic courses, altitude deviations were much more useful than heading deviations could have been. Furthermore, since the aircraft responds to altitude change commands more quickly than airspeed change commands, and since the range of altitudes and potential altitude deviations are much greater, altitude information was better than airspeed data for monitoring flying precision. This altitude error data was converted into Absolute Altitude Error (Feet) and Root-Mean-Square (RMS) Altitude Error (Feet).

Subjects were simply instructed to follow instructions as precisely as possible; thus, they had no indication of what types of deviations would be used as the scored parameter.

In addition, each subject scored a set of five Subjective Workload Ratings at three points during each run. These Subjective Ratings were Activity-Level, Complexity, Difficulty, Stress, and Workload. Ratings were taken at three points rather than taking one overall rating to see if any "point" loading of workload might be occurring and biasing the ratings.

V. Subjects

Four subjects participated in this experiment. All four were highly experienced Air Force pilots and had flown this simulator several times. An experience summary follows:

B: Fighter-Type:	1250 Hours
Jet:	1250
Total:	1250
H: Fighter-Type:	3200
Jet:	2750
Total:	3200
L: Light Aircraft:	550
Fighter-Type:	1000
Heavy Aircraft:	600
Jet:	1000
Total:	2150
W: Light Aircraft:	100
Fighter-Type:	700
Heavy Aircraft:	1300
Jet:	2000
Total:	2100

VI. Instructions

Figure A-4 is a reproduction of the typewritten instructions given to each subject before each session. A few points require emphasis or explanation. Subjects were instructed to fly as "precisely" as possible. Further, all simulated ARTCC instructions were handled verbally between the subjects and experimenter. The CWS switch is the Stability Augmentation Switch mentioned in Section III.

Along with these instructions, two other items were given each subject. A Subjective Rating Sheet (Figure A-5) was provided, and the subjects were asked to examine it and ask questions pertaining to it. They were instructed to consider each scale as continuous rather than discrete. That is, the subdivisions were provided simply as references for the subjects and

experimenter. Each score sheet was used for one day's activities: two runs. Subjects were also instructed that they would give each of the ratings three times during each run, and were to place a 1 at their first rating, a 2 at their second rating, and a 3 at their final rating, as well as give an overall rating (T).

Figure A-6 was also provided, and served as a reference for rating Workload. This "Modified Cooper-Harper" system was adopted from earlier work by Sheridan and Simpson. (See Ref. 18)

VII. Experimental Design

As mentioned in the instructions of Figure A-4, there were two different ground tracks used. Each subject flew both ground tracks during each session. Two different ground tracks were employed in order to minimize the effects of transferring prior knowledge from one run to the next, "learning" the scenario, and consciously or subconsciously anticipating tasks.

Each ground track was flown in two versions. One version was highly loaded with a number of tasks to perform. Most of these tasks were similar to following the instruction, "Climb and Maintain 4000". Such tasks exercise short-term memory because, in executing them, the pilot must constantly remind himself to follow the new parameter. The second version exercised long-term memory by instructing subjects to take some action at some time in the future.

Ground tracks and versions were counterbalanced between and within subjects. Each day's data runs included one run of each ground track and one run of each version (long-term memory and short-term memory).

"Navigational Charts" and Note Pads were provided to enable pilots to record instructions (as in real flight). The Navigational Charts contained Navigational Aid positions, courses, bearings, point identifiers, and distances to and from various points.

Figure A-7 shows such a "Navigational Chart" for the alpha ground track. Subjects began heading 360 degrees at 5000 feet, five nautical miles (nm) due south of VOR #1. They then proceeded to Point A (VOR #1: 021/15.0), VOR #2, Point B (#2: 228/10.0), Point C (#1: 144/5.0), and then headed 045 degrees until intercepting the Localizer for an ILS to Runway 36 (ILS 4).

Figure A-8 shows the nominal alpha ground track flown in its skill - or task-loaded version. Please note how ARTCC directed headings result in significant ground track deviations from the direct course. Figure A-9 pictures the nominal alpha ground track in its memory (long-term memory) version.

Figures A-10, -11, -12 are the corresponding examples for the beta ground track. Referring to Figure A-11, subjects began on a heading of 045 degrees at 5000 feet, Southwest of VOR#2. Then, they proceeded to VOR #2, Point D (#2: 312/22.8), VOR #1, Point E (#1: 156/6.7), and then headed 045 degrees until intercepting the localizer for Runway 36 (ILS 4). Figure A-12 clearly shows the 360 degree turn which is directed at VOR #1 for this version.

The differences between the skill- or task-loaded scenarios (short-term memory) and mentally- or memory-loaded scenarios (long-term memory) is best illustrated by picturing the time history of altitude, heading, and airspeed for each.

Figures A-13 and A-14 illustrate the airspeeds which subjects were commanded to maintain for each version of the alpha groundtrack. Compare task-loaded Figure A-13 with memory-loaded A-14.

Similarly, Figure A-15 can be compared to Figure A-16 for Magnetic Headings. Finally, Figure A-17 can be compared with Figure A-18 for commanded Altitudes.

Every effort was made to make the alpha and beta ground tracks as similar as possible while making the task and memory versions as different as possible. Thus, total Mental Workload Units and Total Physical Workload Units were calculated and plotted for each ground-track/memory-version combination.

The technique used to calculate these "Workload Units" can best be explained with two examples. For a task such as, "Climb 1000 feet", it was assumed that the pilot would climb at approximately 1000 feet per minute. The pilot must respond to the instruction, initiate the climb, monitor his progress in the climb, and execute a level off. For a 1000 foot climb, this entire process was estimated to last 90 seconds. Workload Units were calculated for 30 second intervals, so this task required 1 Workload Unit (WU) for three 30 second intervals, or three Physical WU's. However, in the process of performing this task, the pilot had to constantly update his short-term memory with this immediate goal: climb 1000 feet. Thus, the task also was credited with three memory or mental WU's, and labeled a short-term memory task.

For an example of a long term memory task, assume that ARTCC directs, "Report at Point D". The pilot must respond, usually make some note of the request, keep it in mind until he gets to Point D, and then report arriving at Point D. This also requires both task and memory work. It was assumed that the initial response and copying of the request would be handled in one, 30 second task unit. The same applied to the call to ARTCC at Point D. So this task generated one 30 second task unit at the time of the request, and one unit at the time of fulfilling the request. When receiving the request, the pilot stores it in his memory and hopefully retains it until arriving at Point D. Thus, it required one 30 second mental WU for each 30 second period from the time of the request until arriving at Point D. It also counts as one long-term memory task.

A Time/Workload history was done for each task the pilots were expected to perform for each ground-track. These workloads were then combined for each ground-track/version and plotted against an approximate time-line. Figure 19 is an example of one of these workload plots. Standing alone, these charts are not very enlightening, but they were useful for plotting workload data.

For instance, Figure A-20 shows the Accumulative Number of Physical WU's as a function of time for each type of run. This graph suggests that the physical workload is higher for the skill- or short-term memory versions than the long-term memory versions. Furthermore, it looks like the rate of

physical workload for the alpha and beta ground-tracks are similar within each version.

Figure A-21 is a plot of the Accumulative Mental (Memory) WU's versus time. Again, it appears that within each version, alpha and beta scenarios are similar, and that the overall workload for the skill version is different from that for the memory versions.

Figure A-22 shows the Accumulated number of memory tasks as a function of time. Here, the short-term memory tasks of the skill-or task scenarios balance out the additional long-term memory tasks of the memory versions. Thus, although the physical and mental workloads vary in some details across versions, the total number of mental tasks are roughly equivalent for each.

Figure A-23 breaks out the long-term memory tasks and shows that the long-term memory versions have roughly twice the number of long-term tasks as the short-term memory versions. Notice, also, the good balance between the alpha and beta ground tracks for each version. Comparing Figures A-22 and A-23, one can see that the skill versions must have a higher number of short-term memory tasks than the long-term memory versions.

VIII. Training and Experimental Procedure

After the subjects read the instructions (described in Section VI) and had all their questions answered, they then spent 20 to 30 minutes flying the simulator. This practice consisted of changing headings, altitudes, airspeeds, intercepting courses, and several ILS approaches.

When they felt ready, the subjects were given the Navigational Charts to study (Section VII) and the Charts were explained to them. The data runs then began with the Computer storing x, y, z positions every 10 seconds, and Control Box inputs as they occurred. The runs were frozen at roughly 8 to 10 minutes and 18 to 20 minutes of elapsed time. These two freezes and run termination were used to take the subject's Subjective Ratings.

IX. Results and Comments

The Subjects' Subjective Rating data is summarized in Figure A-24. Each rated category's mean rating and rating standard deviation are given for both alpha and beta ground-tracks, and for the arithmetic combination of alpha and beta.

Student t-tests and F-tests were performed on the data with the following results. For both long- and short-term memory versions, there were no significant difference between alpha or beta ground tracks at the 95 percent confidence level for any of the five categories. This implied that the effort to make the workload levels similar for the two ground tracks was successful from the standpoint of pilot perceptions. For each type of run

(for example, alpha/long-term memory), there was no significant difference between segments 1, 2, or 3 at the 90 percent confidence level. This implied a low likelihood of "point loading" occurring. That is, workload was fairly constant over time.

Student t-tests were performed on the mean subjective ratings to determine if there was a significant difference between the skill and mental versions for each category. There was a statistically significant difference at the 90 percent confidence level for Complexity and Stress. The difference was significant at the 95 percent level for Activity-level, Difficulty, and Workload.

The weaker confidence levels for the Complexity and Stress ratings can possibly be explained. All runs were performed manually, that is, with the autopilot off. Thus, the "complexity" changed little. The relative weakness in the Stress rating may be due to the relatively low workload level. Future experiments, run at greater workload levels, may show greater sensitivity for this rating category.

The Skill or Short-term Memory version was consistently rated higher (harder, more difficult) than the Mental or Long-term Memory version. This was a bit surprising since the average total (physical and mental) workload for the long-term memory version was greater than that for the task version. (218.5 WU vs. 187 WU: 116.8 percent)

Since other tests gave good confidence in the validity of this "workload unit" technique, several possible explanations come to mind. The 17 percent difference in workload units may not be significant at these workload levels. (One should keep in mind that the mean workload ratings were only in the three to five range on a ten-point scale.) Second, because subjects were "busier", doing a greater number of relatively simple tasks, this may have translated into a perception of greater workload.

Figure A-25 shows the Root-Mean-Squared (RMS) Altitude Deviations and the Mean and Standard Deviation of the Absolute Altitude Deviations. (Altitude Deviations were not measured during climbs and descents.) This information is given for each subject and across all subjects. It is also broken down, giving values for alpha and beta scenarios, and combined alpha-beta scores for the short-term memory and long-term memory versions.

Student t-test analyses of these errors for short-term memory versus long-term memory indicates a significant difference between these versions. Mean Absolute Altitude Errors are significant at an 80 percent confidence level and RMS Altitude Errors are significant at a 70 percent confidence level.

The relative weakness in differentiating the two versions may be due to the fact that there was no "baseline" version. Both versions were designed to be difficult, but difficult in different ways. The data only produced small differences between two fairly well-matched versions. Furthermore, both versions were rated only moderately difficult. If subjects are worked harder in future tests, more meaningful distinctions may appear.

Referring to Figure A-25, both the Mean Absolute Altitude Error and the RMS Altitude Error were greater for the long-term memory case than the short-term memory case. This is somewhat surprising since reference to Figures A-17 and A-18 clearly show that the short-term memory case had a much more difficult Altitude profile.

One possible explanation is that subjects became bored during the long-term memory scenario. I reject this hypothesis for three reasons. (1) No individual run lasted more than 30 minutes, and runs were broken by several "freezes" for subjective ratings. (2) Subjects knew that their performance was being measured, increasing interest. (3) The long-term memory version had few "quiet" periods longer than several minutes. Therefore, boredom was unlikely.

Two other, more promising, explanations relate to interest or attention. In the short-term memory or skill version, subjects were repeatedly asked to change airspeed, altitude, and heading. Thus, they probably channelled more effort and attention to these tasks, resulting in smaller deviations. This would also help explain the slightly higher subjective ratings for this version.

Alternatively, another type of prioritizing may have occurred. Given a lower task workload, the subjects may have shifted the task of aircraft control to a lower priority. This would produce a certain level of complacency about altitude, while subjects paid additional attention to memory items.

Mean Absolute Altitude Errors and RMS Altitude Errors were compared with the Subjective Ratings for each of the five Subjective Categories. For all cases, the magnitude of Altitude Error was inversely proportional to the Subjective Rating. That is, task loading resulted in lower Altitude Errors than mental loading, but higher Subjective Ratings.

Figure A-26 gives data on Long-term Memory Errors. (An example of a long-term memory task was given in Section VII). However, this chart further differentiates among long-term memory tasks. Here, these events were divided into "Positional" and "Non-Positional" Memory Tasks. A "Positional" task pertains to some performance required of the aircraft. For example, "Descend to 3000 at Point D." A "Non-Positional" task refers to something required of the pilot. For example, "Report at Point D".

Although it's difficult to generalize because of the small total number of tasks, the percentage of forgotten "Positional" tasks was similar for all versions/ground tracks, and the percentage of forgotten "Non-Positional" tasks was also similar for all versions/ground tracks. The interesting part of this data, however, lies in the fact that, on average, only 12.5 percent of "Positional" tasks were missed, while 40.6 percent of "Non-Positional" tasks were missed.

Professional Pilots are constantly reminded that no matter what happens, maintaining aircraft control should be their top priority. Therefore, this "Positional" information is given first priority. ARTCC requirements for informaton, etc., may be given second, or even third priority. This lower

priority for "Non-Positional" tasks may explain the poorer performance for these types of memory tasks.

X. Findings and Conclusions

1. Alpha and beta ground-tracks were roughly equivalent in perceived workload.
2. During each run, the perceived workload did not vary significantly with time.
3. At a moderate workload level, subjects consistently ranked the task-loaded version more difficult than a memory-loaded version, even though both were designed to be equally demanding.
4. At a moderate workload level, higher subjective workload ratings correlated with lower altitude deviations, possibly due to greater subject interest or attention.
5. Higher Long-term memory workload appears to interfere with, or lower the priority of short-term memory items.
6. Objective measurements (Altitude Error) differentiated between long-term and short-term memory scenarios at a 70 to 80 percent Confidence Level.
7. Pilots systematically weighted information about the physical positioning of their aircraft in space more heavily than other items.
8. Subjects can be worked much harder in future tests.

XI. Follow-up Studies

The next phase of this investigation will build upon these results to further differentiate between task or short-term memory workload, and long-term memory workload.

In an attempt to widen the differences between task workload and memory workload, the following scenarios will be tested:

<u>Workload Type</u>	<u>Aircraft Control</u>
Baseline Scenario	Manual
Task (Short-term Memory)	Manual
Memory (Long-term Memory)	Autopilot
Overload	Manual

The Baseline scenario will be a low workload scenario.

The Task scenario will be similar to the Baseline scenario, but involve many additional tasks: Heading changes, Altitude changes, Airspeed changes. The simulator will be manually flown and long-term memory items will be kept to a minimum.

The Memory scenario will allow the subjects to use the autopilot, freeing them to remember, monitor, and plan. Tasks will be kept to a minimum, but subjects will be repeatedly told to remember certain things for various lengths of time and then perform the directed tasks.

Finally, an "Overload" scenario will attempt to saturate the subjects. Subjects will be forced to fly manually while performing a large number of tasks and told to remember and do a variety of things.

Civilian pilots with less flight experience than the present subject group will be added.

Altitude Deviations, Subjective Ratings, and the percentage of memory items which are missed or not executed properly will be noted.

We postulate the following results:

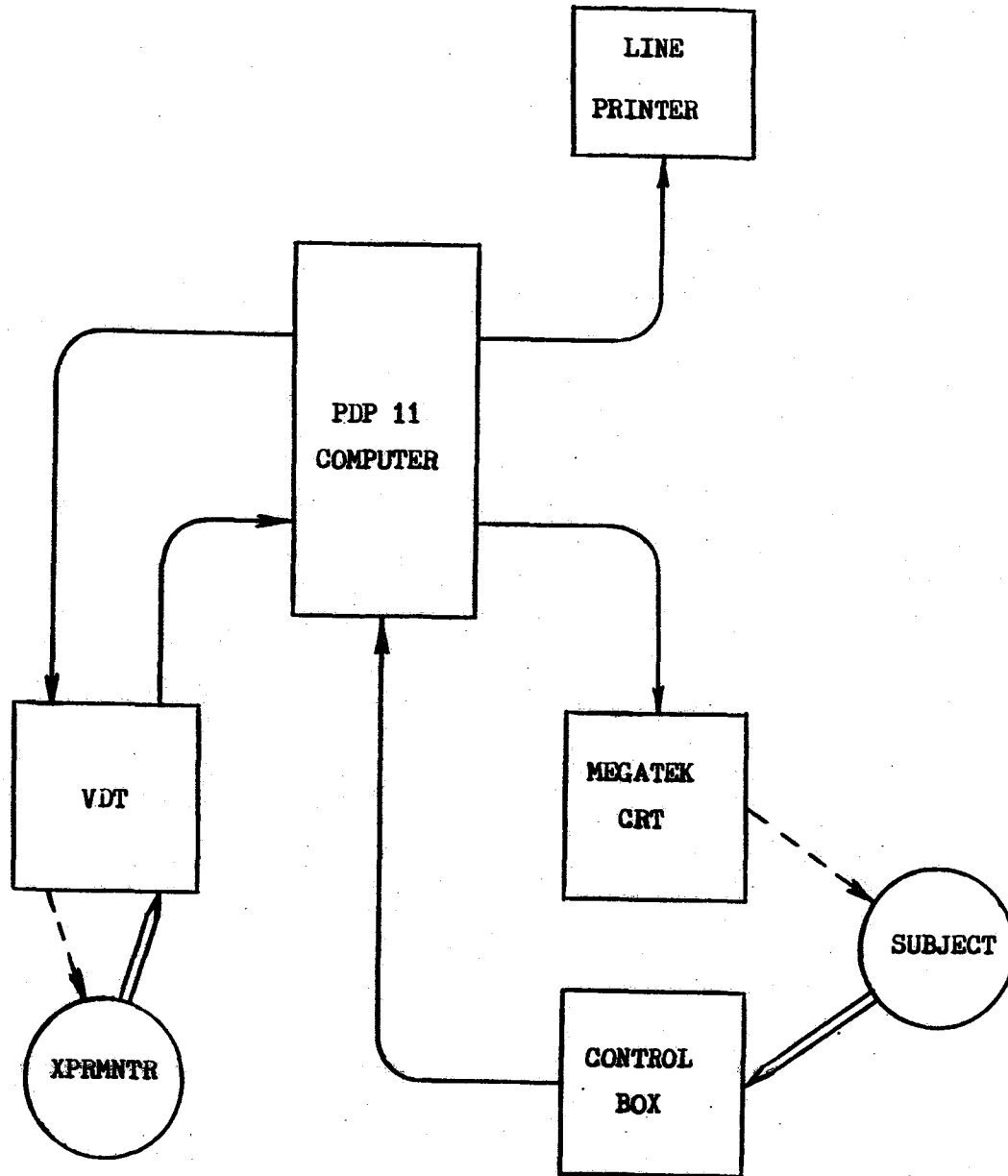
1. Workload ratings will directly relate to prior flight experience.
2. Subjects of all experience levels will do equally well and give similar ratings for the Memory Scenario.
3. There will be a direct relationship between performance, ratings, and experience for the Task and Overload Scenarios.
4. If given enough memory items, workload ratings will be as high in the Memory/Autopilot Scenarios as the Task/Manual Scenarios.
5. As workload ratings approach the high end of the scale, memory errors, or altitude deviations, or both will increase.
6. Subjects will tend to allow the number of memory errors to increase rather than aircraft control to decrease.

XII. References

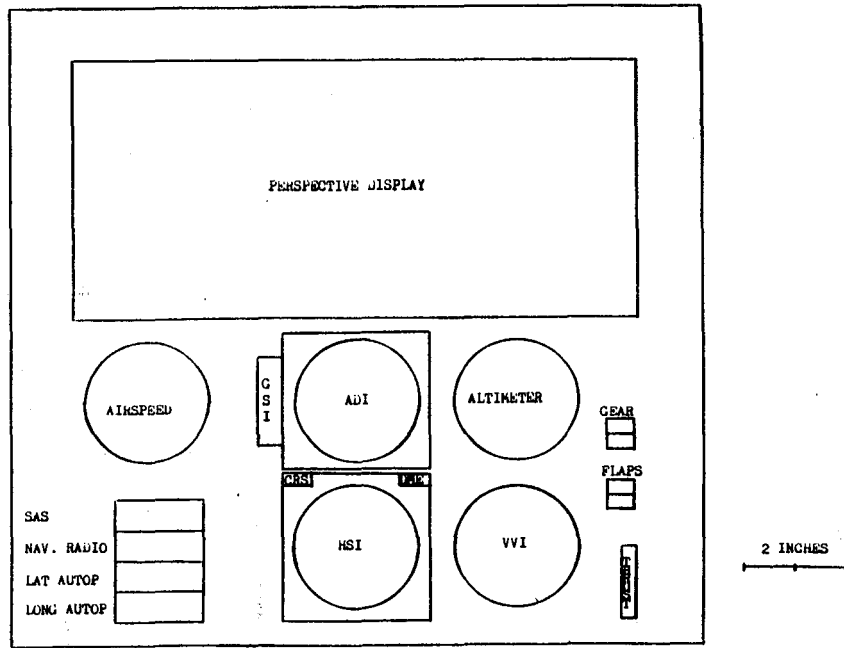
1. Heffley, R. K. and Jewell, W. F.; Aircraft Handling Qualities Data, NASA CR-2144, Dec. (1972).
2. McRuer, D., Ashkenas, I. and Graham, D.; Aircraft Dynamics and Automatic Control, Princeton (1973).
3. Sheridan, T. B.; Simpson, R. W.: Toward the Definition and Measurement of the Mental Workload of Transport Pilots, DOT-OS-70055, Jan. (1979).
4. Anon.; Air Traffic Control, U.S. Air Traffic Service, Federal Highway Administration, 7110.65c, Jan. (1982).

5. Anon.; A300FF Minimum Flight Crew Determination, Airbus Industries, AI/VF-AI/TE4-023/82, Jan. (1982).
6. Soulsby, E.P.; Human Operator Workload; A Survey of Concepts and Assessment Techniques, Technical Report of EECS TR-82-4, School of Engineering, University of Connecticut (1982).
7. Johannsen, G. and Rouse, B.; Studies of Planning Behavior of Aircraft Pilots in Normal, Abnormal, and Emergency Situations.
8. Sheridan, T. B.; Supervisory Control: Problems, Theory and Experiment for Application to Human-Computer Interaction in Undersea Remote Systems, Man-Machine Systems Laboratory, Technical Report, MIT, March (1982).
9. Sheridan, T. B.; Mental Workload in Decision and Control, Proc. of IEEE (1979) pp. 997-982.
10. Bird, K. L.; Subjective Rating Scale as a Workload Assessment Technique, Proc. of 17th Annual Conf. on Manual Control (1981).
11. Hart, S. G.; Childress, M.E. and Bortolussi, M. R.; Defining the Subjective Experience of Workload, Proc. of the Human Factor Society 25th Annual Meeting (1981) pp. 527-531.
12. Hart, S. G.; Effect of VFR Aircraft on Approach Traffic with and without Cockpit Displays of Traffic Information, Proc. of 18th Annual Conf. on Manual Control (1982).
13. Hauser, J. R., Childress, M. E. and Hart, S. G.; Rating Consistency and Component Saliency in Subjective Workload Estimation, Proc. of 18th Annual Conf. on Manual Control (1982).
14. Hart, S. G., Childress, M. E. and Hauser, J. R.; Individual Definitions of the term 'Workload', Proc. of 1982 Psychology in the DOD Symposium (1982).
15. Childress, M. E., Hart, S. G. and Bortolussi, M. R.; The Reliability and Validity of Flight Task Workload Ratings, Proc. of the Human Factor Society 26th Annual Meeting (1982) pp. 319-323.
16. Rasmussen, J.; Outlines of a Hybrid Model of the Process Plant Operator, in Sheridan, T. and Johannsen, G. Eds., Monitoring Behavior and Supervisory Control, Plenum, N.Y. (1976).
17. Sheridan, T. B.; Proposal to Study the Relationship between Aircraft Control Automation, Mental Workload and Pilot Error in a Laboratory Simulator.
18. Tanaka, K.; Buharali, A.; and Sheridan, T. B.; Mental Workload in Supervisory Control of Automated Aircraft, Proc. of 19th Annual Conf. on Manual Control (1983).

Electrical →
Visual - - - →
Manual ==>

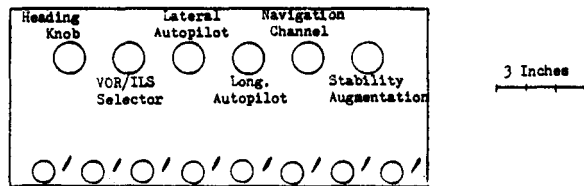
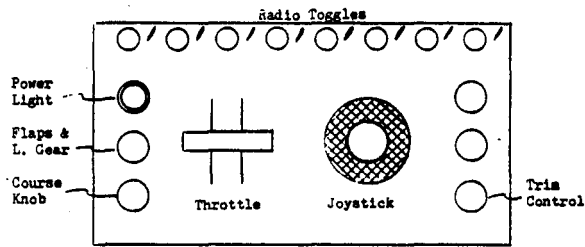
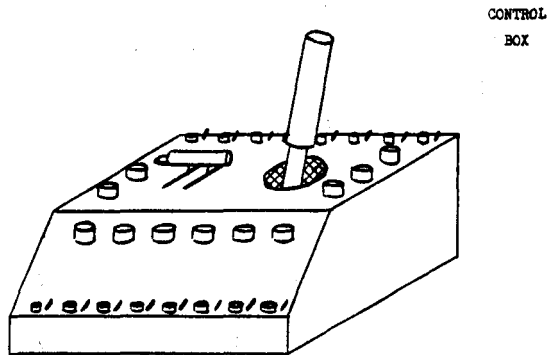


EXPERIMENTAL SET-UP



MEGATEK CRT DISPLAY

A-2



A-3

INSTRUCTIONS

The experiment you are participating in will provide information on pilot workload. The experiment consists of four "flights": two now and two another day. On each day you will fly two different ground tracks, terminating in an ILS approach. For each flight, the number of normal and mental tasks will be varied.

Your task is to fly as precisely as possible while following instructions to the best of your ability.

Ignore any ATC statements or instructions which appear on the display. All instructions and ATC statements will be handled verbally. However, when contacting a new "Controller", toggle off (away) the old radio and toggle on (toward) the new channel. Since all flights will be performed manually, you can ignore the two autopilot controls. In addition, the Trim and CWS switches are best left as set.

You will use 3 Navigation aids: VOR 1, VOR 2, and ILS 4. ILS 4 provides an ILS for Runway 36. Please note that the signal is only received within 10 miles of the runway. So, when on a dogleg to the ILS, hold heading until the Course Deviation Bar comes off the stops or the Glide Slope Indicator shows movement.

The "nominal" airspeed for these runs is 200 kts. Final approach will be flown at 150 kts, with Gear and Flaps down. Usually, a throttle position near center will maintain a stable airspeed.

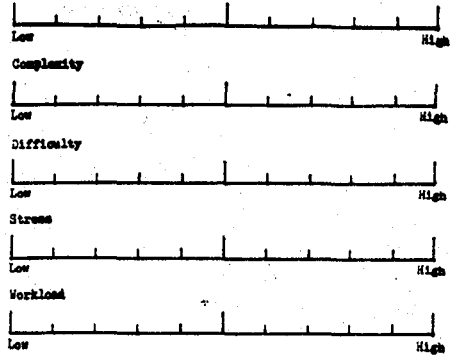
You can expect the following level flight attitudes:

- 200 kts: Clean -2 deg
- Flaps -5 deg
- Gear & Flaps -2 deg
- 150 kts: Clean 0 deg
- Flaps +2 deg
- Gear & Flaps +6 deg

During and after each run, you will be asked to make several subjective ratings. Thank you for your time and effort.

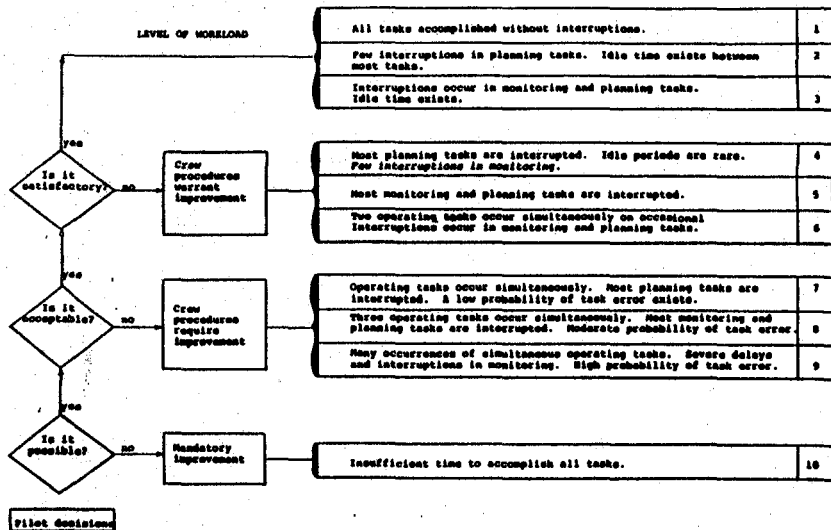
A-4

Activity Level (Busy-ness)



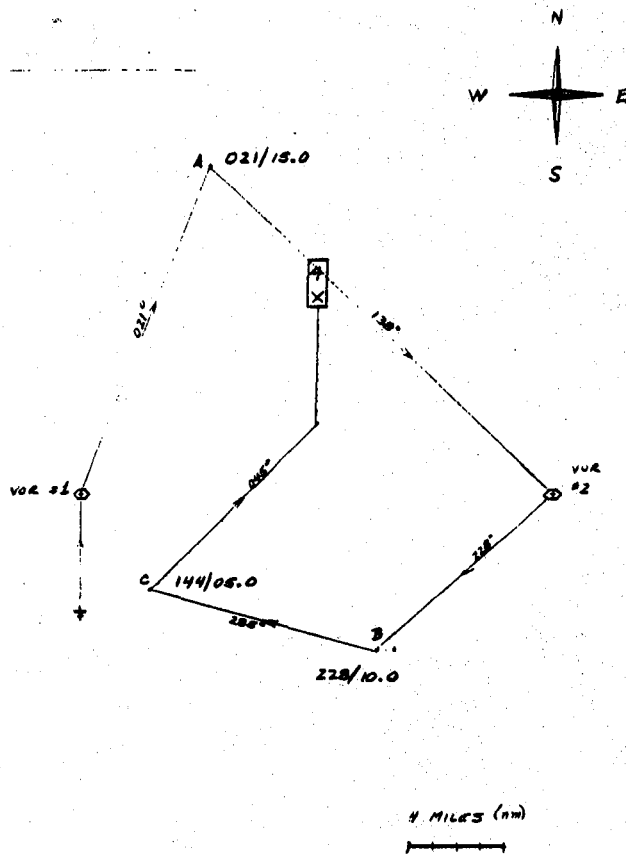
Subjective Rating Sheet

A-5



WORKLOAD RATING SYSTEM

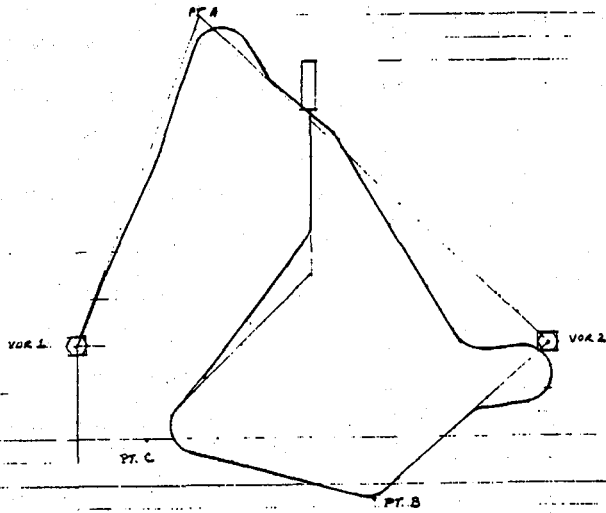
A-6



NAVIGATIONAL CHART: ALPHA SCENARIO

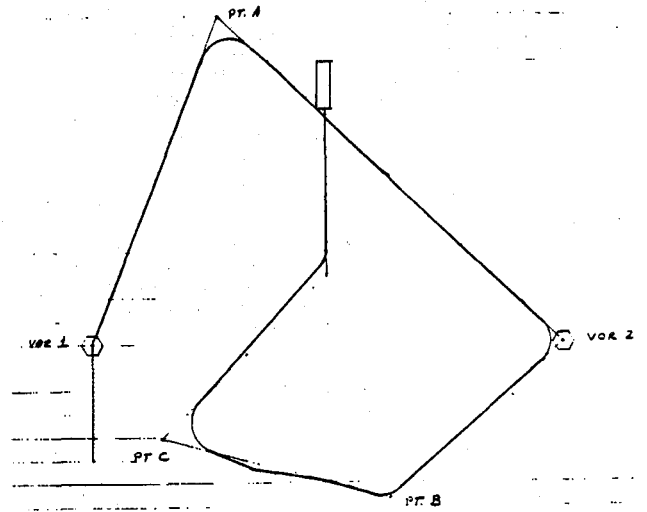
A-7

NOMINAL GROUND TRACK: ALPHA SCENARIO
SHORT-TERM MEMORY



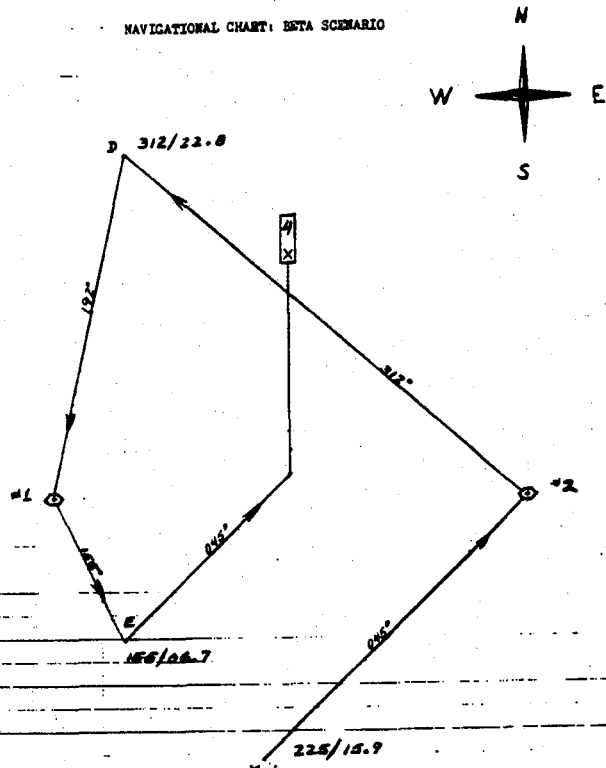
A-8

NOMINAL GROUND TRACK: ALPHA SCENARIO
LONG-TERM MEMORY



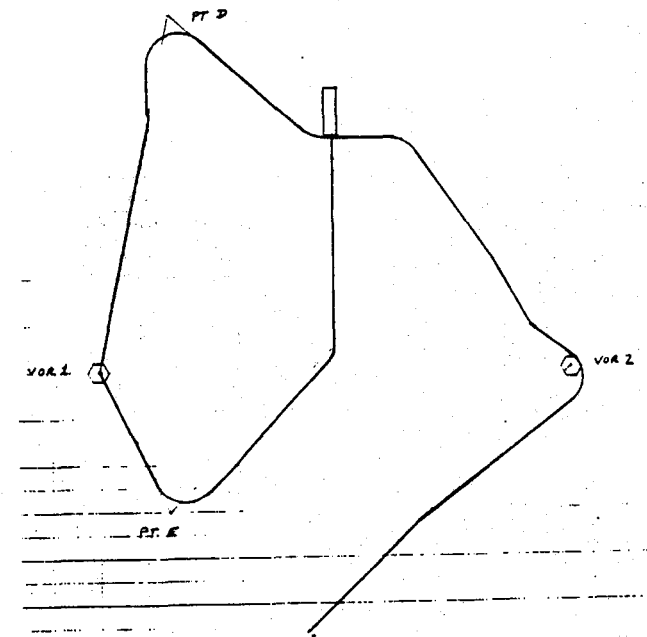
A-9

NAVIGATIONAL CHART: BETA SCENARIO



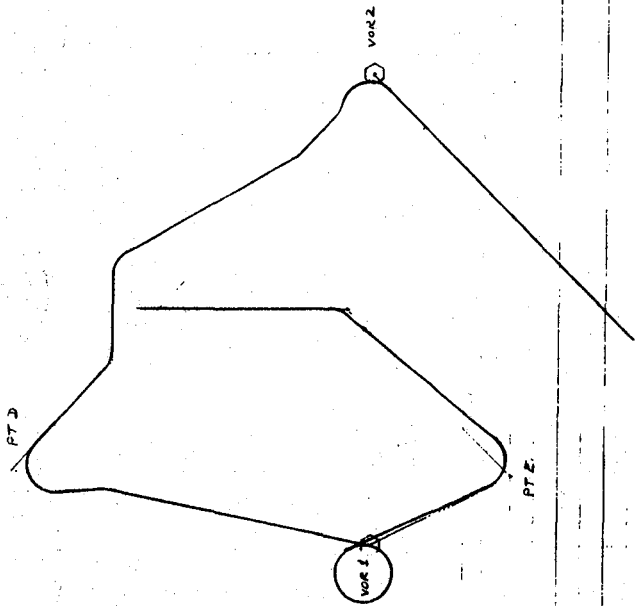
A-10

NOMINAL GROUND TRACK: BETA SCENARIO
SHORT-TERM MEMORY



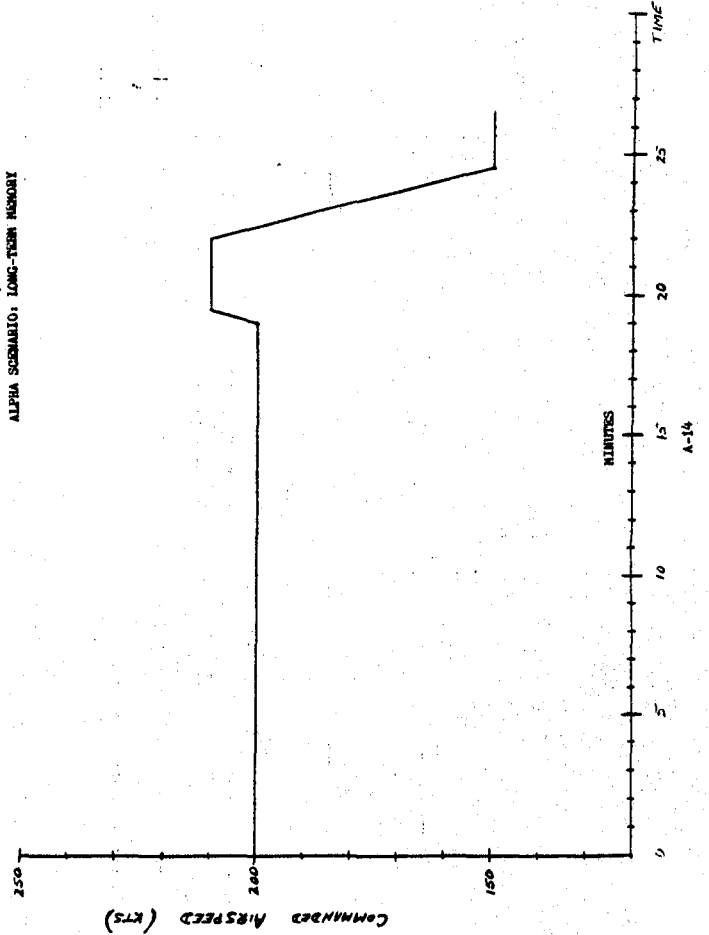
A-11

NOMINAL GROUND TRACK; BETA SCENARIO
LONG-TERM MEMORY



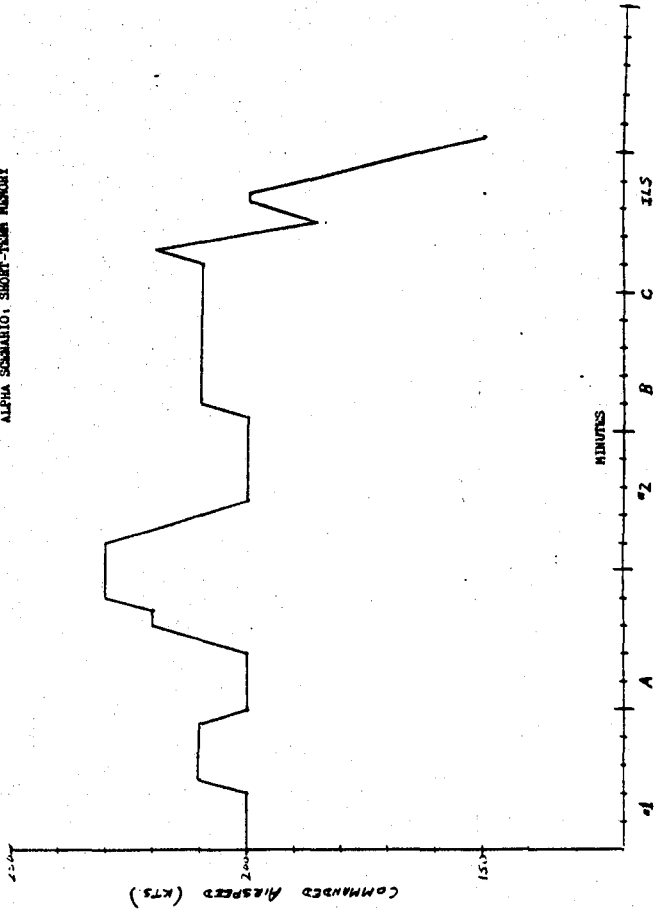
A-12

ALPHA SCENARIO; LONG-TERM MEMORY



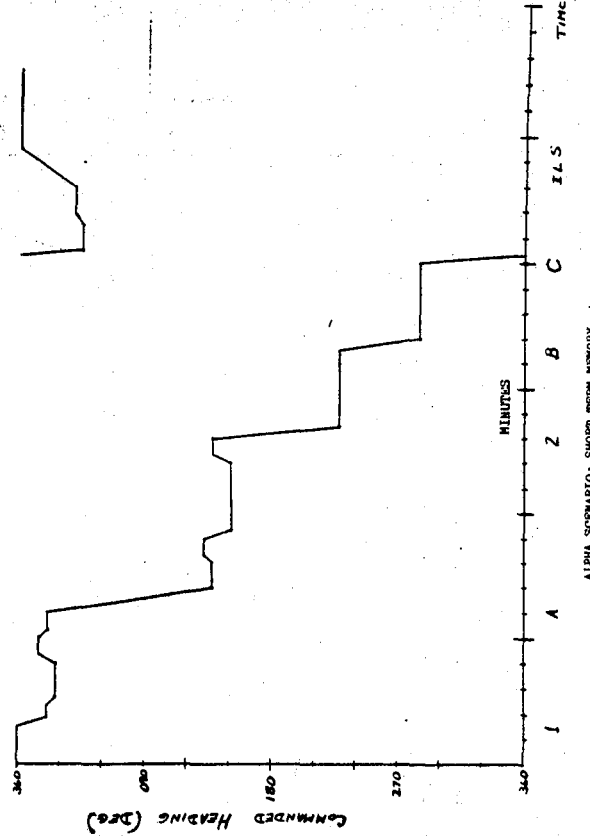
A-14

ALPHA SCENARIO; SHORT-TERM MEMORY



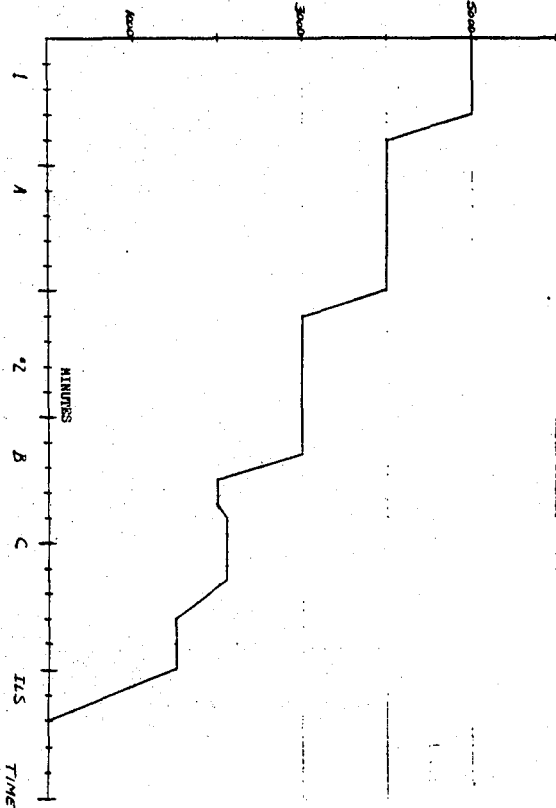
A-13

ALPHA SCENARIO; SHORT-TERM MEMORY



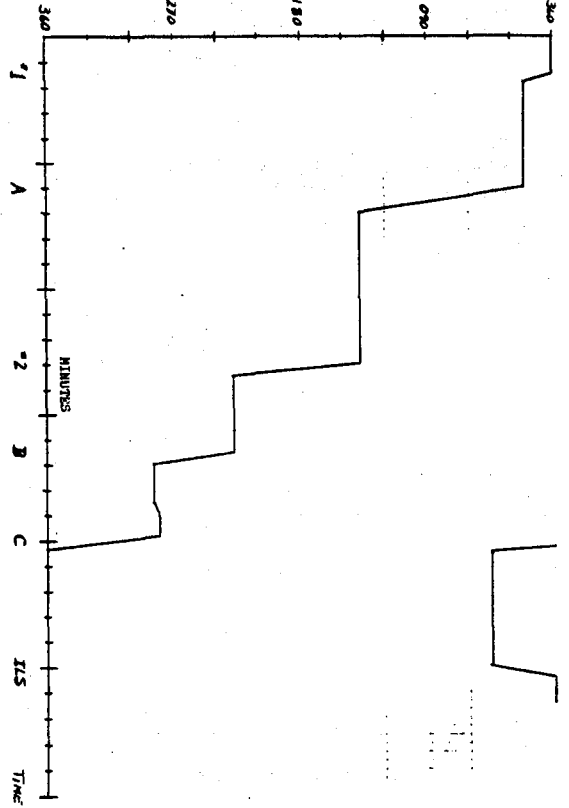
A-15

COMMANDED ALTITUDE (FT.)



A-18

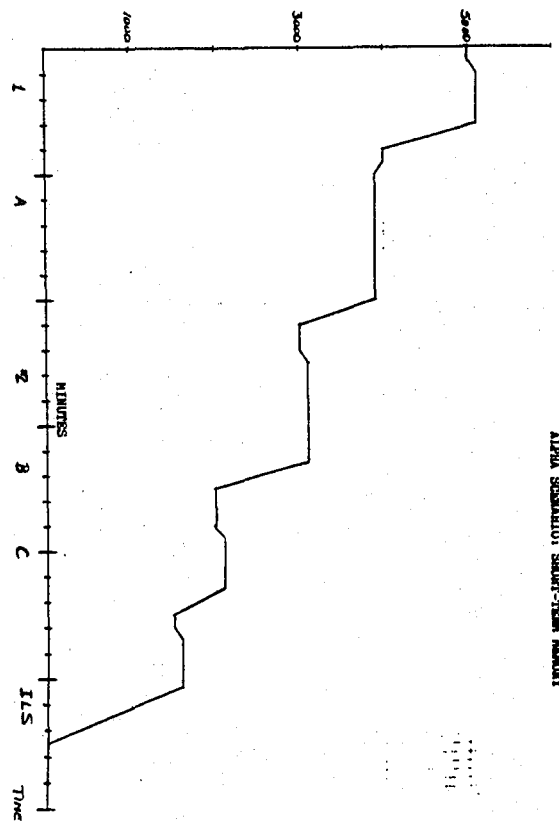
COMMANDED HEADING (DEG.)



A-16

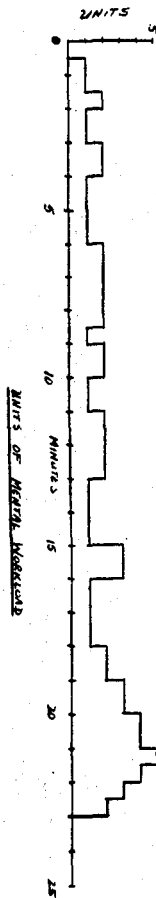
ALPHA SCENARIO: LONG-TERM MEMORY

COMMANDED ALTITUDE (FT.)

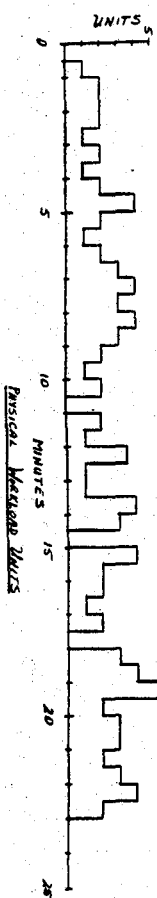


ALPHA SCENARIO: SHORT-TERM MEMORY

A-17



ALPHA SCENARIO: SHORT-TERM MEMORY



PHYSICAL MEMORY UNITS

A-19

VOR 1

PT A

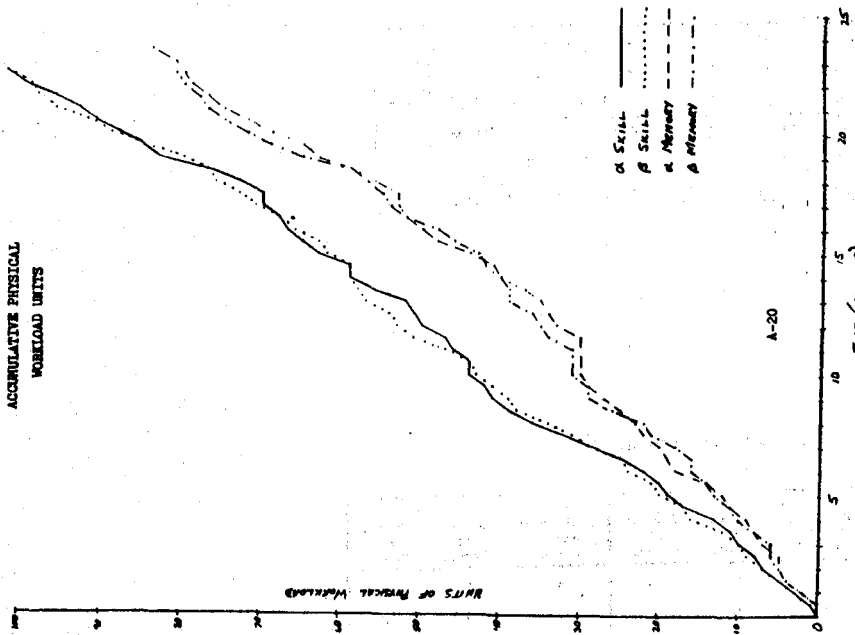
VOR 2

PT B

PT C

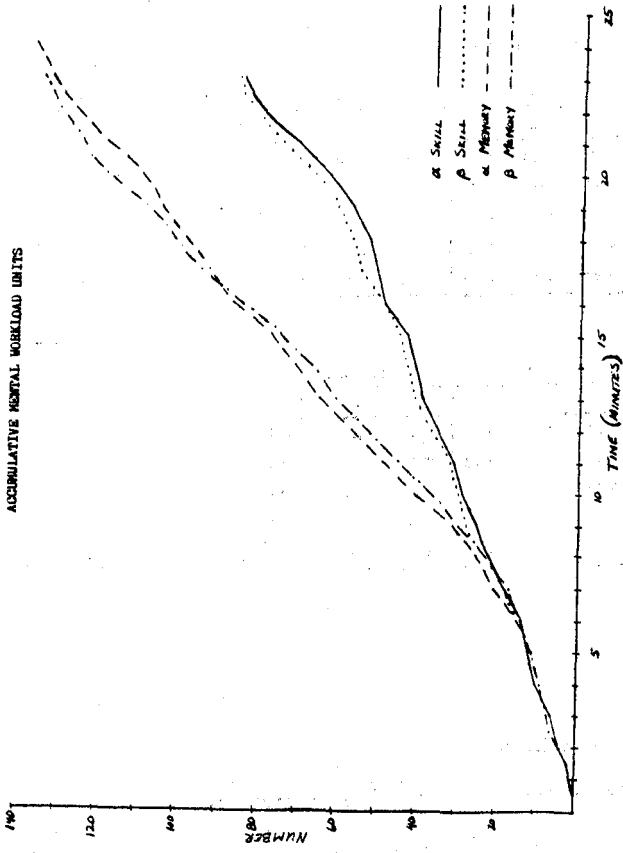
ILS

ACCUMULATIVE PHYSICAL
WORKLOAD UNITS



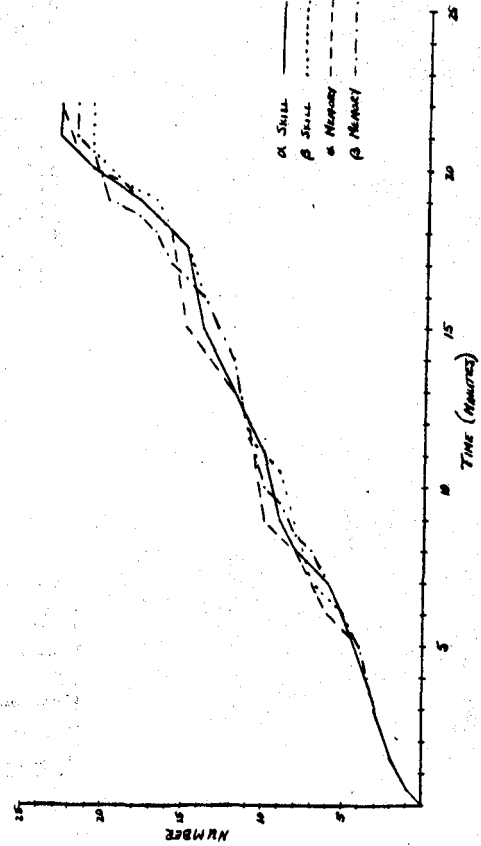
A-20

ACCUMULATIVE MENTAL WORKLOAD UNITS



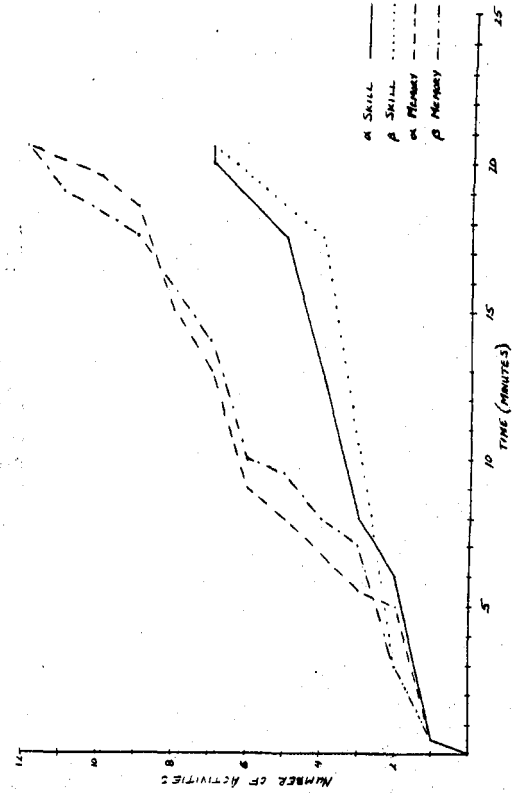
A-21

TOTAL NUMBER OF MEMORY TASKS



A-22

LONG-TERM MEMORY TASKS



A-23

OVER-ALL SUBJECTIVE RATINGS

	mean			std. dev.		
	a	B	a&B	a	B	a&B
Activity-Level	4.93	4.35	4.64	1.42	1.60	1.54
Complexity	4.17	4.18	4.18	1.70	2.15	1.93
Difficulty	5.04	4.33	4.69	1.54	2.15	1.90
Stress	4.54	4.13	4.33	1.82	2.12	1.98
Workload	5.00	4.47	4.73	1.68	2.37	2.07

LONG-TERM MEMORY

A-24

	mean			std. dev.		
	a	B	a&B	a	B	a&B
Activity-Level	3.85	3.18	3.52	1.68	1.20	1.50
Complexity	3.49	2.78	3.13	2.03	1.39	1.78
Difficulty	3.98	3.14	3.56	1.79	1.54	1.72
Stress	3.53	3.11	3.32	1.80	1.82	1.82
Workload	3.79	3.18	3.49	2.19	1.66	1.97

LONG-TERM MEMORY

ALTITUDE ERRORS

		B	H	L	W	AVG.			
SHORT-TERM MEMORY	a	mean	59.0	77.3	95.71	69.4	74.36	mean	
		σ	36.11	71.4	124.78	59.32	79.97		σ
		rms	69.17	105.19	157.26	91.30	109.20		
B	mean	60.69	95.0	59.13	99.06	74.59	76.19	σ	
	σ	53.33	57.6	60.68	105.80	71.05	106.53		rms
	rms	80.79	111.08	84.72	144.94	103.01			

A-25

		B	H	L	W	AVG.			
LONG-TERM MEMORY	a	mean	81.77	163.06	72.41	121.60	103.06	mean	
		σ	49.70	134.80	57.48	94.39	90.38		σ
		rms	95.69	211.55	92.45	153.93	137.08		
B	mean	93.2	122.3	74.38	53.41	84.14	81.08	σ	
	σ	70.86	80.60	61.63	45.69	69.48	123.82		rms
	rms	117.08	146.48	96.59	70.28	109.12			

LONG-TERM MEMORY ERRORS

Scenario/Version	Positional Memory Events	Number of Errors	Percent of Errors	Non-Positional Memory Events	Number of Errors	Percent of Errors
Short-term a	0	0	0	4	2	50
Short-term B	0	0	0	4	2	50
Long-term a	8	1	12.5	12	5	41.7
Long-term B	8	1	12.5	12	4	33.3

A-26