FAA-75-19

REPORT NO. FAA-RD-75-170

HUMAN FACTORS EXPERIMENTS FOR DATA LINK
Final Report

Edwin H. Hilborn

NOVEMBER 1975
FINAL REPORT

DOCUMENT IS AVAILABLE TO THE PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161

Prepared for

U.S. DEPARTMENT OF TRANSPORTATION
FEDERAL AVIATION ADMINISTRATION
Systems Research and Development Service
Washington DC 20591

TECHNICAL REPORT STANDARD TITLE PAGE

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FAA-RD-75-170		
4. Title and Subtitle HUMAN FACTORS EXPERIMENTS FOR DATA LINK Final Report		5. Report Date November 1975 6. Performing Organization Code
7. Author(s) Edwin H. Hilborn		8. Performing Organization Report No. DOT-TSC-FAA-75-19
9. Performing Organization Name and Address U.S. Department of Transportation Transportation Systems Center Kendall Square Cambridge MA 02142		10. Work Unit No. FA513/R6116 11. Contract or Grant No.
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Aviation Administration Systems Research and Development Service Washington DC 20591		13. Type of Report and Period Covered Final Report January 1972 - June 1975 14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

This report describes the results of a series of experiments to evaluate cockpit Input/Output devices for Data Link as Phase I of a larger project to explore all facets of the digital transmission of air traffic control information.

Following preliminary laboratory experiments to investigate optimum means of formatting and presenting information, prototype hardware was fabricated and evaluated on GAT-1 and GAT-2 simulators. Deficiencies in the prototypes were noted, and new improved hardware was fabricated for further evaluation on the GAT-2 and airline simulators (B-727 and DC-9).

Data Link as a concept was found to be generally acceptable, and there was widespread agreement that it would reduce crew workload. Airline pilots were concerned by the loss of information resulting from the selective address capabilities of Data Link, so that they did not hear messages intended for other aircraft. Pilots also questioned the acceptability of Data Link for ground and local control.

Visual displays were preferred for all but emergency messages; for these, the use of synthetic speech (Vosyn) is a likely candidate. In-cockpit printout of clearances, weather and ATIS information was favorably received, and the use of keyboard input for downlink message generation was found to be acceptable.

17. Key Words	18. Distribution Statement
Air Traffic Control, Data Link, Displays, Synthetic Speech, Message Coding, Simulators, Human Factors Engineering	DOCUMENT IS AVAILABLE TO THE PUBLIC THROUGH THE NATIONAL TECHNICAL INFORMATION SERVICE, SPRINGFIELD, VIRGINIA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	134	

Form DOT F 1700.7 (8-69)

PREFACE

This, the final report of Phase I of the Data Link project, describes the evaluation of cockpit I/O devices for the digital transmission of air traffic control information. Phases II and III of the project, which describe the evaluation of the controller-computer interface and the link characteristics, respectively, are covered in separate reports. The project was sponsored by the Federal Aviation Administration, Systems Research and Development Service.

During the 3-1/2 years of effort on Phase I, contributions were made by approximately 250 people in concept formation, equipment design and fabrication, logic design, computer programing, participation in experiments, data reduction, and program documentation. The acknowledgement of the precise contributions of this number of people is obviously impossible; without their combined efforts, progress in the program would have been equally impossible.

Interim Report No. 6 in this series was prepared by Mr. J. M. Diehl of ARINC Research Corp., and material from his report has been incorporated, paraphrased, or abstracted in the preparation of this final report.

The following Raytheon Service Company personnel assisted in the preparation of this report: Ronald Karr, Technical Editor; Conni Segerstedt and Vipapan Owtrakul, technical typists; and the Art Department, Eugene Adelizzi, Director.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (CONTINUED)

LIST OF TABLES

LIST OF TABLES (CONTINUED)

LIST OF TABLES (CONTINUED)

1. INTRODUCTION

The present overcrowding of radio communication links for air-ground-air Air Traffic Control (ATC), along with anticipated future increases in air traffic, led the Systems Research and Development Service of the Federal Aviation Administration to assign to the Department of Transportation/Transportation Systems Center (DOT/TSC) at Cambridge, Mass., some three years ago, the task of investigating the feasibility of digital transmission of ATC information.

The project, as developed at TSC, involved three concurrent phases: (I) the study of cockpit I/O (Input/Output) requirements, (II) the study of the role of air traffic controllers in such an automated system, and (III) study of the link characteristics. The present report is concerned only with Phase I of this project, the interaction between air crews and Data Link I/O devices.

The cockpit I/O portion of the project has already been documented thoroughly in six Interim Reports (listed as References 1 through 6 in the present report). Additionally, an extended summary was published in April 1974, covering the material in Interim Reports Nos. 1 through 4, as FAA-RD-72-82. (Reference 7). For completeness, the material in the referenced Extended Summary, with some slight changes for greater clarification, is included as Sections 2 and 3 of this report. Readers completely familiar with the contents of the Extended Summary may accordingly proceed directly from this introduction to Section 4.

The Human Factors aspects of the cockpit portion of the project are concerned with (1) how to display Data Link information, (2) the determination of what information requires display and what its format should be to make it most meaningful, and (3) downlink requirements. While flight simulators are required for the investigation of the "how" aspects, this has been supplemented by laboratory studies to obtain information on the "what" aspects. Study of this second aspect of the problem is inefficient in a flight simulator, since a realistic simulation can provide data points only every 60-90 seconds. In the laboratory, however, it is possible to accumulate data points every few seconds if proper procedures are employed. Investigation of the "what" aspects of Data Link was accordingly concentrated on laboratory experiments.

2. LABORATORY TESTS OF MESSAGE CODING AND FORMATS

To investigate schemes for message coding and formatting, TSC Human Factors laboratory personnel flashed slides onto a projection screen while requiring a viewer to make appropriate responses. The shortage of prime panel space on the flight deck of commercial airliners made it imperative that messages be as brief as possible while still remaining meaningful. Brevity may be accomplished in two ways; either by using short abbreviations, or by avoiding spaces between words or abbreviations.

In one experiment to evaluate these variables, slides were prepared containing 25 typical short ATC messages in four different formats: (1) short abbreviations without spaces, (2) short abbreviations with spaces, (3) longer abbreviations without spaces, and (4) longer abbreviations with spaces. Thus, a message such as "Resume Speed" could be presented as "RESS," "RES S," "RESSPD," or "RES SPD." The slide sequence was randomized and presented individually to 12 TSC engineers, all of whom had some knowledge of air traffic control terminology and had been given the opportunity to memorize the abbreviations in their various forms (No. 2 - Sect. 1)*. Subjects were told to depress a response button as soon as they recognized the meaning of a message. Depression of the response button blanked the screen to prevent further reference to the stimulus material, and the subjects were then asked to verbalize the meaning of the message. The same slides were later presented individually to eight FAA-NAFEC Test Pilots who participated in the GAT-1 simulator tests (No. 3 - Sect. 3).

Table 2-1 presents the mean response time in seconds and the total errors for the two subject groups for the several conditions. An error was recorded when either the subject failed to respond or failed to respond correctly.

*The (No. 2 - Sect. 1) and similar notations which appear elsewhere in this report refer respectively to the number and section of the Interim Report where additional details are provided.

TABLE 2-1. REACTION TIMES (IN SECONDS) AND ERRORS OF TWO SUBJECT GROUPS TO FOUR DIFFERENT MESSAGE FORMATS

Stimulus Material	TSC Engineers		Test Pilots	
	Time	Errors	Time	Errors
Short abbrev. No spaces	2.08	19	2.68	16
Short abbrev. With spaces	1.61	9	1.98	6
Long abbrev. No spaces	2.20	22	2.33	5
Long abbrev. With spaces	1.58	13	1.82	2

With both subject groups, reaction time was reduced by approximately one-half second when spaces were used, and the error rate was approximately one-half. The use of longer and supposedly more meaningful abbreviations, on the other hand, made little difference. Statistically, the differences resulting from the use of spaces were significant at the .999 and .95 levels for the engineers and test pilots, respectively; differences resulting from the use of long versus short abbreviations were non-significant.

Figure 2-1 represents the mean reaction times for the four message types as a function of message length; "message length" in this case being defined as the number of characters in the shortest version of a particular message. Here, it should be noted that with short abbreviations with spaces, reaction time remains relatively constant, regardless of message length. On the other hand, long messages, particularly those using long abbreviations without spaces, can be comprehended only slowly and with difficulty. The message appears as so much "alphabet soup." The data depicted here are for the TSC engineers; data for the test pilots were comparable.

Since all short or long abbreviations were not equally short or long, reaction times were plotted as a function of the number of "information units" in a message ("information units" in this context meaning an abbreviation for a word or group of digits) and are presented in Figure 2-2 for the TSC engineers. Again, it should

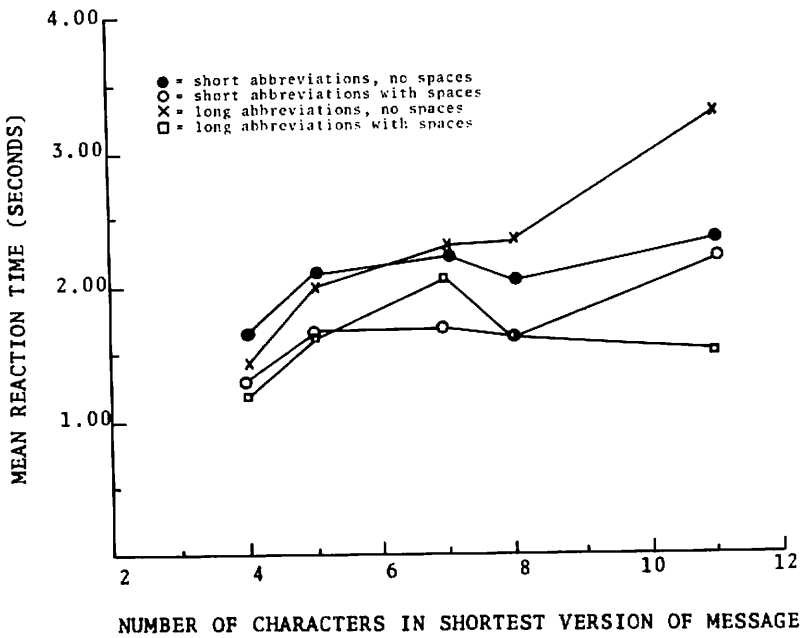

Figure 2-1. Mean Reaction Time as a Function of Message Length

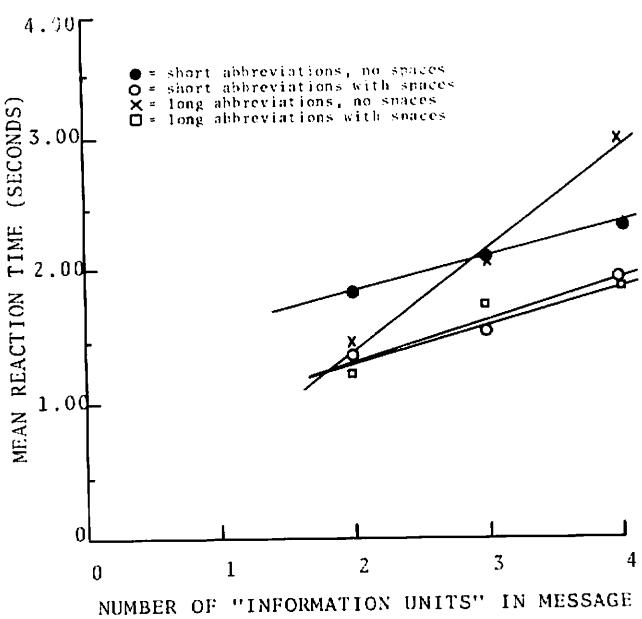

Figure 2-2. Mean Reaction Time as a Function of Number of Information Units in a Message

be noted that increases in message length produce only a modest increase in reaction time when spaces are used.

The relatively low error rate achieved by both subject groups in these experiments makes it impossible to plot smooth curves for error rate versus message length for the separate experimental conditions. Figure 2-3 depicts error rate as a function of message information units for the lumped experimental conditions. Since error rate increases sharply as the amount of information increases, messages should be as short as possible if correct interpretation is to be achieved with a high degree of regularity.

For a second and somewhat more ambitious series of laboratory experiments, 144 slides were prepared containing ATC commands and advisories of several types using 3 different type fonts and a variety of coding and formatting schemes. As the slides were presented in random sequence, experimental subjects were required to select the appropriate responses from among multiple choices according to the type of information being presented. For this they used the response box depicted in Figure 2-4. The four buttons in

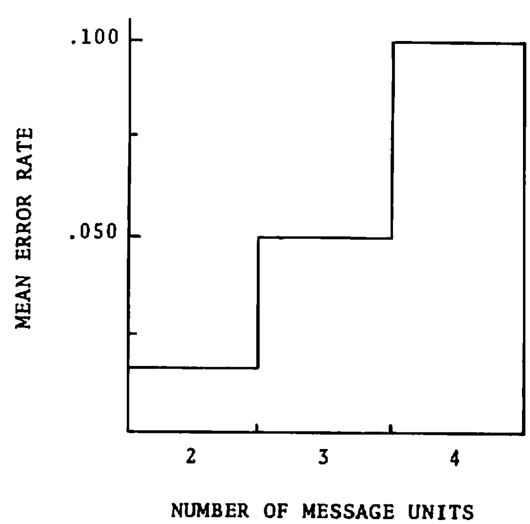

Figure 2-3. Mean Error Rate as a Function of Number of Message Units

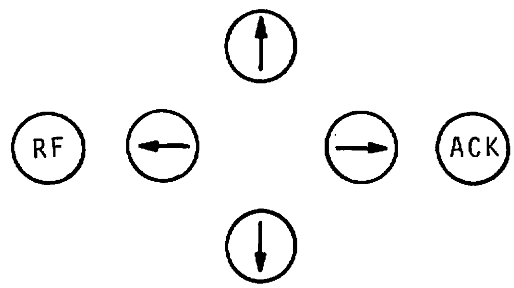

Figure 2-4. Response Box Layout

the center, arranged in a diamond-shapped pattern, represented the "controls" for "fly left," "fly up," "fly down," and "fly right." The button on the extreme left represented the "control" for radio frequency settings, and that on the extreme right acknowledgement of receipt of a message which did not fit any of the other categories.

The legends on the response buttons of Figure 2-4 are to aid the reader in understanding these functions. Blank buttons were used during the experimental runs, and the subjects were given a cardboard facsimile indicating every possible symbol or abbreviation that might be used with each button.

Each message was reproduced with 3 different type fonts, simulating a 16-segment array, a 5x7 dot matrix array, and stencil-type characters similar to those produced by a Charactron CRT.

During a first running of the experiment, data were obtained from 10 TSC engineers (No. 1 - Sect. 3). A second run was later made, again using 10 TSC engineers, with the procedural variation that when the subject responded, the information was automatically removed from the screen; the subject was then required to verbalize the meaning of the message (No. 2 - Sect. 2). In a third replication, the eight NAFEC test pilots were run using this same procedure (No. 3 - Sect. 2).

The results of the three experiments are summarized in Table 2-2. In general, the three type fonts were found to be equally readable. Three-line messages were easier to read than extended single-line messages. New information should preferably be presented at the top or left of a display. For an emergency situation, a single word or error should first be presented, such as "CLIMB" or "↑," followed later by the desired altitude. In the majority of cases, arrows and words were found to be equivalent, except in messages such as "HDGR160," where the "R" was difficult to detect. Here, the use of arrows drastically reduced reaction time. The general agreement of the results from engineers and test pilots validates the use of engineers for an experiment of this type.

TABLE 2-2. COMPARISON OF EXPERIMENTAL CONDITIONS (REACTION TIME IN SECONDS)

Variable	Typical Message	First Run		Second Run		Third Run	
		Reaction Time	Signif. of Dif.	Reaction Time	Signif. of Dif.	Reaction Time	Signif. of Dif.
Type Font Differences: Dot Matrix Segmented Stencil		1.24 1.31 1.21	n.s	1.33 1.38 1.27	n.s.	2.09 2.10 1.89	n.s.
Type Font Differences with Buried Arrows: Dot Matrix Segmented Stencil	A↓120	1.24 1.47 1.31	.001	1.56 1.55 1.41	.02	2.39 2.57 2.41	n.s.
Single Words Arrows	LEFT ←←	0.93 0.79	.02	0.96 0.84	.05	1.16 1.11	n.s.
Words + Numbers Arrows + Numbers	DOWN120 A↓120	1.06 1.00	n.s.	1.19 1.16	n.s.	2.05 2.19	n.s.
3-Line Messages using words 3-Line Messages using arrows	TURN RIGHT 180 vs TURN →→ 090	1.02 0.96	n.s.	1.15 1.07	n.s.	1.56 1.58	n.s.
Arrows Alone Arrows with Numbers	→→ →120	0.79 0.92	.02	0.84 1.11	.001	1.11 1.57	.001
Text with Numbers Text without Numbers	LEFT220 RIGHT	1.04 0.93	.05	1.17 0.96	.01	1.81 1.16	.001
Linear Messages 3-Line Messages	*HDG 070* ALT 190 SPD 210 HDG 160 *ALT 110* SPD 165	1.64 1.33	.001	1.75 1.58	.01	2.73 2.54	n.s.
New Info. on Left New Info. "Buried"	*HDG 070* ALT 190 SPD 210 HDG 260 *ALT 140* SPD 220	1.62 1.73	.05	1.76 1.81	n.s.	2.54 2.88	.001
New Info.on top in 3-line format New Info. "buried" 3-line format	*HDG 180* ALT 160 SPD 200 vs HDG 090 *ALT 080* SPD 165	1.29 1.40	.05	1.45 1.52	n.s.	2.42 2.76	.001
"Buried" Arrows "Buried" "L" or "R"	HDG→080 HDGR080	1.02 1.67	.001	1.12 1.59	.001	1.72 2.77	.001

n.s. = a nonsignificant difference

3. SIMULATOR STUDIES OF SHORT MESSAGE ATC VISUAL DISPLAYS

With the exception of clearances, weather reports, and ATIS, nearly all ATC commands and advisories can be presented on a visual display limited to some reasonably small number of characters, or using short phrases presented via synthetic speech. Since such short messages constitute a large percentage of ATC transactions, and since these messages place more demands for short reaction times by flight crews than do longer messages such as clearances or ATIS, the emphasis during the early experiments in this series was on their evaluation.

3.1 PRELIMINARY EXPERIMENT

In a preliminary experiment run on the Transportation Systems Center's GAT-1 simulator* (No. 1 - Sect. 1), an attempt was made to validate the concept that Data Link could substitute for a voice channel in presenting a limited repertoire of ATC commands. A simple display capable of presenting heading, altitude, and speed commands was fabricated and then was evaluated by eight TSC pilots on simulated flight paths around the Boston area, ending in a landing at Boston's Logan Airport. On other runs, a MacDonnell-Douglas voice synthesizer having 128-word vocabulary was used. With both the visual display and voice synthesizer, commands were generated in real time via keyboard. Each subject made four simulated flights under the conditions of: (1) visual display without auditory alarm, (2) visual display with buzzer to alert to new information, (3) visual display with voice synthesizer, and (4) voice synthesizer alone. The time to acknowledge commands was measured with a stop watch. In the case of the voice synthesizer, timing was from completion of transmission of the message to the subject's reply; with the visual display, timing was from the transmission of the message to the light up of the "Acknowledged" light on the experimenter's control box.

*The GAT-1 simulates a light, single-engine aircraft with a one-man crew.

All subjects completed all runs successfully. Pilot opinion of the visual display was universally favorable. The intelligibility of the voice synthesizer was, in general, considered marginal. (Several of the words were later reprogrammed and testing indicated that intelligibility was considerably improved.) It was further noted that response times were longer when messages were presented both visually and auditorially; subjects apparently prefered to compare both sensory modalities before making a response. The visual display is depicted in Figure 3-1, the Voice synthesizer in Figure 3-2, and its keyboard input in Figure 3-3.

3.2 EVALUATION OF ADDITIONAL DISPLAYS ON THE GAT-1

The success of this very preliminary experiment justified the fabrication and testing of additional displays having capabilities for the presentation of a somewhat greater variety of short ATC messages (No. 3 - Sect. 1). Four such additional displays were fabricated and are depicted in Figures 3-4 through 3-7.

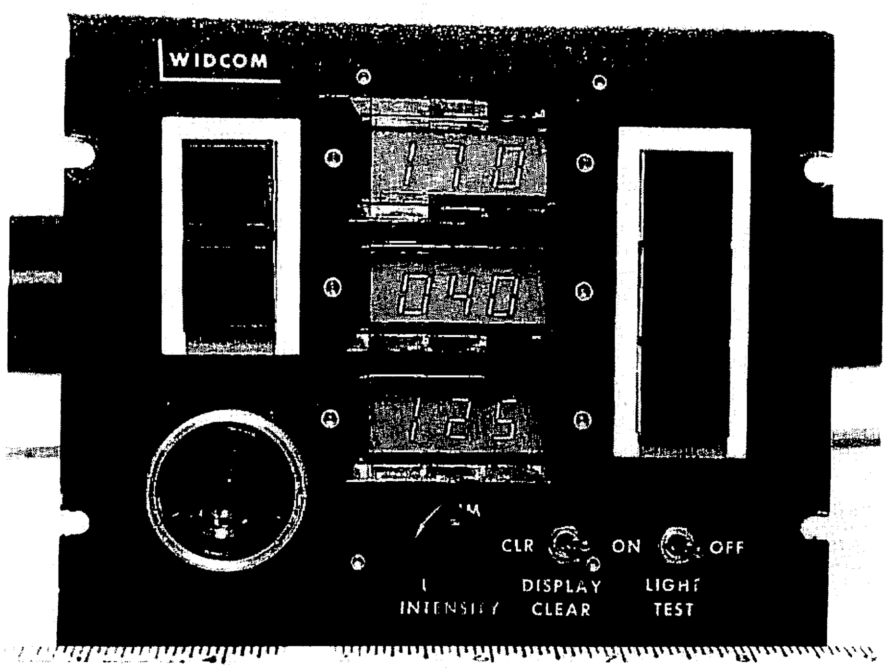

Figure 3-1. The WIDCOM Visual Display

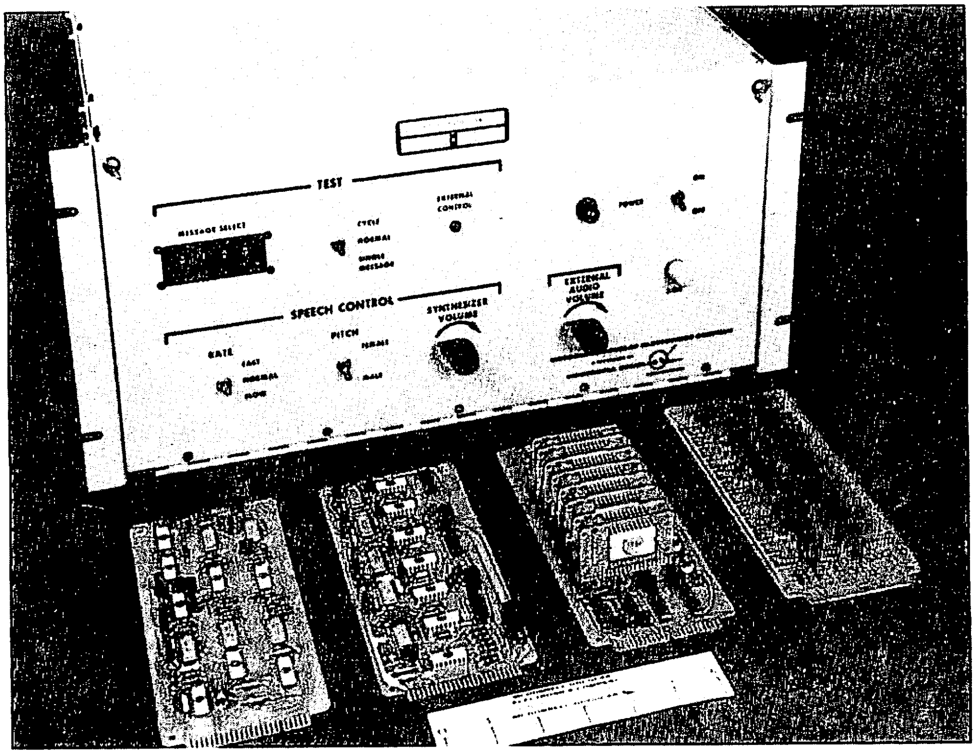

Figure 3-2. The McDonnell-Douglas Voice Synthesizer

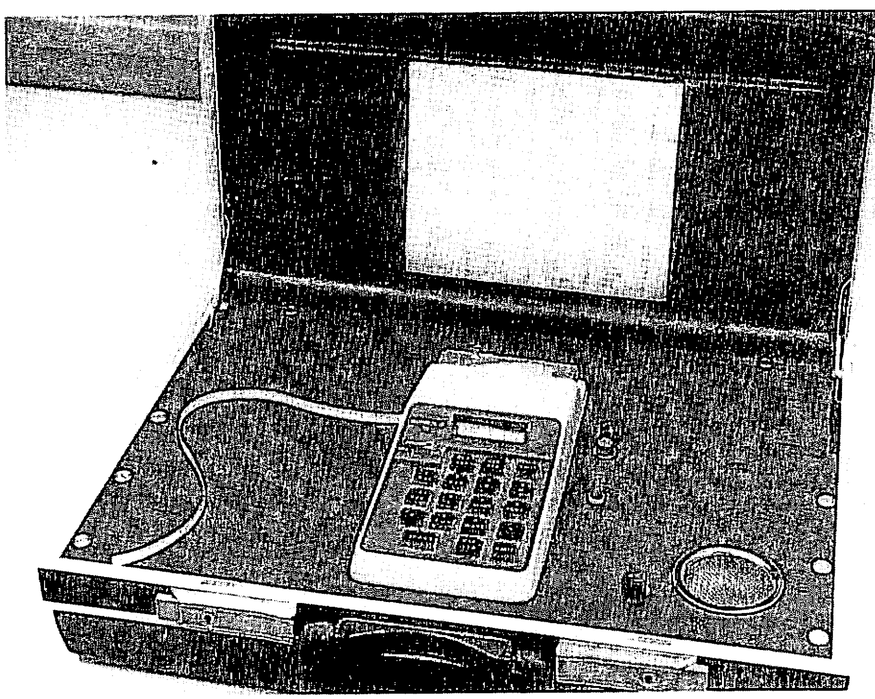

Figure 3-3. Keyboard Interface for the Voice Synthesizer

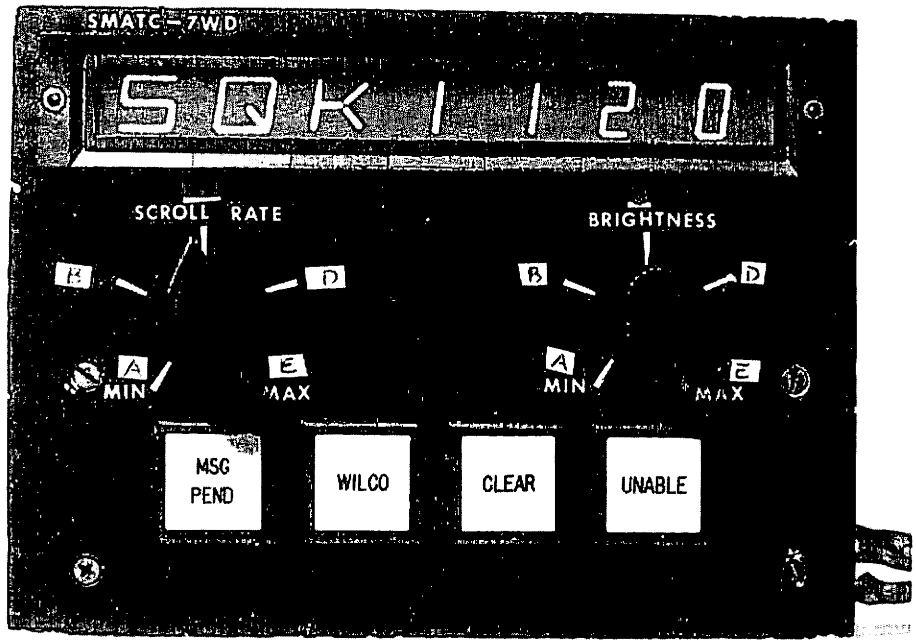

Figure 3-4. The 7-Window Display

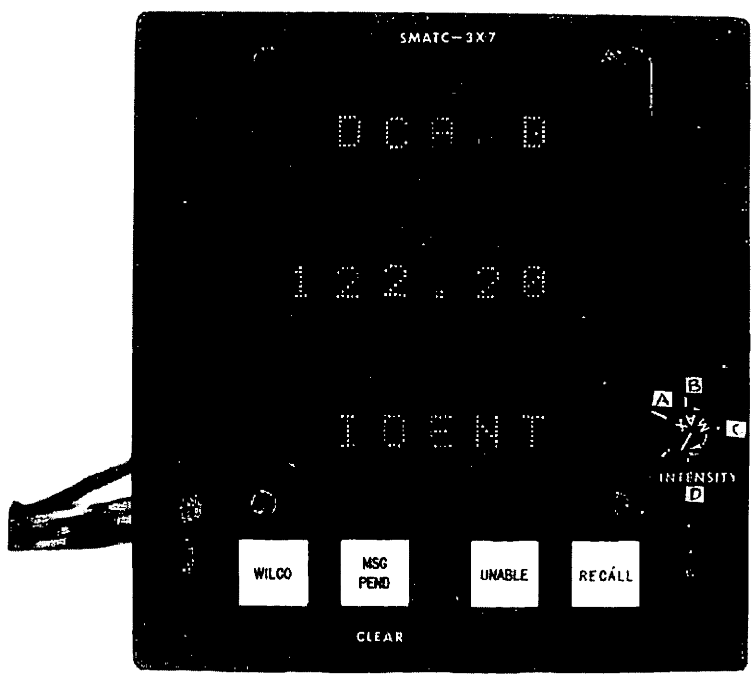

Figure 3-5. The 3x7-Window Display

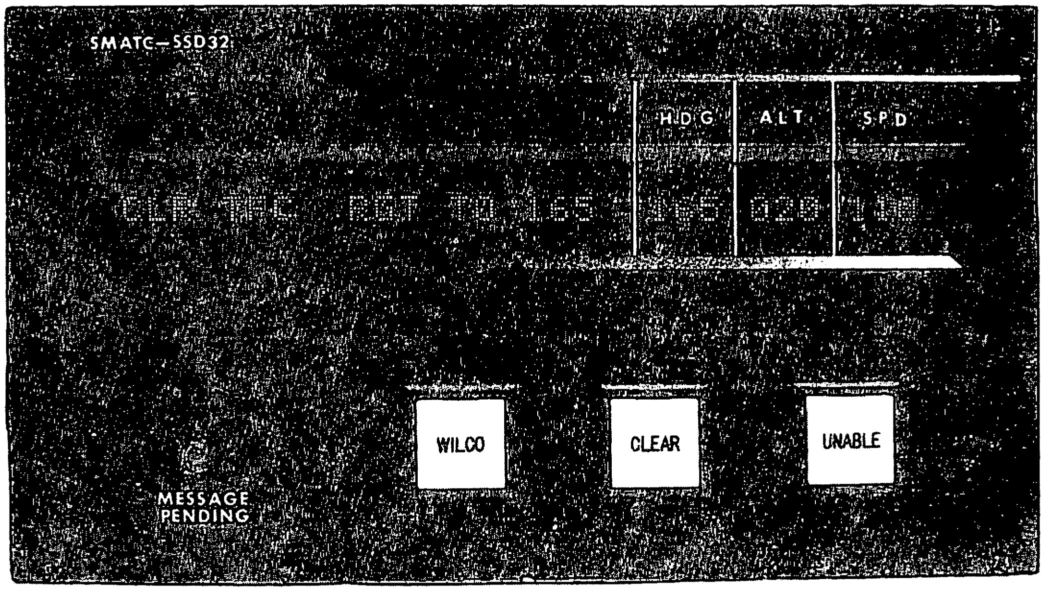

Figure 3-6. The 32-Window Display

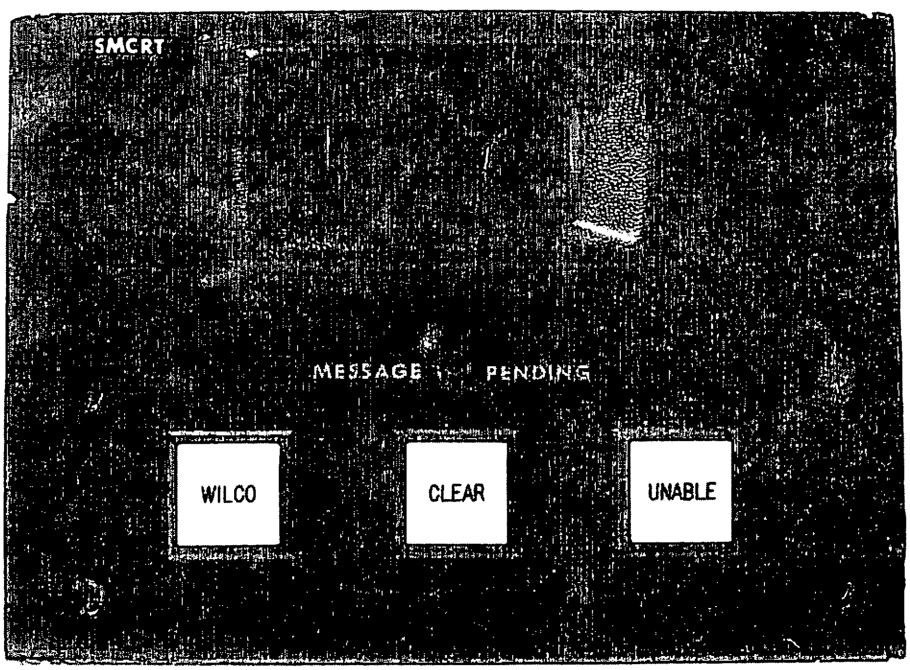

Figure 3-7. The NIMO Display

The 7-Window display, shown in Figure 3-4, used 16-segment alphanumeric readouts from Master Specialities Company as the means for character generation. Characters were generated using incandescent lamps and fiber optics, with a height of 0.42" and a typical brightness of 400 foot-Lamberts. Storage registers were provided for three seven-character messages, which could be accessed sequentially by depression of a "Message Pending" pushbutton located below the readout.

The second prototype display, depicted in Figure 3-5, presented three lines of seven characters each, using Monsanto red light emitting diodes (LED). Characters were based on a 5 x 7 dot matrix, character height was 0.35", and brightness typically 300 foot-Lamberts.

Figure 3-6 illustrates a plasma display utilizing a Burroughs Self-Scan Panel 8.50" wide by 2.25" high. On this panel, it was possible to present a linear array of 32 characters, each in a 5 x 7 dot matrix format with characters 0.20" high. Nominal light output per dot was 25 foot-Lamberts.

The final prototype, depicted in Figure 3-7, utilized a special NIMO tube, a variety of miniature Charactron CRT produced by Industrial Electronic Engineers, Inc. The tube contained an array of 64 cathodes and 64 areas of metal stencil, so arranged that any cathode and its associated stencil mask area could produce a character or a message on a 3/4-inch square area on the end of the tube. No deflection circuitry was employed; the position at which information appeared on the tube face was entirely a function of the geometry of the cathodes and their associated stencil mask areas. For the TSC application, a special mask was employed which made it possible to present messages to a maximum of six characters on each of three lines. Additionally, certain mask positions were reserved for individual digits at specific locations, so that by time-sharing cathodes and mask positions at a flicker-free rate, it was possible to display messages along with any required numerical values.

In addition to the display proper, each of the above units provided pushbuttons to permit the pilot to make "Wilco" and "Unable" responses to the previous message. On each display, the appearance of a new message was accompanied by an audio alert consisting of a "beep" repeated three times per second at 50% duty cycle until a response was made.

Each of the four displays along with its required drive circuitry was packaged in the same standard-size chassis to permit the displays to be installed and interchanged in a common location on the panel of the GAT-1 simulator. Panel installation was at pilot eye level and approximately one foot to the left of his centerline. Figure 3-8 is a block diagram of the experimental setup. Indicated here is the teletype/paper tape input for messages and the readouts to permit the experimenters to monitor flight parameters and pilot performance.

Eight FAA NAFEC test pilots each made a total of eight simulated flights to evaluate each of the displays under simulated daylight and night conditions. Four different scenarios, each

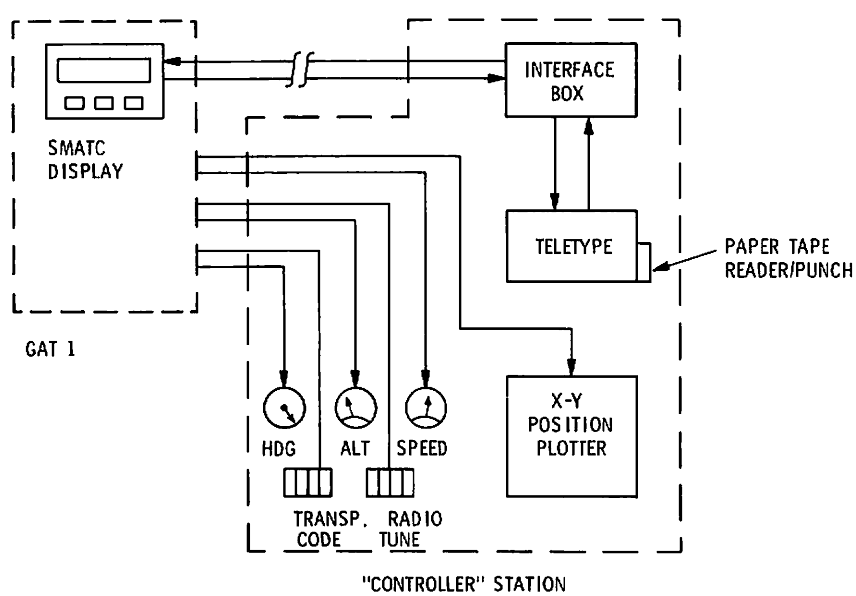

Figure 3-8. Block Diagram of Experimental Setup

involving simulated flights around the Boston area, were employed in a counter-balanced experimental design. Pilot response times to each message were recorded during the runs.* Table 3-1 shows the mean response time to each of the displays under simulated day and night conditions.

TABLE 3-1. MEAN REACTION TIMES TO THE DISPLAYS (IN SECONDS)

DISPLAY	DAY	NIGHT	MEAN
NIMO	3.31	3.36	3.33
7-Window	6.04	6.10	6.07
3x7-Window	4.71	4.66	4.69
32-Window	4.67	4.12	4.39
Mean	4.68	4.56	4.62

Here, it should be noted (1) that response time to the 7-Window display was slowest because of the requirement for scrolling this display to view the separate portions of those messages which required more than 7 characters, (2) that response time to the 32-Window display was appreciably slower under simulated daylight conditions than under night conditions, confirming that the readability of this display in daylight is marginal, and (3) that response times to the NIMO were the fastest of the 4 displays. Part of this difference is attributable to the fact that only a single message could be presented at any one time on the NIMO, whereas multiple commands such as heading, altitude and speed

*While response time is not the only indicator of the "success" of a display, it is easily measured and certainly gives some indication as to how readable and understandable a message may be. It also provides useful data for the design of an associated ground system.

could be presented on the other displays. When reaction time obtained with each of the displays was plotted as a function of the number of "information units" present (an "information unit" in this case being defined as a discrete message), it was first found that reaction time to 2-unit messages was somewhat greater than for 3-unit messages for the 3x7-Window and 32-Window displays. However, it was noted that this apparent discrepancy resulted from the extremely long reaction times to two-unit messages involving radio frequency plus transponder code settings. Here, because of the number of digits involved, pilots were reluctant to "Wilco" a message prior to making the appropriate settings because of the possibility of losing this information should a new message arrive. With these two-unit messages eliminated, there was a linear and only very modest increase in reaction time as the number of information units increased with 3x7-Window and 32-Window displays, as indicated in Figure 3-9. However, it should also be noted that with the requirement for manual scrolling of the seven-Window display, response times increased markedly for two-unit messages and would probably not be acceptable for three-unit messages.

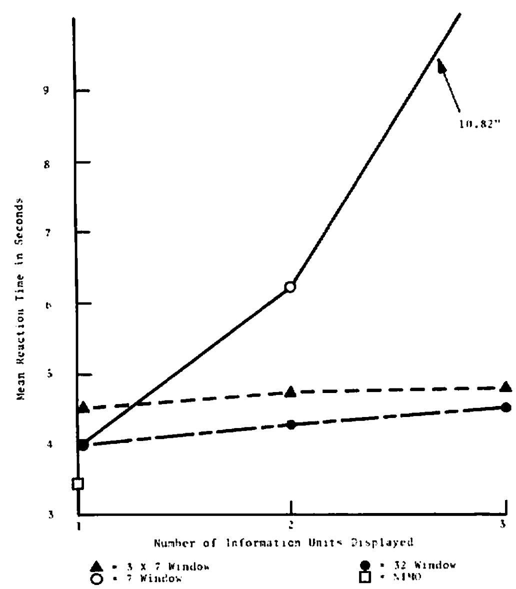

Figure 3-9. Reaction Times as a Function of Number of Information Units in Messages – Two-Unit Messages of Radio Frequency and Transponder Code Settings Omitted

Pilots were informed prior to the start of the experimental runs that impossible commands would occasionally be issued to prevent routine pressing of the "Wilco" buttons without first interpreting the meaning of a message. Typical impossible messages asked for a speed of 900 knots or a heading of 540 degrees. A total of 44 such messages were introduced into the 64 experimental runs. Pilots responded with an immediate "Unable" on 30 of these, with "Wilco" immediatedly followed by "Unable" on 5, and failed to detect erroneous "Wilco" responses to 9 of the impossible messages.

Only two of the eight pilots made no errors. Mean response time for the messages correctly detected as erroneous was 5.71 seconds as compared with a mean value of 4.62 seconds for all messages.

At the completion of the experimental runs, a 28-item questionnaire was administered to each of the pilots to elicit additional information. There was general agreement on only a few of the items. Pilot opinion of Data Link as a concept was generally favorable and a majority felt that it would reduce pilot workload. Pilot comments included: "It's great to get rid of that incessant chatter." The scratchpad capability of the Data Link also drew favorable comments. There was, on the other hand, complete lack of agreement as to the preferred color for the Data Link displays, with comments ranging from "red" to "anything but red." When asked to rank-order the 4 displays, and with weighting of these rank orders, the 32-Window display proved an overwhelming favorite, the 3x7-Window and NIMO displays appeared usable and acceptable, and the 7-Window display (in its current version) appeared unacceptable.

3.3 TESTING OF THE FOUR PROTOTYPE DISPLAYS ON THE GAT-2

The encouraging results from this series of simulated flights prompted further exploration of Data Link using two-man crews in the GAT-2* simulator at FAA/NAFEC (No. 4 - Sects. 1-3). Planning

*The GAT-2 simulates a light, twin-engine aircraft with a two-man crew and space for an on-board observer.

for these newer experiments included certain modifications of the displays used in the GAT-1 tests. An automatic scrolling feature was incorporated into the seven-Window display so that up to three lines of seven characters each could be presented sequentially without a requirement for pilot intervention. A storage register was built into the 3x7-Window display to provide scratchpad callup of heading, altitude, and speed commands on demand. Dimming controls were added to all of the displays. Twelve character spaces on the right hand side of the 32-Window display were reserved for the continuous display of heading, altitude, and speed commands, thereby limiting other messages to a maximum of 20 characters. Also, a "Wilco" button was provided on each simulator control column as well as on the display itself. As in the previous experiment, messages were generated from punched tape through a TSC-constructed interface box which provided the required decoding, storage, and control functions.

Eight two-man crews of FAA/NAFEC test pilots each made four simulated flights, two on each of two scenarios, one involving a flight from JFK to Atlantic City, and the other from Philadelphia to LaGuardia (No. 4 - Sect. 1). The experimental design was replicated for eight additional crews consisting of airline and ALPA Pilots (No. 4 - Sect. 2). A single crew of AOPA pilots also participated (No. 4 - Sect. 3). The greater complexity of the experiment over that of the earlier experiment made it possible to make a much greater number of comparisons between and among data points than had been previously possible. In addition to differences among crews, performance differences within crews were also measurable. Performance on each of the displays could be measured, as well as differences between "Wilco" responses made on the control column versus those made directly on the display panel, day versus night differences, and first versus second runs on a scenario. Response times as a function of message type and length were, of course, measured, as well as responses to "Unable" messages. Comparisons among groups (test pilots versus airline pilots versus AOPA pilots), "old" (test pilots who had previously participated in the tests on the GAT-1) versus "new" (test pilots with no previous Data Link experience), and between the results of the tests conducted on the GAT-1 and GAT-2 were also possible.

The separation of the data from the pilot groups was made because the FAA/NAFEC pilots represented a more homogeneous population than the Airline/ALPA pilots. The age range for the NAFEC pilots was 45 to 53, as compared with an age range of 30 to 56 for the airline pilots. Flying experience of the NAFEC pilots varied from 4,200 to 20,000 hours, while the experience of the airline pilots ranged from 800 to 25,000 hours. On the average, the NAFEC pilots had some 2,200 extra hours of experience over that of the airline pilots. The fact that the NAFEC pilots were test pilots by profession and thus experienced in the evaluation of new equipment in a variety of aircraft provided further justification for the separation of the data from the two groups.

Due to an oversight in data collection, no record was kept as to which crew member made the responses on the display panels; only those responses made on the control column "Wilco" buttons can be attributed to a specific crew member. This defect was rectified in future experiments.

For those data in which the crew member making the response can be identified, it can be seen that different crews handled Data Link differently. Under usual commercial procedures, the copilot is responsible for the majority of communications transactions, while the pilot handles control of the aircraft. During the experiment, the crews were given no instructions as to which crew member had responsibility for Data Link, and different crews interpreted Data Link on their own as either a control or communication function, even though they alternated as pilot and copilot on successive runs. A majority favored the copilot communication function. Despite the alternation of cockpit seating, crew members were remarkably consistent in having shorter response times while serving as copilot. For the NAFEC crews, mean response time for pilots was 7.2", as compared with 5.5" for copilots. For the airline crews, the times were 6.4" and 5.5", respectively. For both groups, responses made on the display panel were measurably longer, with means of 7.8" and 6.7", respectively. Some definition of Data

Link function is indicated in the instructions for crews in future experiments.

Mean reaction times of the NAFEC and airline crews for the four displays are presented in Table 3-2. For both groups, mean response times were fastest for the 3x7-Window and NIMO displays, somewhat slower with the 32-Window display, and slowest for the 7-Window display. However, with each of these pilot groups, there was wide variability in the response of individual crews to the four displays. The ratios of the reaction time of each crew on each display to the mean reaction times of all crews to that particular display are presented in Tables 3-3 and 3-4. The more disparate performances are underlined. Among the test pilots, at least one crew performed poorly on each of the displays, and two crews (B and G) performed poorly on two of the displays (but on different displays). Only one crew performed better than average on all four of the displays. Among the airline pilots, much greater consistency in performance was observed on the NIMO and 3x7-Window displays. One crew performed poorly on the 7-Window display, and the same crew (M) performed poorly on the 32-Window display, with a response time twice as great as the mean. On the other hand, crew "L" had a response time on the 32-Window display half of the mean value for all crews. These results indicate the need to use several crews in the evaluation of any given display if meaningful data are to be obtained. The use of fewer than eight crews does not seem to be advisable in future experiments, since some of the conditions tested proved to be excessively difficult or easy to only one out of the eight crews who participated.

If variability among crews on their performance with a given display is to be minimized, it is important that the scenarios be equated for difficulty insofar as possible, since with an incomplete block experimental design as in the present experiments, all crews did not evaluate each display with each scenario. Mean reaction times for the NAFEC crews on Scenarios "A" and "B" were 6.8" and 6.4" respectively; for the airline crews, 6.1" and 6.1", indicating little difference in difficulty between the two scenarios. There was, however, a noticeable practice effect between first and second runs on a scenario, as indicated in Table 3-5. These data

TABLE 3-2. MEAN REACTION TIMES OF TEST PILOTS AND AIRLINE CREWS TO THE FOUR DISPLAYS (IN SECONDS)

DISPLAY	FAA CREWS	AIRLINE CREWS
NIMO	5.9	5.4
7-W	8.5	6.9
3x7-W	5.9	5.1
32-W	6.4	6.8
Mean	6.7	6.1

TABLE 3-3. RATIO OF THE MEAN RESPONSE TIME OF EACH NAFEC CREW ON EACH DISPLAY TO THE MEAN RESPONSE TIME OF ALL NAFEC CREWS ON THAT PARTICULAR DISPLAY

DISPLAY	CREW								
	A	B	C	D	E	F	G	H	RANGE
NIMO	.71	1.08	1.00	1.02	.73	.81	1.46	1.14	0.75
7-W	.98	1.46	1.26	.69	.94	.84	.69	1.18	.71
3x7-W	1.07	.86	1.17	.63	.76	.75	1.58	1.14	.95
32-W	.94	1.40	.84	.75	1.06	.97	1.11	.91	.65

TABLE 3-4. RATIO OF THE MEAN RESPONSE TIME ON EACH AIRLINE CREW ON EACH DISPLAY TO THE MEAN RESPONSE TIME OF ALL AIRLINE CREWS ON THAT PARTICULAR DISPLAY

DISPLAY	CREW								
	I	J	L	L	M	N	O	P	RANGE
NIMO	1.07	1.00	.89	.96	1.15	1.06	.96	.89	0.26
7-W	1.13	1.14	.81	.68	1.65	1.14	.75	.71	.97
3.7-W	.90	1.06	1.16	.80	1.12	.94	1.12	.92	.36
32-W	.91	1.04	1.03	.50	2.00	1.18	.59	.72	1.50

TABLE 3-5. MEAN REACTION TIME OF CREWS ON 1ST AND 2ND RUNS ON EACH SCENARIO (IN SECONDS)

CREWS	SCENARIO "A"		SCENARIO "B"	
	1st Run	2nd Run	1st Run	2nd Run
FAA	6.9	6.6	7.3	5.4
Airline	6.9	5.1	7.0	5.2

indicate that a given scenario should definitely not be run more than twice by any given crew, and that preferably it should be used only once.

For the "daylight" condition in the tests run on the GAT-1, a high intensity light source was used, which produced an uncomfortable glare. Limitations of the layout in the GAT-2 area did not permit similar high intensity lighting, and the "daylight" condition instead approximated that of a well lit office. While differences in performance were found between the "night" and "day" conditions on the tests of the GAT-2, these differences were inconsistent, as indicated in Table 3-6.

TABLE 3-6. REACTION TIME DIFFERENCES BETWEEN DAY AND NIGHT CONDITIONS (IN SECONDS)

DISPLAY	FAA CREWS			AIRLINE CREWS		
	DAY	NIGHT	MEAN	DAY	NIGHT	MEAN
NIMO	5.9	5.7	5.8	5.3	5.6	5.4
7-W	8.2	9.2	8.7	7.8	6.4	7.0
3x7-W	4.9	6.8	5.8	4.7	5.5	5.1
32-W	6.4	6.9	6.7	7.8	5.7	6.7
Mean	6.4	6.9	6.7	6.4	5.8	6.1

As with the GAT-1 tests, certain of the messages presented on each of the displays (except the NIMO) contained more than one unit of information. The use of such multiple units of information for Data Link could increase its efficiency, since the long string of characters required to establish sync, aircraft ID, parity, etc., would require transmission only once, while the message per se could provide multiple units of information. Such an increase in transmission efficiency must, however, be predicated on the ability of crews to assimilate such multiple messages readily and without large increases in response times. Table 3-7 indicates that response time is increased only modestly when two units of information are presented in a single message.

TABLE 3-7. RESPONSE TIMES AS A FUNCTION OF NUMBER OF INFORMATION UNITS IN MESSAGES (IN SECONDS)

DISPLAY	FAA CREWS		AIRLINE CREWS	
	# OF INFO. UNITS 1	2	# OF INFO UNITS 1	2
NIMO	5.2	---	5.0	---
7-W	9.1	11.8	6.3	5.9
3x7-W	5.5	5.5	4.9	5.2
32-W	5.9	7.4	4.8	6.3
Mean	6.3	8.3	5.3	5.9

For the tests using the GAT-2, the differences in response times to two-unit messages containing combined radio frequency and transponder code settings and other two-unit messages were trivial, probably because the two crew members could share the load of tuning the radio and transponder.

Again, as in the tests on the GAT-1, responses to "Unable" messages were longer than for "Wilco." For the NAFEC crews, mean "Unable" response time was 7.7 seconds, as compared with 6.3 seconds for "Wilco." For the airline crews, these figures were 10.0 seconds and 6.1 seconds, respectively.

The two pilot groups handled Data Link responses in a somewhat different manner. With the test pilot group, a majority of the responses were made by the copilot using his control column response button; with the airline pilots, the largest number of responses were made directly on the display panel. Despite this, and the different age and experience makeup of the two groups, they were remarkably consistent in their relative ranking under the several conditions studied. Table 3-8 compares the two groups for a number of the parameters which were studied.

TABLE 3-8. SUMMARY OF RESPONSE DIFFERENCES BETWEEN TEST AND AIRLINE PILOTS (IN SECONDS)

	FAA Pilots	Airline Pilots
% Responses:		
By Pilot	15.4	15.7
By Copilot	51.1	31.6
On Display	33.5	52.7
Response Times:		
Mean	6.7	6.1
NIMO	5.9	5.4
3x7-W	5.9	5.1
32-W	6.4	6.8
7-W	8.5	6.9
By Pilot	7.2	6.4
By Copilot	5.5	5.5
On Display Panel	7.8	6.7
Scenario A: 1st Run	6.9	6.4
Scenario A: 2nd Run	6.6	5.1
Scenario B: 1st Run	7.3	7.0
Scenario B: 2nd Run	5.4	5.2
Daylight	6.4	6.4
Night	6.9	5.8
1 Msg. Unit Messages	6.3	5.3
2 Msg. Unit Messages	8.3	5.9

A single pilot was used for the tests run on the GAT-1, and somewhat different results might therefore be anticipated from those obtained on the GAT-2 (No. 4 -Sect. 1). Responses were appreciably faster on the GAT-1 tests; probably this was because the single crew member felt no requirement for waiting until he was certain that the other crew member had also absorbed the information before acknowledging, a delay which avoided the possibility of losing that information. The differences in relative response times to the 4 displays on the 2 series of tests probably resulted from: (1) the newly installed dimming capability on the 3x7-Window display, which made it possible to avoid some of the glare present during the tests run on the GAT-1; (2) A tendency for certain crews to scan the entire length of the display before making a response, which may have developed because 12 characters on the right hand side of the 32-Window display were reserved for the continuous display of heading, altitude, and speed, even though the heading, altitude, and speed information on the right hand side of the display frequently duplicated the information presented on the left; and (3) the location of the displays midway between the crew members in the GAT-2, which made it difficult to read the small characters on the NIMO; no similar difficulty was present with the larger characters of the 7-Window and 3x7-Window displays. The data from the tests on the GAT-1 and GAT-2 are compared in Table 3-9.

TABLE 3-9. MEAN RESPONSE TIMES (IN SECONDS) FOR TESTS ON THE GAT-1 AND GAT-2

Response Times	GAT-1	GAT-2
Mean	4.6	6.4
NIMO	3.3	5.7
3x7-Window	4.7	5.5
32-Window	4.4	6.6
7-Window	6.1	7.7
Day	4.7	6.4
Night	4.6	6.4

Eight of the test pilots who served as subjects in the GAT-2 experiment had previous familiarity with the Data Link equipment. This did not, however, result in performance faster than that of the "new" test pilot group. Mean reaction time for the "old" pilots was 6.8 seconds, as compared with 6.6 seconds for the "new" pilots.

The data obtained from the single crew of AOPA pilots do not permit broad generalization, since counterbalancing of the sequence in which displays were evaluated, scenarios, and day versus night conditions was not possible. As might be expected with pilots used to flying their own aircraft, frequently alone, the copilot functioned mainly in the role of observer on all four simulated flights made by this crew. The crew member serving as copilot made no responses using his control column. Two responses were made directly on the display panel and the remaining 157 responses by the pilot on his control column.

Mean response time for the two AOPA pilots was 6.1 seconds, comparable with that of the airline pilots. The rank ordering of the response times to the several displays was comparable with that of the other pilot groups; these data, however, could be subject to modification if a larger number of AOPA crews were used in a counter-balanced experimental design, such as that employed for the NAFEC and Airline Pilots.

The larger number of pilots available to answer a questionnaire after the tests on the GAT-2, as compared with those on the GAT-1, made possible an even greater diversity of opinions. Both groups were highly but not universally favorable to Data Link, with the test pilots appearing to be slightly more enthusiastic than the airline pilots. Thirteen out of the 16 test pilots and only 9 out of the 16 airline pilots thought that Data Link would reduce pilot work load. Airline pilots expressed concern that a visual communications system would interfere with other visual tasks, and that the information presently heard during the ATC communications transactions with other aircraft would be lost. Scratchpad was a highly popular feature. The airline pilots, however, preferred the recall feature of the 3x7-Window display to the continuous scratchpad capability of the 32-Window display by 9 to 6, whereas the test pilots preferred the 32-Window scratchpad by 9 to 4. White was the favorite display color, preferred by seven test pilots and five airline pilots. Red, though it had some adherents, aroused the most opposition. The limited data from the two AOPA pilots did not produce opinions which differed from those expressed by one or more members of the test pilot and airline pilot groups. The rank order of preferences for the several displays by the several pilot groups are presented in Table 3-10.

TABLE 3-10. NUMBER OF PILOTS EXPRESSING THEIR RANK ORDER OF PREFERENCE FOR EACH DISPLAY

Display & Pilot Group	# of Pilots Ranking Display as:			
	Best	2nd	3rd	Worst
Test Pilots:				
NIMO	0	2	3-1/2	10-1/2
7-Window	3	3	6-1/2	3-1/2
3x7-Window	2	9	5	0
32-Window	11	2	1	2
Airline Pilots:				
NIMO	0	0	2	14
7-Window	0	3	11	2
3x7-Window	6	7	3	0
32-Window	10	6	0	0
AOPA Pilots:				
NIMO	0	0	1	1
7-Window	0	0	1	1
3x7-Window	1	1	0	0
32-Window	1	1	0	0
All Pilots:				
NIMO	0	2	6-1/2	25-1/2
7-Window	3	6	18-1/2	6-1/2
3x7-Window	9	17	8	0
32-Window	22	9	1	2

4. TESTS UTILIZING A FULL DATA LINK COMPLEMENT: EXPERIMENTAL CONDITIONS

The experiments described thus far have concentrated on display requirements (both hardware and formatting) for short uplink ATC messages and permitted the formulation of a design for a display for such messages which could be packaged in a standard ATI-3 case, so as to present the appearance of a standard aircraft instrument. Although short ATC messages comprise a high percentage of ATC transactions, simulation of a complete Data Link environment additionally required that means be provided for ground-to-air simulated transmission and display of longer messages such as ATIS, clearances, and weather reports, and for the generation and simulated downlink transmission of a much larger repertoire of messages than the simple WILCO and UNABLE responses of the earlier experiments. Such complete uplink and downlink capability was provided in Phases 1-C and 1-D of the Data Link program, and has been reported in detail in Iterim Reports Nos. 5 and 6. This section of the present report provides an extended summary of these last two Interim Reports. The gradual increase in scope and complexity of the simulation experiments throughout the course of the program is indicated in Table 4-1.

4.1 EXPERIMENTAL I/O DEVICES

For the Phase 1-C and 1-D experiments, a variety of devices were fabricated or purchased, so that those selected for a particular experimental run could form complements ranging from minimal for complete Data Link to those capable of providing redundant information on multiple devices. Each of the individual devices is briefly described below.

4.1.1 16-Window SMATC (Short Message ATC Display)

Figure 4-1 shows the 16-Window SMATC, which was configured as two lines of eight characters per line. While pilots had previously stated that they preferred a color other than red, red dot-matrix

TABLE 4-1. SUMMARY OF TESTING USING FLIGHT SIMULATORS

Interim Report No.	Uplink Equipment Visual	Uplink Equipment VRS*	Downlink Equip.	Simulator	Pilots	Scenarios
1	Widcom	McDonnell-Douglas	Voice	GAT-1	TSC	Approaches & Landings
3	3x7W (or) 7-W (or) 32-W (or) Nimo	None	Wilco/ Unable only	GAT-1	Test Pilots	Four flights around Boston area
4	3x7W (or) 7W (or) 32W (or) Nimo		Wilco/ Unable only	GAT-2	Test Pilots + Commercial	PHL-LGA JFK-ACY
5	2x8W + Printer	Vortax	CDU** or SCDU	GAT-2	Gen. Aviation, Commercial & Test Pilots	ILG-ABE TTN-ILG TEB-PNE WWD-Wings
6	2x8W + Printer	Vortax	CDU	727 DC-9	Commercial + Test Pilots	LAX-SFO SFO-LAX KCI-ORD

*VRS = Voice Response System
**CDU = Control and Downlink Unit: SCDU is a simplified version
"W" refers to the # of characters possible in the display.

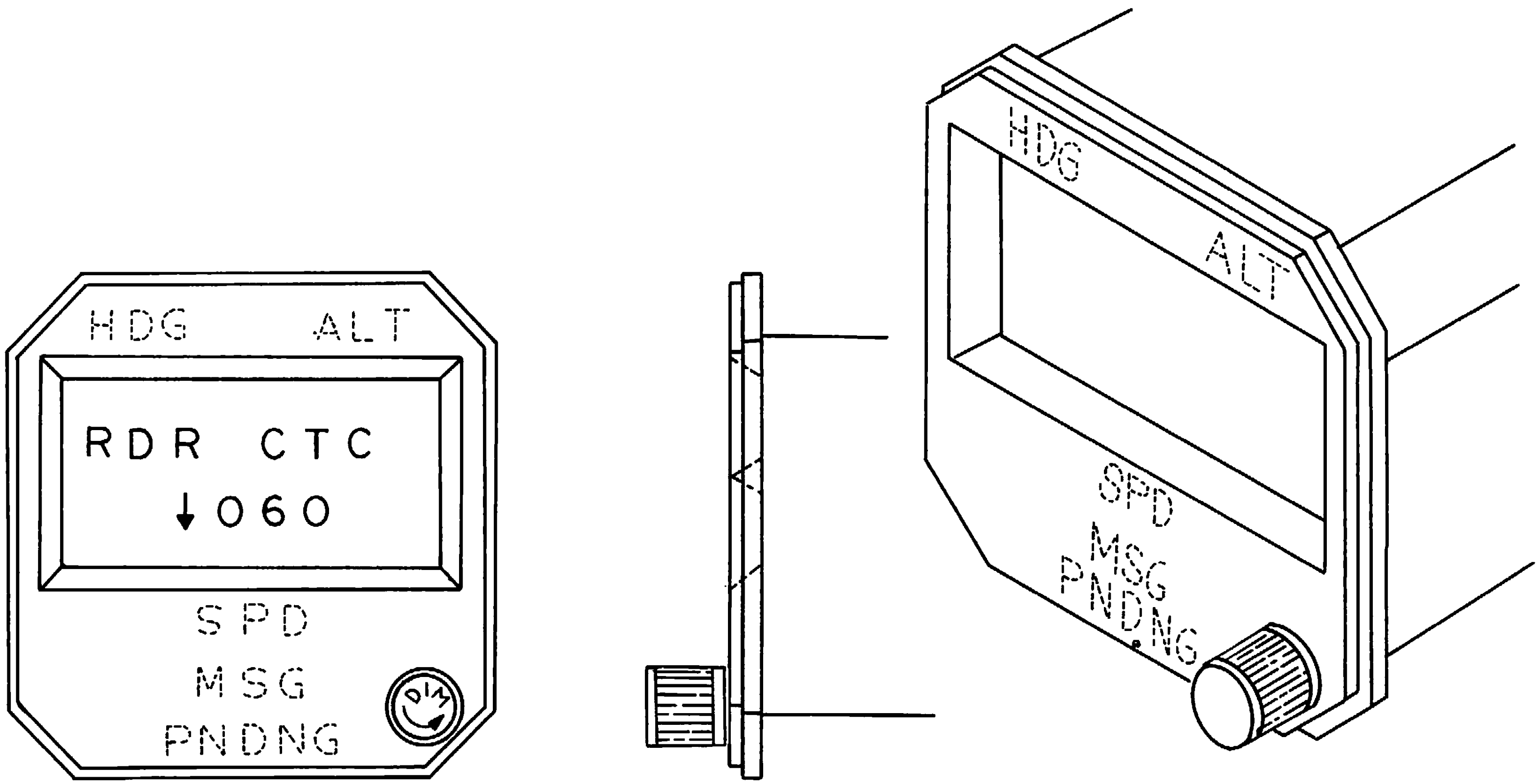

Figure 4-1. SMATC Display

LED (Light Emitting Diode) elements were used, under the assumption that when and if Data Link is implemented, yellow or green LED matrices will be available. As indicated above, the unit was housed in an ATI-3 case. Capability was provided for the storage and recall of the latest heading, altitude, and speed commands (with appropriately illuminated labels when this presentation was called for), as well as for the temporary storage of any other single message for later recall. (Previous experiments had indicated that the character size (0.35" high) was easily readable by flight crews, and that 16 characters permitted the presentation of the majority of short ATC messages without requiring cryptic abbreviations.) An additional illuminable indicator on the display alerted the crew to the availability of a new message when they had not indicated compliance with or unacceptability of a previous message. A single, centrally-located SMATC was used on the GAT-2 tests of Phase 1-C, and was located as indicated in Figure 4-2. For the later Phase 1-D tests on airline simulators, two SMATC's were provided, with the locations indicated in Figure 4-3.

4.1.2 Voice Synthesizer (Vosyn)

The Vosyn, a Model 6 Votrax* from the Vocal Interface Divsion of Federal Screw Works, generated synthetic speech by providing an audible output of its various stored phonemes upon command by the minicomputer. A 256-word vocabulary was stored in the minicomputer to provide all of the required ATC messages for the Phase 1-C experiment. Modifications to the vocabulary for the Phase 1-D tests consisted primarily of changing airport and navaid names to correspond to the different territories over which the simulator was being "flown." Figure 4-4 shows the Votrax and its programming unit.

A Vosyn volume control and a Vosyn repeat button were mounted on the CDU (Control and Downlink Unit, to be described later). The Vosyn was capable of repeating the current heading, altitude, and/or

*Trademark of Vocal Interface Division, Federal Screw Works, Troy Michigan.

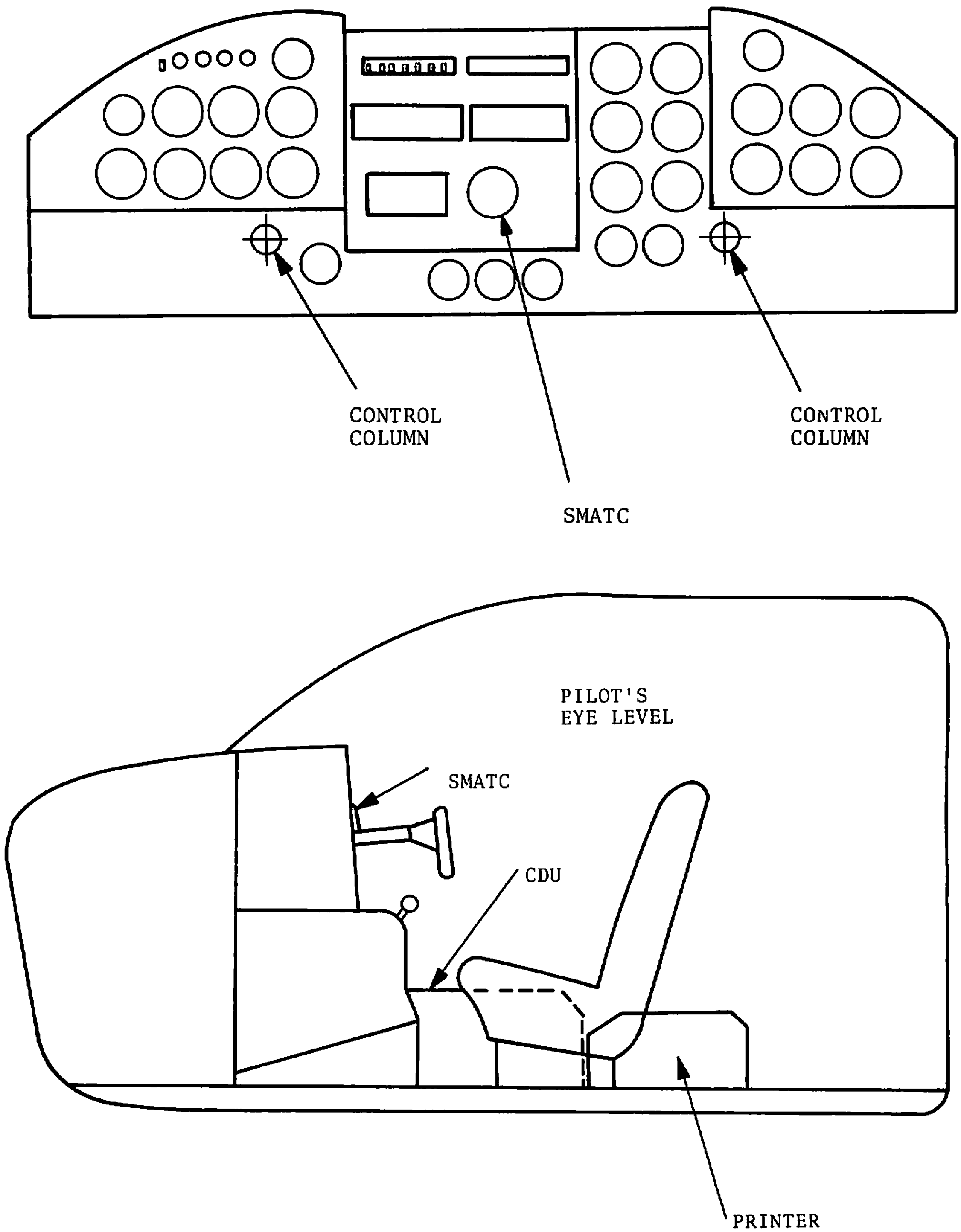

Figure 4-2. Equipment Locations in GAT-2

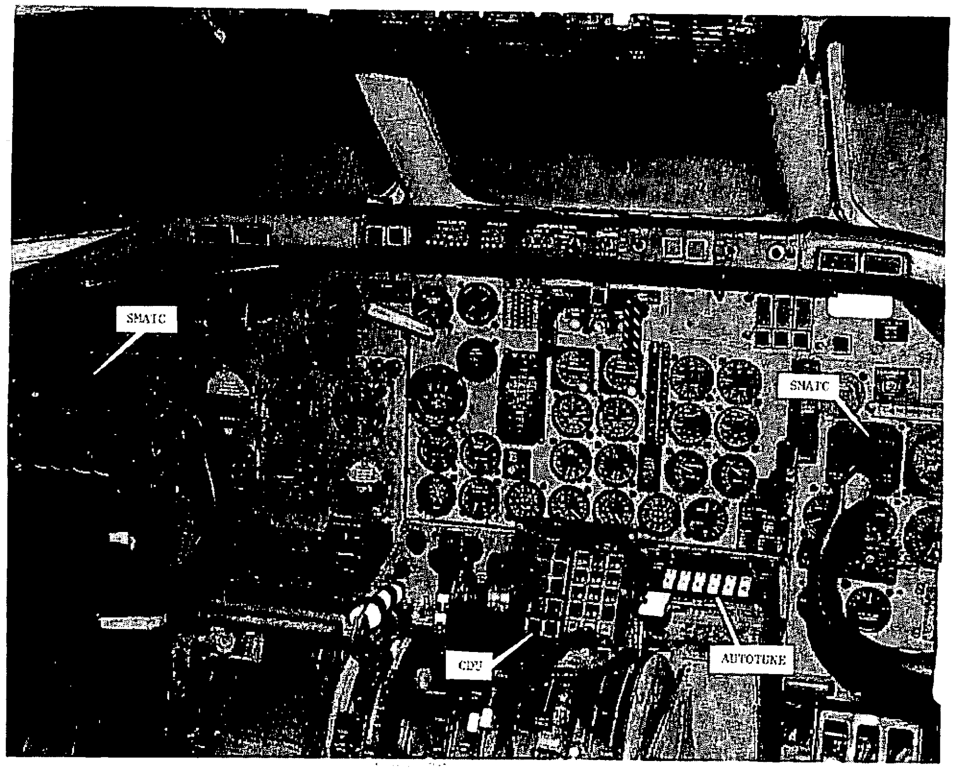

Figure 4-3. I/O Device Locations in the DC-9 Simulator

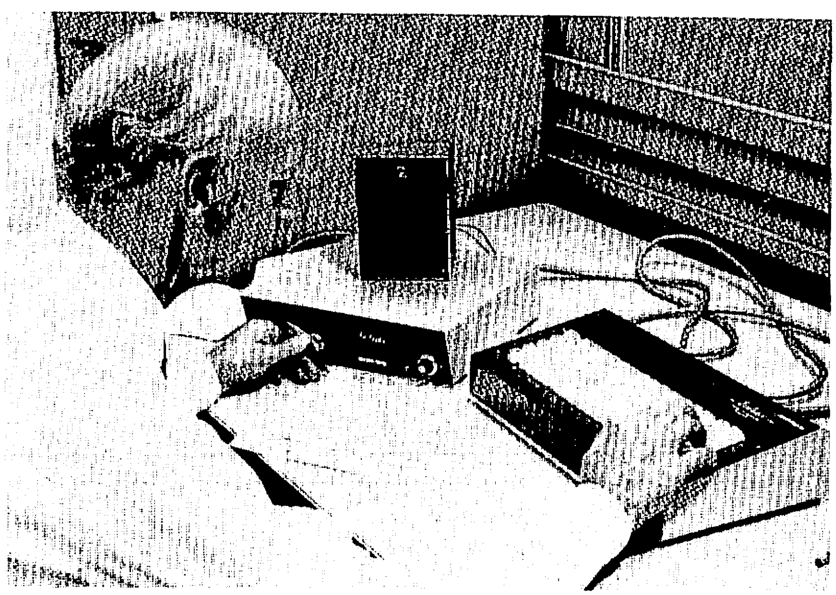

Figure 4-4. Votrax Voice Synthesizer with Programming Unit

speed commands when the proper request was made via the CDU, or any other single message which the crew might have stored temporarily. The Vosyn was mounted in an equipment rack with the minicomputer outside the simulator, and its audio output was introduced into the simulator.

4.1.3 Printer

A 21-column, 2-color (red and black) Anadex impact printer was used as a hard copy display (Figure 4-5). Printing speed was 25 characters per second. For the GAT-1 tests, the device was placed on the floor of the simulator cockpit, behind the pedestal and between the seats, as indicated previously in Figure 4-2. For the Phase 1-D tests on the B-727, it was placed to the right of the Second Officer's workspace, as indicated in Figure 4-6; for the DC-9 tests, it was located to the right of the First Officer, at floor level, as indicated in Figure 4-7. All messages during Phase 1-C were printed in black. On Phase 1-D, red printing was reserved for company messages.

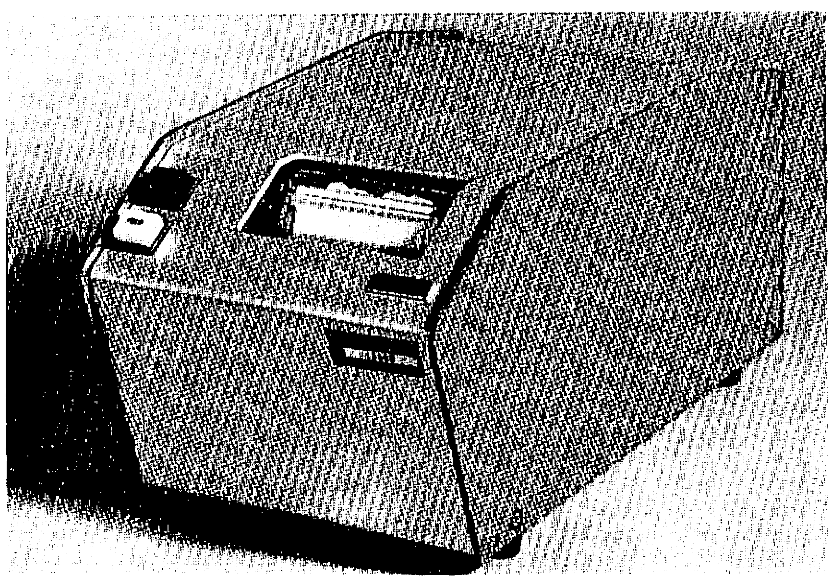

Figure 4-5. Anadex Model DP-751 Printer

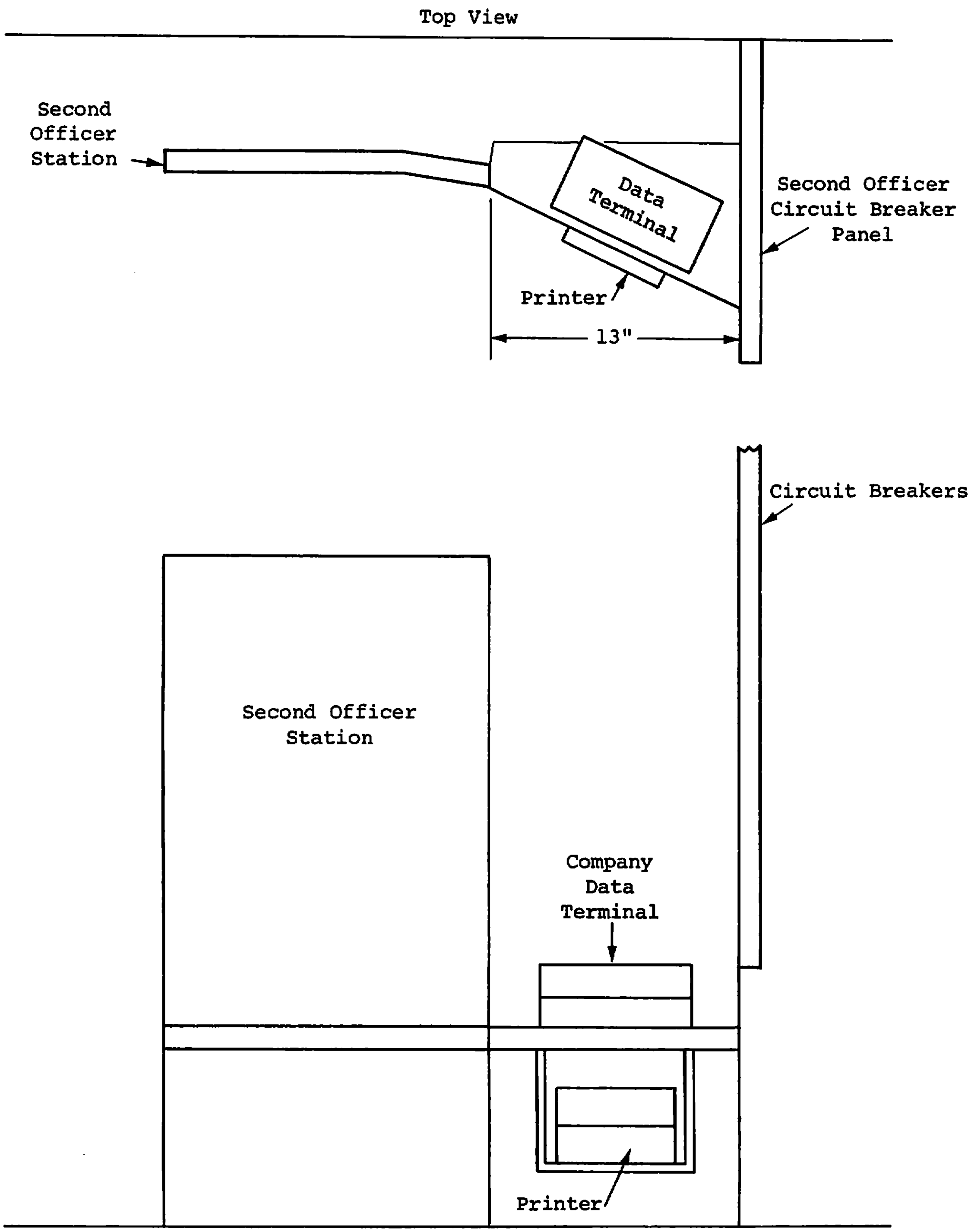

Figure 4-6. B-727 Printer Installation

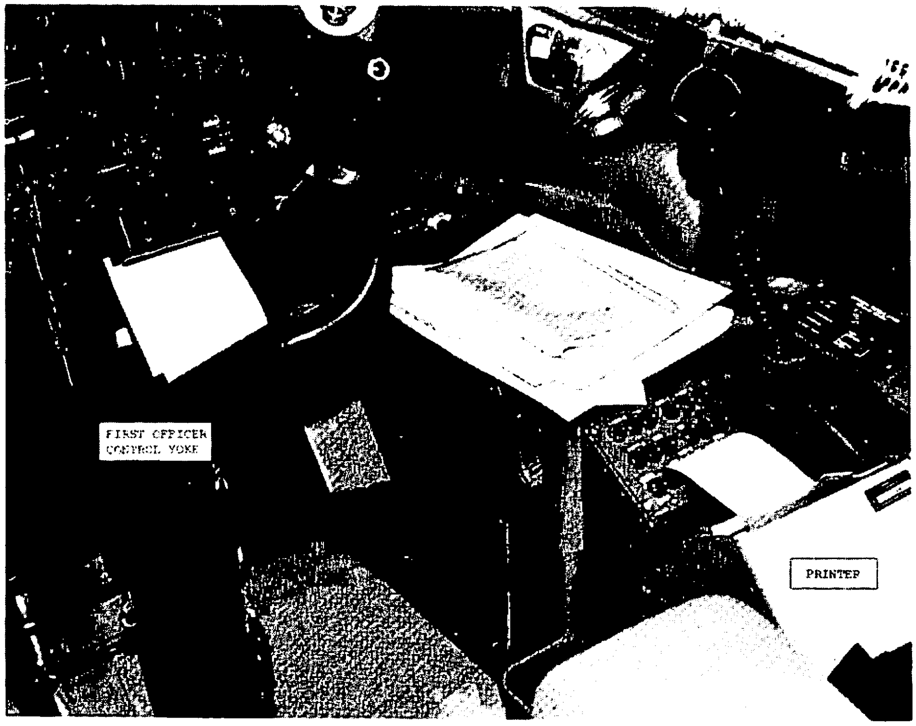

Figure 4-7. Printer Location in the DC-9 Simulator

4.1.4 <u>CDU (Control and Downlink Unit)</u>

The CDU (Figure 4-8) contained all the buttons and controls required to allow the crew to operate the cockpit I/O system and generate downlink messages. Its functions included:

a. Function buttons for generating the most commonly-used downlink messages by a single button depression.

b. An alphanumeric keyset for adding letters and/or numbers to downlink messages when these were required.

c. WILCO, UNABLE, and STANDBY buttons to permit crew response to uplink messages.

d. "Heading, Altitude, and Speed Recall" button for use with the SMATC and Vosyn.

e. A "Vosyn Repeat" button and "Vosyn Volume" control.

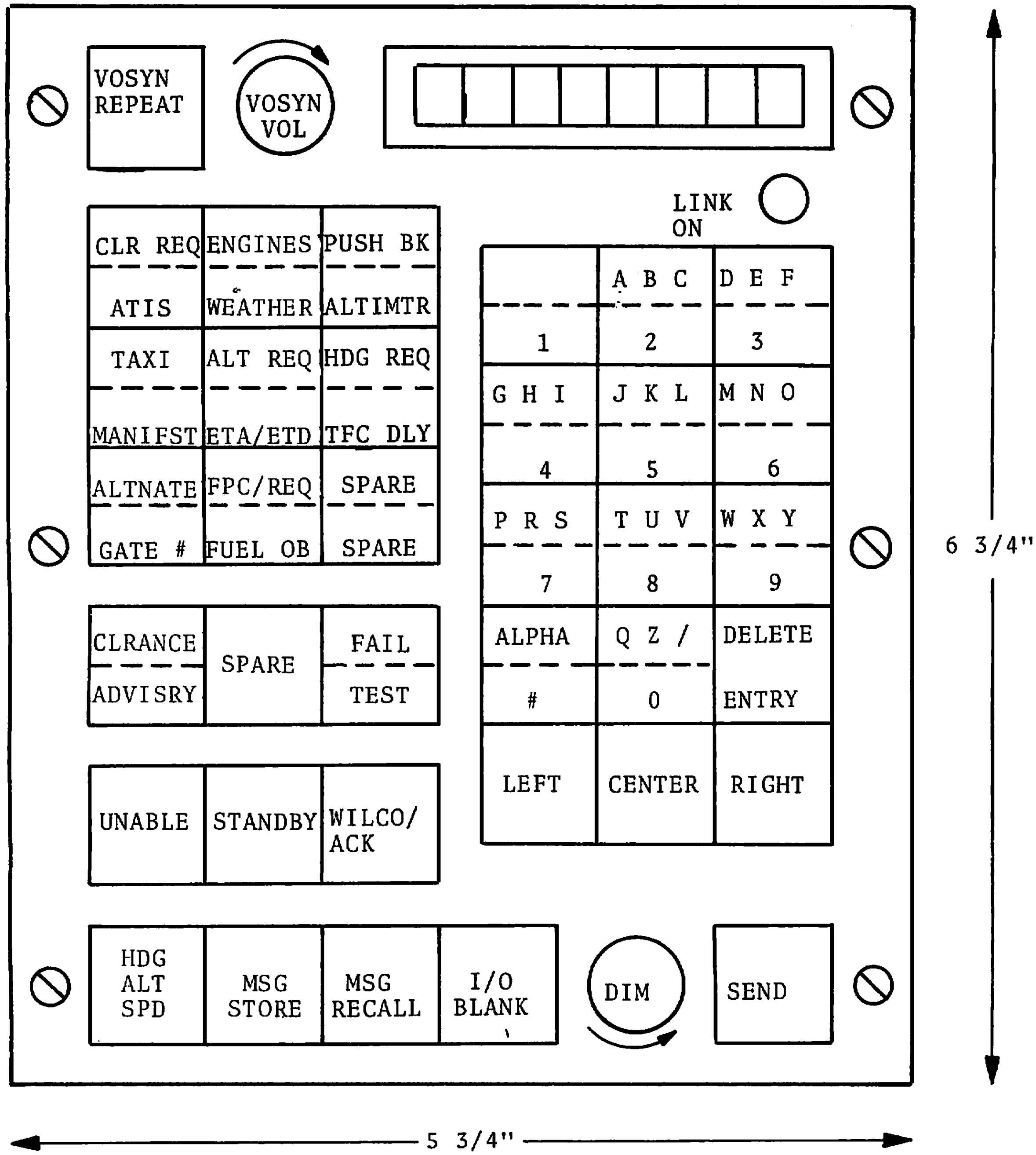

Figure 4-8. Full CDU Panel

f. "Message Store" and "Message Recall" buttons for retaining and later recalling any single message via the SMATC or Vosyn.

g. "I/O Blank" button for blanking the SMATC. This was an alternate action button: a second depression returned the display to view.

h. "Send" button for dispatching a downlink message.

i. A "Link On" indicator to simulate active or inactive polling situations.

j. An eight-window scratchpad display for reviewing downlink messages before depressing the "Send" button.

k. A "Fail/Test" button/indicator for simulating an I/O system self-test function.

Buttons were color coded to indicate functional groupings. Operation of the CDU has been described fully in Interim Reports Nos. 5 and 6. In particular, it should be noted that many of the buttons provided two functions, and only the function that was enambled at a particular time was illuminated and readable. Figure 4-8 indicates the configuration that was used for the GAT-1 tests. For the airline simulator tests, buttons marked "Spare" in this figure provided company business functions.

The CDU was mounted in the center pedestal between the two seats; Figure 4-9 indicates the location in the B-727 simulator. It should be noted here that in this location it was not accessible to the Second Officer. The WILCO button on the CDU was supplemented by additional WILCO buttons mounted on the control columns of the GAT-2 and DC-9 simulators, and by pendant switches in the B-727 simulator.

4.1.5 The SCDU (Simplified CDU)

In planning for the Phase 1-C tests, it was felt that the complexity of the full CDU, coupled with the short training time available to the experimental subjects, might cause certain operational difficulties, and for this reason, a simplified CDU (Figure

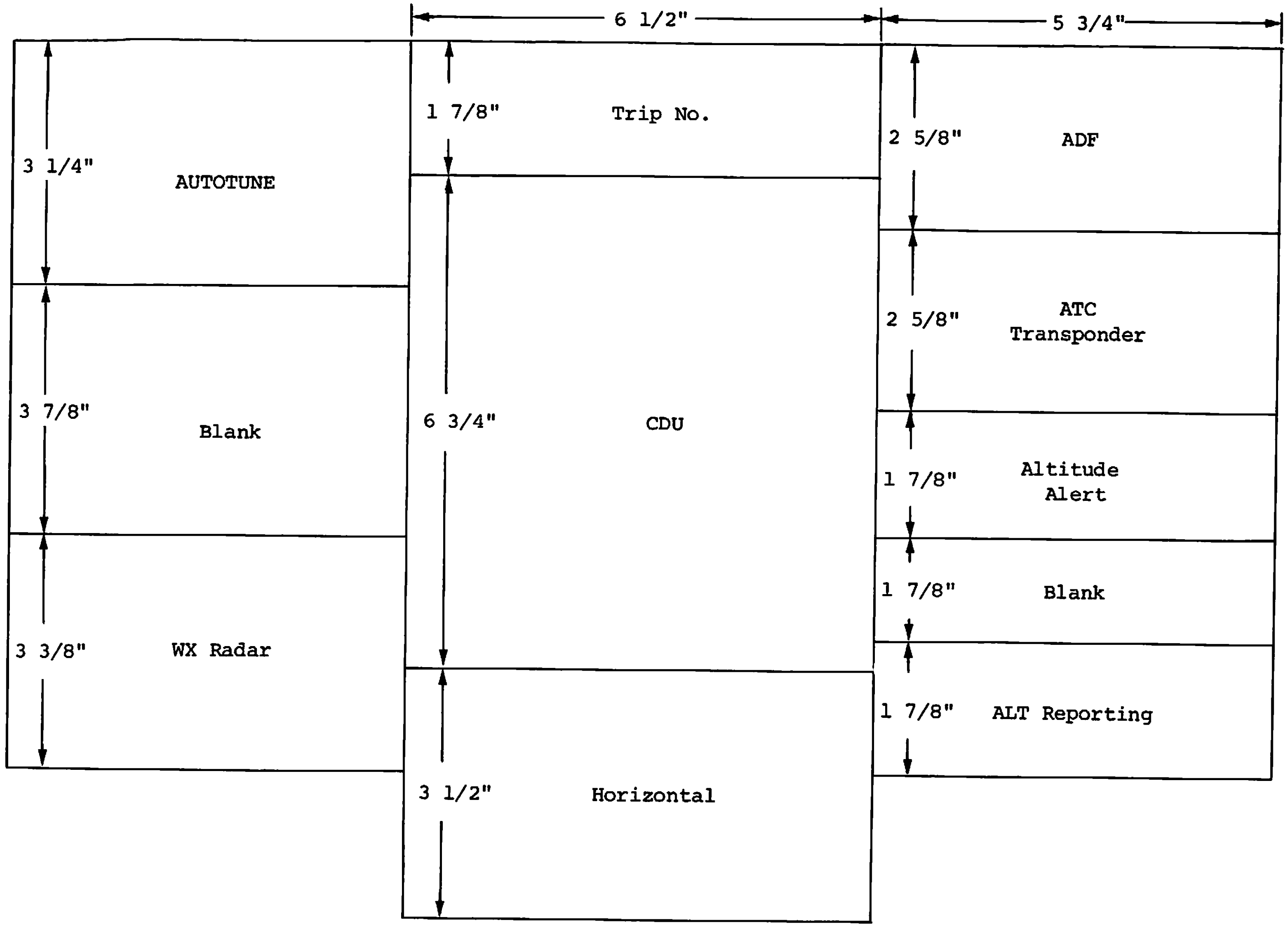

Figure 4-9. B-727 Forward Pedestal

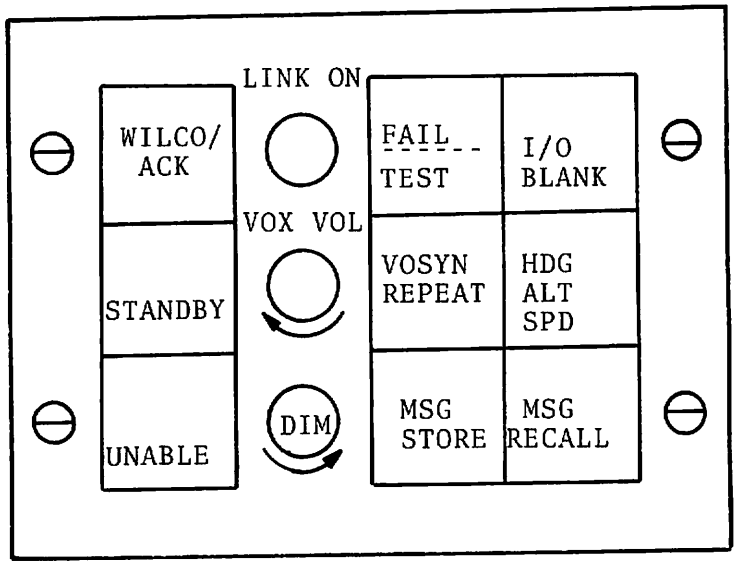

Figure 4-10. Simplified CDU Panel

4-10) was used on certain of the Phase 1-C tests. The SCDU contained only the buttons and controls necessary for operating the cockpit I/O system, and for generating WILCO, UNABLE, and STANDBY downlink responses. It did not provide the capability for generating other downlink messages. Specifically, it included functions c, d, e, f, g, i, and k described for the full CDU.

4.1 6 AUTOTUNE Device (Automatic Frequency Selector and AUTOTUNE Indicator)

For the Phase 1-D tests using airline simulators, one additional device was present, namely AUTOTUNE. The AUTOTUNE device, illustrated in Figure 4-11, simulated the automatic tuning, via Data Link, of the voice and data transceivers. Two displays at the top of the unit displayed the current frequencies of the Data Link and voice radios, respectively, with a resolution of 25 KHZ. The lower single element was a scratchpad for pilot entry. These display elements were visible in direct sunlight. Two selector knobs were provided. The outer portion of the left-hand knob provided manual

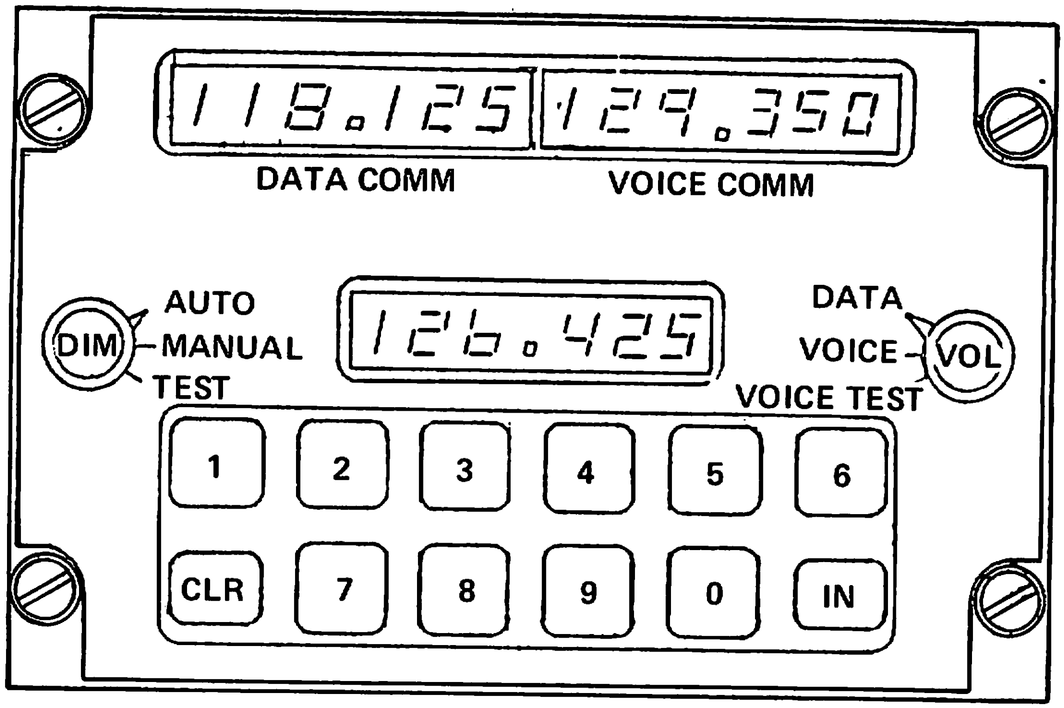

Figure 4-11. AUTOTUNE Indicator Panel

or automatic frequency selection, and the center portion provided a dimming control. The outer portion of the right-hand knob allowed selection of data or voice radio tuning for manual entry, and the center portion provided the volume control for the voice radio.

In the automatic mode, the frequencies were automatically inserted by an uplink message from the computer, after they had first been presented on the SMATC and/or Vosyn and had been WILCO'd by a flight crew member. The autotune location is indicated in Figure 4-9.

4.1.7 Minicomputer System

The minicomputer system for control of the output devices and for recording data concerning the use of the input devices consisted of (1) a TI 950A minicomputer with 16K of memory, and (2) a data terminal that included a standard teletype keyboard, an 80 character-per-line silent printer, and a twin magnetic tape cassette read/write unit. Each scenario/device complement was programmed into a single cassette; the other cassette position was used for data collection. The printer allowed the operator to observe the progress of the scenario as each message was printed out. The keyboard gave the operator control over the system operation, enabling him to dispatch each message in sequence, skip or go back one or more messages, add messages, and initialize various parameters prior to a given experimental run. The computer was mounted in an equipment rack outside the simulator, along with the Vosyn (and in the case of the B-727 simulator, a separate CRT and keyboard for communication with the Second Officer), as indicated in Figure 4-12.

4.1.8 Ancillary Equipment

Differences in simulator configurations and in the amount of information available to the console operator necessitated certain modifications of the manner in which equipment was interconnected for different experiments (as well as changing the role of the

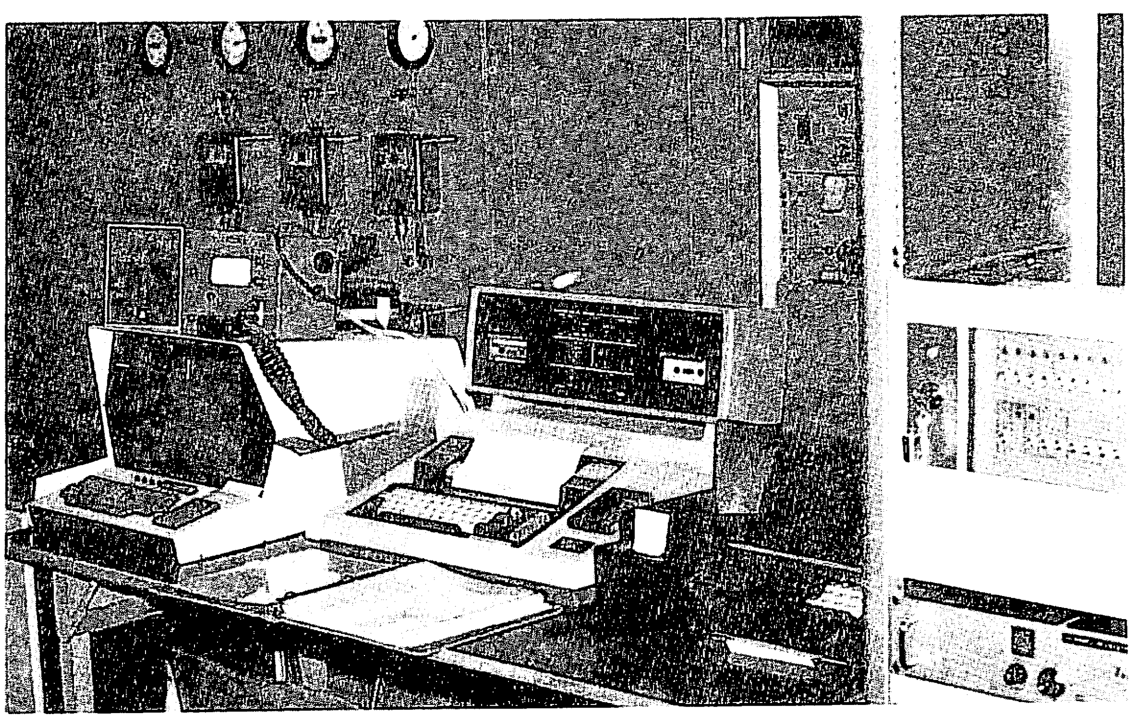

Figure 4-12. Test Equipment Configuration, B-727 Simulator

on-board observer). With the GAT-2, sufficient information from on-board instrumentation was duplicated at the console operator's station that he knew aircraft location and altitude at all times, and thus knew the appropriate time for dispatching the next ATC message. With the B-727 simulator, the console operator required co-ordination with both the on-board observer and with a remotely located radio aids operator in order to determine the proper time for message dispatch. On the DC-9 simulator, a plotting board visible to the console operator provided him with a rough indication of aircraft location during the simulated flights, but co-ordination with the on-board observer was required when messages requiring precise timing were to be dispatched.

For the runs on the B-727 simulator, the Vosyn output was introduced into the regular aircraft audio system; with the DC-9 and GAT-2 simulators, separate speakers were used in the cockpit for Vosyn output. No out-the-window display was available with the GAT-2; such displays were available on the B-727 and DC-9 simulators, but since they were time shared with other simulators, they were available only during takeoff and landing phases of flight.

The presence of a Second Officer during the B-727 runs necessitated further modification of system interconnections. While Figure 4-13 depicts a generalized overall simulator layout, it obviously cannot include all of the above-listed minor modifications.

4.2 SIMULATOR TEST FACILITIES

The GAT-2 simulator, located at FAA/NAFEC, Atlantic City, N.J., consists of a cockpit, an operator's station with control and display panel, strip chart recording equipment, and an X-Y plotter. The GAT-2 is a two-man cockpit simulator which simulates a light twin piston-engine, retractable gear aircraft. It is fully instrumented, including dual nav/com, ADF, DME, glide slope, and transponder. A third seat is provided for an observer (or instructor), and the communications channels provide a simulated radio link with an outside "controller/operator." The GAT-2 has

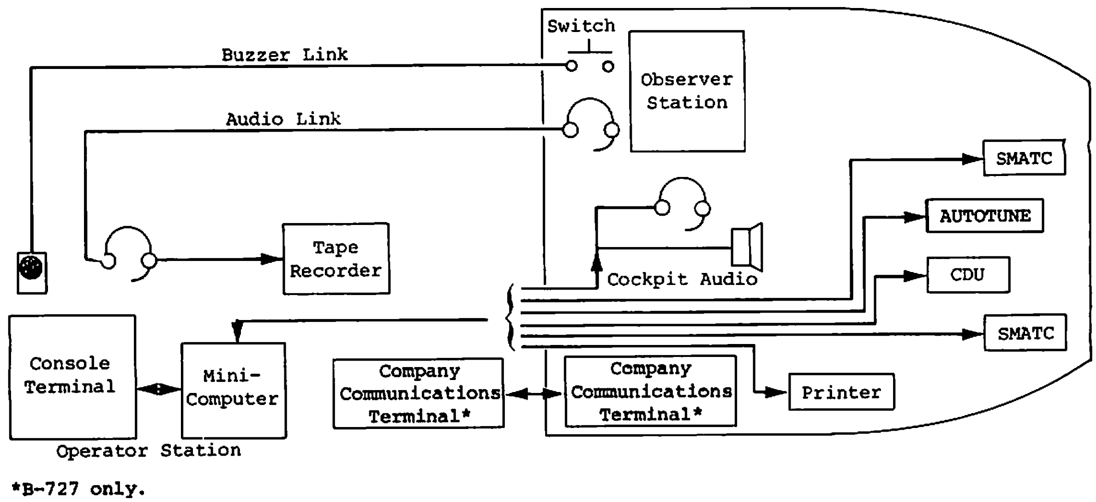

Figure 4-13. Overall Simulator Layout

limited pitch and roll motion, but is fixed with respect to the yaw axis. The simulator is set up to "fly" anywhere within a 12,500 square mile area in the vicinity of Atlantic City.

The outside operator's station is primarily a panel containing all the necessary controls and displays for simulating nearly any flight conditions. This includes controls for wind velocity and direction, turbulence, oil temperature and pressure, cylinder head temperature, barometric pressure, field elevation, and fuel remaining. Displays are provided for heading, altitude, air speed, turn coordinator, rate of climb, omni/loc/glide slope, nav/com frequencies,and localizer deviation. An X-Y recorder with chart scaled to the area over which the simulator can "fly" permits a quick look at any major deviations from the planned flight path.

The B-727 simulator used during the experiments is located at the United Airlines Training Center at Denver, Colorado. Cockpit configuration and system operation are identical in all significant respects to the B-727-222 aircraft in the UA fleet. The fuselage contains an instructor's station from which environmental conditions can be set or varied. The simulator incorporates a hydraulically powered three-axis motion system.

The simulator's visual system utilizes color television projected on a large screen in front of the cockpit. The image is developed from a scaled 10 by 5 mile earth model surrounding an instrument landing system (ILS) runway. Environmental controls for the visual model include visibility variable from 0 to 9-3/4 miles, ceiling variable from 0 to 2,000 feet, time-of-day variation, and approach lighting control. During the experimental runs on the simulator, the ability to vary surface and altitude winds was a major factor in keeping the scenarios on schedule.

The DC-9 simulator used during the experiments is located at the TWA Training Facility in Kansas City, Mo. It is a facsimile of the DC-9 aircraft flight compartment, containing exact replicas of the DC-9 crew stations and duplicating all portions of the external cockpit structure normally visible to the captain and first officer. An instructor's station is located aft of the captain's station.

A visual-display monochromatic closed-circuit television system is installed on the DC-9 simulator. Ceiling, visibility, and lighting conditions are variable under the control of the instructor. The cockpit provides independent motion in three degrees of freedom.

4.3 EXPERIMENTAL DESIGNS

The first tests using the previously described equipment were scheduled for the GAT-2 simulator at NAFEC, to be followed by the tests on airline simulators. Since the airline simulator tests, by definition, practically required simulation of regularly scheduled flights in commercial airline operations, it was determined that the GAT-2 tests should study operations by a somewhat different pilot population, namely the instrument-rated pilot who might fly for his own personal convenience or pleasure, or who might be engaged in charter operations.

With the available I/O devices, a number of different combinations were possible; from these, seven were selected for the differences in capability which they provided for the flight crew, and which also might represent the equipment selections of some future cost-conscious user. Additionally, since some of the volunteer pilots might have limitied familiarity with the flying qualities of the GAT-2 simulator, it was decided that one experimental run should use existing voice-only procedures in order to provide baseline data. The complements selected for study were identified by the letters "A" through "H" as follows:

A: Voice only

B: Vosyn (short messages), Printer (long messages), SCDU

C: Vosyn (all messages), SMATC (Hdg., Alt. & Speed Info only), SCDU

D: Vosyn (long messages), SMATC, Full CDU

E: SMATC, SCDU (Voice for long messages)

F: Vosyn (all messages), Full CDU

G: SMATC, Vosyn (all messages), Printer (all messages), Full CDU

H: SMATC, Printer (all messages), Full CDU

With eight experimental conditions to be evaluated, the ideal experimental design would have involved the use of eight crews and eight different scenarios, flown in a counterbalanced order. Because of the inpracticality of generating 8 scenarios and programing 64 different control tapes, it was decided to use only 4 different scenarios, to pair each scenario with 2 different device complements, so that only 8 program tapes were required, and to have different crews fly the complements in different order and with variations in day versus night conditions. The expermental design used is presented in Table 4-2, with "I," "J," "K," and "L" representing the four scenarios, and the final "D" or "N" indicating day or night conditions, respectively.

Anticipated difficulties in the use of the full CDU by crews having only brief training in its use failed to materialize, and it was therefore possible to eliminate the small CDU from

TABLE 4-2. EXPERIMENTAL DESIGN FOR GAT-2 TESTS

RUN	CREW							
	1	2	3	4	5	6	7	8
1	AID	BJN	CKD	DLN	ELN	FKD	GJN	HID
2	BJN	AID	DLN	CKD	GJD	HIN	ELD	FKN
3	CKN	DLD	AIN	BJD	HID	GJN	FKD	ELN
4	DLD	CKN	BJD	AIN	FKN	ELD	HIN	GJD
5	GJN	HID	ELN	FKD	BJD	AIN	DLD	CKN
6	HIN	GJD	FKN	ELD	CKD	DLN	AID	BJN
7	FKD	ELN	HID	GJN	DLN	CKD	BJN	AID
8	ELD	FKN	GJD	HIN	AIN	BJD	CKN	DLD

consideration for the tests on airline simulators. Additionally, the cost of owning, operating and maintaining a well equiped airline simulator often approaches several hundred dollars per hour, such that the major constraint in developing an adequate design for these later tests was an economic one, and the elimination of further testing of one device, the SCDU, provided some relief in program cost.

An Analysis of Variance of the anticipated experimental design for the airline simulators indicated that, within the program constraints, two simulators, three scenarios,and three device complements could be adequately evaluated. Two main criteria were taken into consideration in designing this evaluation:

1. Interaction between causal factors (simulators, scenarios, complements, crews, turbulence conditions, light conditions and practice effects).

2. Variations in measures due to factors not controlled in the experiment, such as crew background.

The first criterion was handled by obtaining responses while varying all of the factors simultaneously rather than varying only one factor at a time; the second was attained by randomizing the factor combinations to the experimental units (i.e., individual crew sorties).

Nine aircrews were tested at each of the two simulator sites and three scenarios were employed. The three complements combined with the three scenarios yielded nine distinct combinations. Each crew flew one four-hour test session consisting of three flights. This provided a total of 27 flights per simulator, or 54 overall. Each flight experimental run in Table 4-3 was made under smooth or turbulent conditions, and under day or night conditions. This yielded 36 possibilities. To accommodate this number with nine crews in two blocks (simulators), a fractional factorial design was used in which the occurrence of complement and scenario was completely balanced with the crews, the turbulence, and the daylight conditions.

TABLE 4-3. CREWS BY TEST CONDITIONS: AIRLINE SIMULATORS

Scenario and Condition		SMATC		VOSYN		SMATC/VOSYN	
		Day	Night	Day	Night	Day	Night
I	Smooth	5		8	3	6	9
	Turbulent	7	1	4			2
II	Smooth	9	6	1	7	3	
	Turbulent	2			5	8	4
III	Smooth	4	8		2	7	1
	Turbulent		3	6	9	5	

Note: The terms SMATC, VOSYN and SMATC/VOSYN indicate the distinctive characteristics of each complement. All complements also included CDU, printer, and Autotune, and on the B-727, a separate keyboard and scratchpad display for the Second Officer

The basic experimental unit for this evaluation was a flight on a particular simulator using a particular complement with a particular daylight and turbulence condition during a particular scenario. Although, normally, all the factor combinations used in the experiment were randomized on the experimental units, in this situation there was a restriction resulting from each crew having to fly three flights sequentially during a single experimental session.

Table 4-3 shows, for each of the simulators, the combinations of conditions each of the crews experienced. As the crews became available, they were assigned to one of the crew numbers in Table 4-4. The order of missions for each of the nine crews in Table 4-3 was randomly selected.

TABLE 4-4. ORDER OF MISSIONS FLOWN: AIRLINE SIMULATORS

	Scenario (S)								
Crew	SMATC			VOSYN			SMATC/VOSYN		
	S-I	S-II	S-III	S-I	S-II	S-III	S-I	S-II	S-III
1	2				3				1
2		2				1	2		
3			2	1				3	
4			3	2				1	
5	3				1				2
6		1				2	3		
7	2				1				3
8			1	3				2	
9		3				2	1		
S-I: SFO-LAX; S-II: MCI-ORD; S-III: LAX-SFO									

Table 4-4 was restructured into five two-factor tables, as presented in Table 4-5. Only in Table 4-5 (a) is there represented a full-factorial (i.e., all factor combinations tested) design with replications, so that the differences among crews can be tested, as well as all the interactions among crews, complements, and scenarios.

Appropriate transformations on the data were made in order to attain variance stabilization and normalization, which are the necessary assumptions in conducting the desired analysis of variance. This analysis is described in Section 5.2.

4.4 THE SCENARIOS

As indicated previously, the scenarios used during the GAT-2 testing were designed to simulate private or small charter operations, while those used on the airline simulators involved typical commercial flights. The several scenarios are summarized below. The complete scenarios, which include the precise format used for each message on each device, are available in the appendixes of Interim Reports Nos. 5 and 6 of this series.

4.4.1 Scenarios Used During GAT-2 Tests

For the voice-only condition with Scenario "I," the simulated responses from pilots of other aircraft were prerecorded on tape using a number of speakers, such that the experimental subject heard not only the "controller" speak to him, but also to other "aircraft," and heard the replies from the "pilots" or these other aircraft. Scenario "I" presented the commands and advisories for a departure from Runway 1 at Greater Wilmington, Delaware, airport via Newcastle to a landing on Runway 9-Right at Philadelphia International, followed by a departure on Runway 9-Left and via Pottstown 1 departure and East Texas to an eventual landing on Runway 6 at Allentown.

Scenario "J" presented the commands and advisories for a departure from Trenton-Robbinsville (an uncontrolled airport with simulated clearance via land line) to Philadelphia International

TABLE 4-5. CREWS BY FACTOR COMBINATIONS: TWO FACTORS CONSIDERED AT A TIME (NUMBERS LISTED ARE CREW MEMBERS) AIRLINE SIMULATORS

	Scenario by Complement			
	Scenario	SMATC	VOSYN	SMATC/VOSYN
(a)	I II III	1, 5, 7 2, 6, 9 3, 4, 8	3, 4, 8 1, 5, 7 2, 6, 9	2, 6, 9 3, 4, 8 1, 5, 7
	Light Condition by Complement			
	Light Condition	SMATC	VOSYN	SMATC/VOSYN
(b)	Day Night	2, 4, 5, 7, 9 1, 3, 6, 8	1, 4, 6, 8 3, 5, 7, 9, 2	3, 5, 6, 7, 8 1, 2, 4, 9
	Turbulence by Complement			
	Turbulence	SMATC	VOSYN	SMATC/VOSYN
(c)	Smooth Turbulent	4, 5, 6, 8, 9 1, 2, 3, 7	1, 2, 3, 7, 8 4, 5, 6, 9	1, 3, 6, 7, 9 2, 4, 5, 8
	Light Condition by Scenario			
	Light Condition	Scenario I	Scenario II	Scenario III
(d)	Day Night	4, 5, 6, 7, 8 1, 2, 3, 9	1, 3, 8, 9 2, 4, 5, 6, 7	4, 5, 6, 7 1, 2, 3, 8, 9
	Turbulence by Scenario			
	Turbulence	Scenario I	Scenario II	Scenario III
(e)	Smooth Turbulent	3, 5, 6, 9, 9 1, 2, 4, 7	1, 6, 3, 7, 9 2, 4, 5, 8	1, 2, 4, 8 3, 5, 6, 7, 9
	Turbulence by Light Condition			
	Turbulance	Day	Night	
(f)	Smooth Turbulent	1, 3, 4, 5, 6, 8, 9 2, 4, 5, 6, 7, 8	1, 3, 6, 7, 8, 9 1, 2, 3, 4, 5, 9	

via V157. After landing on Runway 27-Right and departure on Runway 27-Left, the flight continued, with a landing on Runway 32 at Greater Wilmington, Delaware.

Scenario "K" involved a simulated flight from Teterboro, N.J., to North Philadelphia. Simulated turbulence and thunderstorm activity in the vicinity of Trenton, N.J., required flight crews to request heading and altitude changes.

Scenario "L" provided the commands and advisories for a simulated flight from Cape May County Airport, N.J. (another uncontrolled airport) to Wings, Pa. Simulated bad weather conditions at Wings forced a change to an alternate destination, Mercer County Airport.

During all of the simulated flights on which a full CDU was available, flight crews had access to weather information from Philadelphia, Allentown, Wilmington, Teterboro, North Philadelphia, Trenton, Atlantic City, Millville, Newark, Reading, and Cape May County. ATIS information was available only from Philadelphia International.

4.4.2 Scenarios Used During the Airline Simulator Tests

The three scenarios, which were used at both simulator sites, were as follows:

Scenario I - San Francisco to Los Angeles
Scenario II - Kansas City to Chicago
Scenario III - Los Angeles to San Francisco

Each of these scenarios described a typical flight of about one hour ten minutes between two high-density terminals. Approximately 60 messages were contained in each scenario: roughly 50 were ATC messages; the remaining 10 were company messages. Situations that required the crew to use the CDU for downlink-message composition were provided. Each scenario was coded into three message scripts, one for each device complement. The use of at least one international scenario had been desired. However, the geographic detail stored in the simulator computer, as well as the lack of suitable navigation capabilities in the simulators, prevented this. Runway visual systems were used for all takeoffs and landings to simulate realism and note any distracting effects of the I/O devices.

Three varying degress of company automation were employed in the three scenarios to obtain qualitative pilot opinion on the acceptability of the varying amounts of button-pushing required in composing these messages. These three scenarios are illustrated in the following subsections.

4.4.2.1 Scenario I - San Francisco to Los Angeles - Scenario I involved a nonstop flight from San Francisco International Airport to Los Angeles International Airport. The crew initially requested Automatic Terminal Information Services (ATIS) by using the CDU. Next they requested clearance. Prior to takeoff, but after gate departure, the flight was advised of a 10-minute delay and asked to forward a delay report to the company. The flight took off, was vectored through its departure procedure, and was instructed to climb to altitude. (The altitudes were somewhat different for the two simulators, consistent with the altitudes at which the aircraft would actually file.) An impossible altitude command was given to the crew in order to require the use of the UNABLE button on the CDU. Following the resolution of the difficulty on the voice channel, the flight was instructed to proceed to its proper altitude.

Clear Air Turbulence (CAT) was encountered en route, forcing the crew to request a change to a smoother altitude. The flight was cleared to a smoother altitude and eventually descended for its approach into Los Angeles. The flight was radar-vectored for an instrument approach at Los Angeles. The assigned approach differed from that mentioned in the ATIS information. The aircraft landed and taxied to the gate. The flight concluded with the crew filing its arrival report with the company.

A certain degree of automation was assumed for company messages in this scenario. It was assumed, for the purpose of Gate Departure, Airborne, Landed, and Gate Arrival reports (respectively

OUT, OFF, ON, and IN ("OOOI")), that on-board flight devices automatically inserted the flight number and applicable station identifiers. The crew entered only the proper function (e.g., DEPARTED) and the applicable time and fuel-weight digits, and then depressed the SEND button.

4.4.2.2 <u>Scenario II - Kansas City to Chicago</u> - Scenario II involved a routine flight from Kansas City International Airport to O'Hare International Airport. As in Scenario I, the crew requested ATIS and ATC clearance via the CDU.

The aircraft took off, was radar-vectored to intercept the proper departure course, and climbed to its flight-planned altitude. An impossible transponder code was given to the flight in order to require the crew to use the UNABLE button on the CDU. Once this situation was resolved on the voice channel, the flight was given the correct transponder setting via the Data Link. The flight was instructed to circle and hold prior to entering the Chicago terminal area.

Full automation of company "OOOI" messages was assumed in this scenario. The crew pressed only the applicable function button, followed by the SEND button. On-board sensors and devices caused the flight number, station identifier, applicable time, and fuel-weight digits to be entered.

4.4.2.3 <u>Scenario III - Los Angeles to San Francisco</u> - Scenario III involved a flight from Los Angeles International Airport to San Francisco International Airport. As in Scenario II, this scenario began with the crew requesting ATIS and Clearance via the CDU. The flight was vectored through a departure procedure to its flight-planned route.

The flight proceeded normally through the departure, en route, and initial arrival stages. An unreadable heading command was given to the flight in order to force the use of the UNABLE button. After this situation was resolved on the voice channel, the correct heading was given. On final approach, the flight was told to go around because of a disabled aircraft on the runway. The flight was radar-vectored for another instrument approach and landed. This scenario was concluded with the crew filing its arrival report.

Only the flight numbers were automated in the company messages of this scenario. The crew entered station identifiers, times, and fuel-weight digits for the "000I" messages.

4.4.2.4 Downlink Request Messages - There were two conditions under which crews were instructed to generate downlink request messages via the CDU. The first occurred when they wanted to obtain a clearance, taxi instructions, ATIS, weather, or altimeter setting, or to send company-type messages. The second occurred in some scenarios at predetermined points. At the proper time, the on-board observer informed the crew that a certain situation existed. The observer instructed the crew on the action to be taken in terms of a downlink request, which was predetermined. The downlink messages that fall into this category are "ALT Request," "HDG Request," and "Alternate Airport." A crew member would key in the required downlink request, for which the uplink response was the next sequential message of the scenario.

4.5 EXPERIMENTAL SUBJECTS

For the GAT-2 tests, 17 general aviation pilots and 8 FAA/NAFEC professional test pilots participated. Mean total flying hours for the general aviation pilots was 5300 hours (range 250 to 15,000 hours) For this group, the mean hours of instrument flying was 1020 (range 50 to 6000 hours), and mean multi-engine experience was 2965 hours (range 20 to 13,000 hours).

On the average, the test pilots were somewhat more experienced. Their mean flying time was 9050 hours (range 3200 to 14,000 hours), their instrument flying experience had a mean of 1420 hours (range 187 to 3,000 hours), and their multi-engine experience had a mean of 4900 hours (range 1650 to 8000 hours). It was necessary to use instrument-rated general aviation pilots for this experiment since non-instrument-rated pilots would have had difficulty in

4.6 EXPERIMENTAL PROCEDURES

For the GAT-2 tests, the experiment as originally designed required that eight crews each fly eight times. During early planning, it was assumed that this eight-by-eight experimental design would be flown by four two-man crews and four one-man crews of general aviation pilots, and would be replicated by eight two-man crews of FAA/NAFEC test pilots. It was further assumed that a crew would fly four runs on one day and the remaining four runs on a second day.

Limitations of available simulator time forced a revision in these plans. Consequently, even though it was known that a training session and eight experimental runs with computer reprogramming between runs could not be accomplished in less than a 10-hour day, it was found necessary for crews to fly all eight runs during a single day. It was also found necessary to limit the number of test pilot crews to four. With 14 days of simulator time available, this left only 2 extra days for contingencies.

When subjects reported for participation, they were first given the opportunity to check out the flying qualities of the simulator and to familiarize themselves with the layout of the flight instruments. They then practiced with a brief scenario (also reproduced in an appendix to Interim Report No. 5) which demonstrated the capabilities of the several Data link devices, with emphasis on the use of the CDU for the generation of downlink messages. During this practice scenario, they were not required to fly the simulator. The cockpit observer acted as instructor during this period, answering any questions which arose.

Also prior to the beginning of the experimental runs, pilots were given a copy of an 83-item questionnaire, so that by knowing what questions would be asked after their experimental flights, they would be in a better position to observe during their experimental flights. Most of the pilots completed the questionnaire within 24 hours of the completion of their experimental runs.

flying the simulator under the procedures employed during the experiment.

For the tests using the airline simulators, it was desired to obtain the widest possible cross section of qualified pilots from the air carrier industry. The primary element of this participation was to be the professional pilots and flight instructors on the staffs of TW's and UA's training centers, so as to minimize scheduling conflicts. Additionally, it was desired to obtain volunteer pilot participation from interested user groups in the industry, as well as FAA test pilot participation.

The number of volunteers surpassed all expectations, undoubtedly due to the written request made by the ATA to all member airlines. Pilots currently flying for or on the staffs of the following organizations participated in the evaluation: Airline Pilots Association, American Airlines, Boeing Company, Braniff International, Continental Airlines, FAA/NAFEC, Southern Airways, Northwest Airlines, Ozark Airlines, Pan American World Airways, Society of Automotive Engineers (S-7 Committee), Trans World Airlines, United Airlines, and Western Airlines. The participants could be classified as Line Pilots, Management Pilots (Flight Instructors and Flight Managers), or Engineering and Test Pilots. Table 4-6 indicates certain characteristics of the pilot subjects.

TABLE 4-6. CREW CHARACTERISTICS FOR AIRLINE SIMULATOR TESTS

Simulator	Average Hours	Average Hours in Aircraft	Airline Transport Rated (in Aircraft)	Line Pilots	Management Pilots	Test Pilots
DC-9	9600	1530	17 (11)	6	8	4
B-727	11600	1900	19 (14)	5	16	6

flying the simulator under the procedures employed during the experiment.

For the tests using the airline simulators, it was desired to obtain the widest possible cross section of qualified pilots from the air carrier industry. The primary element of this participation was to be the professional pilots and flight instructors on the staffs of TW's and UA's training centers, so as to minimize scheduling conflicts. Additionally, it was desired to obtain volunteer pilot participation from interested user groups in the industry, as well as FAA test pilot participation.

The number of volunteers surpassed all expectations, undoubtedly due to the written request made by the ATA to all member airlines. Pilots currently flying for or on the staffs of the following organizations participated in the evaluation: Airline Pilots Association, American Airlines, Boeing Company, Braniff International, Continental Airlines, FAA/NAFEC, Southern Airways, Northwest Airlines, Ozark Airlines, Pan American World Airways, Society of Automotive Engineers (S-7 Committee), Trans World Airlines, United Airlines, and Western Airlines. The participants could be classified as Line Pilots, Management Pilots (Flight Instructors and Flight Managers), or Engineering and Test Pilots. Table 4-6 indicates certain characteristics of the pilot subjects.

TABLE 4-6. CREW CHARACTERISTICS FOR AIRLINE SIMULATOR TESTS

Simulator	Average Hours	Average Hours in Aircraft	Airline Transport Rated (in Aircraft)	Line Pilots	Management Pilots	Test Pilots
DC-9	9600	1530	17 (11)	6	8	4
B-727	11600	1900	19 (14)	5	16	6

4.6 EXPERIMENTAL PROCEDURES

For the GAT-2 tests, the experiment as originally designed required that eight crews each fly eight times. During early planning, it was assumed that this eight-by-eight experimental design would be flown by four two-man crews and four one-man crews of general aviation pilots, and would be replicated by eight two-man crews of FAA/NAFEC test pilots. It was further assumed that a crew would fly four runs on one day and the remaining four runs on a second day.

Limitations of available simulator time forced a revision in these plans. Consequently, even though it was known that a training session and eight experimental runs with computer reprogramming between runs could not be accomplished in less than a 10-hour day, it was found necessary for crews to fly all eight runs during a single day. It was also found necessary to limit the number of test pilot crews to four. With 14 days of simulator time available, this left only 2 extra days for contingencies.

When subjects reported for participation, they were first given the opportunity to check out the flying qualities of the simulator and to familiarize themselves with the layout of the flight instruments. They then practiced with a brief scenario (also reproduced in an appendix to Interim Report No. 5) which demonstrated the capabilities of the several Data link devices, with emphasis on the use of the CDU for the generation of downlink messages. During this practice scenario, they were not required to fly the simulator. The cockpit observer acted as instructor during this period, answering any questions which arose.

Also prior to the beginning of the experimental runs, pilots were given a copy of an 83-item questionnaire, so that by knowing what questions would be asked after their experimental flights, they would be in a better position to observe during their experimental flights. Most of the pilots completed the questionnaire within 24 hours of the completion of their experimental runs.

For the tests on the B-727 and DC-9 simulators, similar procedures were employed. However, since the crews were nearly all ATC-rated and flight-qualified for the aircraft which they were to fly, it was felt that the time used on the GAT-2 tests to allow pilots to familiarize themselves with the flying qualities of the aircraft could now better be utilized to extend the briefing on the use of the Data Link equipment as an adjunct to the written instructions which the crews had previously received by mail. In the majority of cases, the person delivering the briefing also later acted as Test Observer (at the instructor's station in the simulator), since the lack of remote instruments at the Test Operator's station made it difficult for him to judge the progress of the flight and to reach decisions as to the proper time for message dispatch without co-ordinating with the Test Observer in the cockpit.

At the completion of the three flights that composed the B-727 and DC-9 tests, crews were immediatedly given a post-briefing session and filled in a questionnaire. Based to some extent on the findings from the GAT-2 tests and the use of a different category of flying personnel, it was possible to reduce the number of questions asked of the airline simulator crews to 45 from the original 83 items.

5. RESULTS OF FULL SCALE DATA LINK TESTS

When an automated information-transfer system such as Data Link communications is integrated into a system in which critical life-protecting decisions are dependent on the quality of the man-machine interface, the subjective evaluations of the operators may be as crucial as the technical measurements of the system. The use of simulators in these experiments was aimed directly at obtaining both qualitative and quantitative data using various complements of Data Link communications equipment. Since the personnel utilized in the GAT-2 tests were representative of pilots and crews who might or might not have varying degrees of financial constraint in their choice of equipment for Data Link, seven different device complements were selected as representative of what might be available in private or charter aircraft. For the airline simulator tests, this was reduced to three, all of which provided full Data Link capability. This difference, again, permitted a reduction in the number of questions asked during the post-briefing session.

5.1 QUALITATIVE DATA

Subjective or qualitative data were derived from three sources during these evaluations:

1. Debriefing questionnaires and discussions
2. Observer comments
3. A supplemental questionnaire (in the case of the airline simulator tests)

Because of personnel and equipment differences between the GAT-2 tests and those run on the airline simulators, the qualitative results are first discussed separately.

5.1.1 Qualitative Data from GAT-2 Tests

During the planning of the GAT-2 tests, TSC staff members had certain misgivings as to pilot acceptability of down-link keying requirements other than simple WILCO or UNABLE responses. The test results dispelled these doubts. Twenty-three participants found it reasonably easy to use with limited training, and only one found it marginal. One pilot stated, "If controllers can learn to use keyboards, so can pilots."

In general, pilots felt that two to three hours of training made them feel completely comfortable in the use of the CDU. Several listed the required time as 15 minutes, and only one pilot suggested "several days." The requirement for left handed keying by the co-pilot proved difficult to only one; however, it should be noted that 7 of the 25 pilots who submitted questionnaires listed themselves as either left handed or ambidextrous, which is a somewhat higher proportion than might be expected in a normal population.

On the question, "Should downlink requirements other than WILCO and UNABLE be eliminated?" 15 replied "definitely not" 4 "probably not," and 4 were unsure. This again is evidence of the acceptability of downlink keying.

Pilots suggested the addition of six discrete function buttons, but none of these were suggested by more than any one pilot. On the other hand, when "EMERGENCY" was suggested as an addition, 13 replied "yes," as did 10 to a TRAFFIC LOCATED button; but here there were also 14 "no" responses.

The button pressure used (2.5 pounds with 0.187 inches of travel), although appearing somewhat stiff to TSC personnel when the CDU was being fabricated, was deemed acceptable under all conditions by 19 of the pilots; 4 found it too high even in turbulence. Daylight readability was found acceptable to 24 of the test subjects, with 1 "no, marginal." Sixteen would like to use the keyboard for other entries, such as transponder code settings and radio frequency settings, and one suggested that a calculator mode should be added. The use of the existing system of split function buttons was preferred over the addition of separate buttons (with a concurrent increase in CDU size) by 23 of the pilots. The present split between "Clearance" and "Advisory" modes was again approved by 23 of the pilots and with no suggestions as to a better possible choice.

The yellow-white output of the lamps in the CDU buttons permitted only a limited choice of tints for color coding of functional areas. Thirteen pilots found that the tints aided them in recognizing functional groupings, while 10 pilots found them of no value.

Only 3 of the pilots felt that downlink keying requirements should be limited to those available on the SCDU; 10 felt restricted when, after having used the CDU, only the small CDU was available. Ten responded "Yes" to the question "Should we consider something midway in capability between the CDU and the SCDU?" but no suggestions were made as to what should be deleted or added.

Although yellow and green single-element light-emitting diodes (LED) are presently available, the 5 x 7 dot matrices required for the generation of alphanumerics can be obtained only with red outputs; these were used in the SMATC. Only 3 out of the 17 general avaiation pilots surveyed found the red color unacceptable, compared with 5 of the 8 test pilots. Color preferences varied widely: green (9), white (5), yellow (4), red (3), amber (2), orange (2), and blue (1).

Aside from the color of the characters, the SMATC was generally acceptable. Twenty-three of the respondants replied "yes" on readability and on maintaining the existing character size, and 20 agreed that the number of characters (16) should remain as is. A suggestion that the SMATC might be installed on the control columns was rejected by 19 of the pilots.

The Vosyn was also generally acceptable. No pilot listed its intelligibility as "unacceptable," and 15 stated that intelligibility improved with practice. No pilot listed the mechanical quality of the speech as "very annoying." Seventeen wanted the Vosyn to be used at all times, and only three favored that its usage be limited to certain flight phases when the pilot's eyes are busiest. Nineteen voted for in-cockpit control of pitch and rate of speech; during the experiment these controls were not available in the cockpit, but instead were set from the experimenter's console at levels preferred by the individual pilots during the practice run.

The major objection to the printer was its poor location (behind the pedestal); 14 of the pilots listed this location as poor or marginal. Ten wanted the printer usage limited to long messages, while 12 preferred it for all messages. Thirteen respondents approved of the lack of abbreviations or symbols, whereas several of the test pilots suggested the use of weather symbols. The time tag on the messages proved of value to only 10. Fifteen indicated the desirability of adding an "end of message" signal so that messages would not be torn off too soon. The question of the need for an auditory signal to indicate the appearance of a new message on the printer could not be answered authoritatively, since noise levels in the simulator did not approximate those in actual aircraft; therefore, the attention value of the noise of the printer in the operation could not be assessed. The use of two-color printing was found unnecessary by 23; 2 test pilots suggested the use of red for urgent messages.

The audio alert:

approximated the inflection of a voice saying "Data Link" twice. Twenty-three of the respondents found this sufficiently distinctive 21 preferred it over flashing of the SMATC or WILCO button as an alerting means. Nineteen pilots preferred the use of the audio alert over having the Vosyn speak the aircraft identification number.

A number of general questions were considered. It has been claimed that pilots using Data Link lose information when they cannot hear ATC transactions with other aircraft. During the present experiment, 14 pilots stated that no information was lost because of this absence, 10 felt that a little information was lost, and only 1 felt that <u>much</u> information was lost.

During earlier experiments, it was found that pilot responses were slower when information was provided both visually and via synthetic speech than when provided to either sensory modality alone. Pilots were questioned as to whether they felt any such tendency; 11 responded "yes" and 10 "no."

For recall of heading, altitude, and speed (HAS) messages, 20 of the pilots preferred use of the SMATC, 5 advocated a dedicated

HAS display, and only 1 preferred the use of the Vosyn. Fourteen felt that even with HAS information always available for recall, there was a need to store more than one other message. Autotune of radio frequency and transponder settings received 20 "Yes" or "worth trying" responses, and the question concerning the advisability of setting heading, altitude, and speed bugs automatically via Data Link received 19 "Yes" or "worth trying" responses.

To the question, "Should a WILCO be required from both crew members?" 19 responded "No." Seventeen felt that a WILCO response should not be required to the message "radar contact." Thirteen did not find the STANDBY useful, while eight did.

Providing redundant information was generally acceptable. Only four pilots felt that too much information was provided when information was presented on the Vosyn, SMATC, and printer simultaneously. Several suggested that the SMATC information might be presented via CRT, and that this might be used when an (Electronic Attitude Director Indicator) EADI CRT was present.

Pilots were asked to select what they considered an ideal device complement and a minimal device complement for various flight phases. Here, pronounced differences between the two pilot groups developed. General avaition pilots selected relatively constant equipment requirements across all flight phases, whereas test pilots wanted less Data Link equipment during takeoff and more enroute. General aviation pilots found some usage for all of the devices, except the Vosyn when it was restricted to long messages only; test pilots, on the other hand, typically selected a SMATC for short messages, a Vosyn and printer for all messages, and a full CDU. The data for a minimal device complement are somewhat meaningless, since several pilots listed only a small CDU and made no provision for any uplink capability.

Possible changes in crew workload are of major importance to the viability of Data Link. Only 2 pilots felt that Data Link would increase pilot workload, 2 felt that it would remain about the same, 16 felt that it would decrease workload, and 2 were unsure. Of equal importance is the confidence factor — the certainty of the pilot that he is receiving all messages intended for him and not reacting to messages intended for other aircraft. Sixteen pilots felt that their confidence would increase, one felt that there would be no effect, three indicated a decrease in confidence, and five were unsure.

One week after the completion of the experiment, eight of the pilots were interviewed in a debriefing session. A number of points were brought out. It should be emphasized that while these points represent the views of one or more of the pilot participants, they should not be construed as necessarily representing recommendations for future experiments.

1. Two of the test pilots stated that they had not seen the briefing sheet and were not aware of its existence until the debriefing session.

2. Several of the pilots stated that they were not aware that their responses were being timed, and thus felt no concern to make prompt responses.

3. Several of the pilots had heard the Univac voice response system and commented on the superiority of its output to that of the system used in the experiment. (The Univac system, however, is not airborne equipment.)

4. One test pilot stated that he had hearing deficiencies, and that regardless of the quality of the synthetic speech that might be presented, he would prefer a visual disply.

5. Pilots felt that a single SMATC was sufficient for the GAT-2 simulator, but felt that two might be required in a commercial jet, particularly the jumbo jets.

6. It was suggested that for the next experiment, the questionnaire should not be limited to discrete responses such as "yes" and "no," but instead should provide a five-point scale between limits.

7. Red, to pilots, indicates "danger." All possible colors should be investigated for the SMATC before a final decision is made. (As pointed out earlier, red is the only color currently available for alphanumeric LED.)

8. Several of the pilots lost some confidence in Data Link because of an apparent software problem in message storage.

9. Pilots would like to be able to turn off the Vosyn during certain flight phases if the same information could be obtained visually.

10. An attempt should be made to integrate the CDU with RNAV input devices, since RNAV will be mandatory above 18,000 feet after 1977.

11. It would be desirable to have a three-posiiton switch, so that pilots could have Vosyn, SMATC, or both.

12. Two pilots were unaware of the presence of the time stamp on the printer messages.

13. A question arose as to what a pilot should do when he had reached his clearance limit and no new Data Link message had appeared.

14. Test pilots would like weather symbology on the printer.

15. During experiments, certain impossible messages should be included to force an UNABLE response.

16. A RECONFIRM button should be added to the CDU.

17. Additions to the SCDU might include WEATHER and ATIS.

18. One pilot thought that the Data Link alert sounded somewhat like the 747 altitude alert.

19. One pilot thought that with a printer available, there was no requirement to store messages in the SMATC.

20. Adding autotune of VHF or transponder does not seem to infringe on the pilot's job.

21. It was suggested that SMATC messages should be presented on the Electronic Attitude Director Indicator so as to avoid a requirement for an extra display.

5.1.2 Qualitative Data from Tests on Airline Simulators

The tests on airline simulators were designed to provide something more than a mere replication of the GAT-2 tests. The scenarios contained the message content which might be anticipated on typical commercial airline flights, and the equipment allowed full uplink and downlink capability (as opposed to certain of the GAT-2 tests where only a SCDU was provided for downlink message). In general, the crews used had extensive line experience; the exception was the use of two FAA/NAFEC crews who were used as a means for obtaining program continuity. The airline simulators selected (B-727 and DC-9) were designed to provide information as to any differences which might result from the use of two- versus three-man crews. Finally, the autotune device, which was not available at the time of the GAT-2 tests, was added to the Data Link complements to provide further information. Because of these several differences from the previous tests, the questionnaire administered to the test subjects was extensively modified.

In the subsequent data analysis, a primary consideration was to determine those instances where response differences might be attributal to the use of two-man crews in the DC-9 simulator versus three-man crews in the B-727 simulator. Application of the Chi-Square test at the 0.05 significance level indicated that differences were present for only three of the questions. These differences are discussed first, since the remaining data may be considered as belonging to a single population. Two of these questions concerned the Vosyn and the third was related to the desirability of the use of abbreviations on the printer.

The Vosyn was found to be more acceptable in the DC-9 than in the B-727 simulator. The B-727 simulator's audio system approximates a worst case condition for voice communication, and because of an impedence mismatch, the introduction of the Vosyn output into this system produced an output which was virtually unintelligible. Consequently, for the DC-9 tests, a separate amplifier and speaker were used, and pilots now found the Vosyn output to be much more acceptable. Since the total cost of the amplifier and speaker used in the DC-9 tests was small, this improved acceptability cannot be attributed to the use of "high fidelity" equipment.

Concerning the matter of use of abbreviations and/or symbols on the printer, it was found that DC-9 crews were much less receptive than were the less overworked flight engineers in the B-727 simulator; this difference was probably further confounded by the fact that only the first officer, who must monitor printer output along with his other duties, had direct access to the printer in the DC-9.

Except for the differences listed above, the remaining qualitative data will be treated as though belonging to a single population.

5.1.2.1 SMATC Display - Without exception, all pilots found the SMATC readable in the locations indicated previously in this report. One pilot commented that in a small cockpit, such as the DC-9, one centrally located device would be sufficient, although he made no mention of the increased reliability which would result from redundancy.

The majority of pilots agreed that there were no confusing abbreviations on the SMATC. (A complete listing of the abbreviations used may be found in Interim Report No. 6 in this series.) However, several of the pilots confused OSI (the identifier for Woodside) with the 051 radial. During the experiments, the TI font for 5 x 7 dot matrix characters was employed. Recent work by Huddleston in England has resulted in the development of a font which, while less satisfying esthetically, greatly reduces confusions of this type. The use of this font is recommended for future displays employing dot matrix elements.

The majority of pilots felt that the SMATC display did not dominate their attention or distract it from any other flight or navigation instrument. Those who commented that the SMATC did distract, or could to a degree, cited the horizontal situation indicator, flight director, altitude, and airspeed instruments as those which could be neglected.

Two pilots cited the red light of the SMATC LED characters as distracting, since this color is usually reserved for high-priority fault messages. During the approach and landing phase, the majority of pilots believed that the SMATC display could cause at least a marginal distraction. Flashing of the display was frequently mentioned as being distracting, as was the audio alert during this flight phase.

The great majority of pilots thought that the 16-character field of the device was adequate, based on the assumption of printer availability in the cockpit, with the SMATC being limited to short ATC messages. Similarly, a majority felt that individual character size was adequate.

Most pilots showed preference for an audio alert when new message appeared on the SMATC, as opposed to flashing the message on and off to gain attention. The choice of having the WILCO button flash was not attractive to the pilots.

Several pilots observed that the SMATC went unnoticed during the busier phases of flight, including the instrument approach. One pilot suggested that a message on the SMATC during this phase of flight should be more attention-getting.

The use of the SMATC display for emergency or time-critical messages, such as minimum-safe-altitude warning or go-around, is not recommended. Even with the audio alert employed it does not adequately gain the crew's attention during heavy workload situations.

5.1.2.2 <u>Voice Synthesizer (Vosyn)</u> - While B-727 crews almost universally condemned the Vosyn because of the poor audio quality of the amplifiers in the simulator, DC-9 crews indicated a much more favorable opinion; 47% rated the Vosyn intelligibility as good, 47% as marginal, and only 6% as unacceptable.

In commenting on the applicability of the Vosyn several pilots observed that it should not be used for routine communications, but only for high priority emergency or semi-emergency communications, such as minimum-safe-altitude warning or missed approach. Pilots found long messages to be distracting and hard to follow because of lack of tonal inflection and space between words and phrases. This was particularly true if others in the cockpit were speaking. It was observed that the use of the Vosyn detracts from one operational advantage of Data Link, namely, that two persons can communicate without devoting full attention to each other at the same time. In contrast to the SMATC, the Vosyn demands the pilot's attention by blocking out other conversation. In general, it was felt that the use of the Vosyn should be limited to selected short messages.

5.1.2.3 Voysyn with SMATC - While a majority of pilots felt that the use of Vosyn and SMATC together was justified, 37% believed that this complement provided too much redundant information. A large majority indicated that it was not confusing to have the same ATC messages presented on both the SMATC and the Vosyn. (The response time analysis, however, shows that this did significantly delay responses, confirming the results of earlier experiments in this series.) Comments in favor of device duplication came frequently from the crews that had been given the low-altitude go-around with only the SMATC device. In these and similar instances, the Vosyn served as an effective attention-getter.

Despite this, a large majority (84%) believed that the Vosyn could be most easily eliminated, with only 5% stating that the SMATC could be eliminated.

5.1.2.4 Printer - Seventy-five percent of the respondents felt that all messages should appear on the printer, despite the paper management problem which this posed. (The printer used produced 21 columns on paper 3-1/2" wide, and during a typical 1 hour 10 minute flight, about 11 feet of printout was generated.) A majority of the respondents either believed that the time-tag which appeared on all printer messages had no value or had no opinion on it. Seventy percent of the pilots felt that there should be a visual or aural alert to inform the crew of impending company messages. Color coding to differentiate ATC from company messages was considered desirable, but it was noted that the red used for company messages during the experiment became unreadable under red cockpit lighting, and that the use of some other color would be required. Printer location at the flight engineer's station in the B-727 was considered satisfactory and convenient, but in the two-man DC-9 cockpit, a different location was required to make the printer output accessible to both crew members.

Crews found the printer especially useful for the longer messages, such as clearances, ATIS, and weather. With a printer available, the duplication of these messages on the Vosyn is difficult to rationalize.

5.1.2.5 General Considerations - Pilot opinion was heavily in favor of having the AUTOTUNE device automatically set the voice and data communications transceivers. Somewhat surprisingly, opinion was almost as heavily against the use of AUTOTUNE to set or tune any other instrument or radio. A few felt that transponder codes and company radio frequencies could be set automatically, but there was, as might be expected, strong opinions against automatic setting of navigation frequencies. Reactions to the question concerning automatic setting of altitude alert or heading and speed bugs were mixed and indifferent. Many pilots thought that the HAS recall capability was useful even though their heading bug and altitude alert had been set.

The possible loss of data which results from the assumed selective-address capability of Data Link, such that only one aircraft receives or transmits the ground-air or air-ground messages intended for it, although it provides a major improvement in transmission effeciency, may pose one of the major liabilities of the Data Link concept. Many pilots view their own mental air-traffic-situation analysis as a necessary part of safe and reliable air traffic control.

The majority of pilots believed that Data Link would reduce crew workload. This view was taken not strictly with respect to air-ground-air communications, but also with respect to other crew duties. As one pilot commented, "The lack of constant talk format over the radios allowed the crew to perform check lists and other command type duties."

Comments concerning procedures for obtaining a voice channel indicated some confusion. The briefing instructions pointed out that an open channel or "hot" mike to ATC was always available. Further, if pilots responded UNABLE to a message, the ATC controller would always inquire over this channel. The "Request Voice" button on the CDU, it was explained, was to be used only for acquiring a company voice channel. Apparently these potentially confusing instructions were either misstated or misunderstood in several cases. Numerous critical comments were received from crews about having to push a button and wait for data-polled voice channel acquisition before they could talk with ATC. This should not be considered a potential liability of Data Link, but rather a result of inadequate crew instruction brought about by the time constraint imposed during the running of the experiment.

A further potential liability of Data Link, however, did arise. Numerous comments indicated a strong reluctance to accept or trust Data Link during ground-proximate flight phases, including local control, approach, and departure. Not only did pilots feel that the visual presentation of messages while the aircraft was in flight but close to the ground could be distracting – they were actually wary of the Data Link in such circumstances. There seems to be a certain reassurance in <u>talking</u> to a human on the other end of the circuit and not to a computer responding to manually pushed buttons.

5.1.3 Observer Comments

The duplication of the GAT-2 flight instruments at the Experimenter's Console made it possible for the experimenter to monitor the progress of the flight and determine the appropriate time for message dispatch. Under these conditions, the main function of the on-board observer was to record which crew member responded to the ATC commands or generated other downlink messages, a task requiring no special qualifications.

With the lack of such duplication of instruments at the Experimenter's station of the airline simulators, the on-board observer assumed a more important role. The four on-board observers used during the experiment were all instrument-rated pilots with considerable experience in large jet in-flight observation. Accordingly, they were not only able to assist the crew in playing the scenario and to coordinate test conduct with the experimenter, but also to observe crew functioning in the Data Link environment and to record pertinent observations. They noted crew errors and areas of crew confusion caused by the data link, and recorded significant crew comments that would otherwise not have become part of the data base.

5.1.3.1 General Observations - Generally, the crews were usually quite rushed. Under optimum conditions, the familiarization run and the three tests could barely be accomplished in the four hours allotted. Any equipment problems, or even a long coffee break, would cause the schedule to be delayed. The observer controlled the amount of simulated tailwind to keep scenarios on schedule, particularly during en route phases of flight.

This hurriedness seemed to generate crew frustration and annoyance in several instance, which hopefully did not bias them against the Data Link concept or specific equipments. As a positive benefit, this pressure provided an atmosphere conducive to mistakes, and, consequently, a more rigorous test than would have occurred under less trying conditions.

Along with the time constraints, the lack of familiarity of some crews with the scenario geography caused confusion. However, in the end, no adverse performance was observed that could be attributed to Data Link.

5.1.3.2 Data Link Concept - The loss of information from the aircraft ahead regarding holding, approaches, descent instructions, and weather was frequently commented upon. One pilot stated that if "exception" type information from the aircraft ahead was available on a common voice channel, this loss might be tolerable.

Many pilots were critical of the use of Data Link below an altitude of approximately 2,000 feet above ground level. One volunteer captain commented, "From clearance to takeoff until above 2,000, and from clearance to land until turning off the runway I am too busy trying to keep from killing myself to push buttons."

5.1.3.3 SMATC Display - The SMATC display seemed to distract pilots during low-altitude phases of flight and could be dangerous under certain circumstances. The use of SMATC for emergency or critical communications appeared to be inadvisable, since it did not command attention. This was the reaction of several pilots. Pilots were quite pleased with the SMATC display of recalled HAS information.

5.1.3.4 The Voysn - Many printer verifications were required for Vosyn-only messages, especially long ones. In the case of ATIS, the message had to be repeated or confirmed on the printer with regularity. Several informal remarks indicated that the Vosyn might be desirable for emergency or critical messages. "Expedite through 190" was misinterpreted at least twice as "Expedite through 150." When the Vosyn was used in combination with the SMATC, low-altitude critical messages such as "go around" were observed often by Vosyn and not by SMATC presentation, although the SMATC was preferred for routine information. Vosyn HAS recall was disregarded when SMATC HAS recall was available.

5.1.3.5 The CDU - Several crews indicated a desire for a CDU capability that would "Request 10 mile DME Leg" when a holding pattern was assigned. With the brief training time available, the CDU seemed somewhat complex and confusing, and different pilots made suggestions for simplifying it, primarily by combining or eliminating buttons. Crews experimented with the CDU at cruising altitude, although they frequently complained of the need for excessive buttons pushing at lower altitudes, and described a feeling of being boxed in by its use on these occasions. The pilot report message suggested as a result of turbulence, "PI MOD TURB 250 SEND," required an average of 43 seconds to generate (based on speed of entry data described in Section 5.2).

One pilot commented on his reluctance to use the control column WILCO rather than the CDU WILCO when the latter was flashing. Others indicated that on final approach a control column WILCO capability was preferable.

5.1.3.6 The Printer - The location of the printer in the DC-9 was poor in that only the right-seat pilot could access it. In general, the DC-9 pilot did not notice company messages, which appeared only on the printer and without aural alert.

5.1.4 Supplemental Questionnaire

Analysis of the results of the original questionnaire revealed several ambiguities. As a result, in an effort to resolve these ambiguities and to achieve an overall ranking of Data Link type systems as compared with conventional voice, a supplemental questionnaire was mailed to the test subjects some three to four months after the completion of the airline simulator tests. The time dalay between the simulator runs and delivery of the supplemental questionnaire was considered beneficial in that it allowed the crews to reflect on their simulated experience with Data Link and to form more carefully considered rather than spur-of-the-moment decisions. Eighty-six percent of the participants responded.

Crews were asked to consider air traffic control by means of (1) conventional voice, (2) SMATC display, and (3) Synthetic Voice (Vosyn). They were asked to rank them 1, 2, or 3, from most to least desirable, for the 6 different phases of flight: (1) ground control, (2) local control, (3) departure, (4) low en route, (5) high en route, and (6) arrival.

The first choice was arbitrarily weighted 2, the second choice, 1, and the third choice, 0. The SMATC was found to be ranked first in the ground phase, with conventional voice second. Conventional voice was a clear first choice in the local phase, followed by SMATC. For both arrival and departure phases, conventional voice and SMATC were ranked equally, with SMATC emerging as a clear first choice in both the low and high en route phases of flight.

The opinion about the effect of a better quality of synthetic speech on the overall ranking was evenly split, with 48% stating that they would have answered differently, and 52% stating that this would have no effect. Those who answered "yes" could be placed in different categories. Several would have exchanged the Vosyn with the SMATC in the ground, local, arrival, and departure phases. Several who had chosen SMATC first in a given phase would have moved Vosyn up to second choice. It appears that the major effect, if any, of improved synthetic speech would be a redistribution of the second and third choices; it does not appear that the first choice would be altered.

In the original questionnaire, only 3% of the respondents felt that the information which they lost by not being able to hear communications addressed to other aircraft was of no value. The supplemental questionnaire attempted to categorize this lost information under the headings of flight safety versus comfort or convenience for weather advisories, knowledge of the position of other aircraft, and terminal area routing information.

In the consideration of flight safety, a clear majority of pilots are concerned about the position of other aircraft, with weather information and terminal area routing information playing a lesser but still significant role. In the area of flight comfort or convenience, the loss of weather advisory information was considered most important, since it might, for example, influence the captain's decision as to when to turn on the seat belt sign.

As a replacement for the lost information concerning the position of other aircraft, the concept of an Air Traffic Information Display was mentioned frequently. Similarly, for weather information, the only acceptable substitute for common-channel voice appears to be "by exception" reporting of weather anomalies on an open channel. This must be specific and current as to location and altitude, since pilots already have synoptic, forecast, and/or station weather. Pilots were additionally asked whether the presumption of computer conflict-prediction backup (i.e., a system free from human error) would have affected their answer on the loss of essential information. The answer was a decisive "no." Pilots are still wary of the reliability of computers and their programs.

On a weighted basis, the relative desirabilities of Data Link features were determined to be:

1. Data Link presentation of ATC commands
2. Data Link presentation of non-navigational ATC commands (e.g., a transponder code change)
3. Preprinted departure clearances
4. Acquisition of en route ATIS information
5. Automation of company reports

Other items mentioned as desirable included the elimination of voice congestion (mentioned frequently), maintenance-data transmission, minimum-safe-altitude warning, and collision-avoidance system (CAS).

5.1.5 Company Communications

On each of the three flights for every crew on the airline simulators, a varying level of automation was assumed for the Out, Off, On, In (OOOI) reports. At the one extreme, as many as 16 button depressions were required to enter the station identifier, out time, off time, and fuel weights (requiring an average of 46 seconds). At the other extreme, the pushing of one button automatically caused all of these variables to be transmitted. As could be expected, the crews preferred the most automated case; i.e., that which required the least button-pushing. However, crew reaction did not seem to be strongly for or against any of the three assumed levels demonstrated.

The use of a data terminal by the second officer in the B-727 for company communications during this simulation appeared both feasible and practical. Although complete automation of OOOI is most desirable, some manual entry by the second officer appears acceptable. In the DC-9, the first officer's workload when using a data terminal, though somewhat changed in nature, appears comparable to present voice procedures. Although the pilot must enter data, he can do so at his leisure, knowing that once the entry has been made it can be forgotten, with the assurance that the data will be properly communicated.

5.2 QUANTITATIVE DATA

Quantitative data collected during the evaluations on both the GAT-2 and airline simulators consisted of response times to uplink messages, and time-logged records of events such as crew-initiated downlink messages. Because of differences among crews, simulators, scenarios, and experimental design, a direct comparison of all data obtained on the GAT-2 versus the airline simulators cannot be made in all cases. Data from the two studies are accordingly first presented separately, followed by a comparison of those data having a common basis.

In these evaluations, as in previous I/O device evaluations, the response times to messages have been used as a measure of the effectiveness of various displays or complements. Response time does provide a measure of the comprehension of each display or complement if we assume that crews respond as rapidly as possible once they comprehend the message. Since the crew was briefed to respond accordingly, and since the previously described experimental designs were adhered to as closely as possible within the constraints of available simulator time, response times are thought to provide a good indication of communications performance.

The records of events other than response times may be used to determine the relative usefulness of various features provided by Data Link. In addition, crew input errors and practice effects in the use of the CDU can be analyzed. This is, in effect, an evaluation of the Control and Display Unit (CDU), since all events must be initiated from the CDU.

5.2.1 Quantitative Data from GAT-2 Tests

As indicated in Section 4.3 of this report, limitations on the availability of simulator time and crew rescheduling problems made adherence to the original experimental design virtually impossible. Furthermore, it was discovered in the debriefing session that some of the crews were not aware of the importance of making prompt responses. All of this reduces the value of response time data.

In order best to fulfill the design intentions, data from crews that made only a limited number of runs have been eliminated, and the remaining data forced into the best compromise. This is shown in Table 5-1 where crews "G" are general aviation pilots, "T," test pilots, and the crew numbers following the "G" or "T" the sequence of trials listed in the original experimental design of Table 2-1.

TABLE 5-1. ACTUAL SEQUENCE OF RUNS FOR EACH CREW

Run	One Man Crews Gen. Aviation				Two Man Crews Gen. Aviation				Two Man Crews Test Pilots		
	G-1	G-2	G-4	G-5	G-3	G-7A	G-7B	G-8	T-5	T-6	T-8
1	AID	BJN	DLN	ELN	CKD	GJN	GJN	HID	ELN	FKD	HID
2	BJN	AID	CKD	GJD	DLN	ELD	ELD	FKN	GJD	HIN	FKN
3	CKN	DLD	BJD	HID	BJD	FKD	FKD	ELN	HID	GJN	ELN
4	DLD	CKN	FKD	FKN	AIN	HIN	HIN	GJD	FKN	ELD	CKN
5	GJN	HID		BJD	FKN	DLD	DLD	**	BJD	DLN	BJN
6	HIN	**		CKD	HID	AID	AID	BJN	CKD	CKD	
7	FKN	FKN		DLN	GJD	BJN	BJN	AID	AIN	BJD	
8	ELD					CKN	CKN	DLD	DLN		

**Defective Data Recording. Crew G-8 contained one test pilot. The last five runs of crew T-5 were made with a single test pilot.

With such incomplete data, it would seem desirable to check for internal consistency. Since a different device complement was used on each run, strong practice effects would not be anticipated; on the other hand, the results of previous experiments have indicated an improvement in performance on the second run with a given scenario resulting from some capability to anticipate certain commands. The mean reaction times of all crews to all short ATC messages as a function of run sequence are as follows:

RUN	RESPONSE TIME	
1.	9.3"	10.3"
2.	10.5"	
3.	11.0"	
4.	10.3"	
5.	8.4"	8.6"
6.	8.4"	
7.	8.0"	
8.	9.5"	

Here, it should be noted that there are no noticeable practice effects during runs one through four, or during runs five through eitht; but the means for runs five through eight are some 1.8 seconds shorter than for runs one through four, indicating the practice effects from second exposure to a scenario.

As a second indicator of data consistency, the rank order of response times for each group of crews (one-man general aviation crews, two-man general aviation crews, and two-man test pilot crews) for each scenario/device complement combination are presented in Table 5-2, again for responses to all short ATC commands.

Here, the high degree of internal consistency should be noted. If one point is assigned to the scenario/device complement yielding the slowest response, increasing to seven points for that yielding for fastest, the rank order becomes that of Table 5-3.

TABLE 5-2. RANK ORDER OF RESPONSE TIMES OF GROUPS OF CREWS TO SCENARIO/DEVICE COMPLEMENT COMBINATIONS

RANK ORDER	CREWS G-1 G-2 G-4 G-5	CREWS G-2 G-7A G-7B G-8	CREWS T-5 T-6 T-8
Slowest	KC	KC	KF
	JB	LD	LD
	KF	KF	KC
	LD	JB	JG
	JG	JG	JB
	IH	LE	LE
Fastest	LE	IH	IH

TABLE 5-3. RANK ORDER OF RESPONSES BY THREE PILOT GROUPS TO THE SEVERAL SCENARIO/DEVICE COMPLEMENTS (HIGHER SCORES REPRESENT FASTER RESPONSES)

SCENARIO/DEVICE COMPLEMENT		SCORE
KC:	Vosyn (all messages), SMATC (HAS only) SCDU	5
KF:	Vosyn (all messages), Full CDU	7
LD:	Vosyn (long messages), SMATC, Full CDU	8
JB:	Vosyn (short messages), Printer (long messages), Small CDU	11
JG:	SMATC, Vosyn (all messages), Printer (all messages), Full CDU.	14
LE:	SMATC, Small CDU (voice for long messages)	19
IH:	SMATC, Printer (all messages), Full CDU	20

The data of the preceding three tables indicate that there is meaningful information in the overall data. At the same time, because of deficiencies in the implementation of the experimental design, and since at least some of the crews were not aware of the importance of making rapid responses, any apparent differences between experimental conditions must be viewed with caution.

5.2.1.1 Data on Responses to Uplink Messages - Even though the experimental procedures used were less than rigorous, the large number of responses obtained (3226 measured downlink responses to short uplink messages, for example) gives some assurance that major variables may be compared in a meaningful manner. Such variables include the Full CDU versus the small CDU, differences among crew types, differences between responses on the CDU and on the control column, differences between audio presentation on the Vosyn versus visual presentation on the SMATC versus a combination of the two, differences among message types, and differences between the use of the printer and the Vosyn for long messages.

Table 5-4 lists the mean response times to the various categories of short ATC messages and for all long messages (ATIS, clearance, and weather reports) for each crew and crew type. Here, it should be noted:

1. That for many crews, traffic messages produced the shortest response, since crews flying a simulator are well aware that there can be no traffic hazard and thus reply automatically without attempting to establish the identity or location of the indicated traffic.

2. That for general aviation crews, short message responses were faster with two-man than with one-man crews, indicating the greater workload imposed on one-man crews.

3. That one test pilot crew was appreciably slower than the other two in their responses to both long and short messages.

TABLE 5-4. MEAN RESPONSE TIME IN SECONDS TO VARIOUS MESSAGE TYPES BY INDIVIDUAL CREWS AND CREW GROUPS

CREW	SHORT ATC MESSAGES				LONG MESSAGES
	HAS	Traffic	Other	All	
G-1	6.8"	5.1"	10.7"	9.3"	16.1
G-2	12.3	5.8	10.7	10.8	21.4
G-4	17.1	8.4	11.3	12.6	28.8
G-5	7.5	5.5	8.1	7.8	17.8
$\overline{X}$	10.9	6.2	10.2	10.1	21.0
G-3	7.2	9.0	7.6	7.6	23.8
G-7A	10.2	12.7	11.6	9.4	34.3
G-7B	7.6	10.0	7.5	7.5	19.6
G-8	8.7	8.8	9.4	9.2	28.9
$\overline{X}$	8.4	9.9	9.2	8.5	26.0
T-5	7.3	7.1	7.4	7.2	17.9
T-6	12.8	7.7	10.2	11.0	36.2
T-8	6.1	5.9	7.7	7.2	12.5
$\overline{X}$	7.9	7.0	9.0	8.3	23.3
$\overline{XX}$	9.4	8.5	9.3	9.1	23.4

Table 5-5 presents the mean response times to all short messages by pilot and copilot for the two-man crews, as well as a comparison of responses on the CDU and on the control column. Earlier data (see Interim Report No. 4 in this series) have indicated that participants reacted faster when acting as copilot than as pilot. The requirement for left handed response by the copilot on the CDU may be responsible for the lack of differences found here. Of the responses, 80.1% were made by the copilot and only 19.9% by the pilot. Despite the apparent easier access to the WILCO button on the control column to that on the CDU or SCDU, 93.1% responses were made on the CDU (or SCDU). This was less true with the test pilots, who made only 84.3% of their responses on the CDU and 15.7% on the WILCO button of the control column; but this difference resulted entirely from the responses of one crew.

TABLE 5-5. MEAN RESPONSE TIME IN SECONDS FOR TWO-MAN CREWS, COMPARING PILOT VERSUS COPILOT AND RESPONSES ON CDU (OR SCDU) WITH THOSE ON CONTROL COLUMN: ALL SHORT MESSAGES

CREW	PILOT	COPILOT	ON CDU	ON COL.
G-3	12.0"	7.4"	7.6"	--
G-7A	5.6	11.6	9.4	--
G-7B	7.7	7.5	7.5	7.3"
G-8	13.6	9.2	9.2	10.0
$\overline{X}$	12.0	8.1	8.4	8.7
T-5	6.2	7.4	7.2	--
T-6	10.8	11.3	12.6	9.0
T-8	6.2	7.5	7.2	--
$\overline{X}$	9.4	8.0	9.0	9.0
$\overline{XX}$	8.8	8.8	8.7	8.7

While it might appear that the greater complexity of the CDU over that of the SCDU might lead ot longer response times, this assumption is not borne out by the data. Table 5-6 lists the mean response times for the three pilot groups for the scenario/device complements which employed the CDU and the SCDU. If we eliminate condition KF, which used the Vosyn only, where the listed time is from the start of the message, and where a response was not possible until the completion of the message (with an average speaking time of 5.3 seconds), the remaining responses were actually somewhat faster with the CDU.

In the first report in this series, it was noted that responses to a simultaneous visual/auditory display were slower than to either alone. At the same time the impossibility of direct comparison of timing was pointed out, since a short visual display can be comprehended almost instantaneously but auditory information is accumulated during the duration of the speaking. The data from the present experiment confirm that simultaneous visual/auditory presentation produces slower responses than visual alone; for short messages, the mean response times in seconds with visual, auditory, and simultaneous presentation were:

SMATC Alone	7.0"
Vosyn Alone	10.6
Both	8.3

TABLE 5-6. MEAN RESPONSE TIMES IN SECONDS FOR THE THREE TEST GROUPS TO THE SCENARIO/DEVICE COMPLEMENTS EMPLOYING CDU AND SCDU, RESPECTIVELY

TEST GROUP	CDU		SCDU	
One-man Gen. Aviation Crews	IH	8.0"	KC	12.4"
	KF	10.5	JB	11.1
	LD	10.3	LE	6.8
	JG	9.1		
	$\bar{X}$	9.5		10.1
	$\bar{X}$ without KF	9.1		
Two-man Gen. Aviation Crews	IH	7.1	KC	10.4
	KF	11.5	JB	9.9
	LD	7.8	LE	5.2
	JG	8.1		
	$\bar{X}$	8.6		8.5
	$\bar{X}$ without KF	7.7		
Test Pilot Crews	IH	7.0	KC	8.5
	KF	11.2	JB	6.7
	LD	9.2	LE	8.2
	JG	7.9		
	$\bar{X}$	8.8		7.8
	$\bar{X}$ without KF	8.0		
	$\overline{XX}$	9.0		8.8
	$\overline{XX}$ without KF	8.3		

With the long ATC messages (clearances, ATIS, and weather reports), a similar comparison is possible between the use of the Vosyn, printer, and the simultaneous combination of the two. These data are presented in Table 5-7.

TABLE 5-7. MEAN RESPONSE TIMES IN SECONDS TO LONG MESSAGES ON VOSYN, PRINTER, AND SIMULTANEOUSLY ON BOTH

DEVICE	N		Σ	
Vosyn	KC	26	638.6"	
		15	315.9	
	LD	26	676.6	
		28	760.1	
	KF	26	509.8	
		31	718.1	
	Σ	152	3619.1	$\bar{X}$ = 23.8"
Printer	JB	23	406.4	
		22	503.4	
	IH	18	442.6	
		20	439.8	
	Σ	83	1792.2	$\bar{X}$ = 21.6"
Printer + Vosyn	JG	11	456.2	
		21	478.4	
	Σ	32	934.6	$\bar{X}$ = 29.2"

No timing data are available to indicate the mean time required by the printer to complete a message at 25 characters per second, but in general the printer completed a long message in one-third to one-half the time required by the Vosyn for speaking the same message; average Vosyn speaking time for the long messages was 33 seconds. We thus see in Table 5-7 that WILCO responses on the average were made before the Vosyn completed a message, and that response time for the Printer + Vosyn was slower than for the Printer alone.

With the long messages, responses by the pilot and copilot were nearly equal in number, but the predominant number of the responses were made on the CDU rather than the control column, as indicated in Table 5-8.

TABLE 5-8. RESPONSE DATA FOR LONG MESSAGES BY PILOTS AND COPILOT, AND ON CDU VERSUS CONTROL COLUMN

RESPONSE	N	Σ	$\bar{X}$
Pilot	130	3130.2"	24.1"
Copilot	137	3215.7	23.5
On CDU	229	5593.5	24.4
Control Col.	38	752.4	19.8

5.2.1.2 Downlink Data and In-Cockpit Control Operations - Somewhat greater reliance may be placed upon the validity of the data concerning control operations using the CDU or SCDU, or for the generation of downlink messages other than WILCO, since these data resulted from volitional actions by the crews.

Table 5-9 presents data on seven such response categories. As might be anticipated, the one-man general aviation crews made fewer responses and control actions than the two-man general aviation crews; the greater workload imposed upon the single pilot prevented him from taking full advantage of the features inherent in Data Link. It should also be noted that, on the average, stored messages were recalled twice.

Comparison of the relative frequencies of the several responses and control operations cannot be made using the data of Table 5-9, since the Vosyn and SMATC were not both available on many of the runs. Data for the three runs in which both devices were present are presented in Table 5-10 and indicate that HAS recalls and I/O blanks accounted for more than two-thirds of the actions, and that unable and standby responses were negligible.

Table 5-11 presents data concerning the usage of the alphanumeric keyboard. Mean keying time per keystroke was approximately

TABLE 5-9. CREW RESPONSES TO SEVEN DIFFERENT CDU OPERATIONS

OPERATION	G.A 1-MAN	G.A.-2-MAN CREWS			TEST PILOTS			OVERALL
		PILOT	CO-P.	TOTAL	PILOT	CO-P.	TOTAL	
HAS Recalls:								
# of Resp.	164	13	169	182	36	154	190	536
# of Runs	23	26	26	26	19	15	19	68
$\bar{X}$ Resp/Run	7.1	0.5	6.5	7.0	1.9	10.3	10.0	7.9
Messages Stored								
# of Resp.	48	5	35	40	8	32	40	128
# of Runs	23	26	26	26	19	15	19	68
$\bar{X}$ Resp/Run	2.1	0.2	1.3	1.5	0.4	1.9	2.1	1.9
Recalls								
# of Resp.	90	5	47	52	14	92	106	248
# of Runs	23	26	26	26	19	15	19	68
$\bar{X}$ Resp/Run	3.9	0.2	1.8	2.0	0.7	6.1	5.0	3.6
I/O Blanks								
# of Resp.	79	23	180	203	36	108	144	426
# of Runs	15	19	19	19	13	10	13	47
$\bar{X}$ Resp/Run	5.3	1.2	9.5	10.7	2.8	10.8	11.1	9.0
Vosyn Repeats								
# of Resp.	29	6	93	99	7	58	65	193
# of Runs	18	20	20	20	13	11	13	51
$\bar{X}$ Resp/Run	1.6	0.3	4.7	5.0	0.5	5.3	5.1	3.8
Unables								
# of Resp.	22	6	16	22	8	4	12	56
# of Runs	23	26	26	26	19	15	19	68
$\bar{X}$ Resp/Run	1.0	0.2	0.6	0.8	0.4	0.3	0.6	0.8
Standbys								
# of Resp.	10	1	1	2	1	1	2	14
# of Runs	23	26	26	26	19	15	19	68
$\bar{X}$ Resp/Run	0.4	0.0	0.0	0.0	0.0	0.0	0.1	0.2
TOTALS								
# of Resp.	442	59	541	600	110	449	559	1601
$\bar{X}$ # of Runs	22.2	25.1	25.3	25.3	18.1	14.4	18.1	65.6
$\bar{X}$ Resp/Run	19.9	2.4	21.4	23.8	6.1	31.2	30.9	24.4

two seconds, and the error rate was surprisingly high; this is difficult to reconcile with the degree of acceptability reported in the responses to the questionnaire, but the limited training given to crews and their anticipation of being able to do better with practice may account for this.

TABLE 5-10. NUMBER AND RELATIVE FREQUENCY OF SEVEN RESPONSES AND CONTROL ACTIONS

RESPONSE	SCENARIO				%
	KC	JG	LD	TOTAL	
HAS Recalls	70	94	76	240	31.5%
Stored	18	18	24	60	7.9%
Recalls	24	14	32	70	9.2%
I/O Blanks	63	120	93	276	36.2%
Vosyn Repeats	23	35	34	92	12.1%
Unables	6	10	4	20	2.6%
Standby	3	1	1	5	0.5%
					100.0%

TABLE 5-11. DATA ON USAGE OF ALPHANUMERIC KEYBOARD

ITEM	PILOT	COPILOT	TOTAL
# of Alpha Messages	48	78	126
# of Keys Depressed	346	476	822
Total Keying Time	894.9"	786.3"	1681.2"
# of Deletions	45	25	70
Mean Keys/Message	7.2	6.1	6.5
Mean Time/Key	2.6"	1.7"	2.0"
Error Rate %	13.0%	5.3%	8.5%

While the "standby" button was used only infrequently, crews invariably failed to follow this with a "WILCO" or "UNABLE" with any degree of promptness, as indicated in Table 5-12. In future experiments it is essential that a more positive alert be given to crews when they have failed to provide a follow-on response.

TABLE 5-12. TIME BETWEEN "STANDBY" AND "WILCO" OR " UNABLE" RESPONSES

ITEM	PILOT	COPILOT	TOTAL
"Standby" Number	12	2	14
"Standby" Time	432.7"	289.9"	722.6"
Mean Time	36.5"	145.0"	51.6"

Depending on the device complement present in the cockpit, certain equipment or procedural variations might be required in lieu of those used in the present experiment. The high transmission efficiency of Data Link permits transmission of ATIS, weather, and flight plan changes on the Data Link frequency. With a printer available, a second radio for tuning to the ATIS channel would not be required, and/or a second radio would be used only to provide redundancy. With only Vosyn and SMATC available as output devices, provision for storage of two messages besides the latest heading, altitude, and speed information might be required, such that a pilot could use the Vosyn repeat to recall a flight plan change and ATIS while simultaneously monitoring his SMATC for new ATC commands and advisories. Similarly, while the present experiment simulated a total Data Link environment, the unstructured nature of local controller commnuications might not make them suitable candidates for Data Link, and an alternate frequency might be used for such communications.

Except for its somewhat awkward location, the printer used during the experiment was found to be entirely adequate. No pilot objected to the 21-column format instead of a page printer. The speed of printing and print quality were exceptional for such a low cost unit, although repackaging would be required if it were to become airborne.

Pilots found no difficulty in learning to use the CDU and SCDU with only brief training. Reaction times using the CDU were somewhat faster than for the SCDU, indicating that pilots found no difficulty in selecting the correct key even with the larger number of keys present on the CDU.

5.2.2 Quantitative Data from Airline Simulator Tests

In marked contrast to the tests on the GAT-2, the smoothness of the operation on the airline simulators permitted the collection of complete data in all cells of the experimental design; this in turn provided the opportunity to employ an Analysis of Variance for data reduction.

Analysis of Variance is a powerful statistical tool that permits the simultaneous study of the effects of several variables, as well as interaction among these variables. In essence, the test is made of the null hypothesis that there is no difference in the expected values of several populations, (e.g., the populations of all SMATC, Vosyn, and all SMATC with Vosyn response times). The observed differences in samples are tested to determine whether they are true differences or random variations due to noise in the data. The use of the technique presumes normally distributed data and similar variance for each population.

For the tests on the airline simulators, the factors of interest were: (1) device complements, (2) simulator differences, (3) scenario differences, (4) differences among flight phases, (5) effects attributable to turbulence, (6) lighting effects (7) practice effects, and (8) effects resulting from differences among crews.

The mathematical treatment of an Analysis of Variance (ANOVA) of an eight-factor experiment is quite complex. In order to reduce the amount of computer programming required, it was decided to conduct an ANOVA based initially on four factors: (1) simulators, (2) scenarios, (3) device complements, and (4) flight phases, since these were thought to be the most significant factors. The results obtained here could then be used to determine how to treat the remaining four factors.

It was anticipated that response times would probably be either normally or log-normally distributed. It was decided to test both distributions in the ANOVA. Further, it was observed upon examination of response time data that while the SMATC had a characteristically quick response time, there were several outlying values (some more than 200 seconds). These outlying values could generally be attributed to periods of known crew confusion (as determined by the observer's notes) or to inadequate crew briefing. The ANOVA, assuming log-normally distributed data, was conducted using only values of less than 30 seconds (96% of the data). This log-normal analysis was not conducted for all values, because of the distortion that could have been introduced by truncating the longer times.

The test was conducted at the 0.01 level of significance in order to avoid making overstatements from the data. The results of the ANOVA are shown in Table 5-13.

TABLE 5-13. SIGNIFICANT EFFECTS DUE TO FACTORS AFFECTING RESPONSE

Distribution	Values Included	Observed (0.01 Level) Significant Differences
Normal	All	• Complements • Phase
Normal	Less than 30 seconds	• Complements
Log-Normal	Less than 30 seconds	• Complements • Phase • Scenarios by phase • Simulator by scenarios by complement • Simulator by scenario by phase • Simulator by complement by phase

The last four significant effects observed in log-normal analysis represent the interactions of the factors shown with either the Phase Factor or Complement Factor. Scenario by Phase, for example, indicates that different combinations of Scenarios and Phases probably have different means. Because of their complexity, no attempt has been made to explain these and higher order interactions. It should be note, however, that even in the ANOVA of the log-normal distribution, Complements and Phases were by far the most significant in terms of sample differences. It can be safely stated, on the basis of the analysis described above, that there are significant differences in the population response times of the three complements (SMATC, Vosyn, and SMATC with Vosyn), and in the three phases of flight considered (departure, en route, and arrival). No significant differences were observed in the sample means of the two simulators or in any of the three scenarios tested.

Table 5-14 is a tabulation of average response times and standard deviations by simulator, scenario, phase, and complement. It is based on all 2471 response time values, and it repeats the tabulation based on the 2356 values that are less than 30 seconds. Table 5-15 classifies the 115 previously excluded values of 30 seconds or longer by phase of flight, simulator, complement, and scenario. The values appear to be uniformly distributed among these factors.

The first message of every trial generally had an abnormally long response time because of crew unfamiliarity (the crew usually had to be prompted to respond to the first message). One crew had abnormally long response times on almost all messages in one scenario, indicating an inadequate briefing. One crew, completely alien to the geographic area covered during the scenario, had two long response times, 169.8 and 238 seconds, during their SMATC trial. The exclusion of these two times alone reduces the overall SMATC sample response time by one-half second. For the reasons cited, the exclusion of outlying values of 30 seconds or more does not seem unreasonable in an attempt to determine the true characteristics of each factor.

TABLE 5-14. MEAN VALUE AND STANDARD DEVIATIONS OF RESPONSE TIMES BY FACTOR

Factor			All Values (Seconds)		Values Less Than 30 (Seconds)	
			Mean	Standard Deviation	Mean	Standard Deviation
Simulator:	1.	B-727	8.49	8.51	7.21	5.00
	2.	DC-9	9.47	12.7	7.27	5.58
Scenario:	1.	SFO-LAX	9.27	9.64	7.49	5.53
	2.	MCI-ORD	8.58	9.40	6.97	5.03
	3.	LAX-SFO	9.04	12.7	7.25	5.27
Complement:	1.	SMATC	8.06	13.1	6.30	5.11
	2.	VOSYN	10.6	9.82	8.69	5.41
	3.	Both	8.26	8.76	6.75	5.03
Phase:	1.	Departure	10.3	14.2	7.66	5.66
	2.	En Route	9.13	9.17	7.52	5.89
	3.	Arrival	7.96	7.73	6.89	4.85

TABLE 5-15. DISTRIBUTION OF RESPONSE TIMES 30 SECONDS OR LONGER

Factor	Occurrences by Simulator	
	B-727	DC-9
Scenario		
1 SFO-LAX	15	28
2 MCI-ORD	14	20
3 LAX-SFO	17	21
Complement		
1 SMAT	8	23
2 VOSYN	19	28
3 Both	19	18
Phase		
1 Departure	21	38
2*En Route	5	9
3 Arrival	20	22

*Phase 2 has approximately one-quarter the messages of other phases.

A histogram of the total sample of 2,471 response times is presented in Figure 5-1. This histogram shows a distribution that builds rapidly and has a long trailing tail. The overall mean of this distribution is 8.97 seconds, with a standard deviation of 10.8 seconds. Approximately 96% of the values are less than 30 seconds.

Visual inspection of the distribution of response times indicates that it may be more of a log-normal than normal distribution. A plot of the Cumulative Probability Distribution drawn on probability paper yielded a piecewise linear composite of three straight lines. This indicates that the plot of Figure 5-1 may be a composite of separate distributions.

Since the ANOVA showed Device Complement and Phase of Flight to be the significant factors, the distribtuions of response times among these factors were examined separately. Figure 5-2 presents the histograms of response-time distributions for each of the three complements. The three distributions are distinct, with the SMATC having a smaller mean than the Vosyn. The combined complement of the SMATC and the Vosyn indicates a bimodal distribution, which shows some characteristics of both individual distributions. It can be inferred from Figure 5-2 that in the combined SMATC/Vosyn trials, the crew members are responding to the SMATC in some instances and the Vosyn in others. A certain amount of this effect can be expected, since long messages such as predeparture clearances, ATIS, and weather were never presented on the SMATC but were presented on the Vosyn. The number of such occurrences, however, is quite small (on the order of 3 out of 50 messages per flight). The delayed peak in the SMATC/Vosyn distribution is higher than the first peak, indicating that pilots were responding more often upon message completion by the Vosyn than by the SMATC when both displays were present.

The apparent effect displayed in the complement with both the SMATC and Vosyn is that the crews were attempting to verify or cross-check one source of information against the other when both were presented. This is further confirmation of a trend which appeared in the data from the original GAT-1 tests.

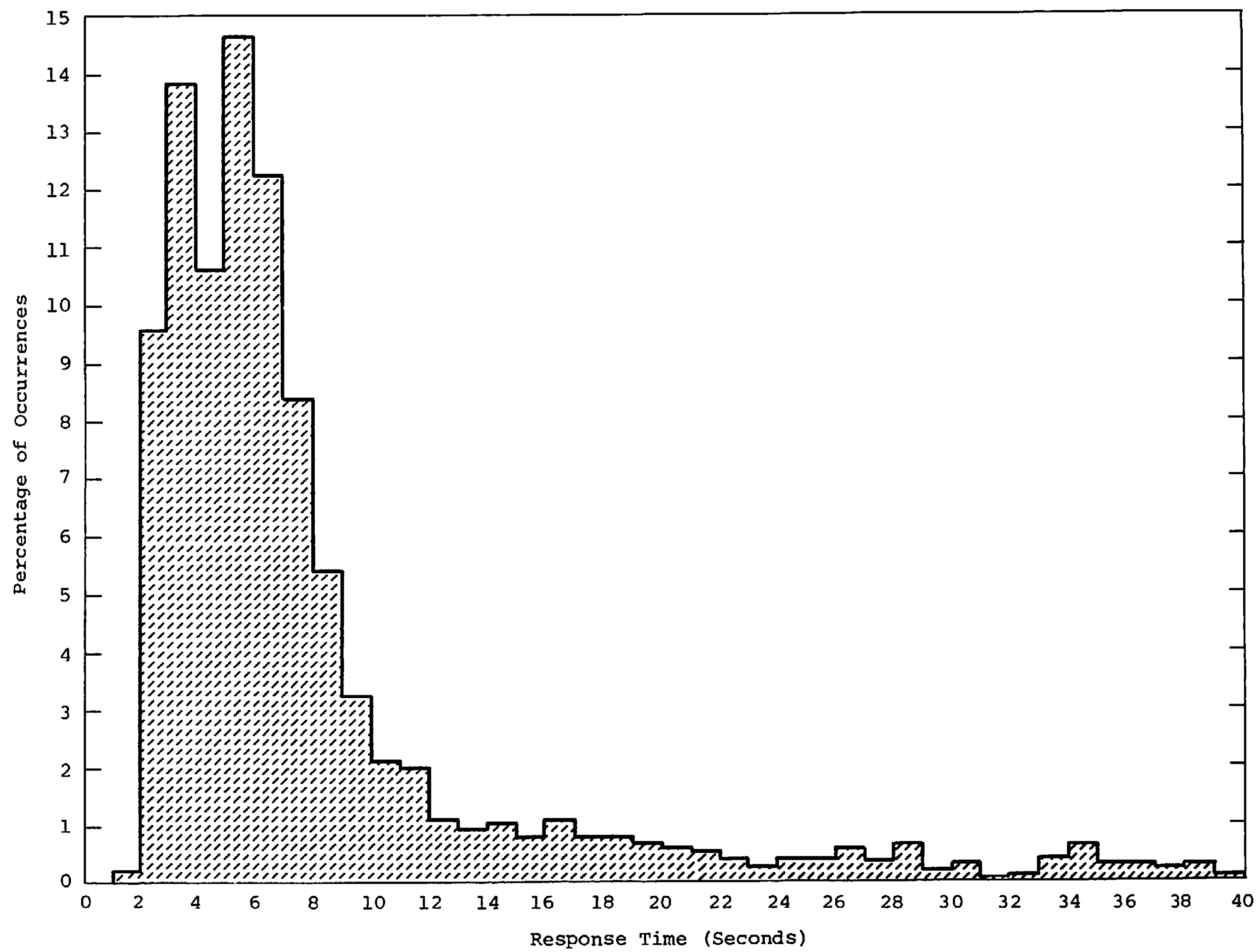

Figure 5-1. Distribution of Response Times (All Responses)

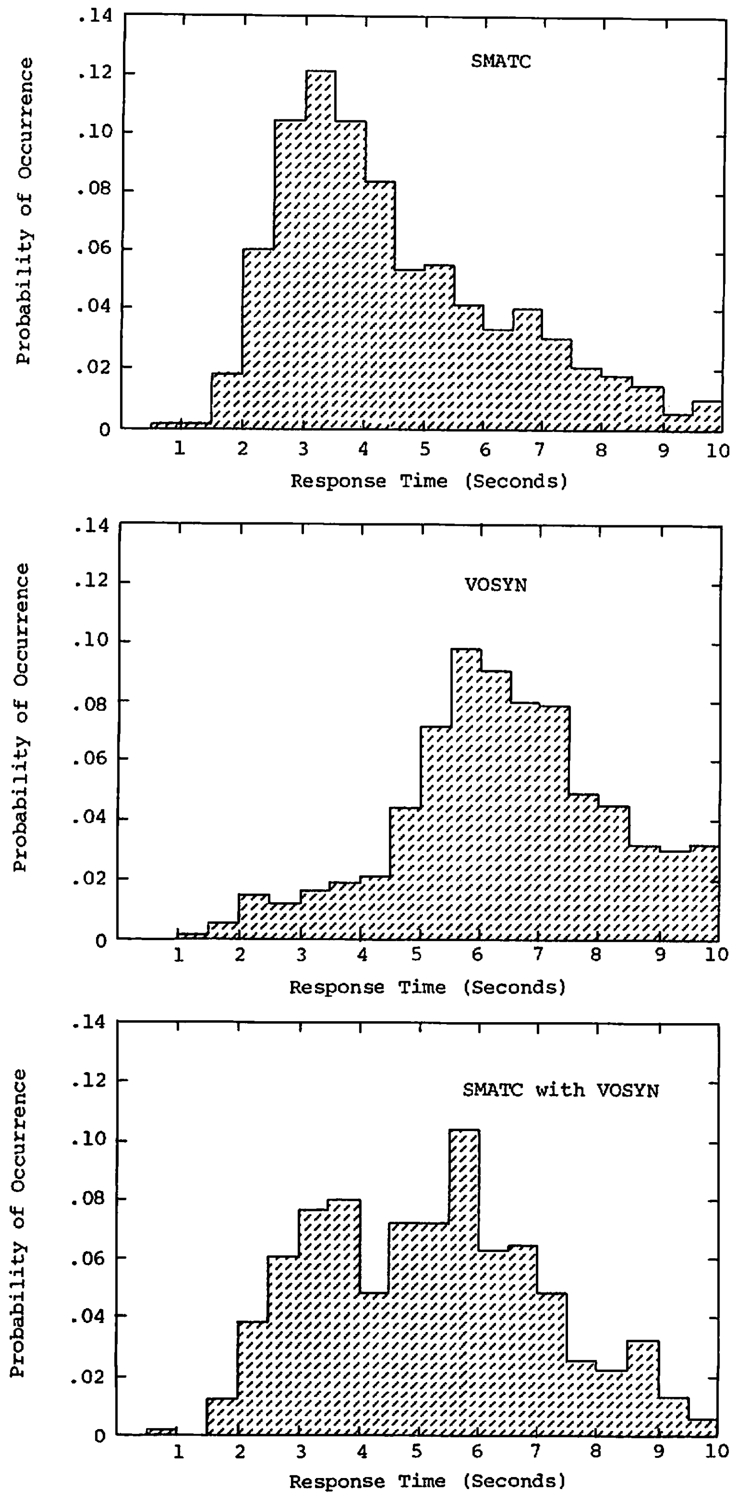

Figure 5-2. Response Time Distribution of Each Complement

Figure 5-3 shows the distribution of all response times as a function of flight phase. Phase 2, the en route phase, is a relatively low-workload period, which shows the pronounced effect of the visual and aural peaks. Phases 1 and 3, on the other hand, are high-workload phases. Here, the two distinct peaks tend to converge toward one. The increased crew workload tends to delay crew response to the visual display, while the need for quick action by the crew tends to make them respond before the aural presentation is completed.

The effects of lighting conditions and turbulence on response times and CDU operation are not determinable from this evaluation. There was very little difference in interior cockpit illumination whether the day or night runway visual system was selected. No glare, direct sunlight, or brilliance effects could be simulated. In addition, the use of visual systems acquired on a time-shared basis with other simulators meant that acquisition was usually lost by 2,000 feet MSL and not regained until the final-approach phase of flight. In effect, all flights were thus conducted under night-time conditions.

The ability to measure effects resulting from turbulence was equally ineffective. The amount of turbulence that could be simulated was at best comparable to what is described as a light chop. It is doubtful that it provided any distraction or co-ordination difficulty to the non-flying officer, who was usually operating the I/O devices.

5.2.2.1 Practice Effects and Crew Differences - "Practice effects" refers to improvements in communications performance as a crew gains experience with the I/O devices. "Crew effects" are significant differences among the 18 participating crews.

The value of the practice effects measurement is somewhat questionable, since one crew member generally communicated on his first and third trials, while the other crew member generally communicated on his second. While there may be some overall crew learning effect between the first, second, and third trials, the effect is certainly not pronounced.

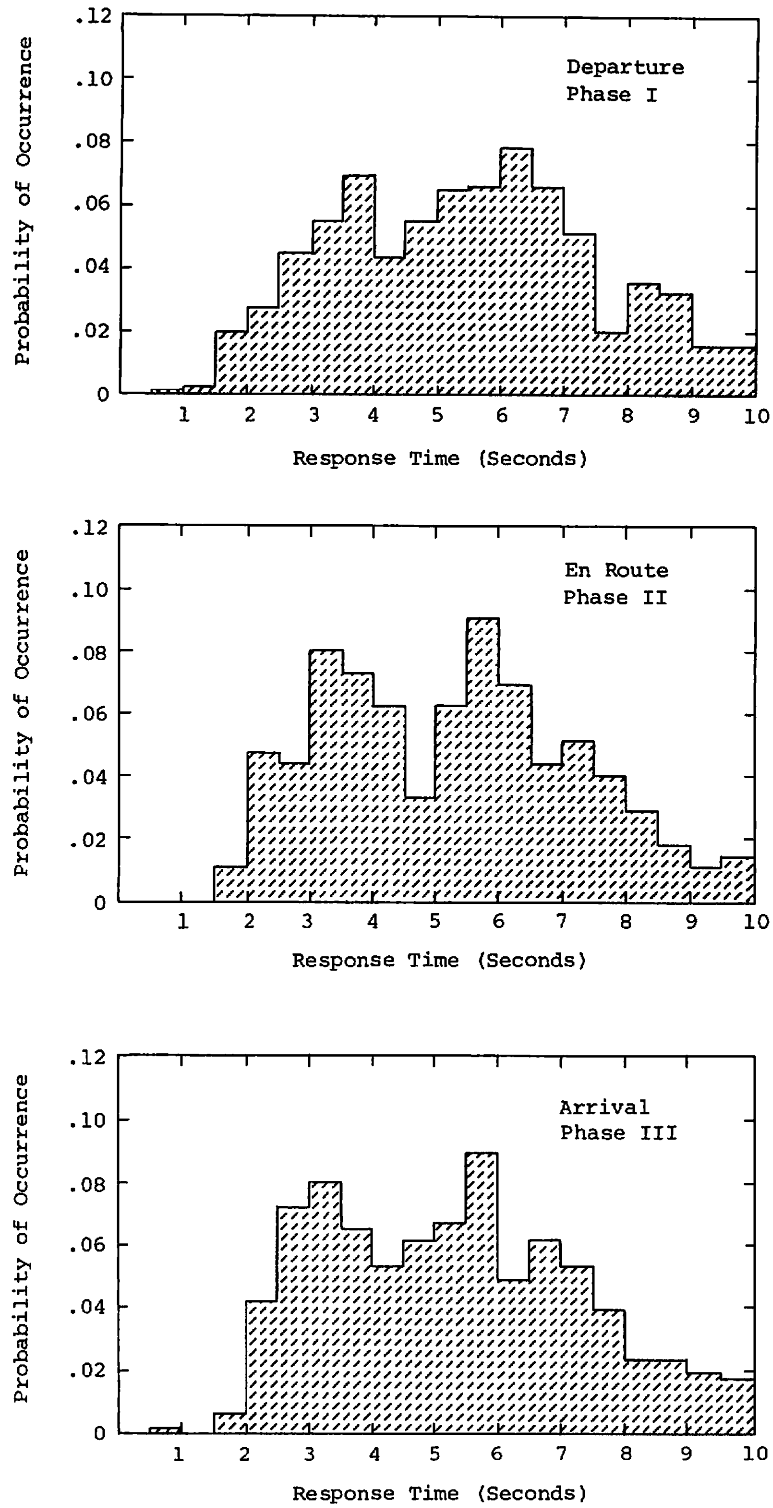

Figure 5-3. Response Time Distribution of Each Phase (Probability Density Function)

A significant test of the interactions among crews, complements, and order (learning effects) was desirable. However, because the data were confounded (i.e., not all factor combinations were represented), the ANOVA was not directly applicable; limitations of simulator time prevented the use of a complete block design. To simplify matters, a two-factor ANOVA of crews and orders was performed. A significance test at the 0.01 level showed both factors and their interaction to be significant.

To conclude, however, that differences among crews and orders are definitely significant would be a statistical overstatement. It must be remembered that in collapsing the four-factor ANOVA into a two-factor ANOVA, the device complement (previously determined to be significant) was confounded with the new factors and, in effect, masked. The fact that crew differences were shown to be significant may simply indicate that the interaction of crew and complement is significant. Likewise, an inference that the order effect is significant may simply mean that the order and complement interaction is significant.

Table 5-16 shows the mean response times and standard deviations as a function of order of flight and of crew.

An ANOVA of response times showed no significant differences among simulators or scenarios; differences among device complements and phases of flight, however, were significant. Differences among crews and order of flights were not determined, but appeared to be insignificant. The mean values for response times by complement and phase of flight, excluding miscellaneous values of 30 seconds or more, are:

Complement Mean Response Time (Seconds)		Phase Mean Response Time (Seconds)	
SMATC	6.30	Departure	7.66
Vosyn	8.69	En Route	7.52
SMATC/Vosyn	6.75	Arrival	6.89

TABLE 5-16. MEANS AND STANDARD DEVIATION OF RESPONSE TIMES GROUPED BY ORDER AND CREW

Factor	Group of Response Times Less Than 30.0 Seconds		Group of All Response Times	
	Mean (Seconds)	Standard Deviation (Seconds)	Mean (Seconds)	Standard Deviation (Seconds)
Order 1	7.88	5.56	9.41	9.10
Order 2	7.28	5.45	9.14	9.94
Order 3	6.54	4.72	8.33	13.00
Crew 1	7.88	5.11	8.56	7.25
Crew 2	8.28	5.40	9.46	8.30
Crew 3	6.47	4.44	6.86	5.47
Crew 4	7.55	5.50	8.92	9.87
Crew 5	8.38	5.16	11.30	11.40
Crew 6	7.86	4.26	9.47	8.93
Crew 7	8.11	5.43	9.47	9.13
Crew 8	5.71	4.19	7.29	7.99
Crew 9	4.94	3.94	4.94	3.94
Crew 10	5.55	5.72	6.40	7.50
Crew 11	7.55	5.53	8.13	6.76
Crew 12	6.91	5.59	7.08	5.92
Crew 13	7.70	5.28	15.00	27.50
Crew 14	6.27	4.56	7.40	7.27
Crew 15	5.87	4.66	6.87	8.80
Crew 16	9.53	6.51	13.20	11.50
Crew 17	7.56	5.26	9.04	8.43
Crew 18	9.08	5.71	12.50	12.60

5.2.2.2 Utilization of Data Link Features on Airline Simulators - The Control and Downlink Unit (CDU) served three purposes in this evaluation: (1) acknowledgment of ground-air ATC messages, (2) generation of air-ground ATC and company messages, and (3) control of data-link operation. Data were collected on the extent of CDU utilization in these areas. Table 5-17 indicates the utilization of link control functions and message-generation functions by simulator, complements, and scenarios. The first five columns are Data Link control functions, and the last six represent air-ground message operation. Not included in this tabulation are company messages (since they did not form part of the experimental design and were included only to add realism to the scenarios) or infrequently used functions, such as requests for engine start, flight plan change, heading report, pushback, and permission to taxi.

By coincidence, exactly 1000 CDU actions are tabulated here, yielding direct conversions to relative percentages of use. Fifty-eight percent of the actions occurred in the B-727, leaving 42% for the DC-9. This difference is not thought to be significant, but rather reflects the somewhat lower workload of the three-man B-727 crew, leaving them a greater amount of time to explore data link features.

By complements, the SMATC complement accounts for 44% of the use, leaving the remainder equally divided between the other two complements. This shows the crew preference for the SMATC and is consistent with other data that rank the three complements. The data are equally divided, as would be expected, among the three scenarios.

HAS (heading, altitude, airspeed) recall capability accounted for 29% of CDU use. The popularity of this feature is somewhat surprising in view of the fact that today assigned altitude is already set in a device as required by regulation, and heading is set on a rotatable reference as a general practice. The HAS feature did not receive undue emphasis during briefing; it was pointed out that it was an on-board control function, as differentiated from a ground-air message. It was used more frequently

TABLE 5-17. CONTROL AND DOWNLINK UNIT UTILIZATION (NUMBER OF OCCURRENCES)

Factor	I/O* Blank	Clearance Request	Altitude Request	Wind Check	HAS Recall	MSG Store	VOSYN** Repeat	Weather	Altimeter	ATIS	Message Recall	Total
Simulator												
B-727	99	44	33	18	171	27	58	16	11	75	32	585
DC-9	63	30	21	6	123	18	16	17	17	76	29	416
Complement												
SMATC	101	29	26	12	154	27	1	8	4	47	35	444
VOSYN	4	23	16	10	61	7	59	18	16	57	9	280
Both	57	22	12	2	79	11	14	7	8	47	17	276
Scenario												
SFO-LAX	51	23	35	5	102	19	12	8	4	47	26	332
MCI-ORD	71	28	11	12	99	11	19	13	13	55	16	348
LAX-SFO	40	23	8	7	93	15	43	12	11	49	19	320
Total	162	74	54	24	294	45	74	33	28	151	61	1000

*Applies to SMATC and SMATC/VOSYN only.

**Applies to VOSYN and SMATC/VOSYN only.

in the B-727, consistent with overall CDU usage in the two simulators. The SMATC complement received more than twice the HAS recall requests than either of the other complements, again indicating the strong SMATC preference for this feature.

Figure 5-4 shows a distribution of HAS recall requests as a function of flight progress. The number of requests is shown as a function both of distance from the departure and destination airports, and of the climb/descent profiles, which were approximately the same in all three scenarios. HAS data were requested most frequently in the departure and arrival phases below 10,000 feet. The unexpected missed approach at SFO and the arrival delay holding pattern of ORD were both large sources of HAS data requests.

I/O Blank, accounting for 16% of CDU usage, ranked second in use among the data link control features. The LED display characters were bright, even with the dimming control in the mid-position; they were located in the primary flight instrument area; and the display characters were red, which suggested an emergency type display to many subjects. The combination of these three factors explains the heavy usage that this function received. The slightly heavier use on the MCI-ORD scenario, which had a somewhat lower message density than the other scenarios, tends to demonstrate that the longer pilots looked at the same message the more inclined they were to erase it.

Vosyn repeats, amounting to about 7% of CDU utilization, was used almost four times more often in the B-727 than in the DC-9 simulator, indicating the need for a clean audio system when Vosyn is used. Message Store and Message Recall accounted for 5% and 6% of CDU utilization, respectively. Both were used more frequently with the SMATC complement than with the other two complements.

The most frequently used Ground-Air request was for ATIS, accounting for 15% of CDU usage. It was normally acquired twice for each of the 54 flights, once at departure and once during arrival. The roughly 50% excess over the required usage indicates unprompted experimentation into this feature on the part of the pilots.

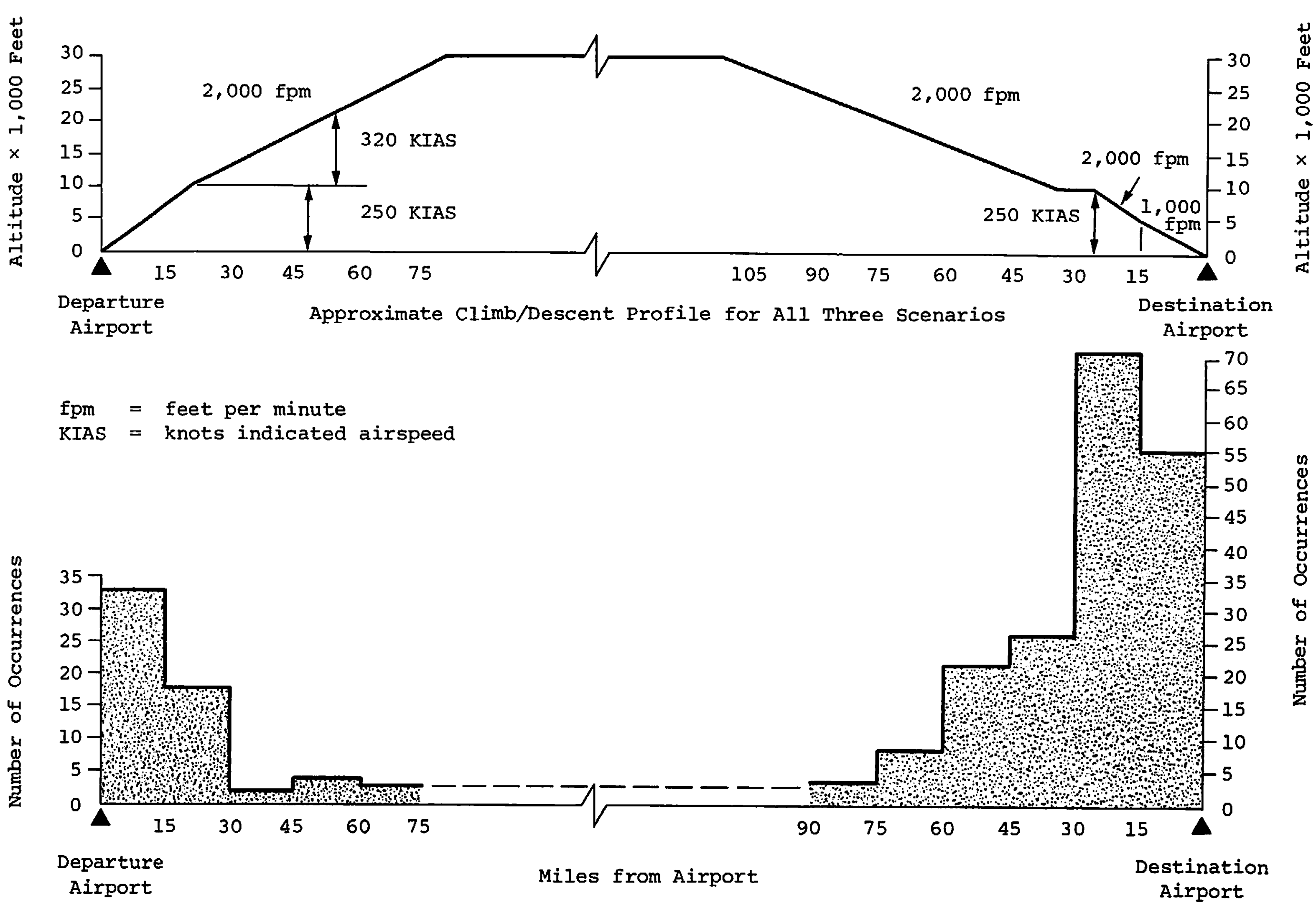

Figure 5-4. Distribution of Heading, Altitude, Airspeed Recall Requests for All Scenarios

The occurrence for all destination ATIS requests was plotted for all three scenarios for both the DC-9 and B-727 crews. Mean distances from the destinations (LAX, ORD, and SFO) were 155, 175, and 92 nautical miles, respectively, at cruise altitude. In today's voice system a typical maximum-range ATIS acquisition can be made 100 miles out, but generally the acquisition must wait until about 50 miles from destination. This earlier acquisition via Data Link would permit crews to select appropriate approach plates during a period of lower workload, and the data indicate that early ATIS acquistion would be preferred by flight crews.

Clearance Request accounted for 7% of CDU utilization. During the course of the 54 experimental runs, 54 requests were required; the additional 20 requests were made primarily in the B-727 and were most often made when the SMATC device was present.

Altitude requests, accounting for 5% of CDU usage, occurred predominantly in Scenario I (SFO to LAX), where turbulence was introduced and crews were directed to request an altitude change. Weather and altimeter setting requests are anomalies in that they appear more often during Vosyn complements than during SMATC. Consistent with this is the higher occurrence of these functions in the DC-9 where a clean audio system was present. Wind check was used most often in the B-727 simulator and during the MCI-ORD scenario. The simulator preference is probably attributable to a briefing variation.

Ground-to-air (uplink) messages in a Data Link system require some form of acknowledgement. The one postulated in the tests called for a WILCO except for "Radar Contact," "Ident," and company messages. Pilots could, however, respond in one of three fashions: (1) WILCO, (2) UNABLE, indicating his inability or unwillingness to comply, or (3) STANDBY, indicating that he wanted time to decide whether or not to WILCO or UNABLE. The incidence of STANDBY responses was essentially zero during the tests; in the real world, STANDBY would probably follow a traffic advisory, during the period in which the crew attempted to locate the traffic.

WILCO responses were analyzed with respect to which crew member made the response and how he made it. Either flying officer could either depress the WILCO button on the CDU or use a switch provided in his own location (a pendant switch in the B-727 or a switch on the outboard side of the control columns of the DC-9). The tabulation of WILCO responses by flight officer and by complement, shown in Table 5-18, indicates that a switch conveniently placed on the control column yoke is preferred over a pendant switch or the use of the CDU WILCO button.

Practice effects in the use of the CDU are described in terms of messages attempted, messages completed, CDU error rate, use of CDU functions, and CDU character-entry speed as a function of the number of trials flown by a crew. Because one officer generally operated the Data Link devices on the first and third trials and the other officer on the second, the repetitive learning effects on each individual crew member cannot be judged. Only the overall effect can be examined.

TABLE 5-18. DISTRIBUTION OF WILCO RESPONSES

Factor	Flying Officer	Non-Flying Officer			Unnecessary WILCO	Total
		Total	Column	CDU		
B-727 Complement						
SMATC	0	408	120	288	10	418
VOSYN	11	415	94	321	14	440
Both	26	424	150	274	8	458
Total						1316
DC-9 Complement						
SMATC	67	383	362	21	6	456
VOSYN	88	344	338	6	22	454
Both	34	372	362	10	6	412
Total						1322
Averages						
B-727	1.3	46.1	13.4	32.7	1.1	
DC-9	7.0	40.7	39.3	1.3	1.2	

Table 5-19 shows the practice effects on the extent of CDU usage. The simulators were comparable with regard to messages attempted, completed, and additional (other than those required) sent. The single exception was the message completion rate, which was generally higher in the B-727 than in the DC-9. (Did other duties in the DC-9 force the CDU user to abandon partially completed messages, or did better briefing on the B-727 account for this?)

Character entry speed seemed to improve slightly during the tests, but not significantly; mean crew entry speed was 2.87 seconds per character. An ATIS request for Los Angeles, for example, required the following nine keystrokes:

ADVISORY
ATIS
JKL)
RIGHT) L
ABC)
LEFT) A
WXY)
CENTER) X
SEND

On the average, this required 26 seconds to enter and send via the CDU.

A tabulation of the most frequently used CDU features indicated that ATIS requests, HAS recall requests, and I/O blank accounted for more than 60% of the CDU usage. A plot of HAS distribution shows the feature to be primarily used in the departure and arrival terminal areas. Geographic plots showed that destination ATIS requests occurred in the high en route phases of flight, earlier than currently encountered under voice. The I/O Blank usage was caused by the undesirable bright red display of the SMATC messages, and a greater range of brightness control is probably required for any future applications.

The remaining CDU functions were used only infrequently, tending to support one comment that the CDU alphas could possibly be eliminated in favor of numbers only, with some special function

TABLE 5-19. PRACTICE EFFECTS ON CDU USAGE

Message Completion					
Trial	Messages Attempted	Messages Completed	Additional Messages Sent	Fractional Completion Rate B-727	Fractional Completion Rate DC-9
1	238	208	122	0.91	0.83
2	229	207	120	0.92	0.88
3	214	180	93	0.85	0.82

Function Usage				
Trial	HAS Recall	MSG Store	MSG Recall	MSG Completions
1	75	10	14	208
2	127	21	28	207
3	90	14	19	180

Average Character Entry Speed (Seconds)				
Trial		B-727	DC-9	All Flights
1		2.7	3.2	Overall Average 2.87
2		3.1	2.7	Button Depressions
3		2.6	2.3	per Trial 67

keys. ATIS requests, for example, could be accomplished by assigning a three-digit number to each of the roughly 500 air carrier airports.

6. CONCLUSIONS AND RECOMMENDATIONS

The Data Link program, as originally conceived, called for the immediate procurement from industry of flight-qualified hardware and its evaluation in an airborne environment. Funding limitations forced cancellation of this approach, and a program of longer duration resulted, one which permitted a gradual evolution of equipment and procedures. The experiments described in Sections 1 through 3 of this report provided background as to how information could be presented to pilots in a manner which promoted the most rapid and error-free comprehension, as well as evaluating a number of hardware concepts in order to select those most acceptable to pilots. The valuable information derived from these earlier experiments, in turn, permitted the fabrication and evaluation of a smaller number of more refined devices which could be combined to form several complements, each capable of simulating a complete Data Link environment.

Evaluation of this latter series of devices in a GAT-2 simulator, using general aviation pilots and FAA/NAFEC test pilots flying scenarios typical of general aviation or small charter operations, indicated a general enthusiasm for the Data Link concept. In-cabin hard copy printout of longer messages, such as ATIS, clearance, and weather information, provided a positive decrease in crew workload, and with only brief training, crews found no difficulty in the generation of downlink messages using a keyboard and special function buttons. Further evaluation of a simplified CDU with more limited downlink capability was thus found to be unncessary.

As might be anticipated, airline crews flying B-727 and DC-9 simulators were more cautious in their evaluation of Data Link concepts, and it thus seems worthwhile to consider their comments, both in general terms and in their opinions of the several Data Link devices which they used.

6.1 GENERAL DATA LINK CONCEPTS

With two exceptions, no significant differences were observed between the operation of Data Link in two- and three-crew-member aircraft. The poor quality of the audio system in the B-727 simulator made the use of the Vosyn almost completely unacceptable; with the use of a separate and better quality audio system in the DC-9, however, only 6% of the pilots found it unacceptable. With the printer, flight engineers found the use of symbols and abbreviations to be more acceptable than did the first officer who received the printer output in the DC-9 simulator.

The selective-address capability of the Data Link in which each aircraft received only transmissions intended for it, caused a loss of information that pilots considered essential. It was stated that the importance of the lost Data Link information affects both the safety of flight, and the comfort and convenience of airline passengers. A majority of pilots believed that the loss of knowledge of the proximity of other aircraft in the same en route or terminal sectors, which is normally acquired on a common-channel VHF system, is detrimental to flight safety. Similarly, a majority of pilots believed that the loss of both terminal-area routing information (such as knowledge of aircraft ahead, holding patterns in use, approaches in use, and anticipated descent instructions) and specific weather-anomaly information (such as the extent, location, and altitude of encountered turbulence) is detrimental to passenger comfort and convenience. The postulation of a ground-computer conflict-prediction system had little effect on these opinions.

The Data Link concept has postulated that an open voice channel would always be available for time-critical downlink messages. Due to misunderstanding during the briefing, certain crews mistakenly thought that voice communications would be on a polled basis similar to that employed for the digital transmission of Data Link information. This they found completely unacceptable.

The concept of Data Link control was somewhat mistrusted during ground-proximate flight phases, including local control, arrival, and departure. The requirement for pilots to use the CDU or receive Data Link instructions during a missed-approach execution caused considerable unfavorable comment.

6.2 SHORT MESSAGE ATC DISPLAY (SMATC)

The SMATC display was found to be easily readable and well located. With the exception of a small number of commands, the abbreviations used on it were not confusing. Except during climbs and descents, the SMATC did not appear to be distracting. With SMATC installation as a standard aircraft instrument, pilots would probably develop a new scanning pattern which would minimize this effect; however, the use of SMATC for critical messages is not recommended, since it did not attract sufficient attention even when combined with an audio alert. Yellow and green LED matrix displays will soon be available to replace the red LED used during these tests, and should eliminate one further source of unfavorable pilot comments. The SMATC was very popular when used as a recall instrument for currently assigned Heading, Altitude, and Airspeed information.

6.3 VOICE SYNTHESIZER

The use of Vosyn should be limited to short messages, particularly of the emergency or time-critical variety. A clean (though not necessarily expensive) audio system is necessary if the Vosyn is to be accepted. The Vosyn seemed to demand attention, whereas the SMATC allowed the crew to communicate effectively without devoting full attention to each other. While not explored during these tests, the Vosyn may find additional use as a replacement for the presently large number of aural warnings of aircraft malfunction, since it can provide specific information requiring no further interpretation on the part of the crew. Duplication of information on the Vosyn and SMATC was not considered desirable, and led to longer response times than did presentation on the SMATC alone.

6.4 PRINTER

Crews found the printer to be especially useful for the longer messages such as clearances, ATIS, and weather. The use of the printer on shorter, more perishable information did not seem desirable, particularly since it contributed to the paper management problem. Restriction of printer access to the second officer (flight engineer) is satisfactory in an aircraft with three crew members, but with two-man crews, the printer must be accessible to both crewmen, since the aircraft is routinely flown from either seat position.

6.5 CONTROL AND DOWNLINK UNIT (CDU)

Although the CDU was somewhat complex, the crew learned to be fully proficient in its use after about two 1-hour flights that had been preceded by a 20-minute training session. The left-center-right method of entering alpha characters on the modified "Touch Tone" key pad was tedious but not entirely objectionable; it was suggested, however, that numbers be assigned to commercial airports so as to eliminate the need for alphas. This, though, would require the assignment of numbers to all navigational beacons if on-board requests for change of flight plan were to be entered; this, in turn, would require the use of an extensive look-up table or revision of aeronautical charts to provide this information. No problems were encountered from left-handed operation of the CDU. The use of WILCO buttons on the control yoke is highly recommended, since this represents by far the highest percentage of CDU usage.

The experiments further pointed out that crews would prefer earlier access to arrival ATIS information than is presently available, so that it could occur during the en route flight stage when workload is lower. The AUTOTUNE feature was considered highly desirable for selection of communications frequencies, but the idea of extending automatic tuning and setting of any other device, with the possible exception of the transponder, was considered undesirable.

6.6 ACCEPTABILITY OF A DATA LINK SYSTEM

Three problems must be solved before air traffic control can be exercised entirely through a digital Data Link system:

1. A replacement must be found for the loss of common-channel information because of selective-address communications.
2. There must be immediate availability of a voice channel if a polled system is used for Data Link messages.
3. A decision must be reached as to the relative merits of Data Link and conventional voice for each of the several phases of flight.

If the ongoing work on horizontal situation displays results in widespread introduction of such devices into cockpits, this may well provide a substitute for the information presently lost when a selective-address Data Link system is used. However, something other than a mere relay of the information appearing on the controller's CRT must be considered. In particular, present controller displays do not present PIREP information. For the safety and comfort of passengers, any display relayed to flight crews from the ground should, as a minimum, include information as to the location and altitude of turbulence reported by other aircraft.

The second potential problem will not exist, providing that Data Link information is transmitted on its own discrete channel. Rather, the use of Data Link for a majority of routine transactions should greatly reduce the present voice channel congestion, making it much easier for a crew to report to a controller when they encounter any sort of emergency.

At the moment, it appears that Data Link is applicable for clearance delivery, and the departure, en route, and approach phases of flight, but voice link remains the prime candidate for local and ground control. Advances in airport ground surveillance systems to provide better control of ground traffic via computer may render Data Link viable for ground control, but it is difficult to visualize an automatic system which could be competitive with a local controller in determining the need for such emergency procedures as landing abort.

6.7 RECOMMENDATIONS

It is recommended that consideration be given to in-flight evaluation of a limited Data Link system. This could probably be a domestic "add-on" to the currently envisioned AEROSAT Test and Evaluation Program and/or the proposed flight tests of a horizontal situation indicator.

The ability of the NAS/ARTS/ARINC ground system effectively to interchange and deliver information under a Data Link concept such as that hypothesized in this project should be evaluated under actual conditions. ARINC Research Corp. has recommended that a limited interchange system should be established between two proximate high-density terminals, such as San Francisco and Los Angeles. A limited number of aircraft flying regularly scheduled turn-arounds on this type of route could yield a significant quantity of cost-effective test data. A parallel study should examine the project in terms of potentially reduced staffs for FAA functions, while assuring that the safety and comfort of passengers is in no way compromised.

REFERENCES

1. E. H. Hilborn, Human Factors Experiments for Data Link: Interim Report I, Report No. FAA-RD-72-150, November 1972.
2. E. H. Hilborn and L. R. Devanna, Human Factors Experiments for Data Link: Interim Report No. 2, Report No. FAA-RD-73-55, April 1973.
3. E. H. Hilborn and R. W. Wisleder, Human Factors Experiments for Data Link: Interim Report No. 3, Report No. FAA-RD-73-69, August 1973.
4. E. H. Hilborn, Human Factors Experiments for Data Link: Interim Report No. 4, Report No. FAA-RD-74-81, April 1974.
5. E. H. Hilborn and R. W. Wisleder, Human Factors Experiments for Data Link: Interim Report No. 5, Report No. FAA-RD-75-14, February 1975.
6. J. M. Diehl, Human Factors Experiments for Data Link: Interim Report No. 6 - An Evaluation of Data Link Input/Output Devices Using Airline Flight Simulators, Report No. FAA-RD-75-160, November 1975.
7. E. H. Hilborn, Human Factors Experiments for Data Link: Extended Summary of Interim Reports 1 Through 4, Report No. FAA-RD-74-82, April 1974.
8. R. W. Wisleder, J. F. Canniff, and E. H. Hilborn, "An Experimental Evaluation of Various Electronic Cockpit Displays for Air/Ground Data Link Communications," Paper presented at AGARD Meeting, November 1974.
9. J. Sabath, "An Experimental System for Evaluating Datalink Cockpit Concepts and Devices," Paper presented at the 1975 International Symposium of the Society for Information Display, Vol. 6, p. 130.