AD A031070

REPORT NO. FAA-RD-76-91

CONTROLLER/COMPUTER INTERFACE WITH AN AIR-GROUND DATA LINK

J. Hagopian

U.S. DEPARTMENT OF TRANSPORTATION
Transportation Systems Center
Kendall Square
Cambridge MA 02142

T. Morgan

Computer Sciences Corpor
8728 Colesville Road
Silver Spring MD 20910



JUNE 1976
FINAL REPORT

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

D
OCT

Prepared for

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

NOTICE

This document is disseminated under the sponsorship of the Department of Transportation in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

NOTICE

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FAA-RD-76-91		

4. Title and Subtitle
CONTROLLER/COMPUTER INTERFACE WITH AN AIR-GROUND DATA LINK

5. Report Date
June 1976

6. Performing Organization Code

7. Author(s)
J. Hagopian (TSC)*
T. Morgan (CSC)**

8. Performing Organization Report No.
DOT-TSC-FAA-75-23

9. Performing Organization Name and Address
*U.S. Department of Transportation
Transportation Systems Center
Kendall Square
Cambridge MA 02142

10. Work Unit No. (TRAIS)
FA513/R6131

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 to May 1975

14. Sponsoring Agency Code

15. Supplementary Notes
**Computer Sciences Corporation
8728 Colesville Road
Silver Spring MD 20910

Under contract to:
U.S. Department of Transportation
Federal Aviation Administration
National Aviation Facilities Experimental Center
Atlantic City, NJ 08405

16. Abstract

This report describes the results of an experiment for evaluating the controller/computer interface in an ARTS III/M&S system modified for use with a simulated digital data link and a voice link utilizing a computer-generated voice system. A modified ARTS III M&S system at the National Aviation Facilities Experimental Center (NAFEC) provided the means for determining which of three candidate control and display modes was the most suitable for the display and dispatch of computer-generated M&S commands in a mixed voice/digital communications environment. The three modes tested were Control-by-Approval/Full Data Block (CBA/FDB), Control-by-Approval/Tabular List (CBA/TAB), and Control-by-Exception/Full Data Block (CBE/FDB). In Control-by-Approval, the controller must approve each command; in Control-by-Exception, the controller must intervene to disapprove commands that are otherwise automatically dispatched. The three modes were tested by NAFEC air traffic control specialists in an M&S scenario simulating south arrivals, single runway only, at Denver Stapleton International Airport. Figures of merit for evaluating the three modes included subjective data in the form of questionnaires from participating controllers and objective data such as message transaction time, command initiation delay, service time, and instantaneous aircraft load. The results of 66 hours of testing with six air traffic controllers show that the fully automated CBE/FDB control mode is preferred because it possesses the best workload, capacity and stress characteristics. Although the CBA/TAB and CBA/FDB modes tend to exhibit shorter communications delays, the former, with its list display, diverts attention; and the latter, with its trackball, increases workload.

17. Key Words	18. Distribution Statement
Air Traffic Control ARTS III M&S Data Link Terminal ATC	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	150	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

A 404082

PREFACE

This report contains the results of an evaluation of the controller/computer interface studied in the Data Link Experiments Program managed by the Transportation Systems Center and sponsored by the FAA/SRDS. The primary objective of this evaluation effort was to test three candidate control and display modes to be used by air traffic controllers in an ARTS III/M&S System modified for operation in a simulated mixed voice/digital communications environment using a scenario based on Denver Stapleton International Airport. The evaluation system was implemented and tested at the NAFEC TATF (Terminal Automation Test Facility).

The authors wish to express their appreciation to those persons who played important roles in various phases of this simulation program. Of those included below, Holly Barlage FAA/NAFEC, is singled out for a special word of thanks for not only contributing in more than one area but also for effectively managing and conducting the laboratory tests.

Program Sponsorship: William Fraser, John Talley, William Hyland, and Thomas Williamson, FAA/SRDS.

NAFEC Program Support: Donald Schlotts and Holly Barlage, FAA/NAFEC.

ARTS III M&S/DL System Functional Requirements: Robert Wiseman, DOT/TSC.

ARTS III M&S Software and I/O Modifications: Dennis Kisby, Richard Connett, and Kay Freitag, Sperry Univac, St. Paul MN.

Data Reduction: John Royal, Kentron of Hawaii, Ltd.

Traffic Sample Design and Development: Mark Connelly, MIT/ESL; Robert Wiseman, DOT/TSC; and Holly Barlage, FAA/NAFEC.

System Support: John Lybeck, Sperry Univac, St. Paul MN.

Controller Operations: Holly Barlage, Donald Bottomly, and James Akers, FAA/NAFEC.

METRIC CONVERSION FACTORS

Approximate Conversions to Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
in	inches	2.5	centimeters	cm
ft	feet	30	centimeters	cm
yd	yards	0.9	meters	m
mi	miles	1.6	kilometers	km
		AREA		
in^2	square inches	6.5	square centimeters	cm^2
ft^2	square feet	0.09	*square meters*	m^2
yd^2	square yards	0.8	square meters	m^2
mi^2	square miles	2.6	square kilometers	km^2
	acres	0.4	hectares	ha
		MASS (weight)		
oz	ounces	28	grams	g
lb	pounds	0.45	kilograms	kg
	short tons (2000 lb)	0.9	tonnes	t
		VOLUME		
tsp	teaspoons	5	milliliters	ml
Tbsp	tablespoons	15	*milliliters*	*ml*
fl oz	fluid ounces	30	milliliters	ml
c	cups	0.24	liters	l
pt	pints	0.47	liters	l
qt	quarts	0.95	liters	l
gal	gallons	3.8	*liters*	*l*
ft^3	cubic feet	0.03	cubic meters	m^3
yd^3	cubic yards	0.76	cubic meters	m^3
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5/9 (after subtracting 32)	Celsius temperature	°C



Approximate Conversions from Metric Measures

Symbol	When You Know	Multiply by	To Find	Symbol
		LENGTH		
mm	millimeters	0.04	inches	in
cm	*centimeters*	*0.4*	*inches*	in
m	meters	3.3	feet	ft
m	meters	1.1	yards	yd
km	kilometers	0.6	miles	mi
		AREA		
cm^2	square centimeters	0.16	square inches	in^2
m^2	square meters	1.2	square yards	yd^2
km^2	*square kilometers*	*0.4*	square miles	mi^2
ha	hectares (10,000 m^2)	2.5	acres	
		MASS (weight)		
g	grams	0.035	ounces	oz
kg	kilograms	2.2	pounds	lb
t	tonnes (1000 kg)	1.1	short tons	
		VOLUME		
ml	milliliters	0.03	fluid ounces	fl oz
l	liters	2.1	pints	pt
l	liters	1.06	quarts	qt
l	liters	0.26	gallons	gal
m^3	cubic meters	35	cubic feet	ft^3
m^3	cubic meters	1.3	cubic yards	yd^3
		TEMPERATURE (exact)		
°C	Celsius temperature	9/5 (then add 32)	Fahrenheit temperature	°F

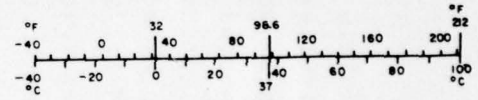


TABLE OF CONTENTS

Section		Page
1	INTRODUCTION	1
2	TEST OBJECTIVES	3
3	ASSUMPTIONS AND CONSTRAINTS	4
	3.1 ARTS III M&S System Modified for Data Link	4
	3.1.1 Basic ARTS III M&S System	4
	3.1.1.1 Full Data Block	6
	3.1.1.2 Tabular List	6
	3.1.2 Modifications for Data Link	6
	3.1.2.1 Data Link Identifiers and Pilot Response Indicators	7
	3.1.2.2 Data Link Keyboard Functions	8
	3.1.2.3 Computer-Generated Voice	8
	3.2 Denver Stapleton Airspace and M&S Geometry	9
	3.3 Control Mode Selection	12
4	System Description	14
	4.1 General	14
	4.2 Operational Description	14
	4.2.1 Keyboard Functions	14
	4.2.1.1 Dispatch	14
	4.2.1.2 Disapprove	16
	4.2.1.3 Disable/Enable	16
	4.2.1.4 VRS Control	17
	4.2.1.5 Handoff-to-Tower Command	17
	4.2.1.6 Control Mode Selection	17
	4.2.2 Control Modes	17
	4.2.2.1 Control-by-Approval/Full Data Block/Voice Response System	18
	4.2.2.2 Control-by-Approval/Tabular List/Voice Response System	21
	4.2.2.3 Control-by-Exception/Full Data Block/Voice Response System	23
	4.2.3 Pilot Keyboard and Display	23

TABLE OF CONTENTS (CONTINUED)

Section		Page
5	TEST CONDUCT	25
	5.1 Controller Procedures	25
	5.2 Controller Observer Procedures	25
	5.3 Pilot Procedures	27
	5.4 Test Data	28
	5.5 Traffic Samples	30
6	TEST DESIGN	35
7	SUBJECTIVE DATA ANALYSIS	38
	7.1 Pre-Experiment Questionnaire	38
	7.2 Post-Training Questionnaire	42
	7.3 Post-Testing Questionnaire	44
	7.3.1 A-Series	44
	7.3.2 B-Series	50
	7.3.3 C-Series	56
	7.4 Post-Series Questionnaire	63
	7.5 Summary Questionnaire	65
8	ADDITIONAL CONTROLLER COMMENTS	67
9	OBJECTIVE DATA ANALYSIS	70
	9.1 Dynamic Time Line Measures	70
	9.1.1 Results of Linear Model Analysis for Command Initiation Delay for Data Link Aircraft	76
	9.1.2 Results of Linear Model Analysis for Message Transaction Time for Data Link Aircraft	78
	9.1.3 Results of Linear Model Analysis for Command Initiation Delay for Voice Link Aircraft	79
	9.1.4 Results of Linear Model Analysis for Message Transaction Time for Voice Link Aircraft	80
	9.2 End of Run Summary Measures	85
	9.3 Other Measures	92
	9.4 Analysis of Voice Run Data	96
	9.4.1 Results of Linear Model Analysis for Command Initiation Delay for Voice Runs	96

TABLE OF CONTENTS (CONTINUED)

Section		Page
	9.4.2 Results of Analysis of Variance for Message Transaction Time for Voice Runs	97
10	ANALYSIS OF RESULTS	100
11	SUMMARY OF RESULTS	116
	11.1 Mode Rankings	116
	11.2 Control and Display Preferences	116
	11.3 Operational Benefits of Data Link	117
	11.4 Other Information	118
12	CONCLUSIONS	120
13	RECOMMENDATIONS	121
	REFERENCES	122
	APPENDIX - TIMELINES FOR EACH CONTROL MODE	123

LIST OF ILLUSTRATIONS

Figure		Page
1	M&S Command Display in Full Data Block.....	5
2	Command Display in Tabular List.....	5
3	Denver Metering and Spacing Geometry (Runway 26L)....	10
4	Controller/Computer Terminal TATF Simulating.....	15
5	Command Life Line.....	19
6	ARTS Display of M&S Commands.....	19
7	Controller Display in CBA/FDB/VRS Mode.....	20
8	Controller Display in CBA/TAB/VRS Mode.....	22
9	Pilot Display Presentation.....	24
10	End-of-Run Summary Printout.....	26
11	Summary Printout of Performance Measures.....	31
12	Time Response of Selected Traffic Sample.....	33
13	Illustration of the Transformation of Dynamic Measures to Eliminate Dependency.....	72
14	General Pattern of Interaction Effect.....	83
15	Data Reduction Program Summary Printout.....	86
16	Voice Link MTT and CID at Full Load - 60 Minutes.....	109
17	Data Link MTT and CID at Full Load - 60 Minutes.....	109
18	Combined Data Link and Voice Link MTT and CID at Full Load - 60 Minutes.....	112
A-1	CBA/Tabular List/VRS-Data Link Aircraft.....	124
A-2	CBA/Tabular List/VRS-Voice Link Aircraft.....	125
A-3	CBA/Full Data Block/VRS-Data Link Aircraft.....	126
A-4	CBA/Full Data Block/VRS-Voice Link Aircraft.....	127
A-5	CBE/Full Data Block/VRS-Data Link Aircraft.....	128

LIST OF ILLUSTRATIONS (CONTINUED)

Figure		Page
A-6	CBE/Full Data Block/VRS-Voice Link Aircraft.....	129
A-7	CBA/TAB/VRS-Data Link Aircraft - Unable or Fail Exception Case.....	130
A-8	CBA/TAB/VRS-Data Link Aircraft - Timeout Exception Case.....	131
A-9	CBA/Full Data Block/VRS-Data Link Aircraft - Unable, Fail, and Timeout Exception Cases.....	132
A-10	CBA/Tabular List/Voice-Voice Link Aircraft.....	133
A-11	CBA/Full Data Block/Voice-Voice Link Aircraft.....	134

LIST OF TABLES

Table		Page
1	RUN SCHEDULE	37
2	TABULATION OF RESPONSES TO PRE-EXPERIMENT QUESTION 5	41
3	REDUCED RESPONSES TO POST-TEST QUESTION A-1.B	46
4	REDUCED RESULTS OF POST-TEST QUESTION A-2.B	49
5	ACCUMULATED RANKS FROM POST-TEST QUESTION B-2	52
6	SUMMARIZATION OF RESPONSES TO SUMMARY QUESTION 1	65
7	RESULTS OF LINEAR MODEL ANALYSIS FOR SLOPES	84
8	RESULTS OF CORRELATION STUDY FOR END-OF-RUN SUMMARY MEASURES	87
9	RESULTS OF ANALYSIS OF VARIANCE FOR END-OF-RUN SUMMARY DATA	89
10	TABULATION OF CONTROLLER GENERATED COMMANDS	92
11	REDUCED *CONFLICT* DATA	94
12	AVERAGE RUN VALUES BY MODE FOR VOICE RUNS	99
13	MODE RANKINGS BASED ON POST-EXPERIMENT QUESTIONS A.1-A AND A.1-B	101
14	OVERALL RANKINGS BASED ON SUMMARY QUESTION 1	102
15	SUMMARY TABLE OF OVERALL MODE RANKINGS	102
16	MODE RANKING BY SUMMARY QUESTION 1 COMPONENTS	105
17	MTT AT MAXIMUM LOAD FOR DATA LINK AND VOICE LINK	105
18	MODE RANKING BY SHORTEST MTT	106
19	CID AT MAXIMUM LOAD FOR DATA LINK AND VOICE LINK	107
20	MODE RANKING BY SHORTEST CID	108
21	MODE RANKING BY AVERAGE QUEUE SIZE	111
22	GATE ARRIVAL TIME ERRORS	113

LIST OF TABLES (CONTINUED)

Table		Page
23	UNDETECTED CONFLICTS	114
24	AVERAGE NUMBER OF CONFLICTS PER RUN - VOICE-ONLY VERSUS DATA LINK	115

EXECUTIVE SUMMARY

S.1 BACKGROUND

The Department of Transportation/Transportation Systems Center (TSC) under the sponsorship of the FAA/SRDS has been engaged in a Data Link Experiments Program aimed at evaluating: (1) the controller/computer interface in a digital data link communications environment; (2) characteristics of a VHF data link between air and ground stations; and (3) cockpit I/O devices that take advantage of the automatic capability of a data link channel. This report is concerned only with the controller/computer interface effort, designated Phase 2B, which was preceded by earlier preliminary work under Phase 2A at the NAFEC Digital/Simulation Facility. The Phase 2B experiment was conducted at the NAFEC Terminal Automation Test Facility (TATF) with an ARTS III M&S system modified for operation with a simulated data link and a voice link utilizing computer-generated voice. The primary objective of the Phase 2B experiment was to ascertain which of the following three candidate control modes identified in the Phase 2A experiment is the most acceptable to air traffic controllers for the display and dispatch of M&S commands:

1. Control-by-Approval/Full Data Block (CBA/FDB)
2. Control-by-Approval/Tabular List (CBA/TAB)
3. Control-by-Exception/Full Data Block (CBE/FDB)

In the two CBA modes, the controller must approve each command; in the CBE/FDB mode, the controller intervenes only to disapprove commands which would otherwise be automatically dispatched. In the CBA/FDB and CBE/FDB modes, up to three commands can be displayed simultaneously in the third line of the ARTS data block; in the CBA/TAB mode, commands are displayed one per line in a list.

S.2 PROCEDURE

Each of six NAFEC air traffic control specialists tested each of the three data link control modes against three traffic mixes consisting of 20%, 50%, and 80% data link aircraft in a scenario based on Denver Stapleton International Airport. In addition, each controller tested the two control-by-approval modes without the simulated data link and voice response system so that potential improvements obtainable via the data link modes could be compared with "pure" voice mode data.

A total of 66 1-hour test runs were conducted during which both controllers and pilots were watched by test observers. Each controller received a 1-hour training run immediately before testing a different control mode. The schedule for each controller's test run was carefully arranged in a master test matrix to balance out sequence effects. The data from each test run consisted of controller questionnaire responses and data reduced and processed by extraction programs which measured the following:

1. Message transaction time, command initiation delay, instantaneous aircraft load, etc.
2. Counts of the number of voice and data link commands, extemporaneous communications, handoffs, tag repositionings, etc.
3. Aircraft position for subsequent input to conflict detection programs.

The traffic samples, consisting of approximately 30 aircraft, were designed to subject the controller to a linear increase in instantaneous aircraft load from 0 to 9 aircraft during the 1-hour test period. This facilitated the continuous measurement of the controller's communication delays, which were expected to increase until his workload limit was reached and then increase rapidly thereafter.

S.3 RESULTS

On the basis of controller preference as manifested in the questionnaires, the "hands off" CBE/FDB mode was ranked first, CBA/TAB, second, and CBA/FDB, last. CBE/FDB had the longest message transaction time (MTT) of the three modes but was ranked first by the controllers, despite some concern for blunder potential, because it possessed the best workload, capacity, and stress characteristics. Longer MTT's are inherent in CBE/FDB because command initiation delay (CID) is prolonged in order to provide command disapproval time. CBA/TAB, with its single-key command dispatch had the shortest MTT, but was ranked second by the controllers because they feel operating from a command list tends to divide their attention, adversely affecting workload and capacity. CBA/FDB had the second best MTT but it was consistently rated as unfavorable in workload, capacity, stress, and blunder potential because of the necessity for continuous trackball slewing action.

S.4 CONCLUSIONS

1. The CBE/FDB mode is the most favored by the controllers despite blunder potential, but some inherent limitations result in longer message transaction times.
2. As was concluded in the Phase 2A experiment, a CBA/FDB mode in which a single command is displayed at one time in one data block for dispatch with only a single keyboard entry is again indicated as a potentially viable mode in a future system. This mode would combine the advantages of both the CBA/FDB and the CBA/TAB modes as currently implemented.
3. The Phase 2B computer-generated voice system was not favorably regarded because of the talking rate and inability to repeat commands.

S.5 RECOMMENDATIONS

Further studies aimed at determining the optimum means for interfacing the controller and the computer in a data link communications environment should be conducted in the following areas:

1. Additional testing of the Control-by-Exception/Full Data Block mode of operation should be conducted with an on-line controller selectable disapproval time. This will allow a controller to decrease MTT if he feels disapproval time need not be as long (6 seconds) as it was in this study.

2. On the basis of the findings of both Controller/Computer Interface experiments (Phase 2A and 2B) it is recommended that a CBA/FDB mode (in which a single command is displayed at one time in an aircraft data block for dispatch with only a single keyboard entry) be incorporated in a future system.

3. Considering the advantage of computer-generated voice in providing commonality of command dispatch in a mixed voice/digital communications environment, it is recommended that future research be devoted to improving the talking rate and in developing a command repeat capability.

1. INTRODUCTION

The Department of Transportation/Transportation Systems Center (TSC), under the sponsorship of the FAA/SRDS, has been engaged in a Data Link Experiments Program aimed at evaluating (1) the controller/computer interface in a digital data link communications environment, (2) characteristics of a VHF data link between air and ground stations, and (3) cockpit I/O devices that take advantage of the automatic capability of a data link channel. This report is concerned with the evaluation of the controller/computer interface portion of the program designated as Phase 2B. The Phase 2B evaluation system was implemented at the NAFEC Terminal Automation Test Facility and consisted of the ARTS III/M&S system modified for operation with a simulated data link and a voice link utilizing a voice response system (VRS). The evaluation system was operated in three controller modes for the display and dispatch of computer-generated M&S commands and the display of aircraft responses. The three modes were Control-by-Approval/Full Data Block, Control-by-Approval/Tabular List, and Control-by-Exception/Full Data Block. In Control-by-Approval, the controller must approve each command; in Control-by-Exception, the controller intervenes only to disapprove commands which would otherwise be automatically dispatched. In the Full Data Block, up to three commands are displayed simultaneously in the third line of the ARTS data block; in the Tabular List, they are displayed one per line in a list.

It should be pointed out that the selection of the three data link control modes for evaluation was based on the results of a previous controller/computer interface experiment conducted at the NAFEC Digital Simulation Facility (DSF). In the DSF experiment, an ARTS III/M&S system operating within a mixed voice/digital communications environment was simulated and evaluated by air traffic controllers for the purpose of identifying the most promising of seven control modes to be subsequently tested in a real ARTS III/M&S system at the NAFEC TATF.

This report describes the results of a program of evaluation using air traffic controllers and simulation pilots "flying" aircraft created by the ARTS internal target generator. Each controller evaluated each of the three control modes in an M&S scenario simulating south arrivals, single runway only, at Denver Stapleton International Airport. In addition to operating the modified ARTS III M&S system with the three data link control modes, each controller operated the system in the Control-by-Approval/Full Data Block and Control-by-Approval/Tabular List modes without benefit of a simulated data link channel or voice response system. In this manner, performance data were obtained for the "pure" voice modes against which potential improvements obtainable via the data link modes could be compared. Three different traffic samples, each having a different ratio of data link to voice link aircraft, were used in the evaluation of each of the data link modes; a different traffic sample was used for each of the two voice modes.

2. TEST OBJECTIVES

The primary objective of this test program was to evaluate three candidate control and display modes to be used by air traffic controllers in an ARTS III/M&S system modified for operation in a mixed voice/digital communications environment using a simulation based on Denver Stapleton Internation Airport. The ARTS III/M&S system was used in this evaluation program because it provided a high volume of the automatically generated ATC messages that comprise a large portion of the controller's total communications workload. The high message rates obtainable with the M&S system permitted a meaningful comparison of controller reaction to each of the three control modes with respect to the display, dispatch, and receipt of data link messages. Analysis of the test results provided subjective and objective data pertaining to the following:

a. The controller's ability to deal effectively with increasing numbers of data link aircraft in traffic situations involving a high volume of control messages.

b. The control mode which is most acceptable to the controller and enhances his traffic handling capacity without degrading his ability to evaluate commands.

3. ASSUMPTIONS AND CONSTRAINTS

As previously stated, this test program was the second stage of a two-stage effort involving TSC and NAFEC. The first stage was conducted at the NAFEC DSF using a simplified model of the metering and spacing traffic pattern in a simulated ARTS III system. The assumptions and constraints governing the conduct of this, the second stage, were as follows:

a. The actual ARTS III prototype M&S system (November 1973 version), modified to provide data link display identifiers and keyboard functions, would be used at the NAFEC Terminal Automation Test Facility (TATF).

b. The Denver Stapleton Airspace Geometry used in previous M&S development testing would be used, but terminal area traffic would be limited to south arrivals through two feeder fixes (Shawnee and Elizabeth) onto a single runway (26L) handled by one air traffic controller.

c. The control modes to be tested would be the few most promising ones identified in the first stage of testing at the DSF.

Each of the above is discussed in more detail below.

3.1 ARTS III M&S SYSTEM MODIFIED FOR DATA LINK

3.1.1 Basic ARTS III M&S System

The existing ARTS III M&S system replaces the controller's mentally calculated heading, altitude, and speed commands with those that are automatically generated by a computer algorithm that regulates traffic flow in the entire terminal area. Commands generated by the algorithm are displayed on the ARTS III Data Entry and Display Subsystem (DEDS) in either the full data block or in a tabular list for voice delivery from controller to pilot. The two methods of displaying commands are described below and depicted in Figures 1 and 2.

A —— EA123 160 25	(FDB without M&S Command in third line)
A —— EA123 160 25 H200	(Fly heading 200°)
A —— EA123 160 25 A120	(Descend and maintain 12,000 feet)
A —— EA123 160 25 H180A100	(Fly heading 180°, descend and maintain 10,000 feet)

NOTE: The second line of the data block is ARTS III display of altitude and ground speed. A is the controller symbol and denotes aircraft position. Blinking commands are encircled with dots.

Figure 1. M&S Command Display in Full Data Block

EA123	S120	(Reduce Speed 120 Knots)
AA155	H180	(Fly heading 180°)
AL924	A080	(Descend and maintain 8000 feet)
TW331	H270	(Fly heading 270°)

Figure 2. Command Display in Tabular List

3.1.1.1 Full Data Block - Combinations of M&S heading, altitude, and speed commands are displayed blinking in the third line of the ARTS III data block. The commands appear blinking for an interval during which they can be voiced by the controller for execution by the aircraft. After the blink interval has elapsed, the commands stop blinking, briefly remain on the display, and are then automatically erased. The controller can use the trackball with a special function key to erase the command after voicing it to the pilot. By so doing, he can better distinguish pending commands from those that he had voiced.

3.1.1.2 Tabular List - Heading, altitude, or speed commands are displayed one at a time, in a separate list. The command timing and voice delivery are similar to those used in the full data block, except that the controller can erase only the topmost command in the list. In such instances, all other commdands would be moved up one line.

3.1.2 Modifications for Data Link

The existing ARTS III M&S system display presentation and operational features were carefully evaluated in order to define the means by which the system would have to be modified (Reference 1) to make addition of a data link capability to the controller's M&S communications environment operationally feasible and acceptable. In determining what modifications were required, this program benefited significantly from the large body of original work done in this area in the design, implementation, and testing of the DSF experimental system.

The principal areas of modification identified included the following:

a. Designation of data link equipped aircraft
b. Indication of pilot responses to M&S commands
c. Incorporation of function keys for command selection
d. Incorporation of computer-generated voice system for voicing commands to non-data link aircraft.

3.1.2.1 Data Link Identifiers and Pilot Response Indicators - The matter of identifying data link equipped A/C was handled by placing an asterisk in the first line of the data block after the ARTS III aircraft identity. Initially, this presented a problem because the aircraft identity field is limited to seven characters. However, it was easily resolved by restricting aircraft identities to five characters and using the 6th character position for the asterisk designating data link. The remaining 7th character position was thus available for the mutually exclusive data link indications of pilot compliance (W), noncompliance (U), nonresponse (X), and link failure (F). The conditions under which these indications were displayed were as follows:

a. A solid W was displayed after a programmed time delay following controller dispatch of a command to a data link aircraft. This delay, which varied ± 3 seconds around a mean of 5, was automatically generated within the ARTS III IOP in order to simulate the pilot's recognition of the command on his cockpit display and his subsequent depression of a WILCO button. Upon display of the W, the command that had been WILCO'd was erased. The W was retained on the display for 3 seconds in order to give the controller a positive indication of the aircraft's compliance.

b. The characters U, F, or X were displayed in a blinking data block if either the pilot was unable to comply (U) with the controller-dispatched command, a failure (F) occurred in the data link, or the pilot did not respond in time (X) to the command. In such cases, both the indicator and the command (nonblinking) would be retained on the display for a period of time (20 seconds) sufficient for the controller to resolve the problem on the voice link.

3.1.2.2 Data Link Keyboard Functions - The experience gained from the DSF experiment provided the basis for determining that the controller would have to be provided with the following function keys in order to make operation of the ARTS III M&S system feasible with a data link:

a. Dispatch - Controller approval of M&S generated commands for both data link and voice link aircraft

b. Disapprove - Controller manual override of M&S generated commands

c. Disable - Temporarily inhibiting the use of M&S commands for a specific aircraft

d. Enable - Restoring the use of M&S commands for an aircraft against which a DISABLE action was taken.

e. Halt - Temporary interruption in the operation of the voice response system

f. Resume - Restoration to operation of the voice response system after being temporarily halted.

It should be noted that the Disapprove, Disable, and Enable keyboard functions could not be used in this experiment because of the inability of the M&S algorithm to dynamically reschedule and resequence aircraft or to update commands for aircraft.

3.1.2.3 Computer-Generated Voice - The other major necessary modification was interfacing the ARTS III M&S system's computer-generated commands with a Voice Response System (VRS) (Reference 2) which would enunciate commands that would otherwise be voiced by the controller. The aim here was to permit the controller to use the same function key for dispatching both voice link and data link commands, thereby providing some commonality in operating procedures in a mixed voice/digital communications environment. Use of the voice response system would also provide the added advantages of speech standardization and, more importantly, off loading of the oral communications workload of the controller to the extent that he might have an opportunity to consider other commands while the voice response system was speaking for him.

3.2 DENVER STAPLETON AIRSPACE AND M&S GEOMETRY

The original airspace structure employed in the development of the ARTS III M&S system (Reference 3) was based on that existing at Denver Stapleton International Airport. As shown in Figure 3, aircraft enter the terminal area through three pairs of feeder fixes. Each pair of feeder fixes funnel the aircraft into an associated inner fix from which they are directed onto a final approach path, through the approach gate, and on to the runway. Traffic flow in the entire approach area is regulated by the M&S algorithm on the basis of flight plan data, radar position, and aircraft performance data. Regulation is achieved by automatic sequencing of aircraft through the feeder fixes and thereafter by path stretching and speed control.

Depending on type, aircraft are scheduled through either a high performance or low performance gate via associated downwind legs. Since only a limited amount of regulation can be achieved by path stretching or speed control, almost all arrivals must be held at the feeder fix to derandomize high density flow.

The algorithm calculates a feeder fix departure time for each aircraft with the expectation that the aircraft will depart within ± 1 minute of that time. The departure is carried out under direct control of the enroute controller but is seldom on schedule because of the algorithm's propensity to change the departure time at any moment.

Careful evaluation of the airspace and the operation of the algorithm within the airspace resulted in the identification of several problems that would have a deleterious effect on the data link application. These problems and the approaches to circumventing them are as follows:

a. The northern portion of the M&S geometry had critical vectoring areas in which the algorithm had difficulty spacing and sequencing aircraft. Therefore, only that portion of the airspace involving south arrivals onto runway 26L through the Shawnee and Elizabeth feeder fixes was used in this experiment.

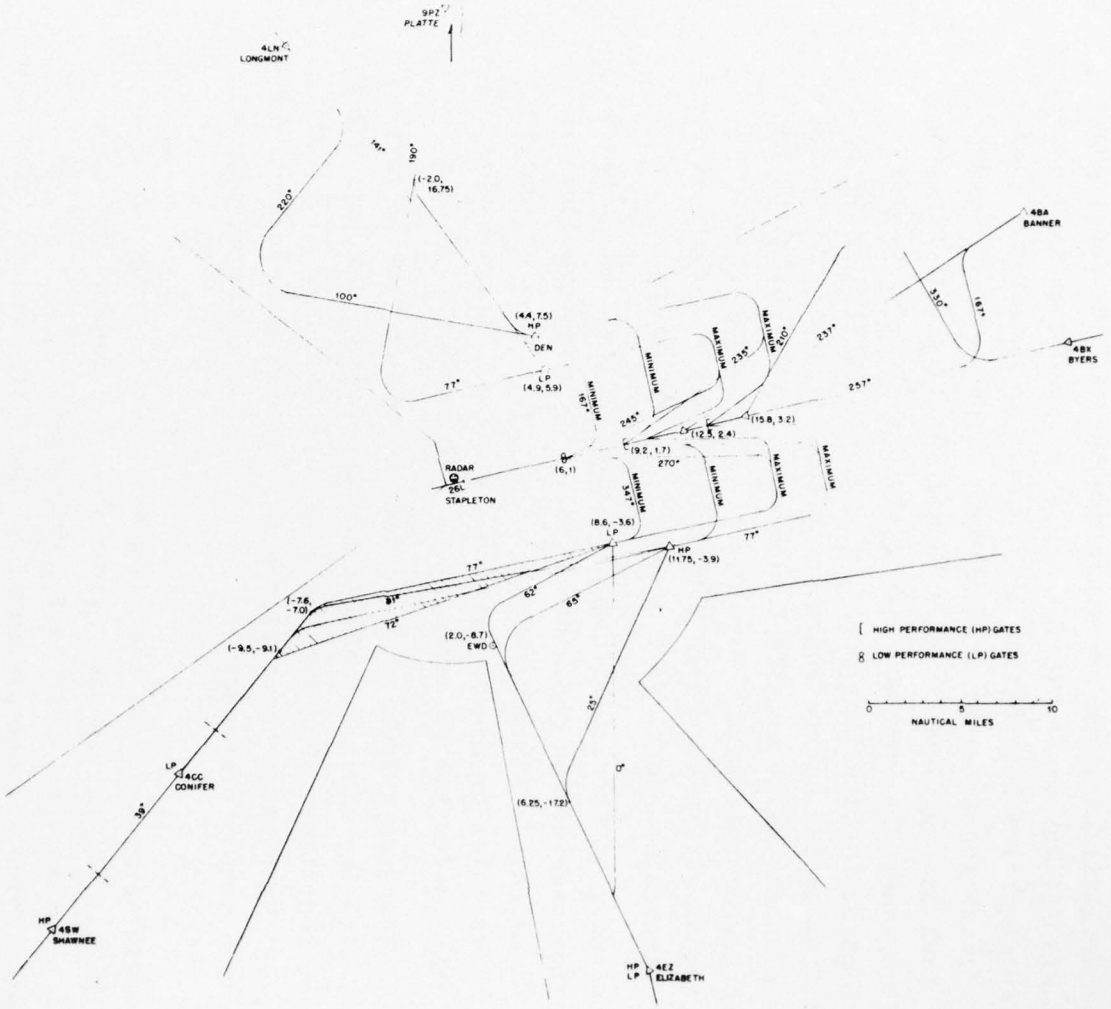


Figure 3. Denver Metering and Spacing Geometry (Runway 26L)

b. The M&S program issued constant time updates in the holding patterns, thus making it difficult to depart an aircraft from the feeder fix within ± 1 minute of a constantly changing departure time. Furthermore, since the fixes are outside the M&S geometry, commands for data link aircraft in holding patterns could not have been dispatched or voiced. Holds were therefore eliminated by careful structuring of traffic samples.

c. By virtue of (a) and (b) above, the role of the enroute controller in the scheduling of feeder fix departures, as called for in the basic ARTS III M&S, was eliminated.

d. The resequencing portion of the M&S program was not able to resequence missed approaches back into the flow of traffic. Subject controllers were therefore instructed to keep aircraft in sequence; otherwise, the algorithm would malfunction, causing conflicts, large spacing gaps, and holds. In addition, provision was made for automatic acquisition of the ILS localizer by simulated aircraft after the final heading command had allowed the aircraft to intercept the localizer.

e. If M&S commands were not followed by controllers, holding would be created and aircraft would get out of sequence. Controllers were therefore instructed to issue all commands, not to maneuver aircraft to avoid conflicts, and not to use the Disapprove and Disable/Enable keyboard functions. They were also instructed to notify observers of situations in which they (the controllers) felt they would normally have intervened. These situations were recorded for future analysis.

f. In the basic ARTS III M&S, the M&S path-stretching geometry was not readily discernible to the controller. Therefore, in this experiment, a video map overlay was included on the controller's display to provide cues on the geometry.

3.3 CONTROL MODE SELECTION

The candidate control modes evaluated in this program were selected on the basis of the findings from a previous controller/computer interface experiment conducted at the NAFEC Digital Simulation Facility (DSF). In the DSF experiment, a simplified version of the ARTS III/M&S system was simulated in order to evaluate seven candidate modes by which air traffic controllers could dispatch automatically generated commands to voice link and data link equipped aircraft in terminal approach traffic. The seven modes tested at the DSF were:

a. Control-by-Approval, Full Data Block, Multiple Command, Voice.

b. Control-by-Approval, Full Data Block, Multiple Command, Computer-Generated Voice.

c. Control-by-Approval, Tabular List, Multiple Command, Voice

d. Control-by-Approval, Tabular List, Multiple Command, Computer-Generated Voice.

e. Control-by-Approval, Tabular List, Single Command, Voice

f. Control-by-Approval, Tabular List, Single Command, Computer-Generated Voice.

g. Control-by-Exception, Full Data Block, Multiple Command, Computer-Generated Voice.

The above modes represented various combinations of the following design alternatives:

1. Display of commands in the full data block or the tabular list.

2. Presentation of active commands one at a time (single) or in multiple fashion.

3. Communication to voice link aircraft by computer-generated voice or by controller voice.

4. Controller approval and dispatch of each command or automatic dispatch without prior controller approval.

The results of the DSF experiment showed that the most promising candidate modes of operation for follow-on testing in the operational version of the real ARTS III M&S system at the NAFEC TATF were as follows:

a. Control-by-Exception, Full Data Block, Multiple Command, Computer Generated Voice.

b. Control-by-Approval, Tabular List, Single Command, Computer-Generated Voice.

c. Control-by-Approval, Full Data Block, Multiple Command, Computer-Generated Voice.

d. Control-by-Approval, Full Data Block, Single Command, computer-Generated Voice.

Although not tested at the DSF, the fourth mode was recommended because it represented all of the alternatives that the test data indicated as being most desirable, particularly the dispatching of commands one at a time without the requirement for time-consuming trackball action. However, only the first three modes were implemented and evaluated in the real ARTS III M&S system at the TATF. The fourth was not included because software modifications would have been required to provide a single command, full data block capability in the M&S program, and it was felt that the design of the existing system should not be altered to accommodate a new candidate mode.

4. SYSTEM DESCRIPTION

4.1 GENERAL

In the real world, an ARTS III M&S System modified for operation with a data link channel would generate and transfer heading, altitude, and speed commands to a data link ground system (DLGS). The DLGS would uplink the commands to the specified data link equipped aircraft and would accept downlinked responses from the pilot. These downlink responses would then be transferred to the ARTS III M&S system for display to the controller.

In the modified ARTS III M&S system implemented in the Terminal Automation Test Facility at NAFEC (Figure 4) all functions other than the controller/computer interface are simulated. All target information is provided by the ARTS internal target generator. An ARTS display with two data entry devices is used by "pilots" to make necessary target maneuver entries for non-data link equipped aircraft in response to commands enunciated by either the controller or by a computer-generated voice system. Responses by simulated data link equipped aircraft to controller-dispatched commands are automatically generated, require no pilot keyboard actions, and are presented to the controller in the appropriate data blocks on the ARTS III Data Entry and Display System (DEDS).

4.2 OPERATIONAL DESCRIPTION

4.2.1 Keyboard Functions

The keyboard entries incorporated in the ARTS III DEDS for use by the controller in data link operations are:

4.2.1.1 Dispatch - The dispatch entry is used in the Control-by-Approval/Full Data Block (CBA/FDB) and Control-by-Approval/Tabular List (CBA/TAB) modes to indicate approval of an M&S generated heading, altitude, or speed command. In the CBA/FDB mode, the Dispatch entry entails slewing the trackball to the selected target symbol on the display and depressing DEDS function 16 key. When

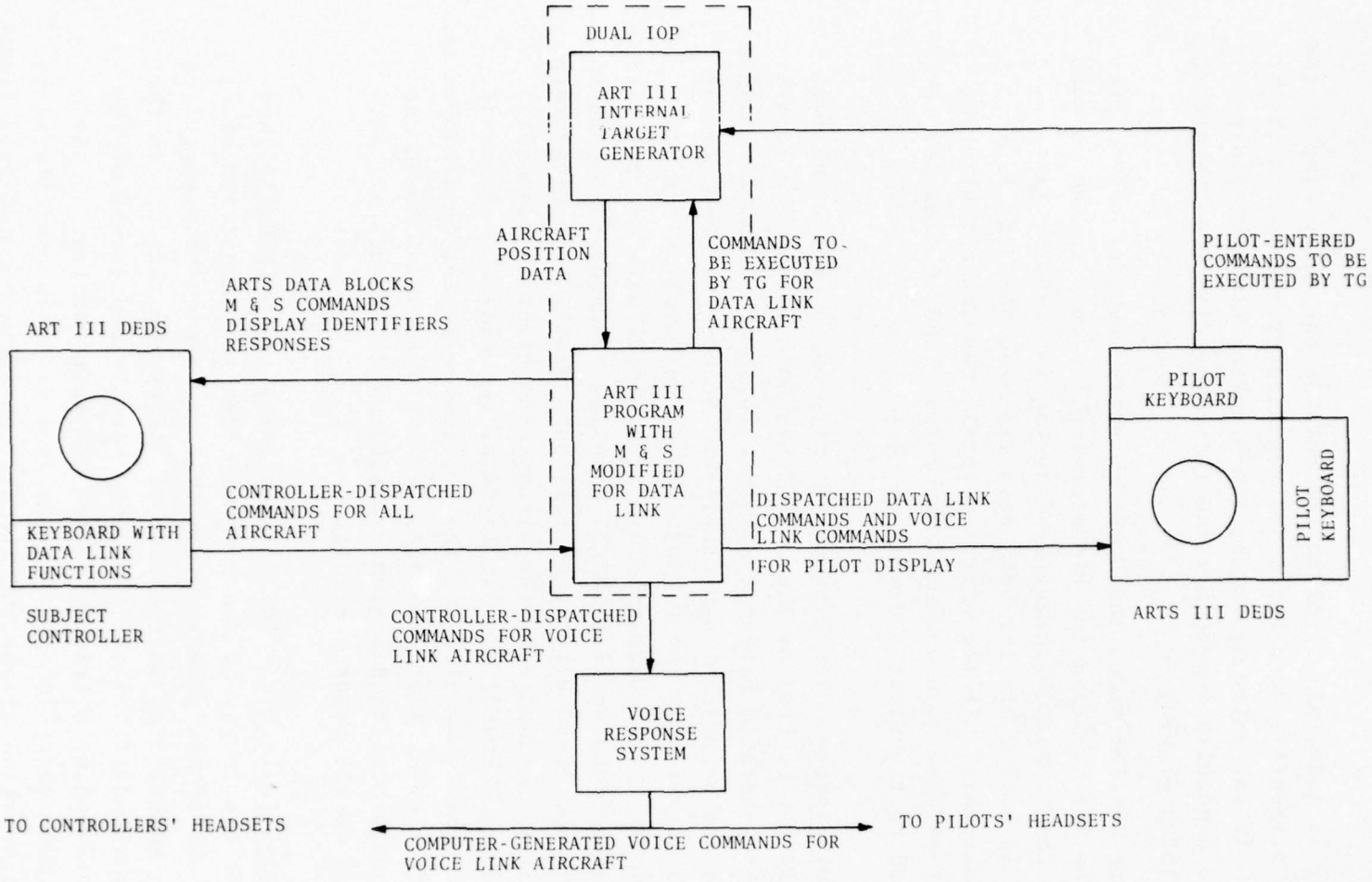


Figure 4. Controller/Computer Terminal TATF Simulating

this entry is made, all blinking commands in the third line of the specified aircraft's data block are regarded by the system as approved. If the entry is legal, the blinking command(s) are changed to a nonblinking format and the command(s) are transferred to the pilot's display.

In the CBA/TAB mode, the Dispatch entry requires only depression of the F16 key, and the Dispatch applies only to the command occupying the topmost blinking position in the tabular list. The entry is acceptable for both data link and non-data link equipped aircraft. If the entry is legal, the topmost blinking command is changed to a nonblinking format, and the command is transferred to the pilot's display.

4.2.1.2 Disapprove - The Disapprove entry was incorporated for use in only the Control-by-Exception/Full Data Block mode to override an M&S generated heading, altitude, or speed command. However, its use in the experiment was disallowed because of shortcomings in the M&S algorithm. The Disapprove action involves the same keyboard entry as for Dispatch (i.e., trackball slew and depression of the F16 key) and applies to all blinking commands in the third line of the specified aircraft's full data block. The Disapprove entry results in erasure of the disapproved blinking command(s). If no blinking commands are displayed at the time the action is taken, the next command(s) for the selected aircraft is disapproved. In the latter case, the command(s) is displayed nonblinking and is thus ineligible for dispatch. Disapproved commands are not transferred to the pilot's display.

4.2.1.3 Disable/Enable - The Disable entry was originally incorporated for use in any of the control modes to inhibit temporarily data link use of M&S generated commands pertaining to a specified aircraft. However, as in the case of Disapprove, its use in the experiment was disallowed. The entry entails depression of the F12 key followed by a trackball slew and enter action. A subsequent F12 slew entry identifying the same aircraft constitutes an Enable action which allows the aircraft to again receive data linked

commands. Commands for a "Disabled" aircraft are displayed non-blinking, are ineligible for Dispatch, and are not transferred to pilot's display.

4.2.1.4 VRS Control - The VRS Control entry is used to temporarily interrupt the VRS, permitting controller use of the voice channel. The entry requires despression of the F13 key for either disabling or reenabling the VRS, depending on its current state as indicated in the controller's preview area by either "VRS ON" or "VRS OFF."

4.2.1.5 Handoff-to-Tower Command - In all modes employing the VRS, a Handoff-to-Tower command can be issued to both data link and voice link aircraft by means of a sequence of keyboard entries: F12, H (for Handoff), Trackball Slew, and Enter action by the controller. For voice link aircraft only, these entries will activate the VRS which will enunciate the following message preceded by the aircraft ID for the aircraft in question: "Now X miles from outer marker, cleared for ILS runway two six left, approach. Contact Denver tower on one one eight point three when over outer marker."

For voice link and data link aircraft, the aircraft's position symbol is changed to a "T," the FDB blinks for several seconds, after which the aircraft is under control of position "T."

4.2.1.6 Control Mode Selection - The Mode Selection entry is used to designate the desired mode of operation: CBA or CBE; Voice or VRS. The entry requires an F12, two characters, and Enter. Acceptable character pairs include the following:

AS for CBA/VRS
ES for CBE/VRS
AV for CBA/Voice

The method of command presentation, full data block or tabular list, is selected on the control panel of the ARTS III IOP.

4.2.2 Control Modes

The sequence of events associated with an M&S-generated com-

mand is shown in Figure 5. The command is generated with the expectation that it will be executed by the aircraft at some future time. Shortly after it is generated, the command is displayed blinking to the controller in either the third line of the data block or in a tabular list, as shown in Figure 6. In the latter case, the aircraft's FDB position symbol is also blinked when the command becomes the topmost blinking command in the list. In either command presentation mode, the blinking commands are eligible for controller Dispatch or Disapprove actions and will continue to blink until such action has been taken or issue time has been attained. When issue time has been reached in control-by-approval modes, the command cannot be dispatched; however, in control-by-exception, the command is automatically dispatched. At issue time or at the time of Dispatch (CBA Mode), the command stops blinking and is retained until termination or until a pilot (WILCO) response is received. There are different command presentation parameter values associated with each of the three control modes tested in this experiment. The three modes are discussed in detail below in terms of their respective parameter values.

4.2.2.1 Control-by-Approval/Full Data Block/Voice Response System (CBA/FDB/VRS) - In the CBA/FDB/VRS mode, the M&S command is displayed blinking in the third line of the aircraft's FDB 12 seconds before the issue time. Figure 7 shows the controller's display in the CBA/FDB/VRS mode of operation with an active altitude command for voice link aircraft AF14 and a W(WILCO) response for data link aircraft TT73. The controller dispatches the command by slewing the trackball to the appropriate controller symbol and depressing the F16 function key. The dispatched command stops blinking and is displayed with the associated ACID in a list. Two or three blinking commands for the same aircraft can be transferred simultaneously with a single dispatch action. Since, from this point on, there are different event sequences for data link and non-data link aircraft, resepctively, each is separately described below.

a. Data Link Aircraft - If the dispatched command is for a data link aircraft, a suitably delayed response (W,U,F, or X) is automatically generated and displayed in the ACID line of the

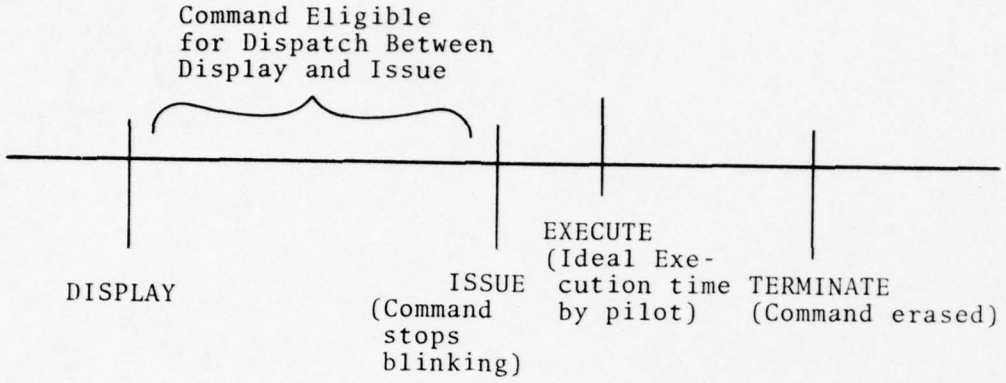


Figure 5. Command Life Line

FDB FORMAT

P AA123*
12Ø 22
H135AØ8Ø

TAB FORMAT

AA123* H135
AA123* AØ8Ø
DL12* H32Ø
EA144 S18Ø

Figure 6. ARTS Display of MQS Commands

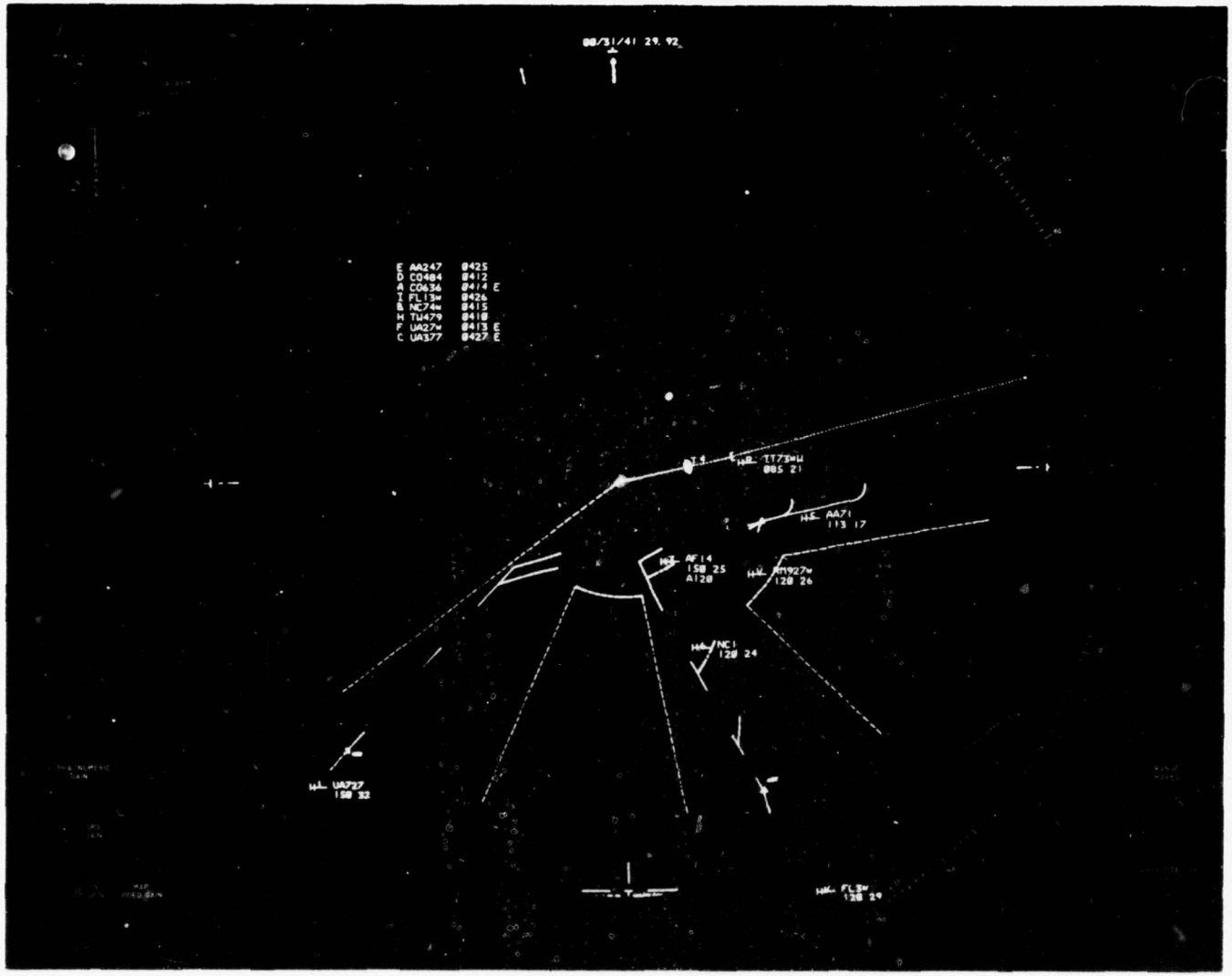


Figure 7. Controller Display in CBA/FDB/VRS Mode

appropriate FDB on the controller's display and adjacent to the appropriate ACID and command on the pilot's display. The W(Wilco) and U (Unable) delays vary ±3 seconds about a 5-second mean, with a standard deviation of 1 second, and are intended to reflect pilot reaction time, equipment delays, sampling delay, and other types of message delays. (These values are based on measurements reported in Reference 4.) An F (link failure) indication was set at one second, an X (timeout) at 14 seconds. Display of a W(Wilco) results in erasure of the command, indicating automatic execution of the command by the internal target generator. The W is erased 3 seconds after it is initially displayed. Display of a U, F, or X response results in the first two lines of the data block blinking for 14 seconds to alert the controller that he must intervene on the voice link to issue orally the vectoring command that has not been or could not be executed by the aircraft. By displaying the command and computer-generated responses for data link aircraft on the pilot's display, the pilot is able to monitor the situation and thus be prepared to communicate orally with the controller when the need arises.

b. Voice Link Aircraft - If the dispatched command is for a voice link aircraft, it is enunciated by the VRS and orally acknowledged by the pilot, who also makes a keyboard entry to allow the target generator to execute the command. The command is retained in the FDB on the controller's display for 20 seocnds after the Dispatch action. It is erased from the pilot's display at the time the pilot makes the keyboard entry.

4.2.2.2 Control-by-Approval/Tabular List/Voice Response System (CBA/TAB/VRS) - In this mode, all new commands are presented blinking at the bottom of the tabular list 12 seconds before issue time. Figure 8 shows the controller's display in the CBA/TAB/VRS mode of operation. When a command is the topmost blinking command in the list, it is eligible for dispatch, and the associated aircraft's position symbol starts blinking to provide the controller with a visual cue for command-aircraft correlation. The controller dispatches the command by merely depressing the F16 function key – a trackball

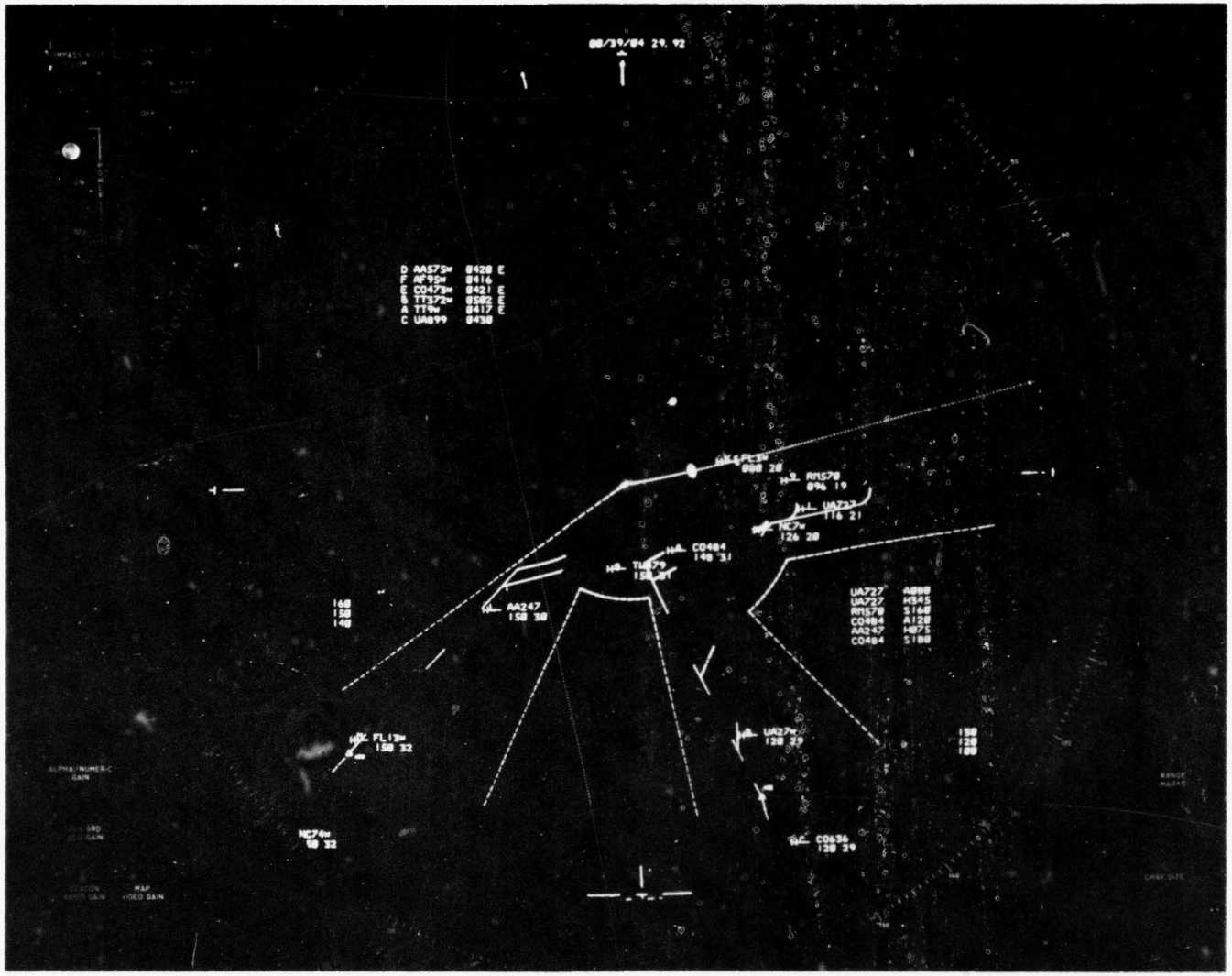


Figure 8. Controller Display in CBA/TAB/VRS Mode

slew action is not required. When the Dispatch action is taken, the command and position symbol stop blinking, and the command is transferred to the pilot's display.

The event sequence following the Dispatch action is the same as in the CBA/FDB/VRS mode; i.e., the command will be erased from the controller's list after 20 seconds for a voice link aircraft or upon display of pilot's WILCO(W) response in the ACID line of the FDB for a data link aircraft.

4.2.2.3 Control-by-Exception/Full Data Block/Voice Response System (CBE/FDB/VRS) - In this mode, commands are displayed blinking in the third line of the FDB and are automatically dispatched 6 seconds later, unless the controller makes a Disapprove entry (trackball slew, F16) during the blinking interval. In the latter case, the command is immediately erased. If the controller made a Disapprove entry with no blinking commands present, the entry will apply to the next command, which will appear nonblinking for 5 seconds before being erased. As previously mentioned in Section 4.2.1.2, use of the Disapprove entry was disallowed in the experiment.

4.2.3 Pilot Keyboard and Display

As previously mentioned, a separate Data Entry and Display Subsystem (DEDS) with two data entry devices was incorporated in the modified ARTS system to accommodate two operators functioning as simulation pilots. For each pilot, the display presented a separate list (see Figure 9) that contains the ACID and a single character identifier. When commands are dispatched by the controller, they appear in the list on the pilot display in association with the ACID and identifier. For data link aircraft, the response indicator is displayed beside the ACID. If the response is a W, the command is erased and the W retained for 3 seconds. If the response is a U, F, or X, the command is retained and the pilot can expect to engage in voice communication with the controller to resolve the problem. For voice link aircraft, the VRS enunciates the command to the pilot, and the pilot must (1) voice his compliance, and (2) enter the single character identifier on his keyboard. Upon entry of the identifier, the target generator executes the command, and the command is erased from the pilot's display.

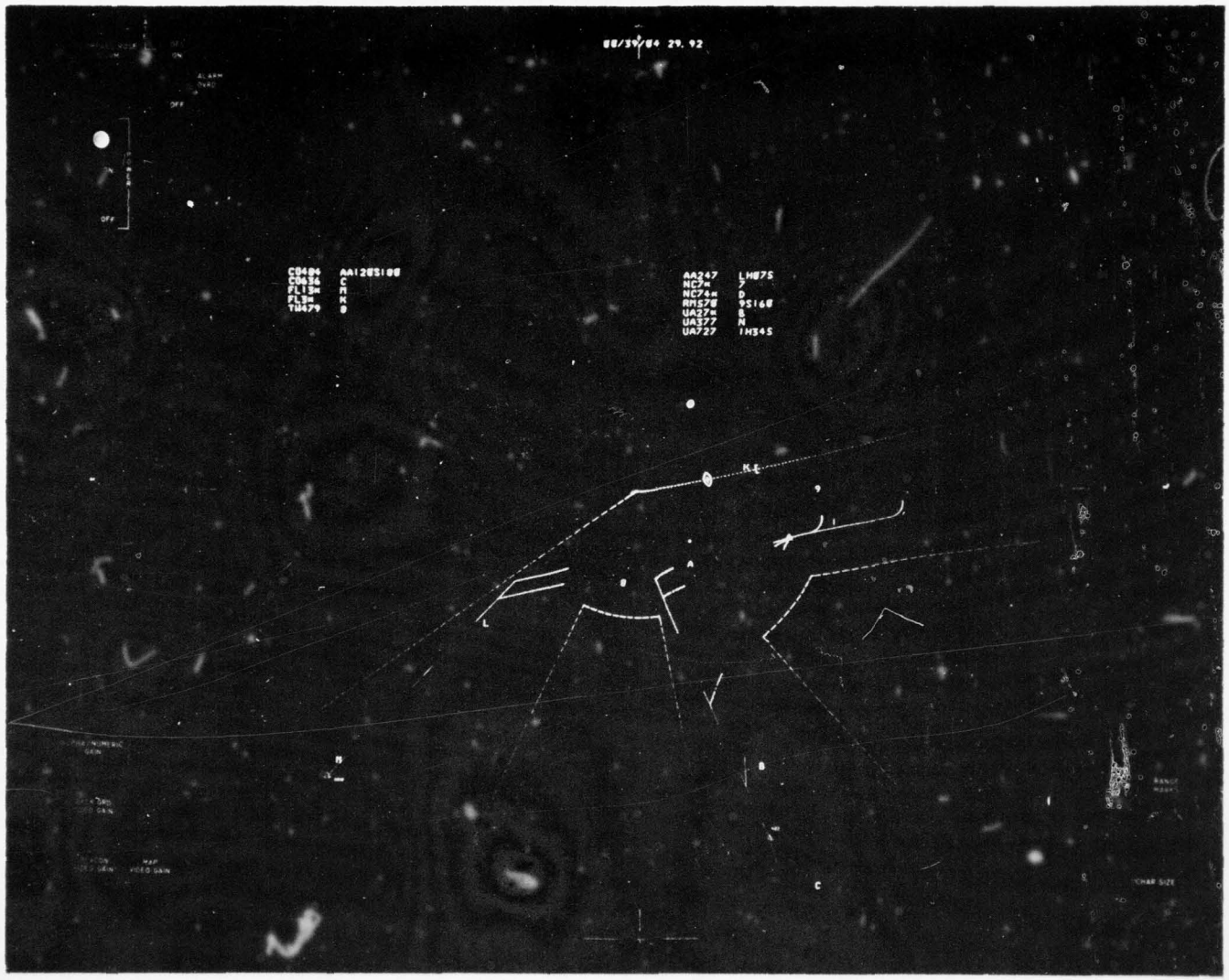


Figure 9. Pilot Display Presentation

5. TEST CONDUCT

5.1 CONTROLLER PROCEDURES

In both test series conducted, each of three controllers tested each of the three data link control modes against three traffic mixes; i.e., 20%, 50%, and 80% data link aircraft. Each controller also tested the two control-by-approval modes without the simulated data link and VRS; i.e., voice-only operation. Prior to the start of the test series, a detailed set of procedures was distributed to each of the subject controllers. The procedures explained the traffic situation and contained step-by-step instructions of the operation in each of the three modes. The subject controllers were also orally briefed on all aspects of the experiment and then allowed to observe the operation of the display during a 1-hour general training session. In addition, each subject received a 1-hour training run immediately before a set of evaluation runs with the mode of operation to be tested. The schedule for each subject's assignment to these runs was kept in strict accordance with the test matrix described in Section 6.

5.2 CONTROLLER OBSERVER PROCEDURES

During each training and evaluation run, the subject was observed by another controller familiar with the simulation design and operation. The function of the observer was to verify that an acceptable level of operation was achieved with respect to all aspects of an evaluation run. The observer made entries in a run log whenever any peculiarities were observed, equipment malfunctions occurred, or actual conflicts were noted by the controller, and when the controller failed to dispatch commands. The observer determined whether or not a test should be rerun if in his judgment a breakdown had occurred in on-line data collection because of the subject controller's failure to observe required procedures. The decision to rerun a test run was confirmed through review of an end-of-run summary printout shown in Figure 10.

DATA LINK SUMMARY CONFIGURATION—CBA-FDB---VRS RUN DURATION 1:00:44.271

AIRCRAFT INITITATED 33 CONTROLLED AT GATE 26 DALAS 16

SAM MESSAGES GENERATED

S	26	COMPONENTS	S	0	H	26	
A	1						
M	225		H	121	A	55	S 60

DAFT MESSAGES RECEIVED

D	0	COMPONENTS	ACTIVATE	0	DEACTIVATE	0
A	223		WILCO	222	UNABLE	1
F	1					
T	1					

KEYBOARD MESSAGES RECEIVED 269

DISPATCH	225	CONFIGURATION	1
DISAPPROVE	0	VRS ON	1
DISAPPROVE NEXT	0	VRS OFF	1
DISABLE	1	ALTIMETER	1
ENABLE	0	HANDOFF/STACK DEPTR	26
FORMAT OFFSET	0		

RESPONSE ANALYSIS

M COMMANDS GENERATED	DISPATCHED OR DISAPPROVED	ACKNOWLEDGE W	U	F	X	RESPONSE DISTRIBUTION - SECONDS MIN.	MAX.	AVE.
260	225					1.8	11.8	4.2
		222				2.6	13.5	5.0
			1			4.8	4.8	4.8
				1		1.0	1.0	1.0
					1			
	VRS OFF	1				8.9	8.9	8.9
	VRS TRANSFERS	135	COMPLETED	134		2.4	10.6	4.0
	PUSH TO TALK			161		.0	8.4	.7

Figure 10. End-of-Run Summary Printout

5.3 PILOT PROCEDURES

Two simulator pilots were employed in each test run to "fly" the aircraft from the horizontal ARTS DEDS. As previously mentioned, the pilot keyboard entries and verbal acknowledgements were limited to non-data link aircraft. Commands to data link aircraft were displayed on the pilot display but were automatically processed (acknowledged) by the simulation software without pilot intervention, thus eliminating uncontrolled pilot errors during a test run. The data link communications environment for the controller was simulated by incorporating the effects of the following two types of delays between dispatch and receipt of an associated response: (1) a real pilot's reaction time to a command on a cockpit I/O device, and (2) uplink and downlink transmission time of commands. The reaction time of a real pilot was simulated by a normal distribution of time delays that were activated at random by each command such that W delays would vary ± 3 seconds about a 5-second mean with a standard deviation of 1 second. Although the transmission delays were generally smaller than the W delays, they were included for realism because they represent the following:

a. Equipment delays in ground and airborne transceivers

b. Delays due to uplinking of other messages (ATIS, weather advisories, etc.) when a high priority command was received by the data link ground systems.

c. Sampling delays during the repeated ground system interrogation of an aircraft for its W response.

For each non-data link aircraft command displayed and enunciated by the VRS, the simulator pilots were required to (1) enter a single alphanumeric ACID character which would allow the simulated aircraft to execute the command, and (2) verbally acknowledge the command with a "Roger." The single character ACID identifier was quickly entered by the pilot so that the controller would hear a rapid pilot acknowledge. The rapidity of the pilot responses

were considered important to a fair comparison of voice and data link message transaction times inasmuch as real pilots tend to abbreviate or hurry their responses when the voice channel is busy.

5.4 TEST DATA

The subjective and objective test data from this experiment were derived from controller questionnaires and test data retrieved on line by the ARTS III computer during each test run, respectively. The questionnaires were categorized as follows:

a. A pre-experiment questionnaire completed after briefing and before any training runs to ascertain the subject's attitudes toward data link and the experiment.

b. A post-training questionnaire answered after the completion of training on a particular control mode, but before any test runs for that mode. This questionnaire sought to obtain controller opinion about training adequacy as well as pre-test opinion about the performance of the mode to be tested.

c. A post-test questionnaire filled out after the completion of each test run to obtain controller opinion about the mode just tested, including overall performance during the test run, and the operational characteristics of the mode tested.

d. A post-series questionnaire given after completion of all test runs for a particular mode to determine if there were differences in controller reaction with respect to the three traffic mixes.

e. A summary questionnaire answered after each controller had completed all the test runs for all the modes, intended to provide data on controller preference and rating of the modes with respect to such aspects as stress, blunders, and workload.

The test data collected by the ARTS III were reduced and processed by data extraction programs after the test runs were completed. The extraction programs measured the following types of data:

1. Time intervals between successive communications actions
2. Counts of significant events
3. Aircraft positions for each scan

The measurements required for the first two types of data were defined by means of constructing event time lines (see the appendix) for every sequence of communications actions that occurred for each of the three modes. The key time intervals associated with the two-way transaction of each message are depicted in the time lines. The time interval values obtained were used in the comparative analysis of the three modes. The key intervals were:

a. Message Transaction Time (MTT): The total time between display of an M&S command to the controller to the completion of a normal pilot response.

b. Command Initiation Delay (CID): The time between command display and initiation of controller action. Initiation was defined as occuring when the command was dispatched.

c. Service Time (ST): The time required to service an active command that had become the topmost blinking command in the tabular list.

d. Queue Delay (QD): The time required for a command to reach the topmost position in the tabular list after being initially displayed.

The time lines illustrate the importance of the command initiation delay as a component of total message transaction time and as a measure of controller workload. Since one important way in which the CID was thought to manifest itself was in the format of a queue of commands, an expanded list of MTT and command queue

measurements were included in a data reduction program summary printout (see Figure 11). The most important of these were the total times for all voice and data link message transactions and the average queue size.

Another set of objective data obtained during each test run consisted of counts of significant events, such as normal voice link and data link command, exception cases, extemporaneous communications, handoffs, etc. The total number of these events was also listed in each summary printout as an aid in determining the validity of a test run. A total count of the number of manual tag repositionings was also obtained to gain a measure of controller workload and distraction from his normal communications task. (Automatic tag offset was disabled by controller preference.) In addition to the above data, the track position each radar scan of all aircraft was extracted for subsequent input to conflict detection and analysis programs that were exercised after the test runs were completed.

The observer controllers recorded the aircraft identifications and the time of occurrence whenever they were alerted by the subject controllers that a conflict situation was developing in which the aircraft would normally be maneuvered. Similarly, the pilots recorded any instances in which the subjects transmitted commands not generated by the M&S algorithm and when commands timed out without controller dispatch.

5.5 TRAFFIC SAMPLES

The prime objective in the design of the traffic samples was to subject the controller to a linear increase in traffic density from zero to nine aircraft over a test period of 1 hour. Objective data were collected for a 45 minute period from this hour. The traffic density in this case was defined as the controller's instantaneous aircraft count in the region between the two feeder fixes and the point of handoff to the tower. The period of 1 hour was deemed sufficient to provide results representative of an operational situation. The increase in aircraft

DATE [illegible]
TIME EXPERIMENT STARTED 1:19 PM
FORMAT [illegible] 50 PER CENT DATA LINK AIRCRAFT
SOURCE CONTROLLER NO. 2
FORMAT NO. [illegible] FOR THIS CONTROLLER 6 C50E
OBSERVER NO. 1
DATA RUN-TEST SERIES 2
TRAFFIC SAMPLE NO. 0

TOTAL NO. OF:

NORMAL VOICE LINK COMMANDS	114
NORMAL DATA LINK COMMANDS	125
COMMANDS WITH UNABLE REPLY	1
COMMANDS WITH LINK FAIL	1
COMMANDS DISAPPROVED	0
COMMANDS WITH ACKNOWLEDGE TIMEOUT (X)	
VOICE LINK AIRCRAFT	0
DATA LINK AIRCRAFT	1
COMMANDS DISCARDED	
VOICE LINK AIRCRAFT	0
DATA LINK AIRCRAFT	0

TOTAL NO. OF EXTEMPORANEOUS COMMUNICATIONS
CONTROLLER 0 PILOT 18 VRS 0

TOTAL NO. OF QUEUE CHANGES 498
MAX. QUEUE SIZE 4
AVE. QUEUE SIZE .45485

MAX. AIRCRAFT LOAD 9

TOTAL NO. OF:

CONTROLLER-GENERATED COMMANDS	2
HANDOFFS TO TOWER	28
ALTIMETER SETTINGS	1
STACK DEPARTURES	0

TOTAL TIME FOR MESSAGE TRANSACTIONS (SEC)
VOICE LINK 1298.6 DATA LINK 1336.8 CBE/FDB/VRS 50%
AVE. MESSAGE TRANSACTION TIME (SEC)
VOICE LINK 11.4 DATA LINK 10.7
MAX. MESSAGE TRANSACTION TIME (SEC)
VOICE LINK 22.1 DATA LINK 12.6

AVE. COMMAND INITIATION DELAY (SEC)
VOICE LINK 5.8 DATA LINK 5.8
AVE. QUEUE DELAY (SEC)
VOICE LINK .0 DATA LINK .0
AVE. SERVICE TIME (SEC)
VOICE LINK 5.6 DATA LINK .0

TOTAL NO. OF AIRCRAFT ARRIVING AT GATE 27
AVE. ERROR IN ARRIVAL TIME AT GATE (SEC) 6.1
STD. DEV. OF ARRIVAL TIME ERRORS AT GATE (SEC) 6.0

TOTAL NO. OF VRS DISABLED-ENABLED SEQUENCES 1
TOTAL TIME HALTED (SEC) .8

TOTAL NO. OF VRS COMMAND MESSAGES LOST 0
TOTAL NO. OF OTHER VRS MESSAGES LOST 0

TOTAL EXTEMPORANEOUS COMMUNICATION TIME (SEC)
CONTROLLER .0 PILOT 8.6 VRS .0

TOTAL NO. OF:

AIRCRAFT LOAD CHANGES	61
ILS CAPTURES	27
TAG REPOSITIONINGS (MANUAL)	5
AIRCRAFT DISABLED	0
AIRCRAFT ENABLED	0

TOTAL NO. OF VRS MESSAGES

HANDOFFS TO TOWER	14
ALTIMETER SETTINGS	1
STACK DEPARTURES	0

TOTAL NO. OF TIME CONFLICTS AT GATE 0

Figure 11. Summary Printout of Performance Measures

load to nine during the hour was intended to facilitate the continuous measurement of the controller's communication delays which, it was theorized, would slowly increase until his workload limit was reached and then increase rapidly thereafter.

To develop a traffic sample which provided a linearly increasing traffic load as a function of time, an initial sample consisting of 30 aircraft was prepared which approximated the desired pattern. The initial sample had to be changed experimentally until the desired results were obtained. This could not be done analytically because the M&S algorithm adapted to changes in the traffic sample by changing times at which commands were issued or by issuing speed commands. Operational parameters also had to be changed. For example, the minimum allowable separation had to be reduced from 3 to 2 miles to accommodate the maximum of nine aircraft for the controller's instantaneous load. Also, to circumvent operational deficiencies of the M&S algorithm, the aircraft types were limited to high performance jets (707) only; jumbo jets and low performance aircraft were specifically excluded.

The procedure used in developing a master traffic sample consisted of conducting numerous planning runs during which the entry and exit times of each aircraft were recorded. From that information a chart of the aircraft load as a function of time was constructed. This chart was examined and changes in the aircraft entry times were made to improve the linearity of the increasing aircraft load. This process was repeated many times until a point was reached where the recommended changes causes adaptations by the M&S algorithm which adversely affected linearity. The traffic sample selected was the one which immediately preceded the occurence of adverse effects (charted in Figure 12). The portion of the sample from 10 to 55 minutes was used for data collection purposes.

Since it was impractical to create 11 different traffic samples by this process, the same basic sample was employed in all experiment runs by each controller. To minimize learning effects, the aircraft identifications and the occurrence of data link failure events were changed from run to run.

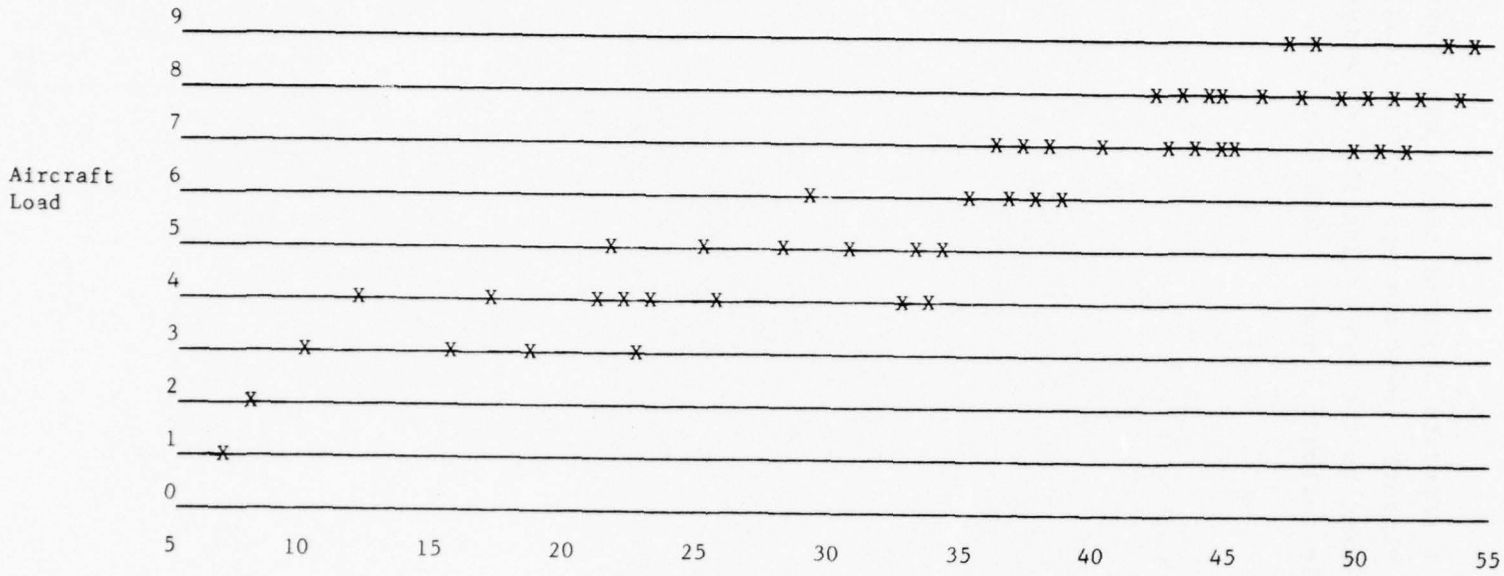


Figure 12. Time Response of Selected Traffic Sample

In addition, a separate training sample was constructed. This sample had a slightly lower average traffic load, and while it generally increased as the run progressed, no attempt was made to make that sample increase linearly with time.

6. TEST DESIGN

This section discusses the determination of the test settings and the sequence of runs scheduled for the simulation study. The objective of the study was to determine differences between the three candidate modes as a function of the percentage or mix level of data link vs. voice link equipped aircraft. A secondary consideration to provide voice-only data against which the other testing could be compared was added late in the planning stages of the experiment.

The development of the experimental plan was influenced by several factors:

a. Rather than limit the experiment to one level of air traffic or increase the number of runs by introducing load as an additional factor, it was decided to use a dynamic density variation that increased linearly with the test time.

b. Each test mode was to be handled by as many subject controllers as possible to prevent the relative evaluation of modes from being distorted by differences in the skills of individual controllers.

c. The order in which the modes and percentages were tested by the subject controllers would contaminate the results unless the plan was set up to balance out such effects. For example, if each controller tested three modes, it must be assumed that his performance on the third test would be better than his performance on the first because of his experience on tests one and two. To compensate for this, the sequence of tests was assigned so that each mode and percentage was tested by one controller for whom it was the first, by one controller for whom it was his second, etc. In statistical terminology this technique is called BLOCKING in the sense that the run sequence is treated as a block effect.

d. A basic unit of three controllers testing all the nine combinations of the three modes and three percentage levels would permit blocking of the mode and percentage sequences. This basic unit could be repeated with a different set of controllers to obtain greater statistical power in the analysis.

e. While the traffic situation simulated represented a fairly standard ATC situation, the manipulation of the function keys and slewball varied from mode to mode. To insure that each controller was operating at an acceptable level of proficiency and to familiarize him with the specific procedures for the mode being tested, sufficient training had to be provided. To minimize the training time all the percentage runs for one mode were to be grouped and preceded by a one hour training run.

f. To handle the addition of the voice-only runs in a structured fashion would greatly enlarge the required number of runs for a basic unit. If the analysis were limited to simple comparisons between voice-only and percentage run averages, then appropriate runs could be inserted at random to the groups of runs for each mode.

The above considerations dictated the use of a factorial design for the test variables (i.e., mode, percentage, and subject) with the percentage and mode sequences blocked in a Greco-Latin square. Using this arrangement, the test runs were sequenced and the training and voice-only runs inserted, as shown in Table 1.

TABLE 1. RUN SCHEDULE

SUBJECT	TIME	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	8:15	TT	TA	A00J	A20A	A50D	A80G	TB	B50E	B80H	B00K	B20B	TC	C80I	C20C	C50F
2	9:30	TT	TB	B80G	B20A	B50D	B00J	TC	C20B	C50E	C80H	TA	A50F	A00K	A80I	A20C
3	10:45	TT	TC	C50D	C80G	C20A	TA	A80H	A20B	A00J	A50E	TB	B20C	B50F	B80I	B00K

Each entry is coded as follows: MXYN or TM where

M = Test modes
XY = Percentage level
N = Traffic sample
T = Training run

Test Modes

A = CBA/FDB
B = CBA/TAB
C = CBE/FDB
T = Training

Traffic Samples

A, B, C } 20%

D, E, F } 50%

G, H, I } 80%

J, K } 00%

7. SUBJECTIVE DATA ANALYSIS

Analysis of the subjective (questionnaire) data for the Data Link Phase 2B experiment consisted of a tabulation of the responses to each question by mode and percentage mix. These tabulations were analyzed to determine if there were any meaningful differences between the modes or the mixes. In addition, where similar questions were given at different points in the experiment, those data were analyzed to determine if the responses differed.

The results of this analysis are reported by question. Each question is given exactly as stated in the questionnaire. Following each question is a discussion of the findings. The responses to each question were subjectively reviewed to determine if there was any consistent, sizable (i.e., meaningful) differences attributable either to the mode or percentage effects. Only meaningful differences are reported. If the mode or mix differences are not meaningful, no comment is made, and generally the results are aggregated across the nonmeaningful effect(s). All important findings in this section are underlined. A statistical analysis of these data was not performed due to the limited sample size of six subjects.

The questionnaires consisted of five parts: (1) the Pre-Experiment questionnaire given after briefing and before any operational exposure, (2) the Post-Training questionnaire given after completion of the one training hour for each run, (3) the Post-Test questionnaire given after the completion of each test hour, (4) the Post-Series questionnaire given after completion of all runs of a mode, and (5) the Summary questionnaire given after completion of all runs.

7.1 PRE-EXPERIMENT QUESTIONNAIRE

The Pre-Experiment questionnaire was given after oral briefing and before any operational exposure. The questionnaire consisted of seven questions intended to ascertain the subjects attitudes toward data link and the experiment as a whole.

1. *Do you understand the objectives of the data link display experiments?*

   Yes 100% No 0%

   All of the subjects felt that they understood the objective of the data link display experiment.

2. *Choose a statement describing the degree of understanding of the data link study which you achieved by reading the documentation provided.*

   a. *Complete and comprehensive* 50%
   b. *Workable with minor questions* 50%
   c. *Marginal with major questions* 0%
   d. *Vague and confused* 0%

   All of the subjects felt they had a workable understanding of the study after reading the documentation provided.

3. *Do you feel that the semi-automation of terminal ATC computer generated metering and spacing commands transmitted via data link will eventually be accepted in actual operations?*

   a. *Yes* 83%
   b. *No* 0%
   c. *Undecided* 17%

   Five of the six subject controllers indicated that they felt the semi-automation of terminal ATC computer generated metering and spacing commands transmitted via data link would be accepted in actual operation.

4. *How do you anticipate that the addition of computer generated metering and spacing commands transmitted via data link will affect the current operational system?*

   a. *(+2) Very favorably* 0%
   b. *(+1) Favorably* 66%
   c. *(0) Not very much* 17%
   d. *(-1) Unfavorably* 17%
   e. *(-2) Very unfavorably* 0%

Most of the subjects felt that the addition of computer generated metering and spacing commands transmitted via data link would favorably affect the current operating system. One controller felt the addition of a data link would not have very much effect on the system while another one felt the effect would be unfavorable. Using the scores indicated in parentheses next to the responses gives an average response of 0.5, indicating a mildly favorable attitude.

5. *Rank those aspects of the upcoming test which you feel least confident about (where the aspect about which you are least confident is numbered 1):*

   a. Performance at maximum traffic loads
   b. Mix of data link and voice link aircraft
   c. Metering and spacing algorithm
   d. Display formats
   e. I/O procedures
   f. Preventing the tag overlaps
   g. Maintaining safe separation
   h. Trackball operations
   i. VRS procedures

The responses to this question are tabulated in Table 2. As shown by the rank sum column, the metering and spacing algorithm was clearly the most troublesome aspect of the test prior to experimentation. Performance at maximum traffic load was the second most troublesome aspect, having received first place ranking by four of the subjects and near last place by the other two. The mix of voice and data link aircraft, trackball operations, and VRS procedures were the least troublesome.

6. *Did you participate in the Data Link Phase 2A experiment on the DSF?*

   Yes 83% No 17%

Five of the six subjects had participated in the Phase 2A experiment, indicating that a comparison between the Phase 2A and the Phase 2B results for similar questions is appropriate.

TABLE 2. TABULATION OF RESPONSES TO PRE-EXPERIMENT QUESTION 5

RESPONSE	RANK									
	1	2	3	4	5	6	7	8	9	Rank Sum
a. Performance at maximum traffic loads	4							1	1	21
b. Mix of data and voice link aircraft		2						4		36
c. Metering and spacing algorithm	1	2	2	1						15
d. Display formats				1	3		2			33
e. I/O procedures			1	1	1	3				30
f. Preventing tag overlaps		2				1	2	1		32
g. Maintaining safe separation	1		3				1		1	26
h. Trackball operations				3	1				2	35
i. VRS procedures					1	2	1		2	42

NOTE: Lower score indicates a more troublesome aspect.

7. *Please indicate how recently (in years) you have had experience with the following systems (if no experience indicate with a 0):*

   a. *Prototype M&S on the TATF*
   b. *Operational ARTS*
   c. *Other TATF experiments*

Only one subject had experience with the prototype M&S on the TATF, and that experience was within one year. Only one other controller had operational ARTS exposure, and that was two years ago. One other subject had worked on other TATF experiments two years ago. While only half of the subjects had direct ARTS experience, all the subjects had considerable experience on other DSF experiments which utilize a digital, ARTS-like display.

7.2 POST-TRAINING QUESTIONNAIRE

The Post-Training questionnaire was given after the second training run and prior to any test exposure for that mode. The objective of the questionnaire was to ascertain the controllers' opinions about the adequacy of the training and their prerun opinions about the performance of the mode being tested.

1. *Were the principal features of the mode demonstrated clearly in the training runs?* Yes 100% No 0%

All the subjects felt that the modes were clearly demonstrated.

2. *How much more training do you think you would need to feel comfortable with the procedures?*

   a. *One hour* 100%
   b. *Three hours* 0%
   c. *Five hours* 0%
   d. *Ten hours* 0%

All of the responses indicated that a minimum amount of additional training (1 hour) would have been desirable.

3. *How do you anticipate that this mode of displaying and transmitting commands will effect the current operational system?*

   a. *(+2) Very favorably*
   b. *(+1) Favorably*
   c. *(0) Not very much*
   d. *(-1) Unfavorably*
   e. *(-2) Very unfavorably*

The responses to this question are tabulated by mode below:

RESPONSE	MODE		
	FDB	TAB	CBE
Very favorably (+2)			2
Favorably (+1)	1	1	2
Not very much (0)		2	
Unfavorably (-1)	4	3	2
Very unfavorably (-2)	1		
Average score	-5	-2	+4

As can be seen above, the subject controllers had a favorable opinion of the CBE mode, a mildly unfavorable opinion of the TAB mode, and a strongly unfavorable opinion of the FDB mode prior to the test runs.

4. *How do you anticipate that this mode of displaying and transmitting commands will effect the following characteristics of the current operating system?*

ASPECT	GREATLY DECREASED	DECREASE	WILL NOT CHANGE	INCREASE	GREATLY INCREASE
Workload	(+2)	(+1)	(0)	(-1)	(-2)
Traffic Handling Capacity	(-2)	(-1)	(0)	(+1)	(+2)
Succeptibility to Blunders	(+2)	(+1)	(0)	(-1)	(-2)
Stressfullness	(+2)	(+1)	(0)	(-1)	(-2)

Using the scores indicated in parentheses, the responses to this question have been reduced and are shown below:

ASPECT	MODE		
	FDB	TAB	CBE
Workload	-8	+2	+4
Traffic Handling Capacity	-4	-2	+3
Susceptibility to Blunders	-7	-4	+2
Stressfulness	-8	-4	0
Total	-27	-8	+9

These scores indicate an opinion very similar to the overall grade given in Question 3. The CBE mode is given a mildly favorable response, TAB mildly unfavorable and FDB strongly unfavorable.

7.3 POST-TESTING QUESTIONNAIRE

The Post-Testing Questionnaire was given after the completion of each run. It was composed of three series of questions, A, B, and C.

7.3.1 A-Series

The questions of the A-Series were designed to solicite general controller attitude toward the mode tested.

A-1. A. *How do you feel that this mode of displaying and transmitting commands would affect the current operational system for the same traffic situation?*

a. (+2) Very favorably
b. (+1) Favorably
c. (0) Not very much
d. (-1) Unfavorably
e. (-2) Very unfavorably

The responses to this question were reduced using the scores shown in parentheses next to the responses. These results are shown on the next page.

MODE	PERCENTAGE DL				SUM OF EXCLUSIVE OF 0%
	0	20	50	80	
FDB	1	-5	-5	-5	-15
TAB	-3	0	-3	-7	-10
CBE	XX	3	2	2	7
TOTAL		-2	-6	-10	

These results summed across percentages indicate the same relationship between the modes as did Post-Training Questions 3 and 4. CBE would have a mildly favorable impact on the operational system; TAB unfavorable; and FDB more strongly unfavorable when summed across the modes. The results indicate a trend which shows an increasingly unfavorable attitude as the percentage of data link equipped aircraft is increased.

A-1. B. *For the above, what ratings would you give each of the following aspects?*

ASPECT	GREATLY DECREASE	DECREASE	WILL NOT CHANGE	INCREASE	GREATLY INCREASE
Workload	(+2)	(+1)	(0)	(-1)	(-2)
Traffic Handling Capacity	(-2)	(-1)	(0)	(+1)	(+2)
Susceptibility to Blunders	(+2)	(+1)	(0)	(-1)	(-2)
Stressfulness	(+2)	(+1)	(0)	(-1)	(-2)

Using the scores indicated in parentheses, the subject responses were reduced to give the results shown in Table 3. The mode totals indicate the same relationship shown previously in Post-Training Questions 3 and 4 and Post-Test Question A-1.A. The subjects felt that CBE would have a generally favorable impact on current operations while TAB would have an unfavorable impact and FDB an even more unfavorable impact.

TABLE 3. REDUCED RESPONSES TO POST-TEST QUESTION A-1.B

ASPECT	MODE	00	20	50	80	20+50+80
Workload	FDB	-3	-5	-7	-8	-20
	TAB	-6	-1	-1	0	- 2
	CBE	X	5	3	3	11
Traffic Handling Capacity	FDB	1	-4	-5	-6	-15
	TAB	0	-2	1	0	- 1
	CBE	X	4	5	3	12
Susceptibility to Blunders	FDB	-1	-7	-6	-7	-20
	TAB	-5	-5	-5	-5	-15
	CBE	X	0	0	0	0
Stressfulness	FDB	-1	-7	-5	-9	-21
	TAB	-3	-8	-6	-6	-20
	CBE	X	-1	0	1	0
Total	FDB	-4	-23	-23	-30	-76
	TAB	-14	-16	-11	-11	-38
	CBE	X	8	8	7	23
Grand Total	FDB+TAB	-18	-39	-34	-41	—
	FDB+TAB+CBE	—	-31	-26	-34	—

The favorable rating for CBE came from an expected improvement in the workload and traffic handling capacity aspects, while the susceptibility to blunders and stressfulness aspects as expected was unchanged. The unfavorable reaction to TAB primarily reflected an unfavorable response to the susceptibility to blunders and the stressfulness characteristics. The unfavorable response to FDB was contributed to by unfavorable reactions in all aspects.

The trend of increasingly unfavorable responses with increasing percentage observed in Post-Test Question A-1.A was not as evident, which indicates that the controllers' negative reaction to increasing DL employment could not be explained in terms of the individual control aspects. The voice-only runs were graded somewhat more favorably than the data link/VRS combination runs.

Questions A-2.A and A-2.B are only applicable to modes which employed the data link capability. They were not included in the questionnaires for the voice-only runs.

A-2. A. *Relative to the question just answered, if all of the aircraft had been equipped with a digital data link, how do you feel this mode would have affected the current operational system?*

a. (+2) Very Favorable
b. (+1) Favorable
c. (0) Not very much
d. (-1) Unfavorably
e. (-2) Very Unfavorably

The results of this question, reduced by the scores given in parentheses next to the responses, are given below:

MODE	PERCENTAGE			TOTAL
	20	50	80	
FDB	-4	-5	-7	-16
TAB	-3	-3	-2	-8
CBE	2	1	3	6
Total	-5	-7	-6	-

Once again, the responses to this question indicate the same relationship between modes as shown in Post-Training Questions 3 and 4 and Post-Test questions A-1.A and A-1.B. CBE in an all data link operation was judged to affect the current operation favorably; TAB was judged unfavorably; and FDB even more unfavorably. The magnitude of the responses to this question was very similar to those for Question A-1.A, indicating very little expected difference due to an all data link environment.

A.2 B. *For the above, what ratings would you give the following aspects?*

ASPECT	GREATLY DECREASE	DECREASE	WILL NOT CHANGE	INCREASE	GREATLY INCREASE
Workload	(+2)	(+1)	(0)	(-1)	(-2)
Traffic Handling Capacity	(-2)	(-1)	(0)	(+1)	(+2)
Susceptibility to Blunders	(+2)	(+1)	(0)	(-1)	(-2)
Stressfulness	(+2)	(+1)	(0)	(-1)	(-2)

The responses to Question A-2.B are given in Table 4 reduced by the scores indicated in the parentheses above. The responses to this question indicate the same relationships shown previously in Post-Training Questions 3 and 4 and Post-Test Questions A-1.A, A-1.B, and A-2.A. In an all data link environment, CBE was considered to affect the current operation favorably; TAB was judged unfavorably; and FDB even more unfavorably.

When compared to Question A-1.B, the responses are of similar magnitude indicating that no disadvantage would accrue to an all data link environment. In addition, the contribution by aspect to the favorable or unfavorable reaction on each mode was about the same. For CBE, the favorable response was recorded in the workload and traffic handling capacity aspects, while the susceptibility to blunders and stressfulness aspects were about neutral. For FDB and TAB, the unfavorable response was recorded in all four aspects.

TABLE 4. REDUCED RESULTS OF POST-TEST QUESTION A-2.B

ASPECT	MODE	PERCENTAGE			TOTAL
		20	50	80	
Workload	FDB	-5	-6	-8	-19
	TAB	-3	-3	0	- 6
	CBE	3	1	3	7
Traffic Handling Capacity	FDB	-3	-4	-7	-14
	TAB	-2	0	0	- 2
	CBE	3	3	3	9
Susceptibility to Blunders	FDB	-8	-8	-7	-23
	TAB	-5	-7	-4	-16
	CBE	0	0	-1	- 1
Stressfulness	FDB	-8	-7	-10	-25
	TAB	-6	-5	-6	-17
	CBE	-2	-1	1	- 2
Total	FDB	-24	-25	-32	-81
	TAB	-16	-15	-10	-41
	CBE	4	3	6	13
Grand Total	FDB+TAB+CBE	-36	-37	-36	—

7.3.2 B-Series

The B-Series contained questions which applied to all of the modes. The first three B questions were concerned with the controller's attitude toward his preparation for and performance during the run.

B.1 *Now that you have been tested in this mode, how much more training time do you think you would need to feel comfortable with the procedures?*

a. One hour
b. Three hours
c. Five hours
d. Ten hours

The results of this question have been summarized by mode and shown below:

	MODE		
TIME	FDB	TAB	CBE
One hour	11	16	12
Three hours	1		
Five hours			
Ten hours	8	4	
Average	4.7	2.8	1.0

Only five of the six subjects elected to answer this question. In several other cases, a response was written in which indicated an opinion outside the range of the answers such as "none" or "never." In these cases, the most appropriate box either one hour or ten hours was assigned the response.

The responses show much the same opinion as indicated in Post-Training Questions 3 and 4 and in the A-series questions. The controllers felt very comfortable with the CBE mode and would have required only one additional hour of training. In several instances, in the TAB mode, the controllers felt very uncomfortable

and the average rose to 2.8 hours. For FDB, there were even more cases where the controllers felt uncomfortable and the average rose even higher to 4.7 hours.

B-2. *Rank those aspects of the test run which gave you the most trouble (where the aspect giving the most trouble is numbered 1):*

a. Performance at maximum traffic load
b. Mix of data link and voice link aircraft
c. Metering and spacing algorithm
d. Display formats
e. I/O procedures
f. Preventing tab overlaps
g. Maintaining safe separations
h. Trackball operations
i. VRS procedures

In response to this question, several controllers ranked only those aspects which were troublesome. To normalize the responses, ranked aspects were assigned the average unused ranks; e.g., if aspect a was ranked 1, b ranked 2, and c ranked 3, then aspects d through i would have been assigned the average of ranks 4 through 9 or 6.5. In cases where aspects were given the same rank, they were arbitrarily assigned that rank plus a factor equal to (n-1)/2, where n is the number of co-ranked aspects. For example, if aspects d, e, and f had been assigned rank 2, then that rank would have been changed to 2+(3-1)/2 = 3. If necessary, lower ranks were adjusted to correct for ranks occupied by ties. Following this process the ranks were accumulated to give the results shown in Table 5. The scores for the voice runs are multiplied by 3 since 3 times as many VRS runs were made.

As can be seen from that table, for all modes, the M&S algorithm was a major problem. For the FDB mode, preventing tag overlaps and trackball operations were very significant. For the TAB mode, no additional factors were important. For CBE, the display format and the prevention of tag overlaps were somewhat bothersome.

TABLE 5. ACCUMULATED RANKS FROM POST-TEST QUESTION B-2

ASPECT	FDB		LAB		CBE
	VRS	VOICE	VRS	VOICE	VRS
a. Performance at maximum traffic load	7 103.5	5 97.5	2 83.0	4 88.5	5 89.0
b. Mix of Data Link and Voice Link Aircraft	8 120.5	8 118.5	7 109.5	8 123.0	6 107.0
c. M&S Algorithm	1 56.0	2 51.0	1 33.0	1 54.0	1 48.0
d. Display Formats	4 79.5	5 97.5	3 83.5	2 72.0	3 67.0
e. I/O Procedures	5 92.5	7 102.0	5 89.5	6 90.6	6 107.0
f. Preventing Tag Overlaps	3 62.0	1 39.0	6 91.5	7 96.0	2 63.5
g. Maintaining Safe Separation	6 102.5	4 91.5	4 87.5	3 76.5	4 86.6
h. Trackball Operations	2 58.0	3 66.0	8 110.5	4 88.5	8 108.5
i. VRS Procedures	9 128.5		9 119.0		9 120.0

NOTE: A low score is indicative of a more troublesome aspect.
The number in the upper righthand corner gives the rank of that aspect for that mode.

The rankings and grades for the Voice and VRS runs were very similar except in the FDB mode where the prevention of tag overlaps become very much more bothersome in the voice modes due to the requirement to read the third line in the data block to transmit the command.

In comparison with Question 5 of the Pre-Experiment questionnaire, the responses are very similar except that the concern over performance at maximum traffic load decreases considerably after the run. This result duplicates the findings of the Phase 2A experiment.

B-3. *Which of the following statements best describes your feelings at the end of the test run just concluded?*

a. The aircraft load was relatively easy to handle
b. You could control a small number of additional aircraft safely
c. You were operating at your maximum capacity
d. You were somewhat overloaded
e. You were badly overloaded

The results of this question are tabulated by mode in the following table:

RESPONSE	MODE		
	FDB	TAB	CBE
a. Easy to handle	0.0%	20.9%	33.4%
b. Near capacity	8.4%	12.5%	22.2%
c. At capacity	45.8%	25.0%	22.2%
d. Somewhat overloaded	16.7%	29.1%	11.1%
e. Badly overloaded	29.1%	12.5%	11.1%

Once again, the responses to this question show the familiar pattern of responses by mode where, on the average, CBE is operating slightly below capacity; TAB, slightly above capacity; and FDB, quite a bit above capacity.

The remaining B questions were designed to ascertain the controllers' opinions of engineering or operating characteristics of the experiment.

B-4. *Were any of the characters, symbols, or actions employed in this mode particularly difficult to interpret?*

Yes 47% No 53%

Almost half of the subjects had difficulty with certain characteristics of the modes. This difficulty was not dependent upon the mode but upon the individual. The controllers did not specify which of the characters, symbols, or actions were difficult to interpret.

B-5. *Were you able to evaluate each command carefully before you dispatched it?*

a. Almost always 14%
b. In a majority of cases 20%
c. About half the time 30%
d. In a minority of cases 19%
e. Rarely 17%

There was considerable variation in response to this question. An examination of the mode percentage of differences did not show any pattern. However, in the majority of instances, the controllers did not feel that they were able to evaluate all commands prior to dispatch.

B-6. *Were you able to maintain the "picture" of the overall traffic situation and monitor it for potential conflicts even while transmitting commands at the maximum rate?*

a. Almost always 14%
b. In a majority of cases 13%
c. About half the time 17%
d. In a minority of cases 13%
e. Rarely 34%

As in Question B-5, there was a varied, unpatterned response. However, in this case the responses clustered around the extreme points.

B-7. *How do you rate your confidence in being able to detect and resolve any conflicts that might occur while using this mode?*

a. Extremely high

b. High

c. Moderate

d. Low

e. Very pessimistic

As shown below, the responses to this question show the familiar modal pattern:

RESPONSE	MODE		
	FDB	TAB	CBE
a. Extremely high	13%	13%	33%
b. High	13%	21%	6%
c. Moderate	4%	16%	22%
d. Low	37%	13%	28%
e. Very pessimistic	33%	37%	11%

With the CBE mode, the controllers' confidence in being able to detect and resolve conflicts is high to moderate; with the TAB mode moderate to low; and with the FDB mode, low.

B-8. *Was the M&S conflict list useful in monitoring the traffic situation?*

Yes 17% No 83%

Only 17% of the controllers felt the conflict list was useful. This was no doubt due to the unreliability of the M&S conflict prediction algorithm.

B-9. *Did the M&S conflict list distract you from other more important tasks?*

Yes 17% No 83%

Only one of the six controllers felt that the M&S conflict list was distracting.

B-10. *Did you use the automatic tag respositioning feature?*

Yes 14% No 86%

The controllers used the automatic tag repositioning feature in only 14% of the runs. For the FDB mode, that feature was only used once or 8% of the time.

B-11. *Which character size setting did you use?*

a. 1

b. 2

c. 3

d. 4

Most of the subjects preferred to use the size 2 setting (.136 inch/character); however, during maximum traffic load, many controllers used setting 1 (.084 inch/character). The responses are not tabulated since many of the subjects checked more than one response.

7.3.3 C-Series

The C-Series questions pertain to characteristics of the individual modes and were given to the controllers only after completion of a run of that mode.

Questions C-1 to C-5 pertain only to the CBE and FDB modes, where the command information is contained in the full data block.

C-1. *Did the command line in the full data block increase display clutter to the point that your ability to control traffic was impaired?*

Yes 62% No 38%

Most of the subjects felt that the increased display clutter impaired their ability to control traffic. Prelocation of tags by approach phase was not employed as in the Phase 2A experiment.

C-2. *Did the clutter cause the command line to be obscured to the point where it caused operational difficulties or an excessive number of tag repositioning?*

Yes 83% No 17%

An even greater percentage of the subjects felt that the increased clutter caused by the presence of the command line caused operational difficulties or an excessive number of tag repositionings.

C-3. *If you used the automatic offset feature, did it cause you to miss any of the commands?*

Yes 67% No 33%

This question was answered only three times. In two of these cases the controllers felt that the automatic offset feature did cause them to miss commands.

C-4. *Was the presentation and coding of the commands in the full data block easily interpreted?*

Yes 76% No 24%

Approximately 1/4 of the subjects had difficulty interpreting the presentation and coding of the commands in the full data block. Controllers did not state the reasons why they had difficulty; however, that difficulty was probably caused by the somewhat cluttered display.

C-5. *Did you have any difficulty in recognizing an unable(U), a link-fail (F), or a no response (X) indication to act on it in a sufficiently timely fashion?*

Yes 76% No 24%

Almost 80% of the subjects had difficulty in recognizing the exception condition indications to act on them in a sufficiently timely fashion. Once again, this difficulty was probably caused by the cluttered display.

Question C-6 applies only to the CBA/FDB command mode.

C-6. *When the number of commands waiting to be transmitted builds up, do you feel that it would be acceptable to dispatch several commands in rapid succession without evaluating each command?*

Yes 22% No 78%

Only 22% of the subjects felt it could be acceptable to dispatch several commands in rapid succession without evaluating each command regardless of load.

Questions C-7 through C-9 apply only to the TAB modes.

C-7. *Did the use of the tabular list distract you from monitoring the flow of traffic for potentially conflicting or hazardous traffic?*

Yes 46% No 54%

Over 40% of the subjects felt that the use of the tabular list was distracting.

C-8. *Was it helpful to have the aircraft's position symbol blink in order to show the correlation of the active tabular list command with the appropriate target?*

Yes 54% No 46%

Over half of the subjects felt that it was helpful to have the aircraft's position symbol blink.

C-9. *The commands were placed in the tabular list on a first come/first served basis. Would a priority given to commands around the base leg be preferable?*

Yes 67% No 33%

Two-thirds of the subjects felt that priority should be given to commands around the base leg.

Question C-10 was dropped after the questionnaires were printed. The question was ambiguously worded and the subjects were instructed to ignore it.

Questions C-11 through C-14 apply to the CBE mode only.

C-11. *Were provisions for halting and starting the VRS adequate for maintaining discipline on the voice channel?*

Yes 50% No 50%

Half of the responses indicated that additional provisions for halting and starting the VRS were needed. No detailed information was provided by the subjects.

C-12. *Did you find it disconcerting not to have pre-transmission approval of issued commands?*

Yes 44% No 56%

Almost half of the responses indicated that not having pre-transmission approval of issued commands was disconcerting.

C-13. *Was the 3-second gap between VRS transmissions*

a. Too long 29%

b. About right 53%

c. Too short 18%

Most of the controllers felt that the 3-second gap between VRS transmissions was appropriate.

C-14. *Six-seconds were allowed to disapprove each command. Was this time*

a. Too long 38%

b. About right 54%

c. Too short 8%

Most of the responses indicated that the 6-second disapproval time was about the right length. However, a fairly large percentage felt it was somewhat too long.

Question C-15 applies only to the control-by-approval modes.

C-15. *If a command is not dispatched within 12 seconds, the command stops blinking. Twenty seconds after that, the command is erased. Did this feature cause any problems during the run?*

The responses to this question varied by mode as shown in the table below.

RESPONSE	MODE	
	FDB	TAB
Yes	78%	58%
No	22%	42%

In both modes, the majority of responses indicated that the time-out of commands caused problems. However, in the FDB mode, almost twice as many responses as in the TAB mode indicated a problem. This was probably caused by the increased time required to detect a flashing command and to slew to the aircraft symbol for dispatch.

Questions C-16 to C-19 apply to VRS runs only.

C-16. *Did you find any of the following aspects of the VRS bothersome?*

a. Intelligibility 1
b. Voice quality 2
c. Talking rate 14
d. Inability to repeat command 20
e. Interference with your own extemporaneous messages 9

The frequency with which each aspect was checked is indicated next to each response. As can be seen, the predominant problems were first, the inability to repeat commands and secondly, talking rate. Interference with extemporaneous commands was a mild problem.

C-17. *Do you find the dispatching and monitoring of VRS commands less tiring than voicing the commands yourself?*

Yes 66% No 33%

Two-thirds of the responses indicated that the VRS was less tiring than voicing the commands.

C-18. *Did you listen to the VRS deliver its message and to the pilot's reply before shifting your attention to the next command?*

a. Almost always 17%
b. In a majority of cases 28%
c. About half the time 23%
d. In a minority of cases 13%
e. Rarely 19%

The responses to this question were fairly equally divided among the choices. No pattern was discernible in either the modes or percentages.

C-19. *Commands to all voice link aircraft remain displayed for 20 seconds after a successful dispatch. Should that time be:*

a. Lengthened 0%
b. Remain the same 43%
c. Shortened 57%

Over half of the responses indicated that the commands to voice link aircraft should remain displayed for less than 20 seconds after a successful dispatch.

Questions C-20 through C-22 pertain to the response indication for data link commands and apply only to the VRS/data link mix runs.

C-20. *The WILCO acknowledge (W) symbol was displayed nonblinking for 3 seconds after the receipt of the WILCO. Do you feel that time should be:*

a. Lengthened 6%
b. Remain the same 41%
c. Shortened 53%

Most of the responses indicated that the 3-second nonblinking display of the W symbol after receipt of a WILCO could be shortened. The 6% of the responses for lengthening that time came from the CBE mode; however, even there the majority were for shortening the display time.

C-21. *Was the display of the WILCO acknowledgement (W) symbol necessary or would the erasure of the command be an adequate indicator of successful dispatch?*

a. The WILCO symbol was required 11%

b. Erasure would be adequate 89%

Almost 90% of the responses indicated that erasure of the WILCO acknowledgement would be an adequate indicator of successful dispatch.

C-22. *Considering that link failures, unables, and no responses are exception cases and, as such, generally are more time consuming, do you feel that an excessive amount of time or effort is required to resolve these exception cases?*
Yes 7% No 93%

Only 7% of the responses felt that exception cases required an excessive amount of time or effort to resolve.

Questions C-23 and C-24 pertain to the extrapolation of the operation of the modes under different failure conditions.

C-23. *Envision the following situation: Computer aided metering and spacing with data link suddenly becomes inoperative during the peak load of this test. However, you still have the current ARTS III display and procedures. In your opinion:*

a. Could you continue operations indefinitely at the same traffic level? Yes 48% No 52%

b. Could you land all the aircraft under your control safely if no further aircraft were accepted in the terminal area? Yes 97% No 3%

c. If no further aircraft were accepted could you clear the terminal area of traffic safely? Yes 95% No 5%

d. A hazardous and unpredictable situation would exist?
Yes 9% No 91%

About half of the responses indicated that if M&S with data link failed, they could continue indefinitely at the same traffic level if the current ARTS III capabilities were available. Almost all of the responses indicated that they felt they could clear the terminal area or land all aircraft under those conditions. Less than 10% felt a hazardous situation would exist.

C-24. *Envision the following situation: Computer aided metering and spacing with data link and ARTS III data block suddenly becomes inoperative during the peak load of this test. You would be left with only the primary radar targets. If you had the current capability to see flight strips of your targets and to request squawk ID:*

a. Could you continue operations indefinitely at the same traffic level? Yes 7% No 93%

b. Could you land all the aircraft under your control safely if no further aircraft were accepted in the terminal area? Yes 40% No 60%

c. If no further aircraft were accepted could you clear the terminal area of traffic safely? Yes 38% No 62%

d. A hazardous and unpredictable situation would exist? Yes 89% No 11%

As would be expected, when left with only primary radar, squawk ID, and flight strips, all the subjects felt the situation would be much worse. Less than 10% felt they could continue operations indefinitely, while only 40% felt they could clear the area or land the aircraft even if no further aircraft were accepted. Almost 90% felt a hazardous situation would exist.

7.4 POST-SERIES QUESTIONNAIRE

The Post-Series questionnaire was given after each controller completed all the runs for a mode. The questions were designed to ascertain differences between mix percentages.

1. *Did you feel that there was a significant operational difference between the four runs just completed?* Yes 61% No 39%

Slightly over 60% of the responses indicated that there was a difference between runs. No comments were made by the subjects on the nature of the differences.

2. *If yes, could you rank those runs according to their ease of operation (rank the easiest run 1).*
   a. *First run*
   b. *Second run*
   c. *Third run*
   d. *Fourth run*

The sum of the ranks by mode and percentage are given below:

PERCENTAGE	MODE					TOTAL
	FDB		TAB		CBE	
	WITH 0%	WITHOUT 0%	WITH 0%	WITHOUT 0%	WITHOUT 0%	WITHOUT 0%
0	14.0		12.5			
20	15.0	11.5	16.5	12.5	11.0	35.0
50	16.0	12.5	15.5	11.5	13.0	37.0
80	15.0	12.0	15.5	12.0	12.0	36.0

The modes were ranked in a manner similar to that described in Post-Test Question B-2. As can be seen in the table, the 0% runs are scored somewhat better than the mixed runs. This is probably due to the greater familiarity that controllers have with voice-only procedures. There is no apparent difference between the various levels of mix.

3. *To which of the following factors do you attribute the observed differences?*
   a. *Changes in the traffic sample* 1
   b. *Previous exposure with this mode* 2
   c. *Mix of data link and voice link aircraft* 3
   d. *Different aircraft loads* 0
   e. *Use of the VRS* 1

The number of times each factor was checked is listed next to the choice. The question was asked 18 times; as can be seen, none of the factors seem very prevalent. The mix of data link and voice link aircraft is checked most frequently, but there is no apparent difference from this factor as seen in Question 2, above.

7.5 SUMMARY QUESTIONNAIRE

The Summary Questionnaire was given after each subject had completed all his runs. The objective of this questionnaire was to ascertain the controllers comparison of the modes.

1. *Please rank (1 indicates most preferred) the three modes which you operated on each of the characteristics listed below.*

MODE	OVERALL	WORKLOAD	CAPACITY	SUSCEPTIBILITY TO BLUNDERS	STRESSFULNESS
CBA/TAB/VRS					
CBA/FDB/VRS					
CBE/FDB/VRS					

Using the sum of the ranks as discussed in Post-Test Question B-2, scores for each mode were developed and are shown in Table 6.

TABLE 6. SUMMARIZATION OF RESPONSES TO SUMMARY QUESTION 1

CHARACTERISTIC	MODE		
	FDB	TAB	CBE
Workload	13	13	10
Capacity	13	13	10
Susceptibility to Blunders	13	11	12
Stressfulness	14	11	11
Sum of Above / 4	13–1/4	12	10–3/4
Overall	11	14	11

In terms of the individual characteristics, the CBE mode ranks well in all characteristics. The TAB mode ranks poorly in workload and capacity and well in susceptibility to blunders. The FDB mode ranks poorly in all categories, particularly in terms of stressfulness. Combining the scores of the individual characteristics gives the same order to the modes as seen in most previous responses, CBE best, TAB intermediate, and FDB last.

In looking at the overall rating, the FDB mode and the CBE mode are ranked equally, and the TAB mode worst. This ranking is contrary to all other indications. One possible cause for this rating is that when the controllers were asked for an overall evaluation, the preference for an FDB display overrode their previous operational difficulties.

2. *How do you feel that the availability of a data link in each of the following modes will effect controller workload?*

MODE	(-2) STRONGLY INCREASE	(-1) MILDLY INCREASE	(0) NO EFFECT	(+1) MILDLY DECREASE	(+2) STRONGLY DECREASE
CBA/TAB/VRS					
CBA/FDB/VRS					
CDF/FDB/VRS					

Using the scores indicated in parentheses above, the modes achieve the following scores: FDB=-8, TAB=-10, and CBE=-7. Once again, as in the overall evaluation, the FDB mode is not rated as poorly as in the other questionnaires.

3. *How would you rate realism of the scenario and simulation?*
   a. *Acceptable as is* 50%
   b. *Adequate with minor problems* 17%
   c. *Poor with major problems* 17%
   d. *Totally inadequate* 17%

Half of the subject controllers felt that the realism of the scenario and simulation was acceptable. The other half each checked one of the other responses.

8. ADDITIONAL CONTROLLER COMMENTS

In addition to checking the responses on the questionnaire, the controllers often expressed their attitude toward certain aspects of the test. Many of these comments explained the particular rationale behind a judgment, while others emphasized situations where the subjects felt very strongly about an issue. Several of the comments that typify the general reaction or clearly point out problems or recommended improvements have been selected and are discussed in the following paragraphs. In some cases, the comments have been paraphrased to define the issue being discussed.

In the general evaluation of the modes, the key troublesome issues were clearly pointed out. For the Full Data Block (FDB) mode among the comments were "Trackball functions are too time consuming – completely unsafe for peak traffic conditions!!" and "Too busy slewing and button pushing, occupies both hands – unacceptable for ATC – unsafe." For the Tabular List (TAB) mode, the comments included: "TAB List is no good – constantly requires double checking back and forth from TAB List to aircraft" and "Too busy looking at and keeping up with TAB List – very poor." For the Control-by-Exception (CBE) mode, the controllers found it disconcerting not to have pretransmission approval of commands, "especially when commands were issued which caused a conflict." For the FDB voice-only runs, one subject stated that he "was frequently unsure that commands had been transmitted," indicating a need for a manual action to change the status of voice dispatched commands in some manner.

Very frequently, the controllers commented on being paced by the computer. Comments such as: "Attention basically to machine instructions," "At peak of test, I felt compelled to push buttons without proper command elevation and this bothers me!" "Confusion reigns with four, five, or six simultaneous commands," and, "Controller is racing to keep up with the computer," indicate strong pressure not to let the queue of unattended messages get

too long. These comments were most often expressed for the TAB mode. These sentiments are also reflected in comments such as: "At peak, it was impossible to evaluate command dispatch," "At peak load, I could not keep up," and, "At peak, it was impossible to evaluate the command dispatch status."

Another aspect of the M&S algorithm which enhanced the controllers' feelings of urgency in keeping the queue small was the fact that commands timed out after 25 seconds. The controllers were in strong agreement with regard to the timeout feature. Some of the comments made were: "Commands should blink until the controller gets a chance to activate it – in no instance should it be erased," "On all runs the problem was the timeout feature. I believe the command should blink until action is taken," "Timeout hurried my dispatch to the point that command evaluation suffered," and, "With the knowledge that the commands would time out, command dispatch was hurried to the point that careful evaluation was impossible!" The timeout feature was particularly bothersome for the exception conditions. Controllers included comments such as: "Timeout complicates exception procedures" and "Timeout caused problems during failures."

The M&S conflict list was frequently commented upon. In general, the subjects discounted its utility with comments such as: "I ignored it (conflict list)" and "Sometimes the conflict list didn't work."

Although most of the controllers felt the 3-second gap between VRS transmissions was appropriate, one controller indicated that the gap length should be keyed to the traffic load. He commented that though the 3-second gap was "about right in moderate traffic" it was "too long in crucial situations."

One subject observed that the "cleared for approach" command was not acknowledged in a clear manner. He indicated that a blinking "W" (WILCO acknowledgement) was required in addition to the flashing data block, since the latter normally indicates the aircraft is in a handoff mode. Another subject indicated that the tag clutter on the base leg area along with the flashing handoff-to-tower was annoying and complicated the detection of exception

conditions. One other subject commented, "80% Voice – it was like 'over the shoulder' monitoring – the voice should also mouth the data link commands as an attention director."

9. OBJECTIVE DATA ANALYSIS

This section describes the results of the analysis of the objective (computer-generated) measures in terms of the statistical relationships identified. No attempt is made to infer the operational significance of the findings except as necessary to determine the course of further analysis. Throughout this section, key findings are underlined.

Analysis of the objective data emphasized the analysis of the results for the three modes at the 20, 50, and 80 percent levels. This analysis was divided into three areas: (1) analysis of the dynamic time line measures recorded for each message during the experiment, (2) analysis of selected measures from the end-of-run summary report, and (3) analysis of other data such as the conflict reports or the number of missed commands collected by the observer controller or the pilots during the course of the run. In a somewhat reduced form, these analyses were also produced for the voice-only runs.

9.1 DYNAMIC TIME LINE MEASURES

As previously discussed, one of the major data sources of this experiment was the measurement of time intervals for each message. From a statistical standpoint, the use of all these measurements increases the sample size and thereby the statistical power of the analysis as long as those measurements are independent. This usage was somewhat complicated by the incorporation of aircraft load as a dynamic factor within each hour's run. Therefore, each measurement must be considered as being functionally related to the actual aircraft load at the time the measurement was made. As was learned during the Data Link Phase 2A experiment (Reference 5), this relationship could be expressed in a simple linear form:

$$y = a + bt \quad (1)$$

where y = the output measure

a, b = fitted regression coefficients

t = the time of the measurement (a proxy for load)

Before the analyses of the test modes was started, the data from each experiment was processed using a simple linear regression program (from the Xerox Statistical Subroutine Library) in which the following transformation was made to all data to remove the time effect:

$$y'_t = \hat{y}_{30} + (y_t - \hat{y}_t) \quad (2)$$

where y'_t = new value of the data measure

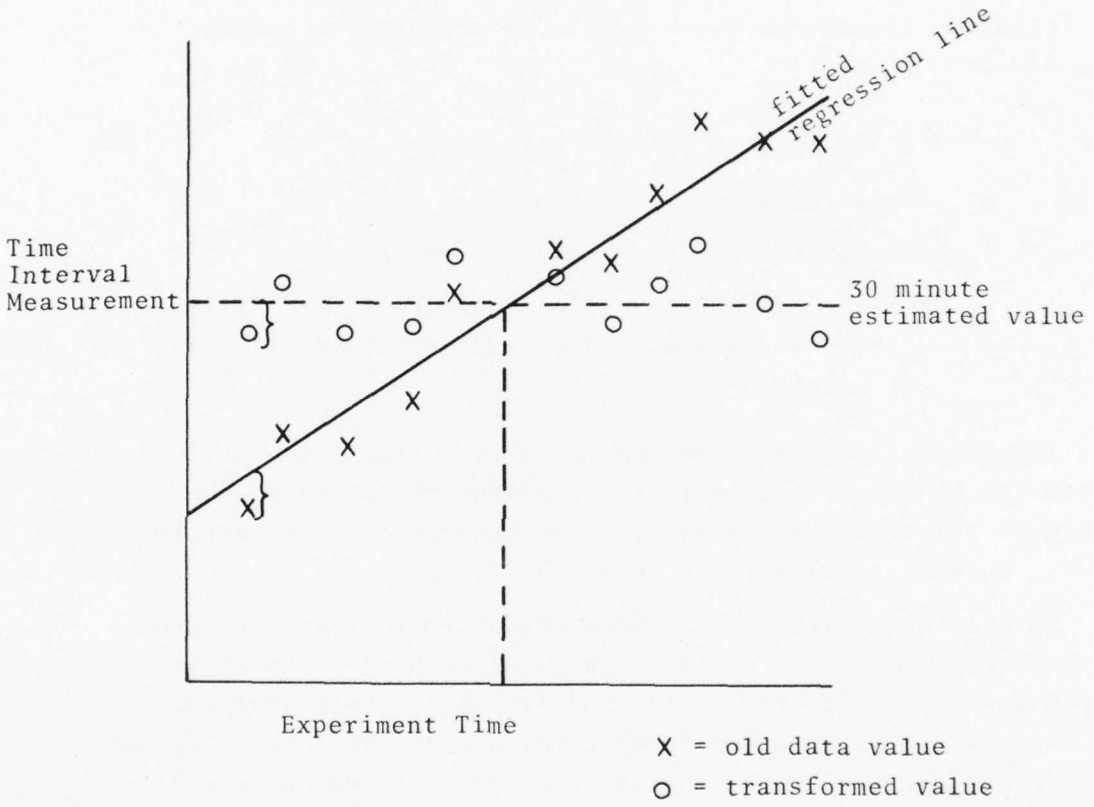
y_t = original value of the data measure

$\hat{y}_t$ = value of the regression equation fitted to the original data and evaluated at subscripted time t minutes

This transformation converted the variations around the best linear fit of the data to variations around the value at the 30 minute point in the run as shown in Figure 13. The computed slopes, b, were saved for subsequent analysis.

Three preliminary studies were conducted to test the assumptions of the techniques employed in the following analysis and to aid in the interpretation of the results of those analyses. The first was a correlation study to determine the interrelationships among the two communications time measures which were defined for all modes — message transaction time and command initiation delay. The results of this study indicate that the two measures are highly correlated. For data link messages, the measures have a .96 correlation coefficient; while for voice link messages, they have a .87 correlation.

The second preliminary study was a test for homogeneity of variance, a basic assumption underlying analysis of variance and the F test, the procedures to be used in the following analyses. In this study, the residual or error variance from each of the



NOTE: The transformation is made such that the differences between the original data and the fitted regression line, d, is maintained as an equal difference between the transformed values and the estimated value at 30 minutes.

Figure 13. Illustration of the Transformation of Dynamic Measures to Eliminate Dependency

linear regression analyses was compared in two ways using the Bartlett and Cochran tests (Reference 6). The Bartlett procedure tests whether all the sample variances can be considered equal (homogeneous). The Cochran procedure tests whether the largest sample variance is different from the others, a common occurrence which can significantly affect the results of the Analysis of Variance. These procedures were performed for the voice and data link message transaction time and command initiation delay measures. In all cases, the tests indicated a high degree of heterogeneity.

The last preliminary study investigated the autocorrelation between successive measurements in a run. Using a procedure developed by Bartlett (Reference 7), the serial correlation (i.e., the correlation between successive measurements) was tested to determine if it was significantly different from zero. In all cases, the serial correlation was significant and positive, indicating a relationship between successive measurements.

A study of the correlation between values of the residual variance and the serial autocorrelation for each experiment run indicates that both the heterogeneity of variance and the positive serial autocorrelation are manifestations of an incomplete elimination of the effects of the changing load. The effects of these factors are to overstate the level of significance in the Analysis of Variance. Therefore, in all subsequent analysis, a very conservative significance level ($\alpha > .99$) will be used in all tests of the dynamic time line measures.

The first step in this portion of the analysis consisted of making Analysis of Variance runs for command initiation delay (CID) and message transaction time (MTT) for each type of aircraft (i.e., data link or voice link equipped). These analyses were performed using the Biomedical Computer Program, BMD11V, Multivariate Analysis of Variance (Reference 8).

As was mentioned previously in Section 6, "Test Design," the experimental design for this study consisted of two replications of a 3 x 3 x 3 factorial with the run sequence blocked in

Greco-Latin Square pattern. For this problem, the model of the contributions of the different variable effects on the measured data can be expressed in linear model form as:

$$y(i,j) = \mu + M_i + P_j + I_{ij} + C_k + D_\ell + S_m + e \qquad (3)$$

where $y(i,j)$ = the measured value

μ = the grand average of the data

M_i = the effect of mode i

P_j = the effect of percentage j

I_{ij} = the interaction effect between mode i and percentage j

C_k = the effects of the sequence in which the subject experienced the modes

D_ℓ = the effects of the sequence in which the percentages were used within each mode

S_m = the effects of subject m

e = normally distributed error

Throughout this discussion, the following coding system was used:

1. For modes - 1 = CBA, TAB List, Single (TAB)
   2 = CBA, FDB, Multiple (FDB)
   3 = CBE, FDB (CBE)

2. For percentages - 1 = 20%
   2 = 50%
   3 = 80%

3. The interaction index indicates the appropriate pair of the mode and percentage indices, for example, I_{23} = the interaction effect between FDB and 80%.

4. For mode sequence - 1 = The sequence mode 1, mode 2, mode 3
   2 = The sequence mode 2, mode 3, mode 1
   3 = The sequence mode 3, mode 1, mode 2

5. For percentage sequence - 1 = mode A has the percentage sequence 1, 2, 3
mode B has the percentage sequence 2, 3, 1
mode C has the percentage sequence 3, 1, 2

2 = mode A has the percentage sequence 2, 3, 1
mode B has the percentage sequence 3, 1, 2
mode C has the percentage sequence 1, 2, 3

3 = mode A has the percentage sequence 3, 1, 2
mode B has the percentage sequence 1, 2, 3
mode C has the percentage sequence 2, 3, 1

6. The subject index indicates six different individuals selected at random from the NAFEC pool.

In this model, the effects of mode and percentage and the mode/percentage interaction are used to estimate the value of the measure under different conditions. The other effects are included to complete the description of key variables and to account for all contributions to the total variance. In presentations of the calibrated parameters, block and subject effects are shown for completeness. However, in estimating the value for a particular mode/percentage combination, they can be given an average of zero since they are of no operational importance. The model given in (3) was fitted to the data for each measure and percentage level using BMD11V.

Hypotheses were tested for each effect, including:

1. Mode effect is significant
2. Percentage effect is significant,

3. Interaction effect is significant,

4. Mode sequence effect is significant,

5. Percentage sequence effect is significant,

6. Subject effect is significant.

In all these hypothesis tests, the null hypothesis (H_0) is that the individual effects are equal to each other and to zero, while the alternate hypotehsis (H_1) is that at least one of the effects is not equal to the others. For example, for the mode effect:

H_0: $M_1 = M_2 = M_3 = 0$

H_1: $M_j \neq M_i$ $i \neq j$ for at least one i and j

The results of each analysis of variance run and of each hypothesis test are as follows:

9.1.1 Results of Linear Model Analysis for Command Initiation Delay for Data Link Aircraft

Model: $CID_D\ (i,j) = \mu + M_i + P_j + I_{ij} + C_k + D_\ell + S_m + e$

where CID_D = Command initiation delay for Data Link aircraft

μ = overall average
M_i = effect of mode, i
P_j = effect of percentage, j
I_{ij} = interaction effect between mode, i, and percentage, j
C_k = effect of mode sequence, k
D_ℓ = effect of percentage sequence, ℓ
S_m = effect of subject, m
e = normally distributed error

Estimated from experimental data:

μ = 4.405 seconds
M_1 = -1.746 seconds
M_2 = .363 seconds
M_3 = 1.383 seconds

} significant

P_1 = -.214 seconds
P_2 = .144 seconds
P_3 = .070 seconds

not significant

I_{ij} in seconds

i \ j	1	2	3
1	-.161	-.797	.958
2	-.027	.921	-.894
3	.188	-.124	-.064

significant

C_1 = -.403 seconds
C_2 = .642 seconds
C_3 = -.239 seconds

significant

D_1 = -.298 seconds
D_2 = .710 seconds
D_3 = -.412 seconds

significant

S_1 = .573 seconds
S_2 = -.918 seconds
S_3 = -1.220 seconds
S_4 = .782 seconds
S_5 = .234 seconds
S_6 = .549 seconds

significant

For example: $CID_D\ (2,1) = \mu + M_2 + P_1 + I_{2,1} + \bar{C} + \bar{D} + \bar{S} + \bar{e}$

$= 4.405 + .363 - 0.0 - .027 = 4.741$ seconds

Where $\bar{C}$ = average mode sequence effect = 0

$\bar{D}$ = average percentage sequence effect = 0

$\bar{S}$ = average subject effect = 0

$\bar{e}$ = average error = 0

9.1.2 Results of Linear Model Analysis for Message Transaction Time for Data Link Aircraft

Model: $MTT_D(i,j) = \mu + M_i + P_j + I_{ij} + C_k + D_\ell + S_m + e$

where MTT_D = Message transaction time for Data Link aircraft

μ = overall average
M_i = effect of mode, i
P_j = effect of percentage, j
I_{ij} = interaction effect between mode, i, and percentage j
C_k = effect of mode sequence, k
D_ℓ = effect of percentage sequence, ℓ
S_m = effect of subject, m
e = normally distributed error

Estimated from experimental data:

μ = 8.429 seconds

M_1 = -1.727 seconds
M_2 = .475 seconds
M_3 = 1.252 seconds

} significant

P_1 = -.149 seconds
P_2 = .141 seconds
P_3 = .008 seconds

} not significant

I_{ij} in seconds

i \ j	1	2	3
1	-.313	-.881	1.194
2	.201	1.034	-1.235
3	.112	-.153	.041

} significant

C_1 = -.213 seconds
C_2 = .610 seconds
C_3 = -.397 seconds

} significant

D_1 = -.450 seconds
D_2 = .660 seconds
D_3 = -.210 seconds

} significant

S_1 = .348 seconds
S_2 = -.817 seconds
S_3 = -.840 seconds
S_4 = .567 seconds
S_5 = .359 seconds
S_6 = .383 seconds

} significant

For example: $MTT_D\ (3,3) = \mu + M_3 + P_3 + I_{3,3} + \bar{C} + \bar{D} + \bar{S} + \bar{e}$
$= 8.429 + 1.252 + .00 + .041 = 9.722$ seconds

9.1.3 Results of Linear Model Analysis for Command Initiation Delay for Voice Link Aircraft

Model: $CID_V\ (i,j) = \mu + M_i + P_j + I_{ij} + C_k + D_\ell + S_m + e$

where CID_V = Command initiation delay for Voice Link aircraft

μ = overall average
M_i = effect of mode, i
P_j = effect of percentage, j
I_{ij} = interaction effect between mode, i, and percentage, j
C_k = effect of mode sequence, k
D_ℓ = effect of percentage sequence, ℓ
S_m = effect of subject, m
e = normally distributed error

Estimated from experimental data:

μ = 4.327 seconds

M_1 = -1.341 seconds
M_2 = -.111 seconds
M_3 = 1.452 seconds

} significant

P_1 = -.245 seconds
P_2 = -.143 seconds
P_3 = .388 seconds

} significant

I_{ij} in seconds

i \ j	1	2	3
1	-.200	-.500	.700
2	-.048	.360	-.312
3	.248	.140	-.388

significant

C_1 = .047 seconds
C_2 = .091 seconds
C_3 = -.138 seconds

not significant

D_1 = -.030 seconds
D_2 = .232 seconds
D_3 = -.202 seconds

significant

S_1 = .323 seconds
S_2 = -.090 seconds
S_3 = -.638 seconds
S_4 = .379 seconds
S_5 = -.019 seconds
S_6 = .045 seconds

significant

For example: $CID_V\ (2,1) = \mu + M_2 + P_1 + I_{2,1} + \bar{C} + \bar{D} + \bar{S} + \bar{e}$

$= 4.327 - .111 - .245 - .048 = 3.923$ seconds

9.1.4 Results of Linear Model Analysis for Message Transaction Time for Voice Link Aircraft

Model: $MTT_V\ (i,j) = \mu + M_i + P_j + I_{ij} + C_k + D_\ell + S_m + e$

where MTT_V = Message transaction time for Voice Link aircraft

μ = overall average
M_i = effect of mode, i
P_j = effect of percentage, j
I_{ij} = interaction effect between mode, i, and percentage, j
C_k = effect of mode sequence, k

D_ℓ = effect of percentage sequence, ℓ
S_m = effect of subject, m
e = normally distributed error

Estimated from experimental data:

μ = 9.010 seconds

M_1 = -1.290 seconds M_2 = -.298 seconds M_3 = 1.588 seconds	}	significant
P_1 = .250 seconds P_2 = -.122 seconds P_3 = -.128 seconds	}	not significant

I_{ij} in seconds

i \ j	1	2	3	
1	-.122	-.393	.515	} not significant
2	.036	.036	-.072	
3	.086	.357	-.443	

C_1 = .116 seconds C_2 = -.028 seconds C_3 = -.088 seconds	}	not significant
D_1 = -.066 seconds D_2 = .564 seconds D_3 = -.498 seconds	}	significant
S_1 = -.072 seconds S_2 = -.053 seconds S_3 = -.087 seconds S_4 = -.016 seconds S_5 = .011 seconds S_6 = .217 seconds	}	not significant

For example: $MTT_V\ (3,2) = \mu + M_3 + P_2 + I_{3,2} + \bar{C} + \bar{D} + \bar{S} + \bar{e}$
$= 9.010 + 1.588 - .00 + .00 = 10.598$ seconds

Examination of the above data shows that the average CID for voice and data link aircraft were approximately equal at slightly under 4.5 seconds. For MTT, messages to voice-link aircraft, with an average of 9.0 seconds, were somewhat longer than messages to data link aircraft, with an 8.4 second average.

The mode effect was always significant. The TAB mode was fastest and was approximately 2 seconds better than the FDB mode, which in turn was approximately 1 second better than the CBE mode in all four cases. The percentage effect was statistically significant only in the case of voice link aircraft for CID, and even then the range of difference was only slightly higher than .5 second. The interaction effect was generally significant and demonstrated the same pattern for all four measures, as shown in Figure 14, although the pattern was somewhat stronger for the CID measures than for the MTT measures. For the TAB mode, the 50 percent level was the fastest and the 80 percent the slowest. The converse was true for the FDB mode in which the 50 percent level was slowest and the 80 percent fastest. For the CBE mode, all percentage levels were approximately equal.

To complete the description of the dynamic data, the slope values, b, computed in the preliminary regression step (1) were analyzed to determine if the relationship with load was in any way affected by the experiment variables. These values were analyzed using BMD11V to fit a model with the same form as shown in (3). Since only one value was computed per run and the homogeneity of variance is not a problem, the tests of hypotheses used a rejection level of .05. As shown in Table 7, there were no significant effects for any of the measures. The magnitude of the slopes indicate that the MTT values were more significantly affected by increasing load, particularly the MTT for voice link aircraft, which increased over five times as fast as any other measure.

The results of the slope analysis can be combined with the analysis of the measures to produce point estimates of the measures at different loads by the use of the model:

$$y(i,j) = \mu + M_i + P_j + I_{ij} + b_{ij} * (t-30) \qquad (4)$$

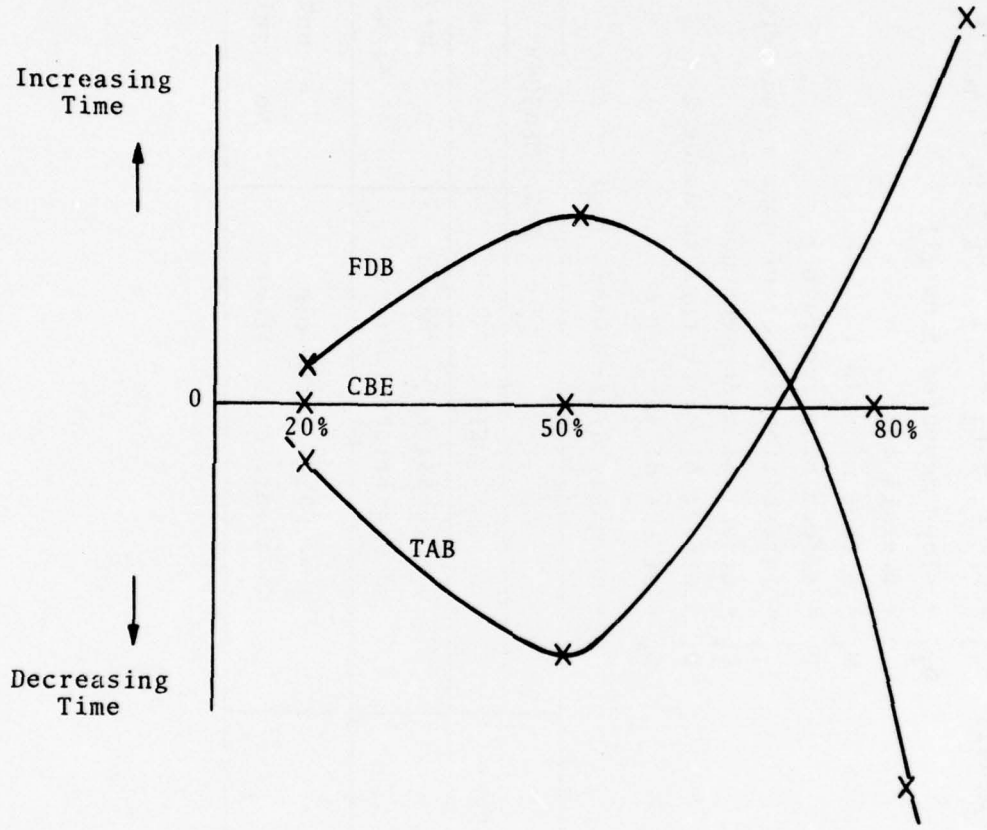


Figure 14. General Pattern of Interaction Effect

TABLE 7. RESULTS OF LINEAR/MODEL ANALYSIS FOR SLOPES

Model: $b_{ij} = \mu + M_i + P_j + I_{ij} + C_k + D_\ell + S_m + e$

b_{ij} = slope computed from (1)
μ = overall mean
M_i = effect of mode i
P_j = effect of percentage j
I_{ij} = interaction effect of mode i and percentage j
C_k = effect of mode sequence k
D_ℓ = effect of percentage sequence ℓ
S_m = effect of subject m
e = normally distributed error

TYPE OF AIRCRAFT	DATA MEASURE	
	CID	MTT
DATA LINK EQUIPPED	μ = .00831 secs./mins. No significant effects	μ = .0128 secs./mins. No significant effects
VOICE LINK EQUIPPED	μ = .00857 secs./mins. No significant effects	μ = .06309 secs./mins. No significant effects

where $Y(i,j)$ = the measure of interest

μ = average effect for that measure

M_i = effect of mode i for that measure

P_j = effect of percentage j for that measure

b = slope for measure of interest

t = experiment time at which load equals desired value (See Section 5.5, "Traffic Samples")

For example, if the value of Command Initiation Delay is desired for voice link aircraft for mode one (TAB) at percentage level two (50 percent) at a load of approximately seven aircraft (i.e., about 42 minutes into the run), then that value can be computed from the values μ, M_i, P_j, and I_{ij} from section 9.1.3 and the value of b_{ij} from Table 7 as follows:

$$CID_V(1,2) = \mu + M_i + P_j + I_{ij} + b_{ij} * (t-30)$$

$$= 4.327 - 1.341 - .143 - .500 + .00857 * (42-30)$$

$$= 2.446 \text{ seconds}$$

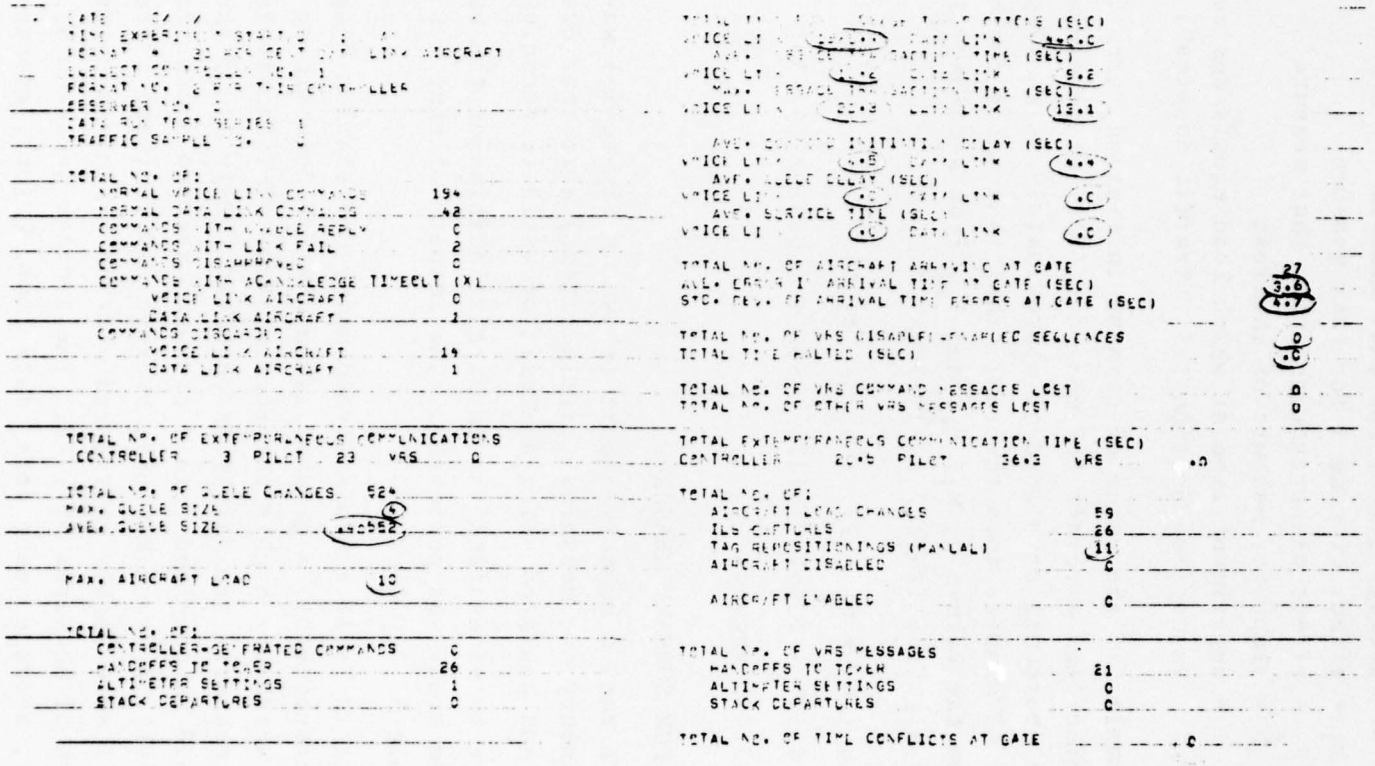
9.2 END OF RUN SUMMARY MEASURES

For each run, a summary report (shown in Figure 15) was prepared. Twenty-one measures (circled items) were retrieved from each report and analyzed using analysis of variance techniques.

In order to understand the interrelationships of the measures, a correlation study was performed. The results of that study are shown in Table 8.

The data in Table 8 was analyzed using Correlation Cluster Analysis (Reference 9), a graphic technique for determining groups of measures which are highly intercorrelated. This analysis showed that the 21 measures are fairly independent. Only 13 of the measures could be combined into 4 groups: average queue size, average MTT(VL), average MTT(DL), average CID(VL), and average CID(DL). All were highly correlated, having a minimum correlation between pairs of measures of .777. The average queue delay voice link, QD(VL), average queue delay data link, QD(DL), average service time voice link, ST(VL), and average service time data link, ST(DL), are correlated at a minimum of .624. Two other pairs of measures, total MTT(VL) and max MTT(VL) at .800, and the number of VRS halt/

Copy available to DDC does not permit fully legible reproduction



COPY AVAILABLE TO DDC DOES NOT PERMIT FULLY LEGIBLE PRODUCTION

Figure 15. Data Reduction Program Summary Printout

TABLE 8. RESULTS OF CORRELATION STUDY FOR END-OF-RUN SUMMARY MEASURES

	Max Queue Size	Avg Queue Size	Max Acft. Load	Total MTT (VL)	Total MTT (DL)	Avg MTT (VL)	Avg MTT (DL)	Max MTT (VL)	Max MTT (DL)	Avg CID (VL)	Avg CID (DL)	Avg QD (VL)	Avg QD (DL)	Avg ST (VL)	Avg ST (DL)	Avg Error in Arrival Time	STD Dev of Arrival Time Error	No. VRS Disable/ Enable Sequences	Total Time Halted	No. Tag Repositioning	No. Time Conflicts
1	1.000	.342	.077	-.138	.209	.054	.296	-.117	.273	.269	.293	-.065	-.054	-.184	-.061	.020	.021	.119	.043	.018	.056
2		1.000	.169	.232	.250	.800	.963	.317	.043	.966	.968	-.599	-.623	-.387	-.705	.493	.073	.105	-.014	.192	.084
3			1.000	-.003	.088	.118	.172	.090	.112	.146	.158	.012	.006	-.082	-.030	.077	.181	.106	-.196	-.048	-.008
4				1.000	-.846	.678	.146	.800	-.321	.252	.202	-.118	-.161	.152	-.190	.176	-.103	.177	.220	.104	-.056
5					1.000	-.206	.370	-.590	.219	.265	.318	-.162	-.131	-.200	-.150	.196	.192	-.128	-.252	.022	.034
6						1.000	.777	.732	-.292	.848	.809	-.476	-.525	-.043	-.600	.580	.090	.204	.099	.161	-.074
7							1.000	.270	-.045	.978	.997	-.611	-.624	-.302	-.709	.598	.093	.083	-.078	.196	.009
8								1.000	-.228	.352	.303	-.131	-.202	.148	-.235	.305	.120	.296	.280	.078	-.092
9									1.000	-.119	-.077	.181	.134	-.479	.121	-.421	.022	-.007	.120	-.021	.320
10										1.000	.984	-.596	-.618	-.259	-.707	.628	.101	.122	-.030	.164	-.041
11											1.000	-.614	-.624	-.281	-.710	.604	.083	.092	-.062	.200	.006
12												1.000	.959	.654	.956	-.469	.030	-.093	-.089	+.002	.112
13													1.000	.642	.946	-.475	.040	-.123	-.083	.038	.125
14														1.000	.675	.078	.143	-.092	-.201	.113	-.122
15															1.000	-.486	.028	-.137	-.112	.013	.117
16																1.000	.196	.087	-.074	-.072	-.572
17																	1.000	.195	-.047	-.050	.038
18																		1.000	.738	-.017	-.006
19																			1.000	-.084	.036
20																				1.000	.305
21																					1.000

resume sequences and the total time the VRS was halted at .738, could also be grouped. The other measures could all be considered as independent.

These groupings reflect the fact that the measures are strongly related and that each measure cannot be considered a separate indicator of system value. The evaluation of these measures should be considered in these groups.

Each measure was analyzed using BMD11V. A model of the form of (3) was fitted to each measure and the significance of the parameters was tested at the .05 level. The results of these analyses are summarized in Table 9.

Study of the group 1 measures, average queue size, average MTT(VL), average MTT(DL), average CID(VL), and average CID(DL), indicates very similar responses, as expected. In all cases the mode effect is significant, with the TAB mode having the lowest values followed by the FDB mode and with CBE having the largest values. The percentage effect is significant only for the average MTT(VL) measure, with that value decreasing nearly linearly for increasing percentage.

The group 2 measures, average QD(VL), average QD(DL), average ST(VL), and average ST(DL), indicate significant values for the mode effects; however, since QD and ST are only defined for the TAB mode, the results should be interpreted as average values of .462 seconds for average QD(VL), .543 seconds for average QD(DL), 7.75 seconds for average ST(VL), and 2.17 seconds for average ST(DL). No other significant effects were observed.

Group 3 measures, total MTT(VL) and maximum MTT(VL), although highly correlated have some differences in the significance of the effects. For total MTT(VL), both the mode and percentage effects are significant with the modes ranked in the normal TAB, FDB, CBE sequence. The total MTT(VL) decreases with increasing percentage of data link aircraft as expected. For the maximum MTT(VL), however, there is no significant mode effect, although the percentage effect also decreases with increasing data link percentage.

TABLE 9. RESULTS OF ANALYSIS OF VARIANCE FOR END-OF-RUN SUMMARY DATA

	MEAN VALUE	MODE EFFECT	PERCENTAGE EFFECT	INTERACTION EFFECT	MODE SEQUENCE EFFECT	PERCENTATE SEQUENCE EFFECT	SUBJECT EFFECT
MAX QUEUE SIZE	4.06	NS	NS	NS	NS	NS	NS
AVG. QUEUE SIZE	.363	1=-.125 2= .031 3= .094	NS	NS	NS	NS	NS
MAX ACFT. LOAD	9.76	NS	NS	NS	NS	NS	1=-.315 4=+.685 2=-.315 5=-.537 3=+.241 6=+.211
TOTAL MTT (VL) (SECONDS)	1251.9	1=-172. 2= -44. 3= 216.	1= 807. 2= -33. 3=-774.	NS	NS	NS	NS
TOTAL MTT (DL) (SECONDS)	1070.7	1=-178. 2= -9. 3= 187.	1=-664. 2= 23. 3= 641.	NS	NS	NS	NS
AVG. MTT (VL) (SECONDS)	9.77	1=-1.54 2= -.11 3= 1.65	1= .91 2=-.02 3=-.89	NS	NS	NS	NS
AVG. MTT (DL) (SECONDS)	9.18	1=-1.59 2= .15 3= 1.44	NS	NS	NS	NS	NS

NS = Not Significant

TABLE 9. RESULTS OF ANALYSIS OF VARIANCE FOR END-OF-RUN SUMMARY DATA (CONTINUED)

	MEAN VALUE	MODE EFFECT	PERCENTATE EFFECT	INTERACTION EFFECT	MODE SEQUENCE EFFECT	PERCENTAGE SEQUENCE EFFECT	SUBJECT EFFECT
MAX MTT (VL) (SECONDS)	20.40	NS	1= 3.56 2= 2.04 3=-5.60	NS	NS	NS	NS
MAX MTT (DL) (SECONDS)	14.20	1= -.06 2= 1.65 3=-1.59	NS	NS	NS	NS	NS
AVG. CID (VL) (SECONDS)	4.30	1=-1.59 2= .09 3= 1.50	NS	NS	NS	NS	1= .011 4= .433 2=-.089 5=-.100 3=-.578 6=+.323
AVG. CID (DL) (SECONDS)	4.32	1=-1.60 2= .12 3= 1.48	NS	NS	NS	NS	NS
AVG. QD (VL) (SECONDS)	.154	1= .308 2=-.154 3=-.154	NS	NS	NS	NS	NS
AVG. QD (DL) (SECONDS)	.181	1= .362 2=-.181 3=-.181	NS	NS	NS	NS	NS
AVG. ST (VL) (SECONDS)	4.46	1= 8.92 2=-4.46 3=-4.46	NS	NS	NS	NS	NS

NS = Not Significant

TABLE 9. RESULTS OF ANALYSIS OF VARIANCE FOR END-OF-RUN SUMMARY DATA (CONCLUDED)

	MEAN VALUE	MODE EFFECT	PERCENTAGE EFFECT	INTERACTION EFFECT	MODE SEQUENCE EFFECT	PERCENTAGE SEQUENCE EFFECT	SUBJECT EFFECT
AVG. ST (DL) (SECONDS)	0.72	1= 1.45 2= -.72 3= -.72	NS	NS	NS	NS	NS
AVG. ERROR IN ARRIVAL TIME (SECONDS)	2.76	1=-2.21 2=-1.30 3= 3.51	NS	NS	NS	NS	NS
STD. DEVIATION OF ARRIVAL TIME ERROR (SECONDS)	6.20	NS	NS	NS	NS	NS	NS
NO. OF VRS DISABLE/ENABLE SEQUENCES	.70	NS	NS	NS	NS	NS	NS
TOTAL TIME HALTED (SECONDS)	3.91	NS	NS	NS	NS	NS	NS
NO. TAG REPOSITIONINGS	5.46	NS	NS	NS	NS	NS	NS
NO. TIME CONFLICTS	.30	NS	NS	NS	NS	NS	NS

NS = Not Significant

For data link aircraft, the two complementary measures, total MTT(DL) and maximum MTT(DL), show somewhat different responses. The total MTT(DL) also has significant mode and percentage effects. The mode effects are in the normal relationship, TAB, FDB, CBE. The percentage effect, however, increases with increasing data link percentage as expected. For the maximum MTT(DL) measure only the mode effect is significant. In this case, the CBE mode has the lowest maximum value, the TAB mode the second lowest, and FDB the worst.

Among the other measures, the only one with any significant effect is the average error in arrival time which has a significant mode effect.

9.3 OTHER MEASURES

As described in Section 5, "Test Conduct," additional information was collected by the observer controller and the simulation pilots during the course of each run. Specifically, the observer controller recorded warnings of imminent conflicts as detected by the subject. This information included the identification of the aircraft pair in conflict and the time the subject issued the warning.

In addition, the pilots kept a handwritten tally of the number of non-M&S commands issued by the subject. The tabulations of such controller generated commands by mode and percentage are given in Table 10. As can be seen, there is no pattern to the responses.

TABLE 10. TABULATION OF CONTROLLER GENERATED COMMANDS

DATA LINK PERCENTAGE	MODE		
	TAB	FDB	CBE
20	10	28	12
50	7	11	8
80	24	11	4

The subject controllers were instructed to inform the observer whenever it was necessary to maneuver the aircraft to provide separation. The observer recorded the aircraft indentifications and the time at which the warning was given. After the simulation run, an analysis of the aircraft tracks was performed to determine which pairs of aircraft actually came into conflict and the time at which those conflicts began.

The two sets of data were compared manually, and counts were made of the number of times in which (1) the controller detected a conflict situation which did not materialize, (2) an actual conflict was not detected by the subject, and (3) the controller detected the conflict more than 15 seconds after it started. The results of this effort are given in Table 11. The missed conflicts were further examined for some indication of severity. Several of the actual conflicts were very minor incidents in which the aircraft were in conflict for only one scan. If these conflicts are removed from consideration, the numbers of missed conflicts are reduced to those indicated in parentheses in Table 11.

Statistical analysis of the conflict data is somewhat complicated, since they are discrete, relatively rare events, and since the number of actual conflicts in any of the simulation experiments was not controlled but was itself a random variable. Assuming that the number of actual conflicts, n, in any hour is a Poisson variable with unknown arrival (occurrence) rate, λ, then

$$\text{Prob } (n) = \frac{\lambda^n}{n!} e^{-\lambda}.$$

If the detection of a conflict is a Bernoulli event with the probability of failing to detect the conflict, θ, the number of missed detections, x, given the number of actual conflicts, n, is then distributed as a binomial with

$$\text{Prob}(x|n) = \binom{n}{x} \theta^x (1-\theta)^{n-x}$$

TABLE 11. REDUCED CONFLICT DATA

MODE	20%			50%			80%			TOTAL		
	E	M	L	E	M	L	E	M	L	E	M	L
FDB	5	2(2)	3	2	4(1)	4	4	6(3)	3	11	12(6)	10
TAB	5	5(1)	5	0	2	4	4	2	4	9	9(5)	13
CBE	3	2(1)	2	2	2(1)	5	2	0	1	7	4(2)	8
TOTAL	13	9(4)	10	4	8(2)	13	10	8(5)	8	27	25(13)	31

E = Conflict detected by controller which did not materialize

M = Actual conflict which was not detected by the controller

L = Actual conflict which was detected at least 15 seconds after the start of the conflict

The probability of z undetected conflict is then the sum of the situations in which the actual number of conflicts, n, is greater than z; of these n, only z are undetected. This can be expressed as:

$$\text{Prob}\ (z) = \sum_{n=z}^{\infty} \text{Prob}(n) \cdot \text{Prob}\ (z|n)$$

$$= \sum_{n=z}^{\infty} \frac{\lambda^n}{n!} e^{-\lambda} \frac{n!}{z!(n-z)!} \theta^z (1-\theta)^{n-z}$$

By algebraic manipulation, this expression can be transformed to

$$\text{Prob}\ (z) = \frac{e^{-\lambda z}(\lambda\theta)^z}{z!} \sum_{i=0}^{\infty} \frac{i_{(1-\theta)^i}}{i!} \qquad i = n - z$$

However,

$$\sum_{i=0}^{\infty} \frac{\lambda^i (1-\theta)^i}{i!} = e^{\lambda(1-\theta)}$$

and

$$\text{Prob}\ (z) = \frac{e^{-\lambda z}(\lambda\theta)^z}{z!} e^{\lambda(1-\theta)}$$

$$= \frac{e^{-\lambda\theta}(\lambda\theta)^z}{z!}$$

which is the expression for the probability of Poisson variable with arrival rate = $\lambda\theta$.

In addition, since the observed average conflict rates were small, it is not valid to approximate the Poisson data by a normal probability distribution. Therefore, classical statistical test procedures which assume that the variables are normally distributed are invalid.

Several alternate procedures were applied to the missed conflict data, both filtered and nonfiltered, to determine whether the data for the three modes exhibited significant differences. A χ^2 test statistic assuming Poisson distributed variables and Friedman's nonparametric two-way analysis of variance both indicated that the effect due to mode was not significant for the missed conflicts in both the filtered and nonfiltered data.

9.4 ANALYSIS OF VOICE RUN DATA

Voice runs for the TAB and FDB modes were added to the experiment to provide a baseline against which to compare the experiment results. The adjustment for load was performed on the dynamic measures for the voice runs as for the VRS runs. The dynamic measures were analyzed using BMD11V to fit a model of the form:

$$y(i) = \mu + M_i + S_m + e$$

where $y(i)$ = the dynamic measurement
μ = overall average
M_i = effect of mode i
S_m = effect of subject m
e = normally distributed error

The results of this analysis for the CID measures are shown in Sections 9.4.1 and 9.4.2. It can be seen that the mode effect is not significant. The CID values are approximately equal to the TAB CID values for the mixed voice and data link runs. The MTT values represent a 1 second improvement for TAB and a 2 second improvement for FDB.

9.4.1 Results of Linear Model Analysis for Command Initiation Delay for Voice Runs

Model: $CID_V(i) = \mu + M_i + S_m + e$

where $CID_V(i)$ = measured command initiation delay

μ = overall mean
M_i = effect of mode i
S_m = effect of subject m
e = normally distributed error

μ = 3.076 seconds

M_1 = -.012 seconds
M_2 = .012 seconds

} not significant

S_1 = .134 seconds
S_2 = -.702 seconds
S_3 = .122 seconds
S_4 = .573 seconds
S_5 = .174 seconds
S_6 = -.301 seconds

} not significant

For example: $CID_V(1) = \mu + M_1 + \bar{S} + \bar{e}$
$= 3.076 - 0.043 = 3.076$ seconds

9.4.2 Results of Analysis of Variance for Message Transaction Time for Voice Runs

Model: $MTT_V(i) = \mu + M_i + S_m + e$

where: $MTT_V(i)$ = measured message transaction time

μ = overall mean
M_i = effect of mode i
S_m = effect of subject m
e = normally distributed error

μ = 6.665 seconds

M_1 = .035 seconds
M_2 = -.035 seconds

} not significant

S_1 = .234 seconds
S_2 = -.790 seconds
S_3 = .369 seconds
S_4 = .731 seconds
S_5 = .162 seconds
S_6 = -.706 seconds

} significant

For example: $MTT_V(2) = \mu + M_2 + \bar{S} + \bar{e}$
$= 6.665 - 0.035 = 6.665$ seconds

With only 12 data points for each measures, the slopes, the end-of-run summary data, and the other measures were not statistically analyzed. The average values for each mode were computed and are shown in Table 12. For the end-of-run summary data, except for the number of tag repositionings, the two modes were virtually the same. The FDB mode had many more tag repositionings than TAB. In terms of the other measures, the TAB mode had less missed commands and controller generated clearances than FDB. However, FDB had less than half the extra and missed conflict warnings.

When compared to the mixed data link and voice link runs shown in Tables 9, 10, and 11, most of the results were reasonably in agreement. The only meaningful differences were: (1) the total and average MTT values are substantially lower for the voice-only runs, and (2) in all cases, most of the other measures, controller clearnaces, extra warnings, and missed warnings were somewhat lower for the voice-only runs.

TABLE 12. AVERAGE RUN VALUES BY MODE FOR VOICE RUNS

END-OF-RUN SUMMARY MEASURES	TAB LIST	FULL DATA BLOCK
1. Max Queue Size	4.0	3.8
2. Avg Queue Size	.338	.366
3. Max Aircraft Load	9.5	9.8
4. Total MTT (VL)	1717.2	1643.7
5. Total MTT (DL)	-	-
6. Avg MTT (VL)	7.27	7.05
7. Avg MTT (DL)	-	-
8. Max MTT (VL)	22.4	21.9
9. Max MTT (DL)	-	-
10. Avg CID (VL)	3.57	3.62
11. Avg CID (DL)	-	-
12. Avg QD (VL)	3.57	undefined
13. Avg QD (DL)	-	-
14. Avg ST (VL)	4.58	undefined
15. Avg ST (DL)	-	-
16. Avg Error in Arrival Time	.88	-.28
17. Std. Dev. Arrival Error	5.47	5.55
18. No. VRS Sequences	-	-
19. Total Time Halted	-	-
20. No. Tag Repositionings	1.00	6.12
21. No. Time Conflicts	.00	.33
OTHER MEASURES		
1. Controller Clearances	.83	1.09
2. Extra Warnings	.42	.17
3. Missed Warnings	.42	.17
4. Late Warnings	.25	.25

10. ANALYSIS OF RESULTS

In this section, the most pertinent subjective and objective data are evaluated and reconciled in an effort to ascertain mode rankings. Data pertaining to the following experimental measures were used in this analysis:

1. Controller opinion questionnaires
2. Time to execute equivalent functions such as Message Transaction Time (MTT) and Command Initiation Delay (CID)
3. Queue behavior (average number of commands awaiting dispatch)
4. Counts of conflicts (separation violations that occurred and the number detected by controller)

There was a very definite recurring pattern in the answers to the Post-Experiment questionnaire. In a vast majority of the questions, the order of controller preference with respect to the control modes was CBE/FDB/VRS, CBA/TAB/VRS, and CBA/FDB/VRS. This was especially evident in such questions as A-1.A and A-1.B, which exhibited significant mode effects. A-1.A asked the subject how he believed "this mode of displaying and transmitting commands would affect the current operational system for the same traffic situation." Question A-1.B asked the subject to rate the mode with respect to workload, capacity, blunder, and stress. By assigning the weight (+2, +1, ..., -2) to every answer to these questions, summing the results algebraically, and dividing the sum by the number of responses, a normalized overall index of merit was computed for each mode. Table 13 lists the figures used in deriving an overall index of merit for each mode and the ranking of each mode on the basis of the index. Not surprisingly, the rankings in Table 13 coincide with those derived more straightforwardly in the Subjective Analysis Section.

TABLE 13. MODE RANKINGS BASED ON POST-EXPERIMENT QUESTIONS A.1-A AND A.1-B

MODE	MIX	SUM OF WEIGHTED REPLIES A.1-A	OVERALL NORMALIZED INDEX (SUMS ÷ 18)	A.1-A RANK	SUM OF WEIGHTED REPLIES A.1-B				OVERALL NORMALIZED INDEX (SUMS ÷ 24)	A-1.B RANK		
					WORKLOAD	CAPACITY	BLUNDERS	STRESS		20%	50%	80%
CBA/ FDB	20%	-5			-5	-4	-7	-7	-.96	3		
	50%	-5	-.83	3	-7	-5	-6	-5	-.96		3	
	80%	-5			-8	-6	-7	-9	-1.3			3
CBA/ TAB	20%	0			-1	-2	-5	-8	-.67	2		
	50%	-3	-.55	2	-1	1	-5	-6	-.46		2	
	80%	-7			0	0	-5	-6	-.46			2
CBE/ FDB	20%	3			5	4	0	-1	+.33	1		
	50%	2	-.44	1	3	5	0	0	+.33		1	
	80%	2			3	3	0	-1	+.29			1

The index of merit ranking derived from A-1.A and A-1.B data was verified by repeating the procedure with responses to Summary Questionnaire Question 1, which asked the controllers to rank all of the modes with respect to workload, capacity, susceptibility to error, stressfulness, and overall performance. The calculation of an overall rank index for each mode under these five criteria is summarized in Table 14 on the basis of responses shown in Table 6 in Section 7 of this report.

All the ranking data presented thus far will now be examined to extract information on:

a. Overall ratings

b. Ratings in each of the categories (workload, capacity, blunders, stress)

c. Variations of ratings with percentage mix.

First, the rankings obtained from question A-1.A and the overall rankings obtained from Summary Questionnaire Question 1 are restated in order in Table 15.

TABLE 14. OVERALL RANKINGS BASED ON SUMMARY QUESTION 1

MODE	SUMMARY QUESTION 1 COMPONENT RESPONSES*				OVERALL NORMALIZED RANK INDEX (SUM ÷ 24)	RANK	OVERALL RATING FROM SUMMARY QUESTION 1	OVERALL NORMALIZED RANK INDEX (SUM ÷ 24)	RANK
	WORKLOAD	CAPACITY	BLUNDER	STRESS					
CBA/FDB	13	13	13	14	2.2	3	11	.46	1
CBA/TAB	13	13	11	11	2	2	14	.58	2
CBE/FDB	10	10	12	11	1.8	1	11	.46	1

*The larger the number, the lower the rating.

TABLE 15. SUMMARY TABLE OF OVERALL MODE RANKINGS

MODE	QUESTION A.1.A & A.1.B RANKING	SUMMARY QUESTION 1	
		COMPONENT RANKING	OVERALL RANKING
CBA/FDB	3	3	1
CBA/TAB	2	2	2
CBE/FDB	1	1	1

Since the rankings from Questions A.1.A and A.1.B are identical with the rankings derived from the components of Summary Question 1 (the latter representing a direct evaluation of the modes relative to each other), and since the Summary Questionnaire responses were obtained after each subject had been exposed to the three modes, the following ranking will be regarded as the final consensus on the overall order of merit:

1. CBE/FDB
2. CBA/TAB
3. CBA/FDB

In arriving at the final ranking above, the difference between the component and overall rankings in Summary Question 1 was not considered to be of sufficient weight to effect a change in the final ranking. However, one explanation for the difference might be that when faced with a general question of choice, the controllers expressed their preference to having the command displayed in the third line of the data block – i.e., trading off lesser workload for better command location.

The results of the Phase 2A DSF experiment indicated that the subject controllers distinctly preferred computer-generated voice when voice link traffic is heavy. In this experiment, however, the subject controllers, while acknowledging that it was less tiring than voice commands, found the inability to repeat commands and the talking rate unsatisfactory. Contrary to the assumption that the controller can look ahead to the next command while listening to the computer-generated voice, the controllers in this experiment did not indicate a predilection to do so. The less than positive reaction to the VRS was obtained despite the fact that the controllers enjoyed the convenience of employing the same keyboard functions to dispatch both data link and voice link commands.

The consistently high ranking of the CBE/FDB mode indicates a strong preference for command location in the full data block as opposed to the tabular list, despite the fact that CBA/FDB is consistently ranked last. The latter is obviously attributable to the

fact that the controllers intensely dislike the trackball slew action required for each dispatch action. The principal advantage of FDB, whether it be CBA or CBE, is that the commands are displayed in each alphanumeric tag, close to the aircraft in question. Correlation of the command with the aircraft's current status is facilitated by displaying the command in the data block. However, the controllers were bothered by having to prevent overlapping data blocks which obscure commands. This was more significant in CBA/FDB than CBE/FDB, probably because command display time was longer. Unfavorable reaction to CBA/TAB stems from the fact that the controllers had to split their attention between the list and the traffic situation. The greater ease of dispatch – i.e., single key depression without trackball slew – and the correlation aid of having the controller symbol blink with an associated blinking command were not sufficient to overcome the unfavorable reaction to a list format.

The most significant aspect of the mode rankings is the consistently favorable reaction to control-by-exception. As shown in Table 16, this mode also ranked first with respect to all the components of *Summary Question 1* with the exception of Susceptibility to Blunders; there it ranked second in a close ranking of all modes. Some concern was expressed at not having pretransmission approval, but this was attributed to controller lack of confidence in the M&S algorithm generating conflict-free commands. Only one controller stated that he could not monitor the commands at peak load; another suggested that, at the 80% traffic mix, the VRS also enunciate data link commands as an "attention director."

The quantitative measurements made after the test runs were conducted to provide a basis for understanding the subjective data. The Message Transaction Time (MTT) represents the total time required for the system to process a given command and is thus the most important measure of control mode performance. Since performance at maximum load is of primary interest, the MTT values corresponding to that point in the test run (T_o + 60 minutes) were computed using the regression analysis results. Table 17 contains the MTT of the three control modes for both data link and

TABLE 16. MODE RANKING BY SUMMARY QUESTION 1 COMPONENTS

RANK	MODES			
	WORKLOAD	CAPACITY	BLUNDERS	STRESS
1	CBE/FDB	CBE/FDB	CBA/TAB	CBE/FDB
2	CBA/TAB	CBA/TAB	CBE/FDB	CBA/TAB
3	CBA/FDB	CBA/FDB	CBA/FDB	CBA/FDB

DASHED LINES INDICATE EQUAL RANKING

TABLE 17. MTT AT MAXIMUM LOAD FOR DATA LINK AND VOICE LINK

20%

MODE	DATA LINK	VOICE LINK
CBA/FDB	8.95 + (30 X .01283) = 9.34	8.99 + (30 X .06309) = 10.89
CBA/TAB	6.24 + (30 X .01283) = 6.62	7.85 + (30 X .06309) = 9.74
CBE/FDB	9.644 + (30 X .01283) = 10.03	10.93 + (30 X .06309) = 12.83

50%

CBA/FDB	10.08 + (30 X .01283) = 10.46	8.63 + (30 X .06309) = 10.52
CBA/TAB	5.96 + (30 X .01283) = 6.35	7.21 + (30 X .06309) = 9.10
CBE/FDB	9.67 + (30 X .01283) = 10.05	10.83 + (30 X .06309) = 12.73

80%

CBA/FDB	7.68 (30 X .01283) = 8.06	8.51 + (30 X .06309) = 10.40
CBA/TAB	7.90 (30 X .01283) = 8.29	8.11 + (30 X .06309) = 10.0
CBE/FDB	9.73 (30 X .01283) = 10.11	10.03 + (30 X .06309)= 11.92

voice data link messages at the three traffic mixes. In Table 18 the modes have been ranked in order of increasing MTT.

The ranking of the modes by MTT values at maximum load does not conform with the ranking based on controller opinion. For example, Control-by-Exception/Full Data Block almost always ranks third in MTT, and consistently ranks first according to questionnaire results. Control-by-Approval/Tabular List ranks first in MTT, with the exception of the 80% data link case, but consistently ranks second in controller opinion. Control-by-approval full data block fares better in MTT ranking than in controller opinion, where it always ranked last.

The MTT rankings of the modes may not be as much at variance with the questionnaire data as would appear upon examination of the command initiation delay (CID) values obtained by regression analysis for the three modes. The Command Initiation Delay is the time between initial command display and the dispatch action and, as such, is a major component of the total message transaction time. Table 19 contains the CID of the three control

TABLE 18. MODE RANKING BY SHORTEST MTT

MODE	MODE RANKING DATA LINK MTT			MODE RANKING VOICE LINK MTT		
	20%	50%	80%	20%	50%	80%
CBA/FDB	2	3	1	2	2	2
CBA/TAB	1	1	2	1	1	1
CBE/FDB	3	2	3	3	3	3

TABLE 19. CID AT MAXIMUM LOAD FOR DATA LINK AND VOICE LINK

20%

MODE	DATA LINK	VOICE LINK
CBA/FDB	4.53 + (30 X .00831) = 4.78	3.92 + (30 X .00857) = 4.18
CBA/TAB	2.03 + (30 X .00831) = 2.53	2.54 + (30 X .00857) = 2.8
CBE/FDB	6.04 + (30 X .00831) = 6.29	5.79 + (30 X .00857) = 6.05

50%

CBA/FDB	5.83 + (30 X .00831) = 6.08	4.33 + (30 X .00857) = 4.59
CBA/TAB	2.01 + (30 X .00831) = 2.26	2.34 + (30 X .00857) = 2.6
CBE/FDB	5.81 + (30 X .00831) = 6.06	5.78 + (30 X .00857) = 6.03

80%

CBA/FDB	3.94 + (30 X .00831) = 4.19	4.29 + (30 X .00857) = 4.55
CBA/TAB	3.69 + (30 X .00831) = 3.94	4.07 + (30 X .00857) = 4.33
CBE/FDB	5.79 + (30 X .00831) = 6.04	5.78 + (30 X .00857) = 6.04

modes for both data link and voice link messages at the three traffic mixes at full load. The mode ranking by shortest CID in Table 20 is identical with the mode ranking by shortest MTT in Table 18, indicating quite clearly that CID is the governing component. It is obvious from Table 19 that CBE/FDB is at a distinct disadvantage in CID with respect to the other control modes because of the nominal 6 second value used for the command display time parameter; i.e., the command could not be automatically dispatched until the command had been displayed for approximately 6 seconds. The intent in using this parameter value in CBE was to give the controller time to not only evaluate the command but also to disapprove it prior to automatic dispatch. However, the test data show that the effect of the 6-second CID was to bias CBE/FDB with respect to CBA/FDB and CBA/TAB, since in the latter two modes, the controller was not constrained in the initiation or the timing of Dispatch actions. Therefore, the third place ranking of CBE/FDB cannot be accepted as valid or

TABLE 20. MODE RANKING BY SHORTEST CID

MODE	MODE RANKING DATA LINK CID			MODE RANKING VOICE LINK CID		
	20%	50%	80%	20%	50%	80%
CBA/FDB	2	3	2	2	2	2
CBA/TAB	1	1	1	1	1	1
CBE/FDB	3	2	3	3	3	3

conclusive; its ranking is totally dependent on the display-time parameter, which many subjects felt could be reduced (Post-Experiment Question C-14).

The ranking of CBA/TAB ahead of CBA/FDB is consistent with the rankings derived from the results of the subjective data analysis. Quite obviously, the higher MTT values for CBA/FDB are attributable to the greater CID values arising from the trackball slewing required by this mode. In the CBA/TAB mode, trackball slewing was not required. Although the controller had to correlate the topmost blinking command in the list with its associated blinking aircraft symbol, commands were dispatched by simply depressing a single key.

There are variations in the MTT of the candidate modes with the three different percentages of data link aircraft employed in the tests. The data link and voice link MTT values for maximum load presented in Table 17 are plotted in Figures 16 and 17 to illustrate the changes that occur with different traffic mixes. CID values associated with each of the three modes are also shown in Figures 16 and 17.

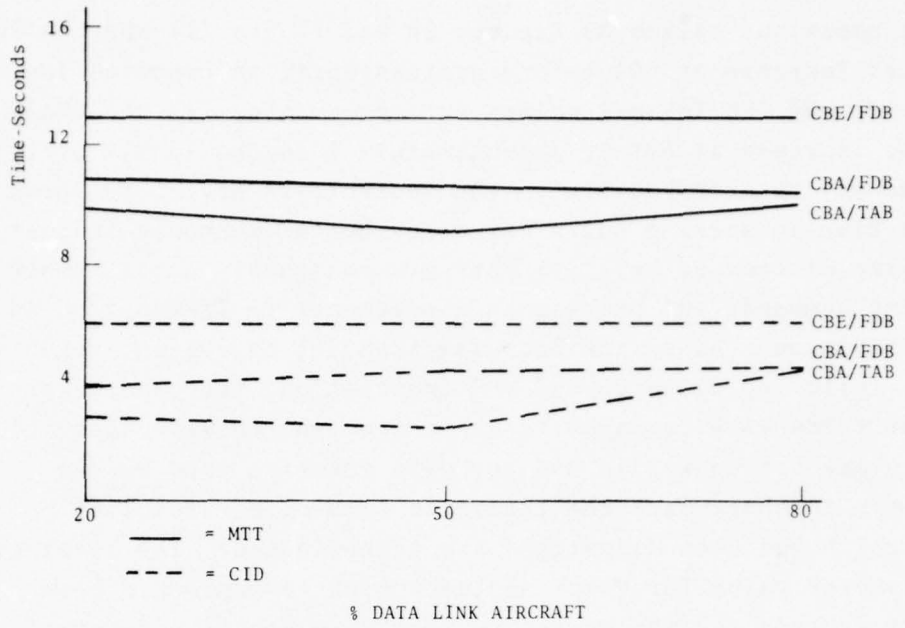


Figure 16. Voice Link MTT and CID at Full Load – 60 Minutes

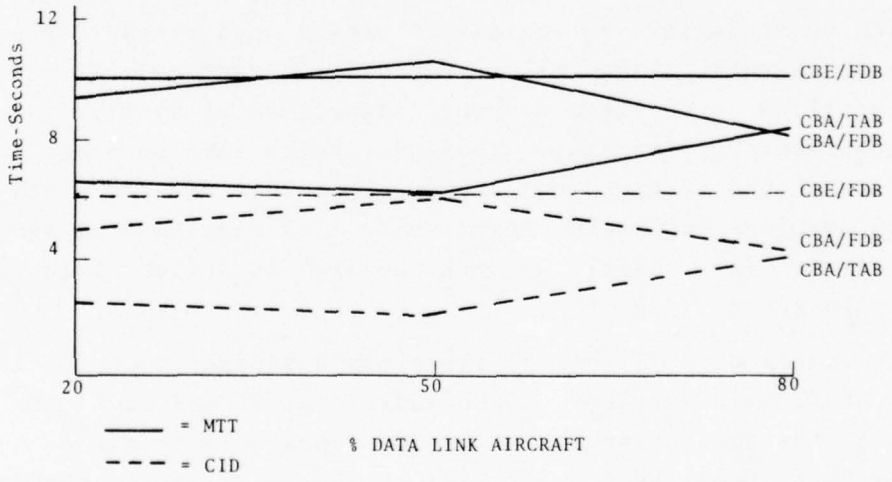


Figure 17. Data Link MTT and CID at Full Load – 60 Minutes

The questions raised by Figures 16 and 17 are (1) why CBA/FDB MTT values increase at 50% before decreasing to an expected low at 80% and (2) why CBA/TAB MTT values act conversely. In the CBA/FDB case, the increase at 50% of approximately 1 second in MTT over the 20% value may be attributable to the controllers having to spend a bit more time in seeking out a larger number of response indicators while being distracted by still having a reasonably large number of voice link commands and pilot acknowledgements to listen to. In the CBA/TAB case, the slight decrease from 20% to 50% is probably because, following dispatch and VRS enunciation, the controllers had fewer voice link commands to erase from the tabular list. The command erase action was an M&S keyboard function used by the controllers to abbreviate the terminate time of a voice link command which had been dispatched and acknowledged. The terminate time parameter value for voice and data link commands had been set to 20 seconds for the tests, so that in cases of off-normal responses (U, F, or X) data link commands would be retained for a period of time long enough to allow the controller readily to correlate them with the response indicator. Unfortunately, there were no provisions for setting different terminate time values for voice and data link commands. Therefore, while dispatched data link commands were automatically erased upon receipt of a W, dispatched voice link commands would not be automatically erased until 20 seconds had elapsed, regardless of earlier acknowledgement by the pilot. Thus, for voice link commands, the controllers found it expedient and convenient to use the erase function to clear each acknowleged voice link command from the tabular list. This additional workload item is reflected in the slightly longer MTT, 20%.

The increase in MTT of approximately 2 seconds from 50% to 80% for CBA/TAB is ascribed to the additional visual workload imposed on the controller by his having repeatedly to glance back and forth from the tabular list to the ACID line of the data block to observe the incidence of the W response indicator for the much larger number of dispatched data link commands.

It should be noted that MTT at 80% for CBA/TAB and CBA/FDB are very comparable, indicating that the disadvantage of trackball slewing is far outweighed by the advantage of having a more integrated command/response display location. A control mode combining the advantages of CBA/TAB with CBA/FDB would seemingly be an ideal solution: that is, a CBA/FDB mode in which commands are displayed in the third line of only one data block at a time and dispatched by a single key depression before subsequent commands in another data block can be displayed and dispatched.

The variations in the overall MTT (data link and voice link) with percentage of data link aircraft for each mode were obtained by averaging the data link and voice link MTT values in the proportions indicated by the mixes. The performance of the three modes, as depicted in Figure 18, again compares favorably with the rankings based on controller opinion, especially if one recognizes, as previously stated, that CBE/FDB was biased out of a possibly higher ranking in MTT.

As one would expect, there is a correlation between message transaction time and average queue size. Indeed, the longer it takes the controller to process a command, the greater the probability that there will be an increase in the queue. Table 21 clearly indicates that mode rankings on the basis of increasing average queue size are consistent with rankings on the basis of increasing MTT. Only a single set of queue size values is shown in Table 21, since queue size does not change significantly as a function of traffic mix.

TABLE 21. MODE RANKING BY AVERAGE QUEUE SIZE

RANK	MODE	AVERAGE QUEUE SIZE
1	CBA/TAB	0.238
2	CBA/FDB	0.394
3	CBE/FDB	0.457

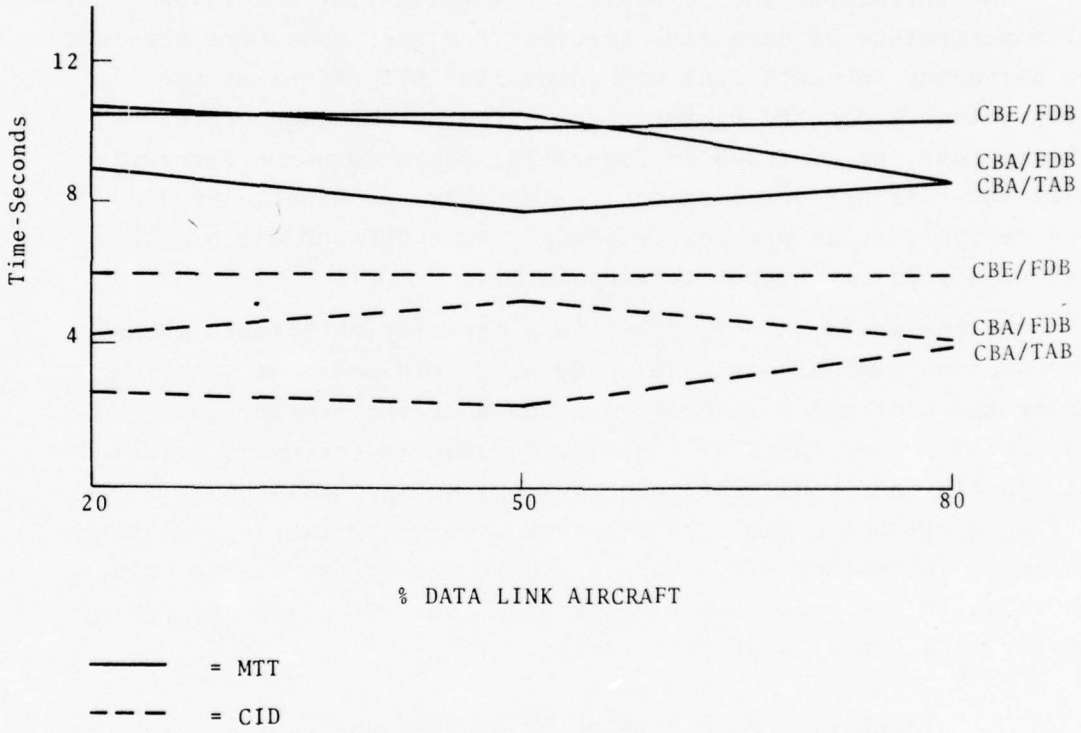


Figure 18. Combined Data Link and Voice Link MTT and CID at Full Load – 60 Minutes

There is also a high degree of correlation between gate arrival error and dispatch of the final heading command. Modal values for arrival time errors at the gate for all three traffic mixes are shown in Table 22. Since they are calculated by subtracting the scheduled arrival time from the actual arrival time, positive values indicate late arrival. The magnitude of error is a function of the timeliness of dispatch of the final heading command to intercept the localizer. Indeed, in the CBA modes, the more rapidly the controller dispatches the final heading command, the more quickly the aircraft executes the command, and the smaller the late arrival error. The difference of almost 1 second between CBA/TAB and CBA/FDB could be attributable to the trackball slewing required in the latter mode. The much longer late arrival error in CBE/TAB is probably due to the fixed command initiation delay of 6 seconds. The absolute magnitude of these errors are of minor concern because they can be virtually eliminated by dynamically adjusting the display and issue time parameters of the final heading command.

TABLE 22. GATE ARRIVAL TIME ERRORS

MODE	AVERAGE GATE ARRIVAL TIME ERROR
CBA/TAB	0.55
CBA/FDB	1.46
CBE/FDB	6.27

Of greater significance than gate arrival time errors is the dispersion of arrival time errors about their mean, since this is a measure of the degree of randomness in command processing time. Dispersion in the delivery and execution of commands contributes directly to dispersion in arrival time at the runway threshold,

a key factor in determining runway capacity. In fact, the errors in timing of the final heading command are effectively doubled at the runway threshold because of the trombone approach geometry. The standard deviation of arrival time error at the gate was 6.2 seconds, independent of mode and traffic mix.

The conflict data collected during the test runs were separated into three categories: (1) detected conflict situations which did not materialize, (2) actual conflicts not detected, and (3) conflicts detected more than 15 seconds after they started. Of greatest concern were the undetected conflicts (category 2), shown by mode and percentage in Table 23. The figures in parentheses represent to what extent the conflict count could be reduced if minor conflicts lasting for only one scan were subtracted from the associated total figures. While the results of the statistical analysis are not statistically significant, the reduced conflicts counts totaled by mode tend to follow the familiar modal ranking pattern.

TABLE 23. UNDETECTED CONFLICTS

MODE	20%	50%	80%	TOTAL
CBA/FDB	2(2)	4(1)	6(3)	12(6)
CBA/TAB	5(1)	2	2	9(5)
CBE/FDB	2(1)	2(1)	0	4(2)

In comparing the conflict data obtained with the data link modes against that obtained with the voice-only modes (see Table 12), the only meaningful comparison is one based on the average number of serious conflicts (more than one scan duration) which occurred per run in the 80% and voice-only modes. The figures shown in Table 24 are the average number of conflicts per run which occurred over six voice-only runs and six 80% data link

TABLE 24. AVERAGE NUMBER OF CONFLICTS PER RUN – VOICE-ONLY VERSUS DATA LINK

MODE	VOICE-ONLY	80% DATA LINK
CBA/FDB	.17	.5
CBA/TAB	.42	.33
CBE/FDB	---	0

runs in each mode. While neither statistically significant nor conclusive, these figures suggest that there is no disadvantage to data link operations in detecting conflicts.

11. SUMMARY OF RESULTS

11.1 MODE RANKINGS

Analysis of the subjective data derived from controller questionnaires shows very definitly that the data link control modes ranked as follows in order of preference:

1. Control-by-Exception/Full Data Block (CBE/FDB)
2. Control-by-Approval/Tabular List (CBA/TAB)
3. Control-by-Approval/Full Data Block (CBA/FDB)

CBE/FDB had the longest message transaction time (MTT) of the three modes, but was ranked first by the controllers, despite some concern for blunder potential, because it possessed the best workload, capacity, and stress characteristics. Longer MTT's are inherent in CBE/FDB because of the fixed command disapproval time which in effect "biases" this mode in Command Initiation Delay (CID), the key component of MTT. CBA/TAB had the shortest MTT but was ranked second by the controllers because they felt that operating from a command list tended to divide their attention, adversely affecting workload and capacity. The low MTT's with CBA/TAB are attributable to the single key dispatch without trackball slew. CBA/FDB had the second best MTT but was consistently rated as unfavorable in workload, capacity, stress, and blunder potential because of the necessity for continuous trackball slewing actions.

It is not possible to rank conclusively all of the modes on the basis of the objective data measures (MTT, CID, etc.) because of the bias factor of command disapproval time in CBE/FDB. CBE will always have a longer CID and MTT than the other two modes. It is noteworthy, however, that CBA/TAB and CBA/FDB can be ranked with respect to each other and that this ranking is consistent with subjective data rankings – i.e., TAB ranked above FDB.

11.2 CONTROL AND DISPLAY PREFERENCES

In looking more closely at the mode rankings on the basis of controller preference, some conclusions can be drawn as to the

display and keyboard input aspects of each mode. CBE/FDB ranks first because the controllers prefer display of the command in the data block with which they constantly deal and are comfortable in using in basic ARTS. The third-place ranking of CBA/FDB does not contradict this conclusion because CBA/FDB entails use of the trackball which is universally disliked by controllers. The single button nontrackball dispatch scheme of CBA/TAB is obviously the principal reason why this mode ranked second despite controller dislike for a command list. Another reason for CBA/TAB's higher ranking may be a veiled preference for single command over multiple commands. The inference that can be drawn here is that a potentially desirable CBA mode would be one in which the commands were displayed one at a time in only one data block at a time with a single button nontrackball dispatch action. These findings are in substantial agreement with the results of the Phase 2A experiment.

11.3 OPERATIONAL BENEFITS OF DATA LINK

Before the experiment began, most of the controllers felt that M&S via data link would be accepted in actual operation and would favorably affect the current operational system in an all data link environment. At the conclusion of the experiment the attitude was negative due primarily to the undesirable features of CBA/TAB and CBA/FDB. To answer the same question on the basis of objective data necessitated extrapolation of 80% data link data to 100% and comparison of the latter with data obtained from non-data link voice-only runs. The nonlinearity of the data across the 20, 50, and 80 percentages indicated that a projection to 100% would be unreliable. The very limited data obtained from the test runs involving the voice-only TAB and FDB mode of operation show an MTT of 6.66 seconds for both modes. In the absence of data on a full data link environment, the only comparison, albeit a gross one, that can be made is to compare these MTT's with those obtained at 80% for the same two modes. In data link TAB, MTT is 8.29; in FDB it is 8.06. While these figures are significantly larger, seemingly suggesting that in terms of communications delays voice link is superior to data link, they may be attributable to greater familiarity with voice-only operations.

11.4 OTHER INFORMATION

Other results of this simulation study indicate the following:

1. The amount of additional training the controllers felt would be necessary to make them feel comfortable with the operational procedures associated with a control mode paralleled the mode rankings – i.e., 1 hour for CBE/FDB, more than 2 for CBA/TAB, and more than 4 for CBA/FDB.

2. Display of the W character and simultaneous erasure of a data link command to indicate pilot compliance is not necessary; simple erasure of the command would be sufficient. A majority of the controllers felt that the 3-second display time of the W character should be reduced.

3. In the control-by-approval modes the commands should not stop blinking (timeout) at issue time but should remain blinking and be eligible for dispatch until the controller can service them. This would require that the algorithm be capable of dynamically updating command values to compensate for delays in dispatching what would otherwise be obsolete commands.

4. Voice commands should be held over in a nonblinking state for a period of less than 20 seconds. This could be achieved by retaining the nonblinking command for a short time after the VRS had enunciated the command. The post-VRS retention time could be based on average pilot voiced acknowledge time.

5. Controllers felt that the workload at peak load was slightly below their maximum capacity for CBE/FDB, slightly above their capacity for CBA/TAB, and far above their capacity for CBA/FDB. They were divided as to their ability to maintain a picture of the traffic situation for conflicts while issuing commands. If M&S and data link failed at peak load, half the subjects felt they could handle the same level of traffic indefinitely. Only 10% thought the situation would be hazardous.

6. Most of the controllers had no difficulty in interpreting commands in the FDB modes despite the fact that there was increased clutter and more manual tag repositioning required. However, 40% of the controllers had difficulty recognizing exception condition indications (U,F,X) in CBA/FDB and CBE/FDB. The majority did not feel that exception cases were time-consuming.

7. A majority of the controllers felt that priority should be given to commands around the base leg in the CBA/TAB mode.

8. Most of the controllers felt that VRS enunciation of commands was less tiring and that the 3-second interval between VRS transmissions was acceptable, but half of them felt that VRS halting and starting procedures needed improvement in CBE/FDB. Controllers were particularly bothered by the VRS's inability to repeat commands and to a lesser extent by the slower VRS talking rate.

9. Half the controllers were bothered by not having pretransmission approval of issued commands in CBE/FDB. On the other hand, more than half felt 6 seconds was sufficient for disapproving a command; over one third thought it too long.

12. CONCLUSIONS

The objective of this study was to ascertain which of three candidate control modes was the most acceptable for the display and dispatch of metering and spacing commands in a mixed voice/digital communications environment. The conclusions that can be reached from the results of this study are as follows:

1. The CBE/FDB mode is the most favored by the controllers despite blunder potential, but it has some inherent limitations that result in longer message transaction times.
2. As was concluded in the Phase 2A experiment, a CBA/FDB mode in which a single command at a time is displayed in one data block for dispatch with only a single keyboard entry is again indicated as a potentially viable mode in a future system. This mode would combine the advantages of both the CBA/FDB and CBA/TAB modes as currently implemented.
3. Contrary to the findings of the Phase 2A experiment, computer-generated voice was not favorably regarded because of the talking rate and inability to repeat commands.

13. RECOMMENDATIONS

Further studies aimed at determining the optimum means for interfacing the controller and the computer in a data link communications environment should be conducted in the following areas:

1. Additional testing of the Control-by-Exception/Full Data Block mode of operation should be conducted with an on-line controller-selectable disapproval time. This will allow a controller to decrease MTT if he feels disapproval time need not be as long (6 seconds) as it was in this study.

2. On the basis of the findings of both Controller/Computer Interface experiments (Phase 2A and 2B), it is recommended that a CBA/FDB mode in which a single command is displayed at one time in an aircraft data block for dispatch with only a single keyboard entry be incorporated in a future system.

3. Considering the advantage of computer-generated voice in providing commonality of command dispatch in a mixed voice/digital communications environment, it is recommended that future research be devoted to improving the talking rate and in developing a command repeat capability.

REFERENCES

1. Data Link ARTS III Program Modifications, Design Data, PX10490, Contract No. DOT-TSC-607, Sperry Univac, Defense Systems Division, St. Paul MN, December 1973.

2. Development Specification, Voice Response System, PX5738, Sperry Univac, Defense Systems Division, St. Paul MN, January 1973.

3. Computer-Aided Metering and Spacing with ARTS III: Phase I Design Study, FAA-RD-70-82, Contract No. DOT-FA70NA-2433, Computer Systems Enginerring, Inc., Billerica MA, December 1970.

4. E.H. Hilborn, Human Factors Experiments for Data Link: Extended Summary of Interim Reports 1 through 4, FAA-RD-74-82 (TSC-FAA-74-6), U.S. Department of Transportation, Transportation Systems Center, Cambridge MA, April 1974.

5. R. Wiseman et al., The Controller/Computer Interface with an Air-Ground Data Link, FAA-NA-74-2, U.S. Department of Transportation, Transportation Systems Center, Cambridge MA and National Aviation Facility Experimental Center, Atlantic City NJ, 1975.

6. W.J. Dixon and F.J. Massey, Jr., Introduction to Statistical Analysis, 2nd ed. McGraw-Hill, New York NY, 1957, pp. 179-180.

7. G.E. Box and G.M. Jenkins, Time Series Analysis Forecasting and Control, Holden-Day, San Francisco CA, 1970, pp. 34-35.

8. W.J. Dixon, ed. BMD: Biomedical Computer Programs, University of California at Los Angeles, Health Sciences Computing Facility, The University Press, Los Angeles CA, 1964, pp.739-751.

9. E.H. Stevens, Correlation Cluster Analysis Documentation, TM-02-016, Contract No. DOT-FA73NA-0102, Computer Sciences Corp., Pomona NJ, December 1973.

APPENDIX
TIMELINES FOR EACH CONTROL MODE

The sequence of events occuring in each of the control modes evaluated in the test series are graphically described in the time lines, Figures A-1 through A-11. Each of the three data link control modes has separate time lines for data link and voice link aircraft. The sequence of actions by the computer, the controller, the voice response system, and the pilot is indicated by coded markers on the timeline. Certain intervals of special importance in evaluating the formats are defined on each timeline. These timelines are referred to in Section 5, "Test Conduct." The symbols used in the timelines are as follows:

◊ = computer action

▲ = controller action

Δ = simulator pilot action

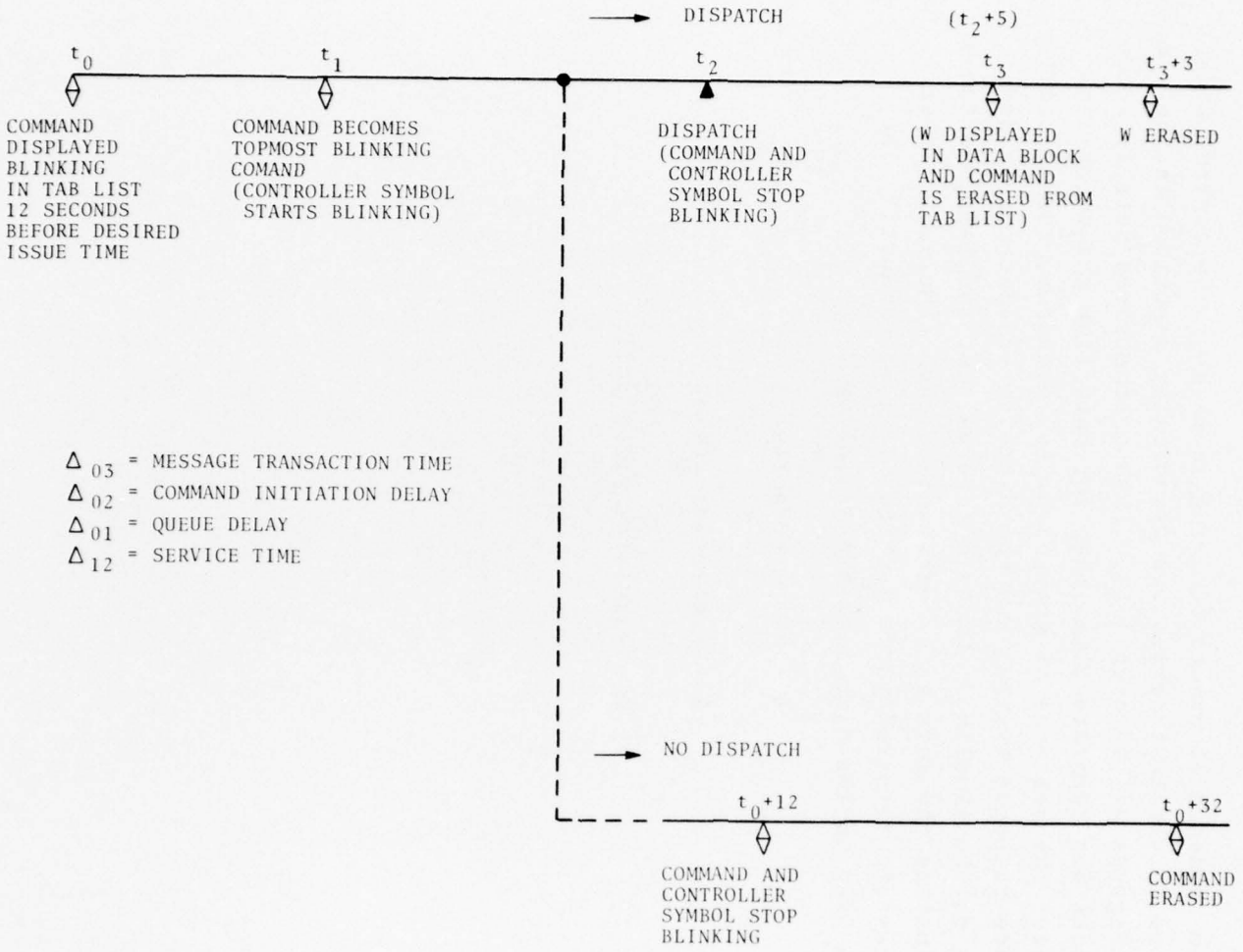


Figure A-1. CBA/Tabular List/VRS-Data Link Aircraft

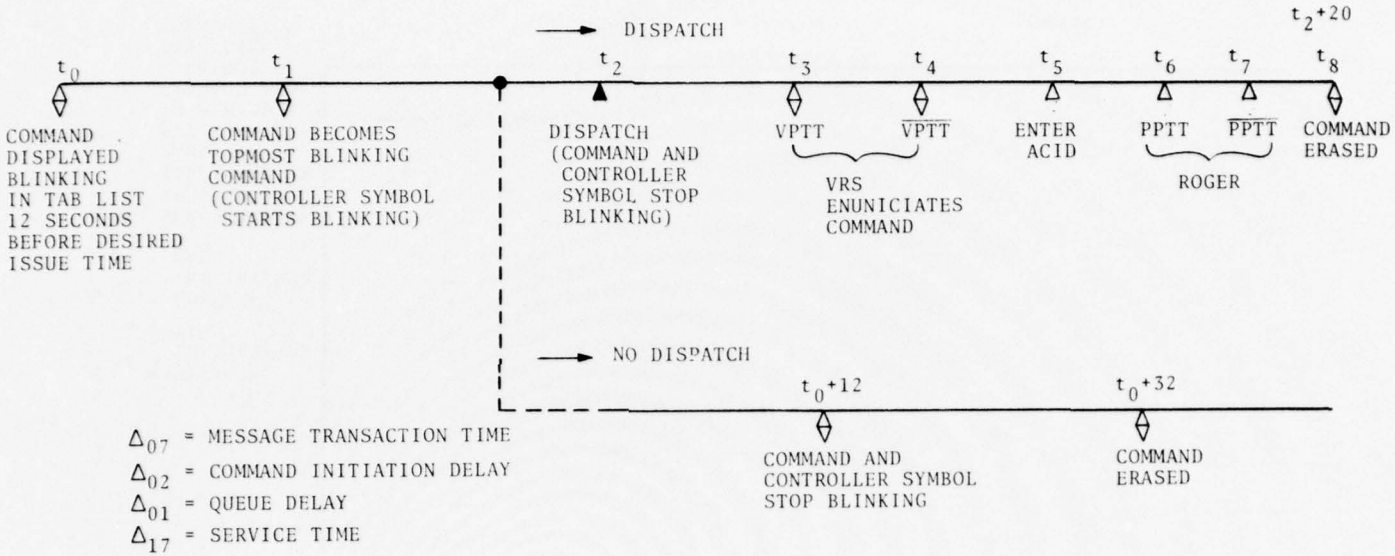


Figure A-2. CBA/Tabular List/VRS-Voice Link Aircraft

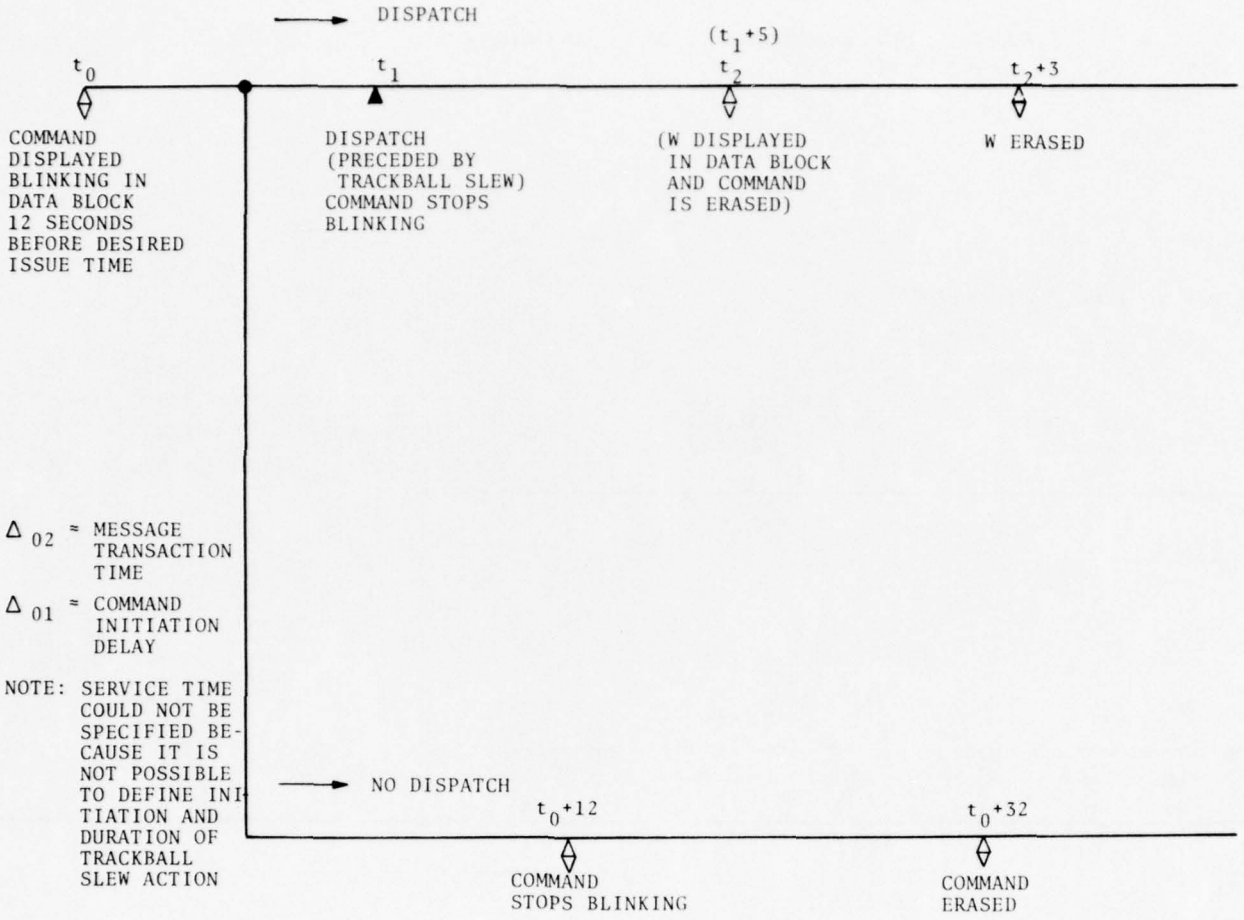
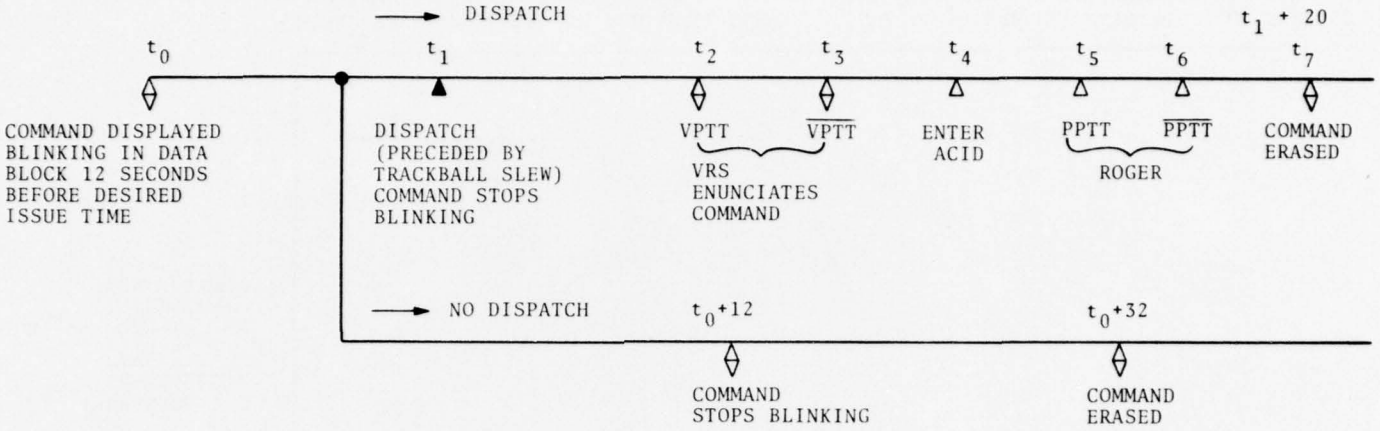


Figure A-3. CBA/Full Data Block/VRS-Data Link Aircraft



Δ_{06} = MESSAGE TRANSACTION TIME
Δ_{01} = COMMAND INITIATION DELAY

NOTE: SERVICE TIME COULD NOT BE SPECIFIED BECAUSE IT IS NOT POSSIBLE TO DEFINE INITIATION AND DURATION OF TRACKBALL SLEW ACTION.

Figure A-4. CBA/Full Data Block/VRS-Voice Link Aircraft

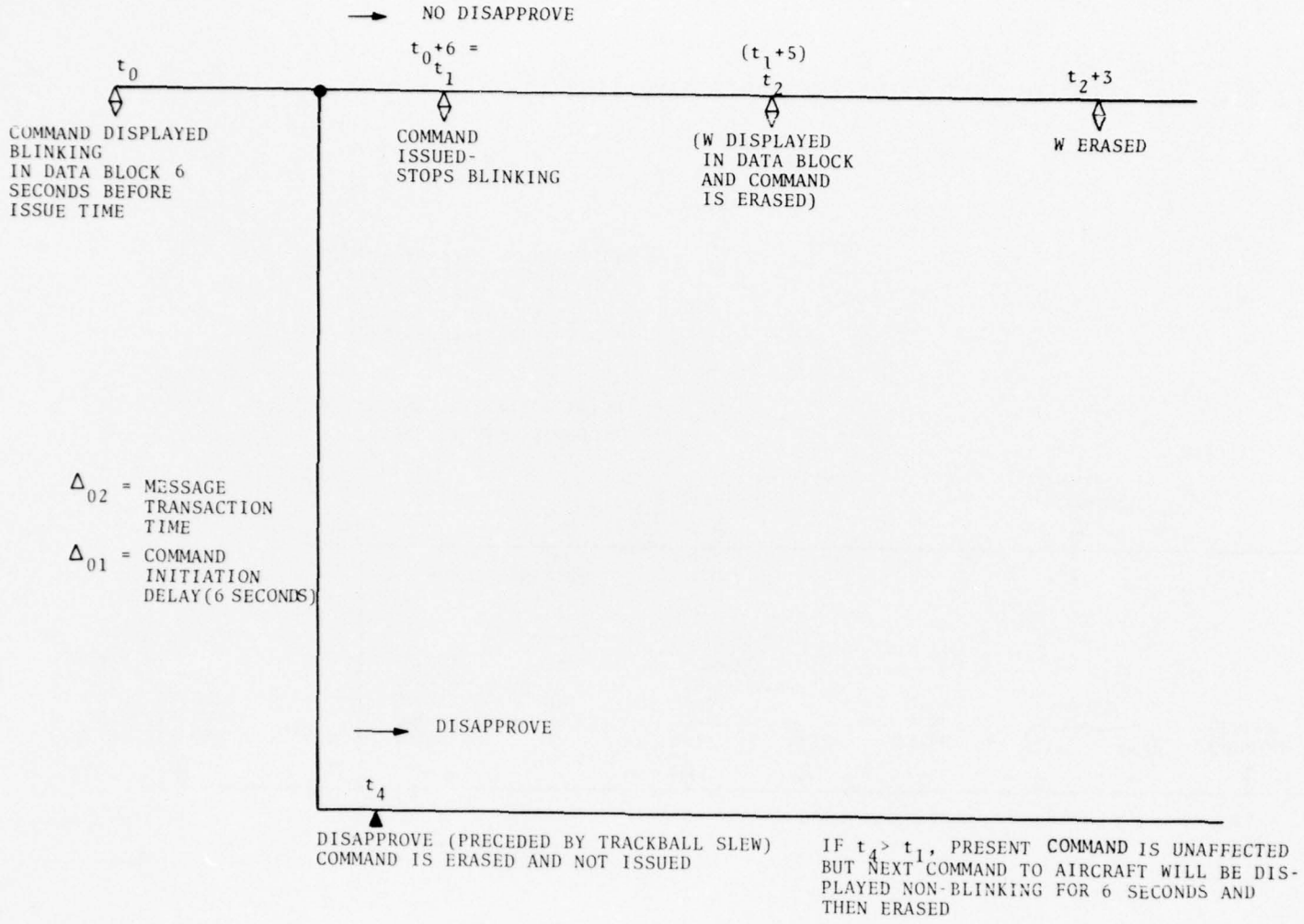
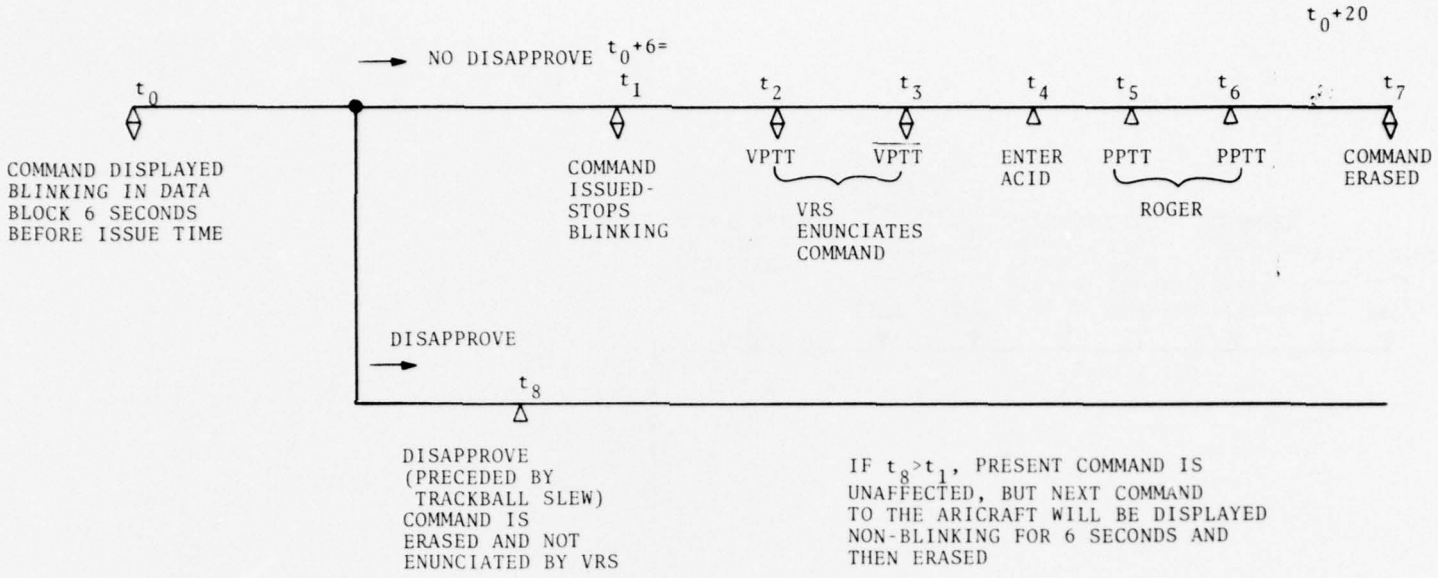


Figure A-5. CBE/Full Data Block/VRS-Data Link Aircraft



Δ_{06} = MESSAGE TRANSACTION TIME

Δ_{01} = COMMAND INITIATION DELAY (6 SECONDS)

Δ_{16} = SERVICE TIME

Figure A-6. CBE/Full Data Block/VRS-Voice Link Aircraft

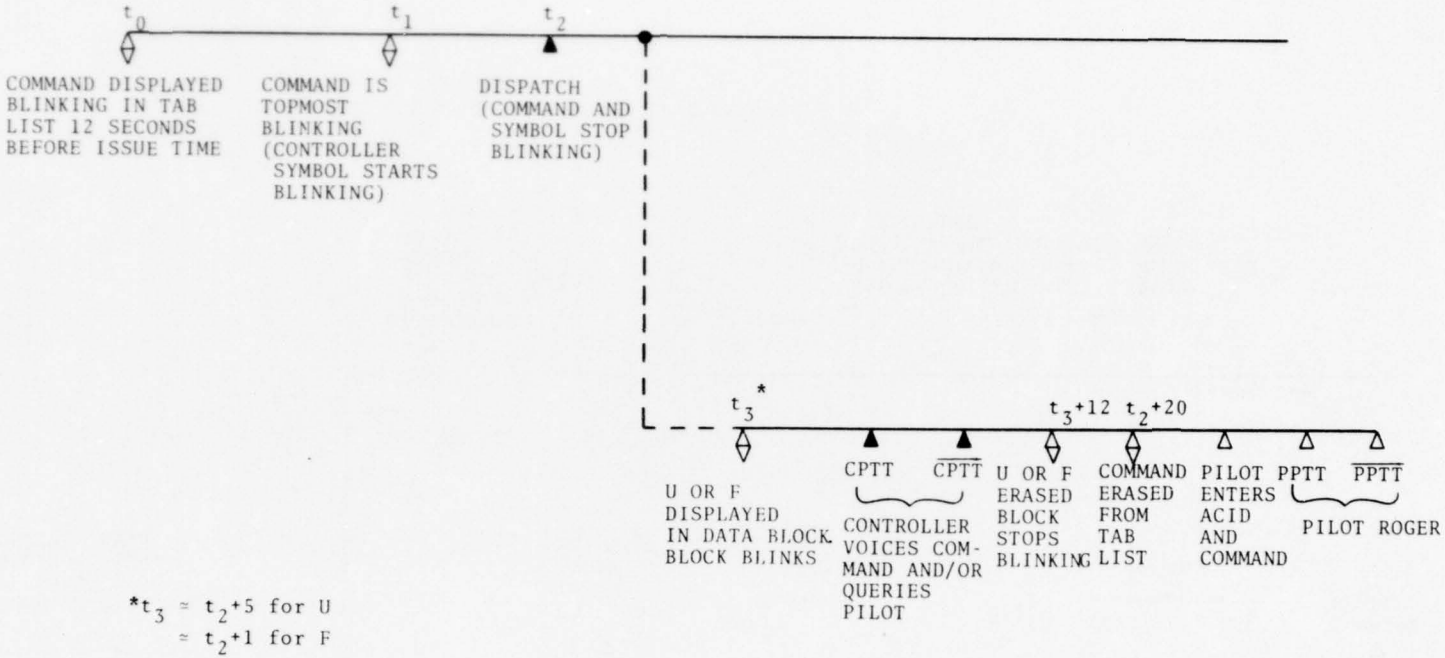


Figure A-7. CBA/TAB/VRS-Data Link Aircraft – Unable or Fail Exception Case

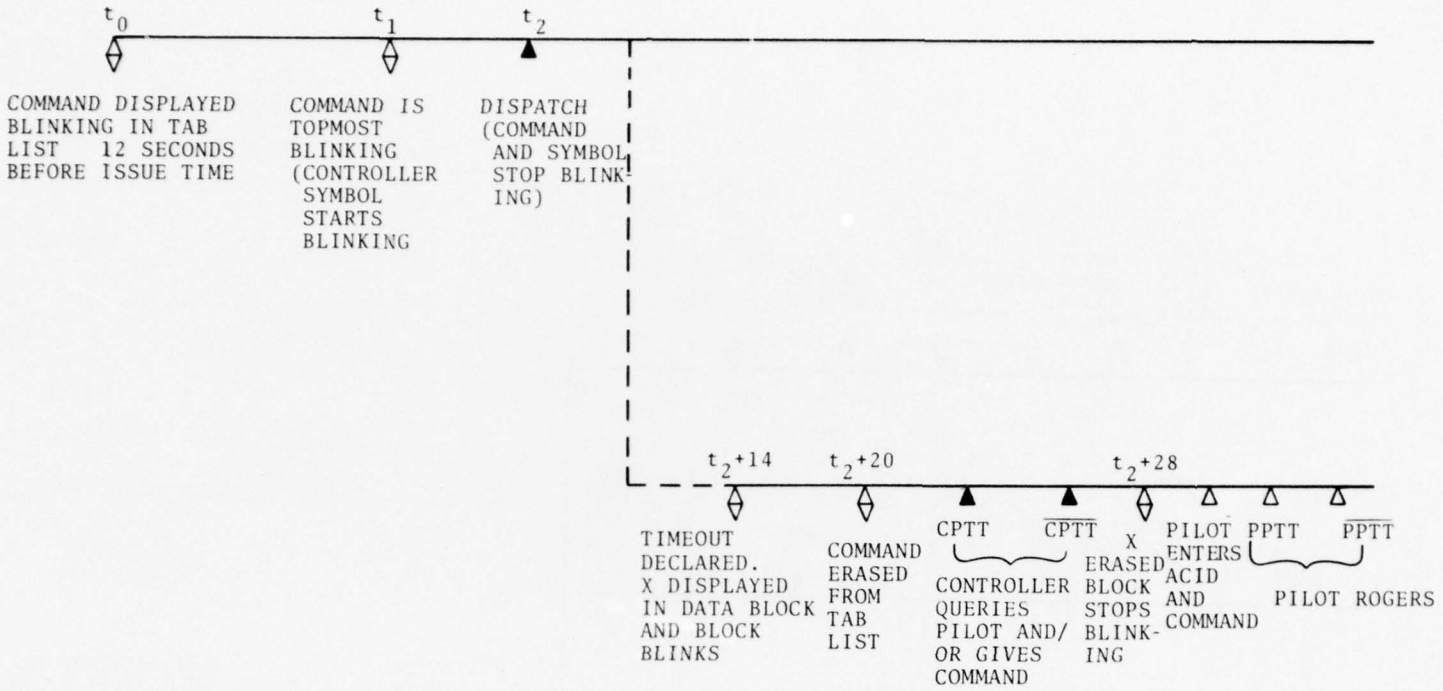
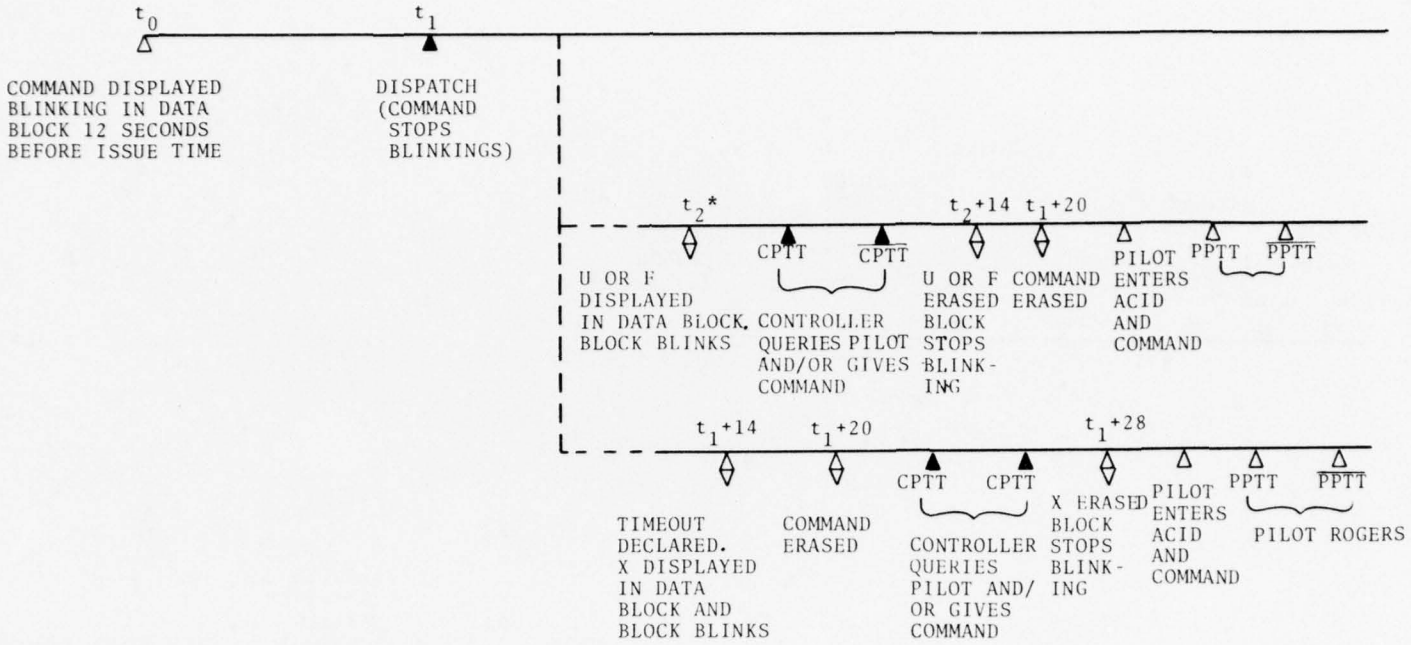
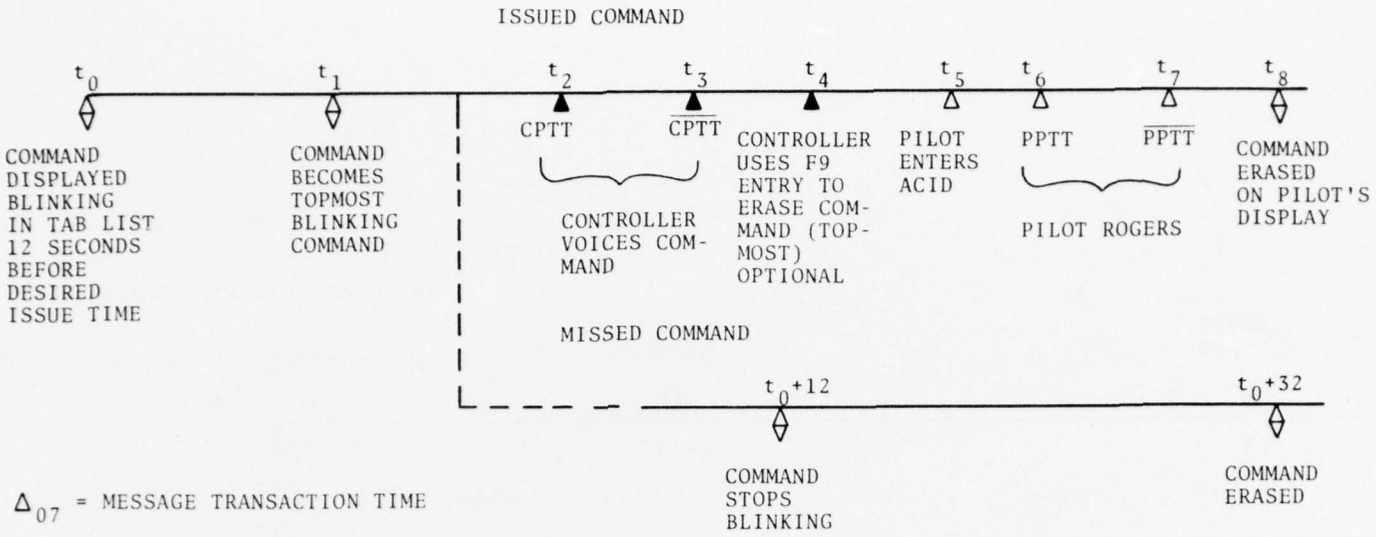


Figure A-8. CBA/TAB/VRS-Data Link Aircraft -- Timeout Exception Case



*$t_2 \simeq t_1+5$ for U
$\simeq t_1+1$ for F

Figure A-9. CBA/Full Data Block/VRS-Data Link Aircraft – Unable, Fail, and Timeout Exception Cases



Δ_{07} = MESSAGE TRANSACTION TIME

Δ_{02} = COMMAND INITIATION DELAY

Δ_{02} = QUEUE DELAY

Δ_{17} = SERVICE TIME

NOTE: ERASURE OF COMMAND ON CONTROLLER'S DISPLAY DOES NOT ERASE COMMAND ON PILOT'S DISPLAY.

Figure A-10. CBA/Tabular List/Voice-Voice Link Aircraft

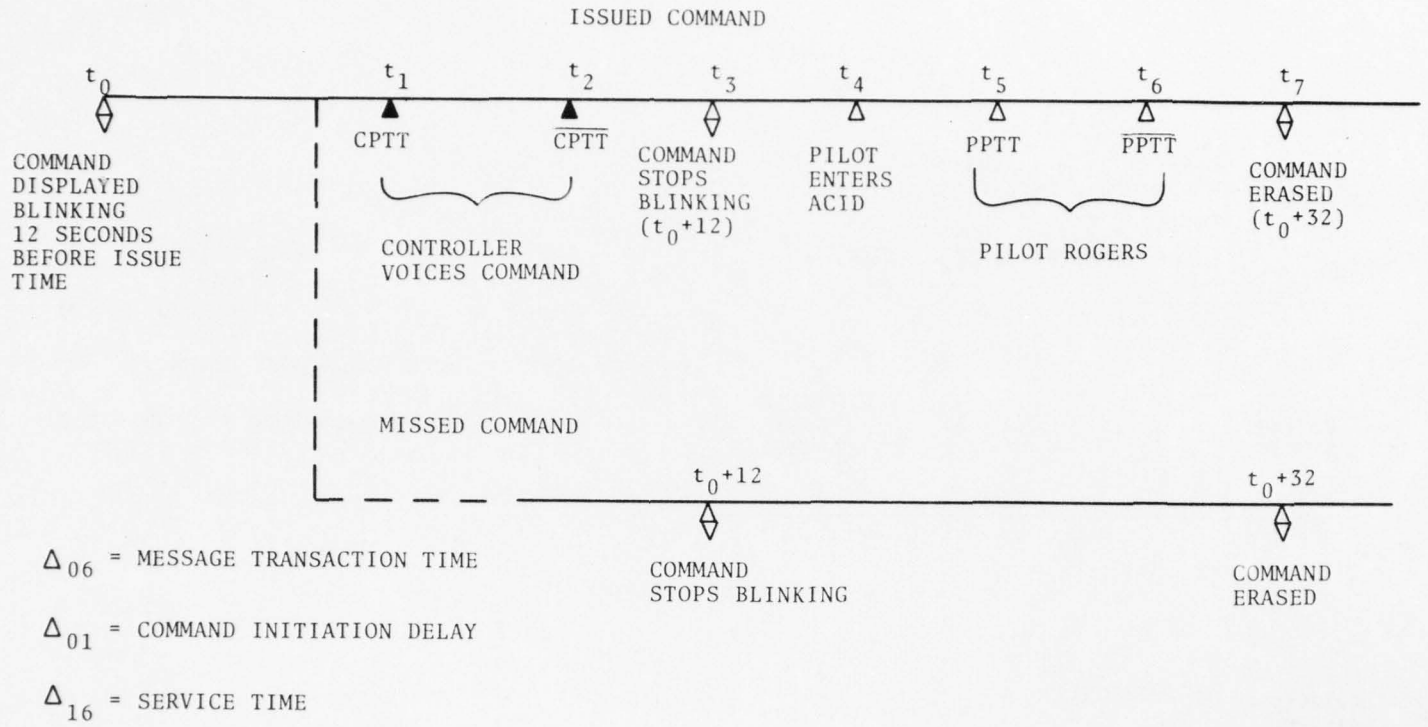


Figure A-11. CBA/Full Data Block/Voice-Voice Link Aircraft