DOT/FAA/CT-92/2, 1

FAA Technical Center
Atlantic City International Airport
N.J. 08405

Controller Evaluation of Initial Data Link Terminal Air Traffic Control Services: Mini Study 2

Volume 1

Nicholas J. Talotta, et al.

April 1992

Final Report

This document is available to the U.S. public through the National Technical Information Service, Springfield, Virginia 22161

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/CT-92/2, I		

4. Title and Subtitle	5. Report Date
Controller Evaluation of Initial Data Link Terminal Air Traffic Control Services: Mini Study 2, Volume I	April 1992
	6. Performing Organization Code: ACD-320
7. Author(s): Nicholas J. Talotta, et al.	8. Performing Organization Report No.: DOT/FAA/CT-92/2, I
9. Performing Organization Name and Address: U.S. Department of Transportation, Federal Aviation Administration, Technical Center, Atlantic City International Airport, NJ 08405	10. Work Unit No. (TRAIS)
	11. Contract or Grant No.: T2001B
12. Sponsoring Agency Name and Address: U.S. Department of Transportation, Federal Aviation Administration, Research and Development Service, Washington, D. C. 20591	13. Type of Report and Period Covered: Final Report
	14. Sponsoring Agency Code: ARD-60

15. Supplementary Notes

16. Abstract

This document details the results of the second Mini Study of the Federal Aviation Administration (FAA) Technical Center investigation and development of initial terminal air traffic control (ATC) services for transmission using Data Link technology. Initial Data Link services were evaluated under part task simulation conditions in order to identify service delivery methods which optimize controller acceptance, performance, workload, and to study the effects of various potential Data Link message response delays.

17. Key Words	18. Distribution Statement
ARTS IIIA Terminal ATC Data Link	This document is available to the U.S. public through the National Technical Information Service, Springfield, Va. 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	99	

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

PREFACE

This report documents the second Federal Aviation Administration (FAA) controller evaluation of an initial group of four terminal air traffic control (ATC) services and functions which are under development for implementation on a Data Link air-ground communications system. The research was conducted at the Federal Aviation Administration (FAA) Technical Center by the Airborne Collision Avoidance and Data Link Systems Branch, ACD-320.

The report is organized in two volumes. Volume 1 contains the main body of the report. It includes a detailed description of the objectives of the study and of the technical approach and testing methods that were used. Volume 1 also presents the results of the research, conclusions, and recommendations for future work.

Volume 2 contains supporting appendixes. These include controller training materials for the test airspace and Data Link procedures, test scenario descriptions, and evaluation forms used for the design review and full scale simulation phases of the study.

ACKNOWLEDGEMENTS

The study reported in this document was conducted at the Federal Aviation Administration (FAA) Technical Center by the Ground Data Link Development Team. The execution of the study was the result of many months of planning and the cooperative efforts of many people.

Mr. Ron Jones was the Data Link Program Manager with overall responsibility for the program.

Mr. Phillip Radzikowski was the Headquarter's Research and Development Associate Program Manager responsible for Technical Center Data Link coordination.

Mr. Nicholas Talotta was the Data Link Technical Program Manager responsible for the overall management and development of the Data Link Services.

Mr. Evan Darby was the test coordinator and provided air traffic control (ATC) system expertise.

Dr. Clark Shingledecker, NTI, provided all human factors research and analysis, and prepared the most significant portion of this report.

Mr. Joseph Lunder and Mr. Michael Headley, both of TMA Inc., designed the ARTS IIIA software.

Dr. Richard Mogford, Mr. David Stahl, and Mr. Terence Fischer of CTA, Inc., provided performance measures analysis and test scenarios.

Mr. George Chandler and Mr. Cuong Le provided test bed engineering and software in the development of these initial services.

Mr. John Van Dongen was technical coordinator.

Mr. Frank Buck, Mr. Preston Cratch, Mr. Dave Sweeney and Mr. Robert Geoghan, all of The MITRE Corporation, provided program support.

Mr. Ephraim Shochet was human factors research coordinator.

Mr. John Evans and Mr. Jim Merel, of UAL, Inc., were ghost controllers.

The Technical Center facility personnel supported this activity during development and testing.

The following members of the Air Traffic Data Link Validation Team (ATDLVT) were active participants in this Mini Study.

Michael Blomquist, Denver ATCT
Michael Bullington, Sacramento TRACON
Joseph D'Alessio, New York TRACON
Harly Jones, Sacramento TRACON
Jerome Karrels, Madison, WI, Tower
David Lister, St. Louis Tower
Jim O'Malley, ATR-320

TABLE OF CONTENTS

TABLE OF CONTENTS (CONTINUED)

TABLE OF CONTENTS (CONTINUED)

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (CONTINUED)

LIST OF TABLES

EXECUTIVE SUMMARY

The controllers who participated in the study expressed a continued belief that implementation of Data Link air traffic control (ATC) services in the terminal environment will have the potential to significantly improve the quality and effectiveness of air-ground communications. Predicted benefits included reductions in frequency congestion and communication errors and improvements in efficiency made possible by shifting lengthy and repetitive messages from the voice channel to the Data Link system.

The design review phase of the study produced a variety of recommended design changes based on the subject controllers' observations in the Data Link Test Bed. Primary modifications included a change in the design of the initial contact service to reduce the number of Data Link transactions and amount of time required to perform this function, and refinements to the designs of the menu text and terminal information service list displays to enhance the speed and accuracy with which items could be selected for uplink.

The full scale simulation phase of the study confirmed the hypothesis that the viability of Data Link services in the terminal ATC environment will be strongly dependent on the duration of the transaction delay between message dispatch and receipt of an aircrew response. While the results must be tempered by the facts that the subject controllers were not trained to operational proficiency levels on the test airspace or on Data Link procedures, and that the Data Link design was not fully mature, they indicate that total transaction times exceeding 20 seconds may significantly increase controller workload, impair performance, and impede the use of Data Link. The severity of the effects of transaction delay were dependent on the type of airspace sector under control and the time criticality of the message. In addition, the findings suggested that the negative effects were partially mitigated by practice over the course of the study as the controllers developed effective strategies for overcoming time delays.

Several of the experimental performance measures evaluated during the study effectively reflected and complemented the controller evaluation data. In addition, requirements for further metric development and training protocol improvement were identified for incorporation into future operational evaluation study plans.

RECOMMENDATIONS.

The results of this second design development study indicate that the initial Data Link services have significant potential to enhance terminal ATC operations. Therefore, it is recommended that continued design, development, and evaluation efforts be pursued to permit implementation of these services at the earliest possible date. The results support the following recommendations regarding the specific nature of these continued activities:

1. The service design modifications identified during this study should be implemented in the Data Link Test Bed for formal evaluation by the Air Traffic Data Link Validation Team (ATDLVT).

2. In preparation for operational evaluation studies, training protocols should be developed to require extensive pretest controller training in order to approximate the level of airspace familiarity and facility in the use of Data Link that would be expected in the operational environment. Additional effort also should be devoted to the development of improved measures and data collection procedures which will accurately and reliably reflect the impact of Data Link on system performance during operational evaluation research.

3. The observed effects of total transaction times on Data Link effectiveness and usability indicate that technical analyses and flight deck Data Link research should be pursued to define the distribution of terminal transaction times that can be expected at the time of system implementation. Parameters derived from this research should be used in future Data Link ATC simulations, and should serve as a baseline for evaluating modifications to Data Link transmission systems and aircrew interfaces intended to reduce message acknowledgement latency.

4. The results showed that controllers appear to spontaneously develop strategies for overcoming transaction delays and improving ATC operations in a combined voice and Data Link environment. Future research should formally document and define the effectiveness of these strategies in order to support the development of optimal Data Link procedures and training materials that will be needed to complement functional specifications for the Data Link service designs.

1. INTRODUCTION.

1.1 PURPOSE.

This document presents the results of a Federal Aviation Administration (FAA) Technical Center investigation of terminal air traffic control (ATC) services developed for transmission using Data Link technology. Based on the results of an initial study conducted in 1990, designs for four ATC services were modified and implemented in the currently operational National Airspace System (NAS) Automated Radar Terminal System (ARTS) IIIA computer and ATC workstation for review and evaluation by air traffic controllers.

The controllers participated in simulated terminal airspace test trials to assess the utility of the Data Link services, recommend requirements for additional service design changes, and provide a preliminary assessment of the combined impact of transmission and flight deck response delays on the utility and effectiveness of Data Link for terminal operations. This study was the second in a planned series of iterative design development tests which will culminate in a full-scale operational evaluation and the production of functional design specifications for an operational terminal ATC Data Link communications system.

1.2 BACKGROUND.

1.2.1 ATC Communications and Data Link.

In response to the phenomenal growth of air traffic in the United States, the FAA has begun to develop and implement a broad range of initiatives aimed at updating and enhancing ATC technology. Many of these efforts are focused on improving the quality and quantity of information that will be needed to increase safety and productivity, and on insuring that this information is reliably and accurately transferred among the computers and humans that form the major components of the ATC system.

One of the primary information transfer problems that constrains the capacity of the current ATC system is the inherently limited communication channel that exists between the air traffic controller and the aircraft pilot. Because this voice radio link operates in a broadcast mode between a single controller and all aircraft operating in the airspace under his control, frequency congestion is a common occurrence when the volume and complexity of air traffic increases. Such saturation of the communications channel affects the performance of the ATC system by preventing the timely issuance of clearances and by restricting the vital exchange of information upon which safe and efficient operation of the NAS depend.

In addition to the limitations that it imposes through frequency congestion, the voice radio channel has been identified as a major contributor to errors in the ATC system. The FAA has noted that as many as 23 percent of all operational errors are caused either directly or indirectly by communications mistakes (New York Times, 1988). Similarly, compilations of voluntary reports provided to the Aviation Safety Reporting System by pilots and controllers have shown that a majority of all potentially hazardous incidents that are filed implicate ineffective verbal information transfer (Billings and Reynard, 1981).

Investigations of the nature of prevalent communications errors demonstrate that they are typically the result of an interaction between the characteristics of the voice radio system and the inherent perceptual and cognitive characteristics of its human users (Shingledecker, 1990). Acoustic confusions, alphanumeric transpositions, misinterpretation due to pronunciation and phraseology problems, poor memory for transient speech presentations of ATC information, and blocking of the radio channel caused by improper keying techniques are common sources of human-induced error found by these studies. In addition, many errors seem to be potentiated by the frequency congestion problem as users experience difficulty in monitoring for relevant messages on the crowded radio channel, and become reluctant to clarify suspected confusions in order to avoid further congestion.

Data Link is a digital communications technology which is being developed as a supplement to traditional voice radio for ATC communications. As shown in figure 1, Data Link communications can be supported by several transmission media. These include very high frequency (VHF) radio, satellite links, and the Mode Select (Mode S) secondary surveillance radar system currently proposed by the FAA for ATC Data Link communications. These multiple links will be integrated within a common Aeronautical Telecommunications Network to provide seamless air-ground communications throughout the NAS.

Regardless of the specific method used to create the channel, Data Link communications are distinguished from traditional voice radio links in two essential ways. First, unlike analogue voice messages, Data Link messages consist of digitally coded information. Thus, data may be entered for transmission either manually, or by direct access to information contained in airborne or ground-based computers. Furthermore, the capability of a digital system to provide automatic error checking of sent and received messages makes Data Link a highly reliable system which is not susceptible to degradation by interfering noise sources.

The second way in which Data Link differs from the voice radio channel is its capability to discretely address individual receivers. Unlike the simplex radio system which permits only a single speaker to transmit on the broadcast frequency at any point in time, Data Link messages can be sent selectively, and transmission rates are not artificially bounded by the effective

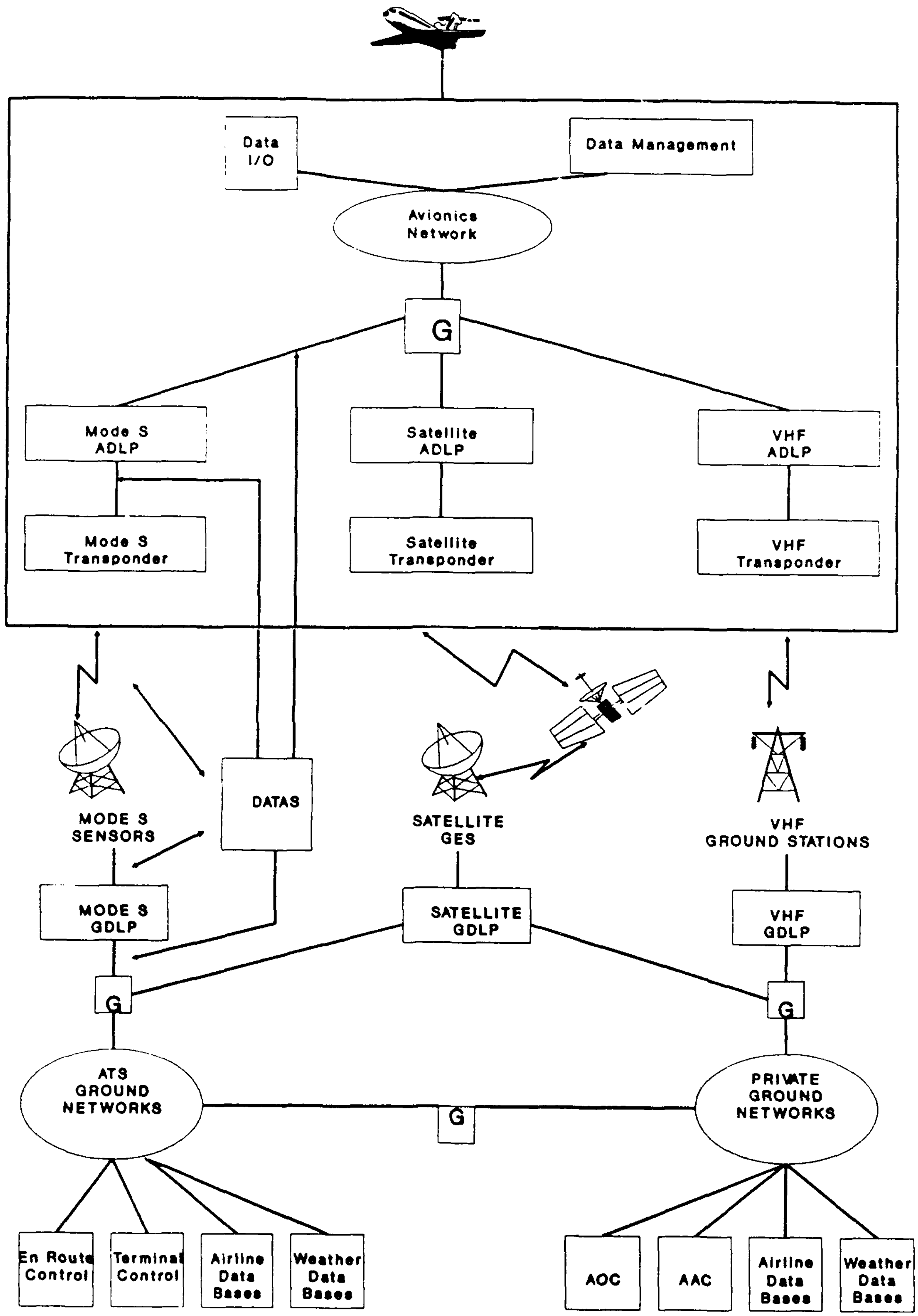


FIGURE 1. PROPOSED DATA LINK SYSTEM

speaking and listening rates of the user. As a result, Data Link channels can have a much higher capacity than voice channels and critical messages sent by a controller are assured of receipt only by the intended aircraft.

These features of Data Link offer significant promise for alleviating both frequency congestion and errors that currently impair air-ground ATC communications. Demands on the voice channel should be relieved in proportion to the number of weather and ATC services that are assigned to the Data Link system. In addition, by automating or simplifying pilot and controller functions in the communication process that are subject to error, Data Link should improve the overall effectiveness of information transfer. For example, by using Data Link it will be possible to reduce ambiguous message transmissions by storing standard clearances in computer memory for simplified uplink to an aircraft; failures to detect messages and accidental acceptance of clearances by unintended aircraft will be eliminated by discrete addressing; interpretation errors should be reduced by the availability of a persistent and recallable visual display of the received data; and the system will automatically verify the integrity of a message without human intervention.

1.2.2 Data Link Research and Development at the FAA Technical Center.

As noted above, the technical characteristics of Data Link have the capability to significantly enhance the safety and productivity of the ATC system. However, Data Link also will introduce a profound change in the way in which ATC tasks are accomplished by controllers, and in the way aircrews will receive and respond to ATC instructions. Because of this, the ultimate success of Data Link will be critically dependent on the extent to which it is employed to create an effective communications system that is thoroughly integrated with its human users and with the full range of tasks that they are required to perform.

Recognition of the need to consider operational suitability and human factors issues as primary drivers of the design process prompted the FAA Technical Center to initiate a program of manned simulation research to guide the development of Data Link ATC services. The overall goals of this research are to: (1) define useful Data Link services, (2) determine the user information requirements for Data Link communications, (3) develop display formats, data entry methods and procedures which promote efficient controller performance, and (4) evaluate the impact of Data Link services on both human and system performance.

The Data Link Test Bed was assembled at the FAA Technical Center to address these goals. The Test Bed is a laboratory facility which uses actual NAS equipment in conjunction with simulation computers to create a system capable of realistically exercising

Data Link applications in an end-to-end fashion. In its current form, the Test Bed is composed of the NAS en route and terminal laboratories, the NAS System Simulation Facility (NSSF), and the Data Link laboratory (figure 2). The NAS laboratory includes the HOST computer system used for en route ATC data processing as well as its primary terminal counterpart, the ARTS IIIA system. Both computers are linked to several suites of their respective operational controller workstations which are used to display radar data and to enter system inputs.

The NAS laboratory is linked to the NSSF through the ATC computers. The NSSF permits the NAS laboratory systems to act as functioning control facilities by providing simulated radar data and voice radio inputs from simulation pilots operating from computer terminals. Alternatively, the ARTS and HOST portions of the NAS laboratory can be used as self-contained simulation systems using the training functions included within the operational systems. In this configuration, pilot functions are performed by simulation operators working at additional controller workstations.

The Data Link laboratory houses a VAX 11/750 computer which acts as an emulation of the future ground Data Link processor. The VAX computer supports digital communication between simulation pilots and controllers. It can also provide two-way communication between controllers and high-fidelity aircraft simulators or actual airborne systems using Mode S or any other installed Data Link technology.

The central thrust of Data Link research in the Test Bed is manned simulation research aimed at defining and testing designs for ATC services. This research follows a three-stage approach originally developed and successfully employed under the en route portion of the Data Link program. In the **Design Verification** stage, engineering tests are conducted in the Data Link Test Bed to insure that preliminary designs for Data Link services are faithfully reflected in operational software and hardware components of the Test Bed simulation laboratories. Following the resolution of engineering issues, a series of manned simulation studies are performed in which air traffic controllers exercise and evaluate the Data Link ATC services.

In the **Mini Study** stage of these experiments, iterative design evaluations are conducted to refine controller procedures, displays, and input requirements. Early studies are completed under controlled, part-task simulation conditions which focus on detailed consideration of basic design issues. As development progresses, simulation exercises are increased in operational fidelity to assess the robustness of the services and to obtain reliable controller judgments of acceptability, usability, and workload effects. A fixed group of Full Performance Level (FPL) controllers from the Air Traffic Data Link Validation Team (ATDLVT)

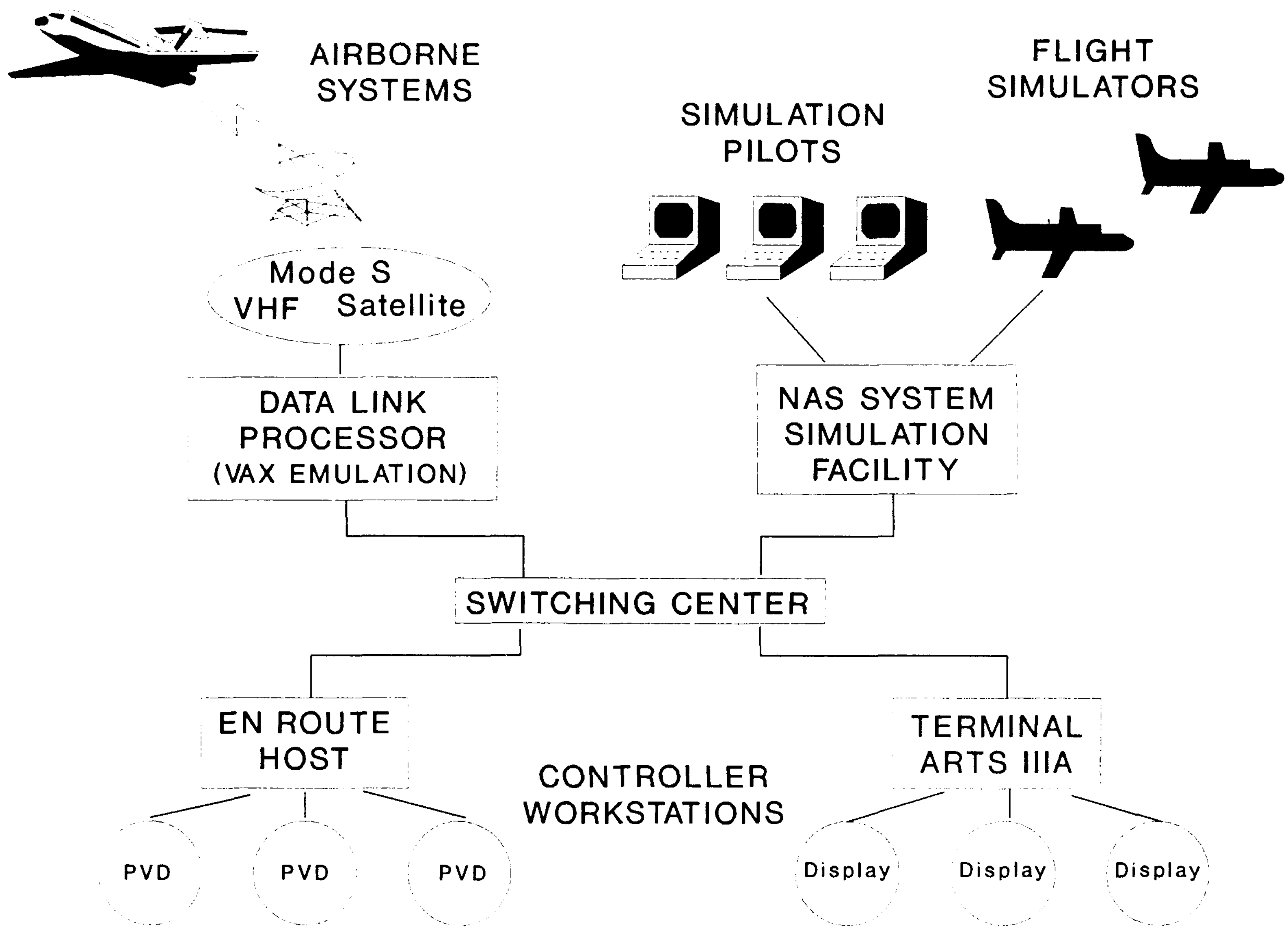


FIGURE 2. THE DATA LINK TEST BED

participated throughout the Mini Study stage to provide continuity in the iterative development process.

The final, **Operational Evaluation** stage of the approach consists of one or more high fidelity simulation exercises in which the optimized service designs are exercised under a variety of realistic operational conditions and air traffic scenarios. For these studies, a new group of controllers with no prior Data Link experience is recruited for participation. Measures of system effectiveness, controller performance, communications efficiency, and workload are used to verify the utility and usability of the Data Link services. The resulting data determines inputs to a Technical Data Package (TDP) that is used to guide the development of operational Data Link software for implementation in the NAS.

1.3 Data Link in the Terminal Environment.

1.3.1 Development Issues.

Research and development using the Data Link Test Bed began with the en route portion of the ATC system. Mini studies were conducted to refine the transfer of communication and altitude assignment services as well as a menu text function for uplinking interim altitudes with crossing restrictions, and an unformatted free text function. These efforts culminated in an operational evaluation which demonstrated beneficial effects of the initial en route services on frequency congestion with no observed reduction in controller performance or increase in perceived workload.

The potential value of Data Link communications technology for terminal ATC operations is at least equal to that predicted by the results of the operational evaluation for the en route environment. At present, the demands at busy airports often can result in a terminal controller engaging in prolonged periods of non-stop verbal communication to convey all of the clearances needed to guide the pilots of arriving, departing, and transient aircraft. In addition, the requirement to convey lengthy advisory messages to aircraft entering the terminal area rapidly expends the limited communication time available to tactically control closely spaced aircraft on the approach and departure flightpaths.

While the need to reduce frequency congestion is similar in the terminal and en route environments, the problem of designing an effective Data Link system for the two is quite different. In general, terminal operations are more sensitive to timing issues than en route operations. Because of this, communications functions assigned to Data Link must be carefully selected, designed, and tested to ensure that transmission delays or display clutter do not interfere with controller performance requirements. In addition, unlike some en route clearances, current terminal procedures do not require the use of keyboard inputs to update an ATC computer data base. For this reason, the need to develop efficient methods for entering and uplinking control messages is essential.

1.3.2 Initial Terminal Services.

Drawing from their experience as terminal controllers and an awareness of the design issues outlined above, the terminal subgroup of the ATDLVT met with FAA engineers and supporting contractors in a series of meetings held in 1989 and 1990 to define an initial group of ATC services suitable for the terminal environment. The following services were identified during these meetings:

a. Transfer of Communication. Transfer of communication is the message sent to an aircraft after track control has been passed to a new sector which instructs the pilot to change radio frequencies in order to communicate with the new controller. Using the designed Data Link service, this message is automatically prepared by the ATC computer and uplinked either automatically or upon a controller input action.

b. Initial Contact. When an aircraft receives a new radio frequency, current ATC procedures require the pilot to contact the new controller and to report the aircraft's currently assigned altitude. With the Data Link version of initial contact tested during this study, after the new controller obtains Data Link communication eligibility, a request for the initial contact report is sent automatically to the aircraft, and the downlinked report is presented to the controller on the radar display.

c. Terminal Information (TI) Service. When arriving aircraft enter a terminal airspace, they are typically given a report of the terminal operating conditions and of the approach clearance that they can expect to receive. Using Data Link, these commonly lengthy messages are stored in a menu and sent by a single manual input which initiates the uplink.

d. Menu Text (MT). MT is a Data Link function which permits the controller to select and uplink a commonly used clearance from a menu list displayed on the radar screen. In the service design tested in this study, such clearances included speed, heading, and altitude change instructions.

1.3.3 The First Design Study.

Development of the initial terminal Data Link services began with demonstration of preliminary designs using a rapid prototyping system at The MITRE Corporation. In 1990, the services were implemented in the ARTS IIIA computer and integrated with the Data Link Test Bed. The first Mini Design study of these services at the FAA Technical Center was conducted in late 1990 to establish a developmental baseline (Data Link Development Team, 1991). Nine controllers from the ATDLVT participated in a design review and compared the current voice-only system to a system supplemented by the initial Data Link capability in a series of simulation trials employing simplified air traffic scenarios.

Controller projections made after the completion of the simulation tests indicated that Data Link would have positive effects on terminal operations if implemented in the existing ATC system. These effects were estimated to include a significant reduction in voice radio frequency congestion, simplified transfer of lengthy advisory messages and clearances, and a reduction in communication errors and controller workload. The results of the design review showed that the preliminary service configurations that were tested might be significantly improved by several enhancements and modifications. These included the addition of transaction status and history lists, improvements in the accessibility of items contained in the MT and TI service lists as well as increased flexibility in combining the messages contained in these lists, and changes in the transaction status information provided in the data block display.

The study described in this document was conducted, in part, to evaluate the effectiveness of the alterations made to the Data Link service designs as a result of these findings.

1.4 ORGANIZATION OF THE REPORT.

The following sections of this report present the research methodology that was used and the findings that were obtained in the second FAA Technical Center controller evaluation study of Data Link terminal ATC services. Section 2 describes the specific objectives of the study and the testing approach that was used to achieve these objectives. Sections 3 and 4 present the detailed results of the testing. Sections 5 and 6 list the conclusions that were derived from the results and offer recommendations for future efforts toward the development of an operational terminal Data Link system.

2. TEST DESCRIPTION.

2.1 OBJECTIVES.

This study was conducted to meet the following major objectives:

a. Evaluate the acceptability of enhanced designs for the initial Data Link terminal services.

As noted above, the results of the first terminal Mini Study included recommendations for changes to the ATC service designs as implemented on the ARTS IIIA equipment in the Data Link Test Bed. Controllers participating in this second study evaluated the modified displays and procedures to determine the adequacy and acceptability of the enhanced service designs. The modifications examined during the study included:

1. Addition of Data Link transaction status and history lists.

2. Expansion and standardization of the data block display of Data Link transaction states.

3. Extended options for automatic and manual operation of the Transfer of Communication (TOC) service.

4. Modifications to permit uplinks of additional messages during the period that a TOC is in a "held" status.

5. Procedures for acquiring Data Link eligibility in the absence of a TOC.

6. New labeling and procedures for designating the default message in the TI list.

7. Addition of combined (speed, heading, altitude) clearance items to the MT list.

8. Addition of capability to combine clearances for uplink using the menu by-pass feature.

9. Capacity to combine any TI message with an MT message.

b. Resolve open service design issues.

A limited number of basic service design and procedural issues were either not addressed during Mini Study 1, or were kept in a pending status until various options could be examined in the Data Link Test Bed. In order to resolve these issues, the present study permitted test subjects to exercise some potential design alternatives during a series of simulation trials and consider others during formalized discussion sessions. The following options and issues were assessed:

1. Graphic character options for the Data Link equipage and eligibility indicators.

2. Location options for the Data Link key (F9 vs. F16).

3. Spacing and labeling options to enhance the controller's ability to locate messages within the TI and MT lists.

4. Operational suitability of free controller access to TI and MT lists for modification, deletion or addition of items.

5. Possible reconsideration of names for TI and MT lists, and rules for assignment of messages to these lists.

6. Feasibility of uplinking final approach clearances.

7. Operational validity of requirement for all designed automatic/manual TOC options.

c. Provide a preliminary assessment of the impact of system delay/pilot response time on the usability and effectiveness of the initial terminal services.

Beyond the consideration of basic design problems, the present study also was used to provide an initial examination of issues affecting the suitability of Data Link services within the operational terminal ATC environment. As a part of the study, test subjects participated in full scale simulation exercises in which current, voice radio procedures were compared to combined voice and Data Link communications. Test runs conducted with Data Link varied in the average time between initiation of an uplink message and the receipt by the controller of a downlinked aircrew response. The purpose of this manipulation was to determine the usability of Data Link at different time delays, and to examine the relationship between usability, delay, and control position type. The primary indicators of delay effects were workload ratings, controller perceptions of the nature and extent of delay effects, and measures of voice radio and Data Link usage.

d. Evaluate the effectiveness of a subject training protocol and the validity of experimental system and controller performance measures for use in future operational evaluation studies.

The final objective of this study was to evaluate subject training procedures and performance measures for possible use in future operational evaluation research. Proven training protocols will be needed for this research in order to insure that controllers inexperienced with Data Link and the airspace used for high fidelity simulations receive sufficient pretest training to permit uncontaminated measurements of the operational effects of Data Link. Advanced performance measures will be required for this research to supplement expert controller opinions with objective indicators of Data Link's impact on ATC communications efficiency, controller error, and overall ATC system performance.

The protocol evaluated during this study included training on procedures and displays for the initial ATC services and training on the Raleigh-Durham (RDU) ATC Tower airspace that were used during simulation tests. The experimental performance measures were collected during full scale test runs and evaluated for sensitivity and applicability to operational evaluation research questions.

2.2 APPROACH.

The approach that was adopted to meet the objectives of this study involved the participation of terminal air traffic controllers in a series of training exercises, test sessions, and structured debriefings. The simulation test trials were conducted at the ARTS IIIA workstations in the Data Link Test Bed. During testing, subjects controlled traffic in a group of ATC scenarios involving aircraft arrivals, final approaches and departures at the RDU ATC Tower.

Early test sessions were devoted to a detailed review of each of the four Data Link service designs as modified by the results of the first Mini Study. Individual reviews were supplemented by debriefings aimed at achieving group consensus.

Later sessions were used to compare voice radio air-ground communications to a system supplemented by Data Link under full scale simulation conditions. Data Link sessions differed in the average time which elapsed between the dispatch of an uplink message by a controller, and the receipt of a confirmatory message from a simulation pseudo-pilot. During these sessions, voice radio and Data Link usage, as well as a number of experimental performance measures, were recorded automatically by the simulation system and through the use of video tape and facilitator observations. Following each run, controller perceptions of workload and delay effects were quantified using rating scales and checklists. Post-test scales were used to obtain controller projections of the effectiveness and acceptability of the tested service designs.

The general rationale underlying the test design was to provide an ATC environment which could be used by the controllers to assess the capabilities and robustness of the service designs against relatively realistic operational conditions, and to generate a preliminary evaluation of the usability of Data Link services under various transmission delay conditions and in differing sections of terminal airspace.

2.3 TEST CONDUCT.

2.3.1 Subjects.

The subjects for this study were six FPL ATC specialists with current terminal radar control experience. Five of the six subjects were drawn from the membership of the ATDLVT and had participated during the conceptual development of the initial Data Link terminal ATC services and acted as test subjects for Mini Study 1. All subjects were males. Their average age was 38.7 years (32 to 47 years). Their experience as FAA controllers ranged from 7 to 24 years with a mean of 16 years. Five of the six had prior military controller experience and one controller held an aircraft pilot's license.

2.3.2 Test Scenarios and Data Link Operations.

The ATC scenarios developed for this study utilized the RDU terminal airspace. The airspace, local ATC procedures, and test scenarios used during training and testing are presented in detail in appendix A of this report and are briefly described below.

Traffic patterns and procedures used in the scenarios were identical to those used at the operational facility. The single exception was that simultaneous approaches to the parallel runways were permitted in the simulation, whereas, staggered approaches are required at RDU. As shown in figures 3 and 4, incoming aircraft were routed through two arrival sectors located to the east and west of the airport. Each arrival controller accepted aircraft handed off from the Washington Air Route Traffic Control Center (ARTCC) over two fixes. Overflight aircraft were given clearances for their destination airports, while RDU arrivals were established on a downwind leg or on headings for final approach before control was transferred to the associated final approach sector. Each final controller merged the two streams of aircraft received from his arrival sector and issued the approach clearances. Controllers in the two departure sectors (north and south departure) each directed aircraft to one of two departure fixes according to their flight plans. In both sectors, one of the departure streams crossed an arrival route. This required the controllers to insure that the departing aircraft met specific altitude restrictions while crossing the arrival route.

The mix of aircraft types used in all scenarios were derived from actual RDU air traffic records. All scenarios used during full scale simulation testing presented traffic loads equivalent to 75 percent of airport capacity, and differed from one another only in the sequence and spacing of arrivals and departures. Lower traffic levels were used during training and during the design review phase of testing.

Pilot functions for this study were provided by the NSSF. Pseudo-pilots received voice and Data Link messages from the controllers and made inputs on specialized computer terminals to control the simulated aircraft radar tracks. When responding to voice radio instructions, the pseudo-pilots acknowledged clearances in the normal fashion with a voice response to the controller. Precise control of the elapsed time between the issuance of a Data Link message and receipt of a confirming response was achieved by using the VAX computer to automatically generate and send the pilot acknowledgement to the controller via Data Link. In order to produce realistic temporal coordination between aircraft maneuvering responses and downlinked pilot responses under extended delay conditions, the VAX computer was programmed to withhold displaying messages to the pseudo-pilots until approximately 8 seconds before the acknowledgement was sent to the controller.

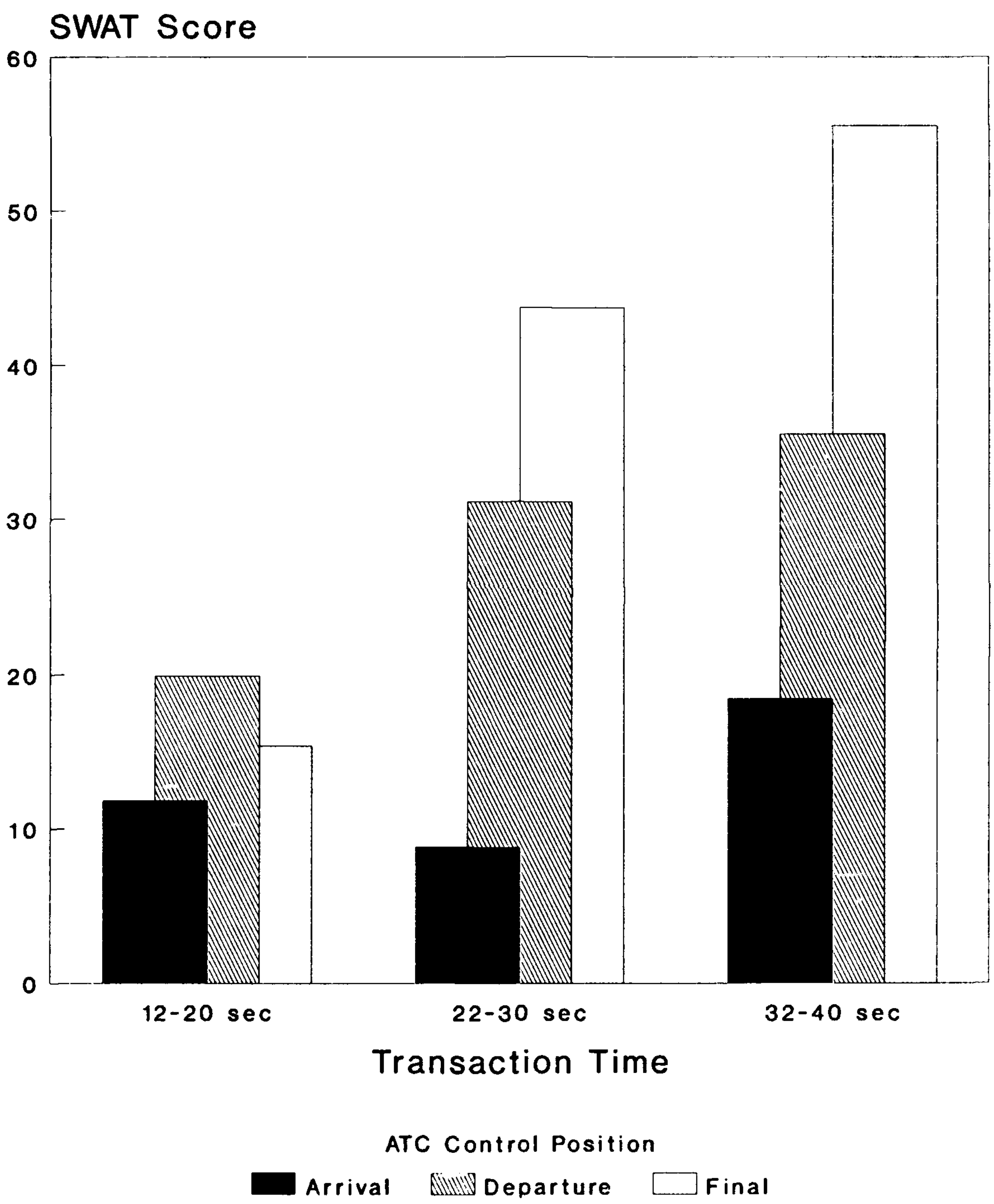


FIGURE 3. EFFECTS OF DATA LINK TRANSACTION DELAY ON CONTROLLER WORKLOAD

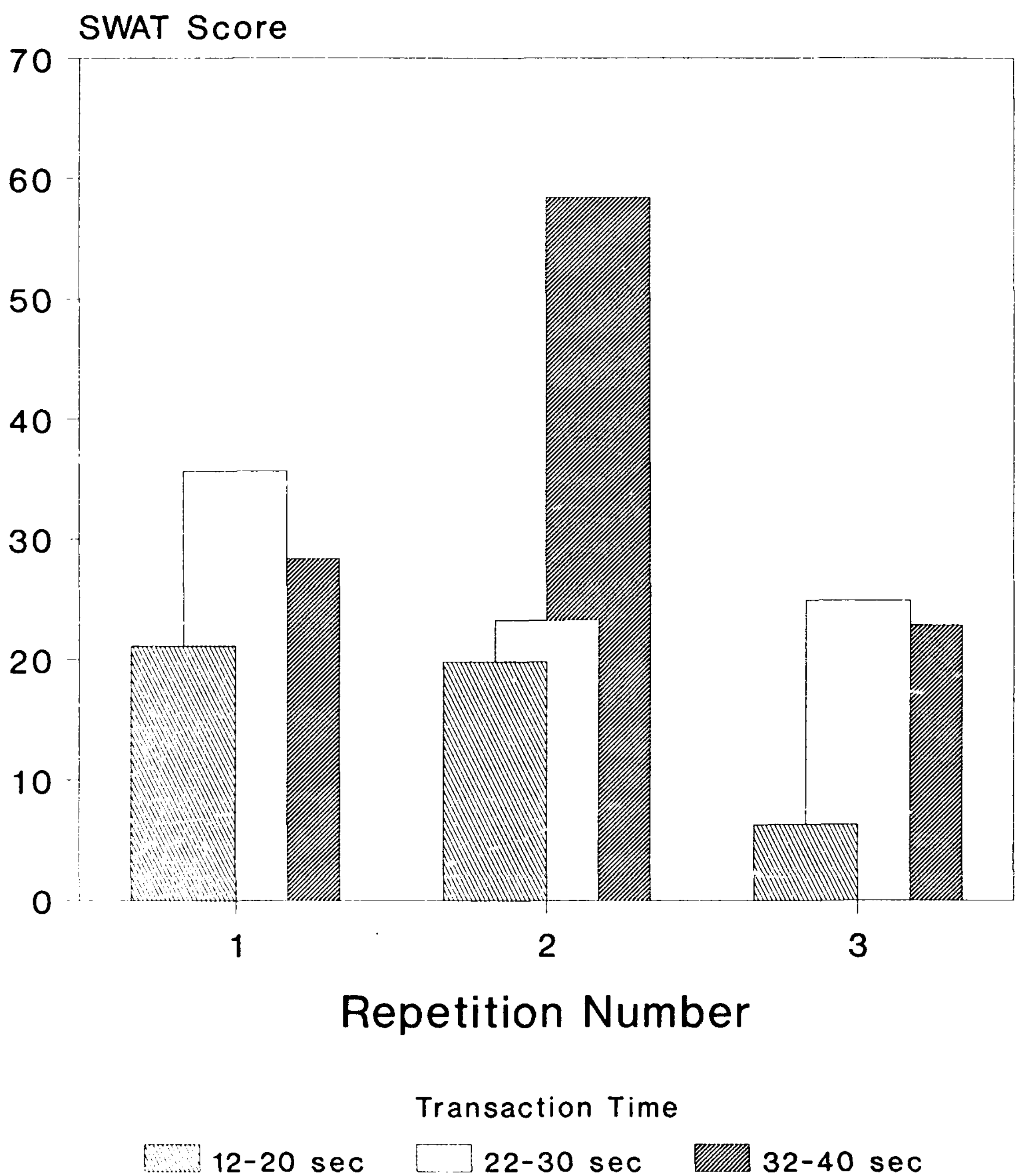


FIGURE 4. EFFECTS OF REPEATED TESTING ON WORKLOAD SCORES

2.3.3 Test Procedures.

This study was conducted over a 6-day period. Portions of the first 2 days were devoted to subject prebriefings and training. Data collection began on the third day with a critical review of the Data Link terminal service designs. On the fourth and fifth days the controllers completed 12 full scale simulation tests under voice-only and Data Link conditions. The sixth day was reserved for final debriefings and structured discussion.

2.3.3.1 Airspace and Data Link Training.

To simplify airspace familiarization, the six subjects were divided into two groups and assigned to either the north or south halves of the airspace. The subgroups of three controllers were required to learn the arrival, final, and departure sectors only for their assigned half of the airspace. These assignments were maintained throughout the experiment.

Training for this study began with a 1-1/2 hour classroom session during which the RDU airspace and procedures were described. The subjects also received copies of the briefing materials for study and reference use in subsequent simulation training exercises. The controllers received practice with controlling traffic in the airspace during a 3-hour session in the Data Link Test Bed which followed the classroom briefing. The scenario used for this session employed only voice radio communications in order to emphasize the airspace training task. The subjects rotated through their assigned arrival, final, and departure sectors to obtain experience with all relevant airspace.

The second phase of training was devoted to the displays, manual inputs, and procedures for using the initial terminal Data Link services. Initially, the subjects received a 3-hour period of classroom briefings on the operation of software functions for each service. These briefings were supplemented with a handout summary of Data Link inputs and displays. The remaining 4 hours of training were conducted in the Data Link Test Bed. For these practice sessions, 75 percent of the aircraft in the test scenarios were equipped with Data Link in order to provide experience with combined voice and Data Link communications.

Appendix B to this report fully describes the protocol used for both the airspace and Data Link training, and contains the briefing presentation materials and summary handouts that were used.

2.3.3.2 Design Review.

Data collection for this study began with a formal review and verification of the ATC service designs as modified by the results of Mini Study 1. The subjects completed the design review in the Data Link Test Bed while seated at the ARTS IIIA workstations and controlling aircraft in a simplified version of the training scenario. All aircraft in the scenario were Data Link equipped in order to maximize the subjects' opportunities to examine the service displays, inputs, and procedures. To permit observation of the failure displays, approximately 5 percent of the attempted uplinks resulted in a failed technical acknowledgement (NAK), 5 percent in a time-out (failure of the pilot to respond to an uplink within 40 seconds), and 5 percent in an unable response from the pilot.

During the design review, the subjects' primary task was to exercise each of the Data Link functions a sufficient number of times to thoroughly evaluate the service designs. Evaluations were made by completing a questionnaire booklet during the simulation runs. The subjects were informed that the object of the simulation activity was to aid them in completing the detailed design review, and that maintaining routine control over the moderate (50 percent) level of air traffic in the scenario was secondary to this task. The subjects were encouraged to exchange control positions as required in order to examine services which were not used in some sectors. In addition, test facilitators assigned to each sector were available to assist the subjects in making keyboard inputs necessary to permit examination of unresolved display and control options. The scenario was repeated throughout the 4-hour laboratory session, providing ample time for the subjects to complete their individual reviews.

Upon completion of the Test Bed exercise, the subjects met with test personnel for a group debriefing session. This 3-hour session was used to perform an item-by-item review of the subjects' responses to the review questionnaire. The emphasis of the debriefing was to identify and resolve disagreements regarding the fidelity and acceptability of the service designs, and to achieve a consensus regarding recommended changes to the service designs. The results of this debriefing, as well as those of additional design discussions held after the full scale simulation tests, were documented in test personnel notes and in an audio tape record for reference during data analysis.

2.3.3.3 Full Scale Simulation and Delay Testing.

The fourth and fifth days of the study were devoted to a series of full scale simulation tests in which the subjects controlled traffic in the six sectors of airspace comprising the RDU terminal area. The purpose of these test runs was to: (1) determine the impact of varying pilot response delays on the usability and acceptability of the Data Link ATC services, (2) evaluate a group of candidate measures of controller and system performance for use in future operational evaluations, and (3) provide the subjects with a realistic simulation experience as a basis for additional expert evaluation of the operational utility and acceptability of the terminal Data Link services as implemented for this study.

The independent variables that were manipulated during these test runs were communication condition and control position. The communication condition variable had four levels permitting comparison of exclusive voice radio air-ground communications with three conditions in which Data Link communication were possible with 75 percent of the aircraft within the test scenarios. The three Data Link conditions differed in the average time delay between initiation of an uplink and receipt of a confirmatory response. The control position variable was manipulated to assess any relationship between the communication conditions and the type of airspace being controlled. The control positions tested were the arrival, departure, and final approach sectors.

The experiment was conducted under a completely within-subjects design with repeated measures on both factors. Thus, each of the six subjects experienced all four communications conditions while working at each of the three control position types. This design required 12 simulation runs, each with an approximate duration of 40 minutes. The subjects maintained their original assignments to the north or south portions of the test airspace and rotated among the position types as shown in table 1.

The order of presentation for the communication conditions was varied, as illustrated in the table, to provide a partial control for sequence effects. Within this ordering scheme, the voice condition and the longest delay condition (DL 3) both occurred first, last, and approximately in the middle of the test sequence for one of the three repetitions of the four conditions.

Six versions of a basic test scenario were used during the 12 test runs. In all cases, the traffic level was raised to 75 percent of airport capacity from the 50 percent level used during the design review session. As noted in section 2.3.2, the six versions of the scenario were matched in terms of the number and type of aircraft involved and the complexity of the control problem. However, the aircraft spacing and sequencing were varied in order to minimize the effects of overlearning on later test trials. As shown in table 1, the traffic scenario was changed after each simulation run so that subjects did not work at adjacent control positions within the same scenario on sequential test runs. In addition, subject assignments to adjacent sectors in the north and south portions of the airspace were varied to minimize consistent pairing. These manipulations were introduced to reduce any contaminating effects of excessive familiarity with specific patterns and to avoid the development of cooperative strategies among controllers.

Under the experimental design shown in table 1, an individual subject experienced each scenario twice, and worked at the same control position in a single scenario in only two of the six cases. In all instances, these exposures to the same control position within the same scenario occurred at least five test runs apart, and were evenly distributed across subjects over control positions and communication conditions.

TABLE 1. EXPERIMENTAL DESIGN FOR FULL SCALE SIMULATION

Scenario	Run Order	Communication Condition	Subject/Position Assignment					
			S1	S2	S3	N1	N2	N3
1	1	Voice	F	D	A	A	F	D
2	2	DL 1	D	A	F	F	D	A
3	3	DL 2	A	F	D	D	A	F
4	4	DL 3	F	D	A	A	F	D
6	5	DL 1	A	F	D	D	A	F
1	6	DL 3	D	A	F	F	D	A
5	7	Voice	D	A	F	F	D	A
2	8	DL 2	F	D	A	A	F	D
3	9	DL 3	A	F	D	D	A	F
4	10	DL 2	D	A	F	F	D	A
5	11	DL 1	F	D	A	A	F	D
6	12	Voice	A	F	D	D	A	F

KEY

Subjects: S# = South
N# = North

Conditions: Voice = Voice Radio Only
DL 1 = Short Pilot Data Link Response Delay
DL 2 = Moderate Pilot Data Link Response Delay
DL 3 = Long Pilot Data Link Response Delay

Control Positions: A = Arrival
D = Departure
F = Final Approach

The mean pilot delay times that were used in the three Data Link test conditions were 10 seconds for the shortest delay (DL 1), 20 seconds for the moderate delay (DL 2), and 30 seconds for the longest delay (DL 3). Within each condition, pilot response delays varied within a range ±2 seconds around the means and were randomly selected from the range for each Data Link transaction. In addition, a system transmission delay time was added to the pilot delay for each transaction from a rectangular distribution with a mean of 6 seconds and a range from 4 to 8 seconds. These values are based on a best current estimate of Mode S Data Link total system delays for the combined uplink and downlink. The sum of the pilot and transmission delays yielded average total delays for the three conditions of 16 seconds (DL 1), 26 seconds (DL 2), and 36 seconds (DL 3). The corresponding ranges of total delays were 12 to 20 seconds (DL 1), 22 to 30 seconds (DL 2), and 32 to 40 seconds (DL 3).

In all Data Link conditions, 2 percent of the attempted uplinks resulted in a technical failure (NAK). All successful uplinks received a positive acknowledgement (WILCO). Thus, time-outs and UNABLE responses were not tested in this portion of the study.

The delay times used in this study were generated by examining current estimates of Mode S transmission times and of potential pilot response times as projected from a limited number of airborne Data Link simulations. However, it should be noted that the subject controllers were not aware of the portions of time attributable to system and pilot delays, and experienced only an undifferentiated total delay time between sending a message and receiving a pilot response. Hence, the results of the study are not limited to any specific transmission technology, and should be applicable to any combination of system and pilot delays which would produce total transaction delays in the ranges that were tested.

The subjects were told prior to each run whether or not the scenario would contain Data Link equipped aircraft. However, to avoid potential judgement bias in post-run ratings, they were not informed of the average duration of pilot delays to be expected during the run.

During each test run, a group of communications measures and experimental system performance measures were collected. Following each run, subjects completed a rating to indicate the perceived level of workload that they had experienced. In addition, after each Data Link run, the subjects estimated the perceived degree of impact of the Data Link delays, and were asked to provide written comments on specific experiences and coping strategies. After all test runs were completed, the subjects were asked to provide projective ratings of the operational suitability and of the acceptability to controllers of each Data Link service as implemented for the present study.

2.3.3.4 Post-Test Session.

The final day of this study was devoted to an extended structured debriefing and discussion. The primary topics addressed during the session were the individual Data Link service designs, global design philosophy issues, training methodology, the perceived impact of time delays on Data Link usage and system performance, and the overall projected effects of Data Link on ATC.

Group discussions on specific topics were initiated by test personnel. Each topic or question was presented on an overhead projection slide. The discussion facilitator elicited differing opinions from the test subjects and attempted to determine whether a group consensus opinion could be achieved in each case. Extensive notes were taken by test personnel in order to accurately represent the findings and any conclusions of the discussions. In addition, audio tape recordings were maintained to resolve any discrepancies among the written notes during data analysis.

The service designs were revisited during the debriefing in order to identify any changes in the judgments made during the design review that had arisen as a result of extensive usage of the services during the full scale simulation tests. Global issues discussed during the session included the perceived appropriateness of using Data Link in the arrival, departure, and final approach sectors, potential sources of controller error detected by the subjects during the full scale tests (e.g., memory failures), the usability of the TI and MT lists, and reconsideration of names for the TI and MT lists.

Discussion of the training protocol centered on the adequacy of the training provided during the present study, and suggested modifications for future testing with inexperienced controllers. Finally, the subjects were asked to consider the way in which Data Link delays had affected their performance during the tests and to project some of the overall effects that Data Link may have on the ATC system.

2.3.4 Data Collection.

2.3.4.1 Design Review Materials.

The questionnaire booklet used for the design review was similar to that developed for the first terminal Mini Study. For the present study, the review booklet was organized in two sections (see appendix C). The first section addressed the designs of the transaction status list and data block status displays, the history list, the initial contact service, the TI service, the TOC, and the MT function used for sending speed, heading, and altitude clearances.

Questions relevant to each of these topics were prefaced by a text description of the operational features relevant to the function or service. Initial questionnaire items verified the fidelity with which the Test Bed service implementations reflected the design modifications introduced for this study by requiring the subjects to judge the correspondence between the descriptions and their actual Test Bed experience Succeeding items assessed the acceptability of the service implementations observed in the Test Bed, and solicited recommendations for any further design modifications.

The second section of the review focused on selected unresolved design issues. Questions addressed preferences for the symbols used to indicate Data Link equipage and communication eligibility in the full data block, optional locations for the Data Link key on the ARTS IIIA keyboard, and potential changes to the MT and TI service list designs.

2.3.4.2 Controller Evaluation Instruments.

Controller workload was assessed immediately after each of the 12 full scale simulation test runs using the Subjective Workload Assessment Technique (SWAT). SWAT was developed in the early 1980's by the U.S. Air Force as a standardized method for obtaining quantified estimates of perceived workload in a broad variety of occupational tasks. The technique has received extensive use in simulation and operational testing environments within the Department of Defense, and was used successfully with air traffic controllers in prior en route and terminal Data Link studies.

SWAT is used to rate the level of workload experienced by a subject during a preceding period of work. Briefly, SWAT consists of three, 3-point rating scales referring to the dimensions of time load, mental effort, and psychological stress. Subjects indicate the workload associated with an activity by marking the appropriate point on each scale.

A unique feature of this workload measurement technique is that the three ordinal ratings that are obtained are converted to single points on an overall interval measurement scale of workload with values ranging from 0 (low workload) to 100 (high workload) using a mathematical analysis method known as conjoint measurement. This method not only yields data which are amenable to powerful, parametric statistical testing, but also tailors the measurement scale to each individual's (or homogeneous group's) concept of how the time, effort, and stress dimensions combine to produce the overall perception of workload.

The interval scale used to interpret the ordinal ratings is created by having subjects generate an ordering of all 27 combinations of the time, effort, and stress levels on the scales which reflects their individual concepts of how the three dimensions combine to produce different workload levels. The card sorting task used during the scale development exercise was completed by all six subjects in this study during a prebriefing session. The SWAT rating scale is included in appendix C along with instructions provided to subjects during the scale development exercise.

Following each of the nine Data Link test runs, the subjects completed a series of rating scales and a checklist designed to elicit their perceptions of the effects of the Data Link delays. The data form used for these items also is contained in appendix C.

The rating scales asked the subjects to estimate the effects of the delays on: (1) their ability to control air traffic, (2) their use of the Data Link services, and (3) their use of voice radio instead of Data Link. A fourth scale required the subjects to estimate the extent to which they would use Data Link in their home facilities if delays as long as those experienced on the previous test run were typical. Finally, a 10-item checklist was provided for the controller to record additional effects of delays that were experienced and any coping strategies used to deal with them.

After completing all 12 simulation runs, the subjects performed a final set of evaluations. The controllers rated each of the four terminal service designs on two factors (see forms in appendix C). The first scale required the subjects to assess the operational effectiveness and suitability of the service design. The data form permitted the subject to rate a service as "not operationally suitable" or on a 7-point scale ranging from 1 "minimally effective" to 7 "highly effective." Instructions for completing this scale asked the subjects to use their experience in the Data Link Test Bed and their prior background in terminal ATC operations to assess how well each service design could accomplish its intended task in the full range of potential field environments.

The second rating was an estimate of the acceptability and preferability of each of the service designs to air traffic controllers. The data form permitted the subjects to rate a service as "completely unacceptable" or on a 7-point scale ranging from 1 "acceptable, but not preferred" to 7 "highly preferred." Instructions directed the subjects to consider the extent to which the displays, input requirements, and procedures used to deliver a service would be usable by air traffic controllers.

Additional explicit instructions for these two scales indicated that the dimensions of effectiveness and preference should be treated independently from one another since a service might include all functions needed to meet operational requirements and still be poorly suited to the way in which controllers perform their functions. Likewise, a design could be easy to use, but have missing features which prevent it from meeting operational needs.

2.3.4.3 Communications Measures.

Because one of the proposed benefits of an air-ground Data Link is a reduction in voice radio frequency congestion, data were collected during the simulation runs to gauge the use of the voice radio and Data Link by the subject controllers. Radio usage was assessed by automatically detecting the occurrence of all push-to-talk activations and deactivations of the controllers' microphone when speaking to the simulation pilots. Recordings of these events yielded measures of the number of controller initiated voice transmissions during a test run and of the amount of time the radio channel was occupied by controller transmissions. The number of Data Link transactions initiated by the controllers were automatically recorded by the VAX computer which acts as an emulation of the future Data Link Processor (DLP) and is responsible for handling all digital communications in the Data Link Test Bed.

2.3.4.4 Experimental Performance Measures.

In addition to the controller evaluation and global communications measures, a candidate group of objective performance metrics were collected during the full scale simulation tests in order to determine their feasibility for use in future operational evaluation. These measures included ATC system performance and additional indices of communications system use by the controllers.

Systems performance indices that were collected included measures of the degree of spatial separation among aircraft and of the efficiency with which aircraft were handled by the controllers and moved through the test airspace.

Separation measures recorded by NSSF computers included the number of times two aircraft came within 1,000 feet vertically or 3 miles horizontally of one another, as well as the amount of time that the aircraft spent within these proximity limits. The computers also recorded two additional separation measures: the Closest Point of Approach (CPA) and the Aircraft Proximity Index (API). The CPA is a nonweighted calculation of the shortest slant range distance between two aircraft within the proximity limits. The API is calculated from an algorithm which compensates for differences in vertical and horizontal separation limits by using a weighted combination of the distances. The API vertical component is weighted approximately 18 times greater than the horizontal component and the resulting score ranges from 0 to 100, with 100 indicating a collision between the aircraft.

Efficiency measures also were collected by the NSSF computer. These included the distance flown by each aircraft within a sector, the time spent within a sector, and the number of controller-initiated path changes. Capacity was handled by each controller and examined by a minute-to-minute assessment of the number of aircraft.

Additional communications measures recorded to examine controller performance included the time needed to issue a Data Link clearance or message, the frequency of callbacks from simulation pilots, keystroke errors, repeats of voice messages, voice contacts with Data Link equipped aircraft, and attempts to contact unequipped aircraft using Data Link. Data Link message input times were assessed for only three of the control sectors using a video tape record of the preview area on the controller's display. The preview area displays the results of each key action in real time and is normally used by the controller as visual feedback for data entries. Keystroke events and input times were extracted manually from the tapes during analysis. Observers seated next to each controller used checklists to record keystroke errors, callbacks, repeated messages, usage of voice with Data Link aircraft, and attempts to contact nonequipped aircraft using Data Link.

3. TEST RESULTS.

3.1 DATA LINK SERVICE DESIGN REVIEW.

Two major objectives of this study were to evaluate the terminal Data Link service designs as modified by the results of the first ARTS IIIA study and to resolve a group of design issues which remained open after the first study. The following subsections of this report present the findings that were obtained with respect to these objectives. The results combine data that were derived from the individual design review booklets, the design review debriefing, and the portions of the post-test debriefing that addressed the service designs after the subjects had completed the full scale simulation testing series.

3.1.1 Status List and Data Block Status Displays.

The original ARTS IIIA implementation of the terminal Data Link services presented information to the controller regarding the status of an ongoing transaction in the third line of the full data block associated with an aircraft track. The results of the first study suggested that it may be necessary to increase the amount of information provided in the display by indicating the precise nature of a transaction failure (e.g., a technical failure to reach the aircraft (NAK), a pilot failure to respond promptly, or a pilot response of "UNABLE to comply"). In addition, it was recommended that a status list be tested which would summarize all ongoing transactions and provide full status data as well as a 4-second display of WILCO as a confirmation that a recently completed transaction had been successfully closed. Both of these modifications were examined by the controllers in the present study.

The design review results indicated that five of the six subjects felt that the status list would not be a required feature for operational implementation. The subjects who felt that the list would not be needed noted that the list was largely redundant with the data block display, and that consulting the list would require an undesirable redirection of the controller's attention. The controller who felt that the list would be required indicated that it provides the persistent WILCO display which was not included in the data block display to reduce clutter, but would be useful in the list if the controller failed to notice the termination of a transaction and was unsure whether the message had been successfully sent and acknowledged. In subsequent discussion, the controllers generally agreed that the status list may have value in limited situations because it can contain more complete information than the status block alone, and should be retained in the design to accommodate any need for cuing the controller's memory as well as the requirements of a predicted minority of users who may rely on the list as a primary source of transaction information. However, because of its redundancy, it appeared that the list could be sacrificed in the final design if system computational or display limitations demanded the deletion of some Data Link features.

All of the controllers felt that the status abbreviations used in the list and the full data block display provided sufficient information and were easy to interpret. The subjects also agreed that the 4-second persistence of the WILCO in the status list was probably sufficient, but that a longer duration should be tested. They noted that, as long as this feature is maintained, the full data block should continue to indicate the successful completion of the transaction by immediately clearing the third line. One subject said that the 4-second display would be sufficient because the completed transaction also could be verified by consulting the history list. Four of the subjects also indicated that if the status list is maintained, it should be upgraded to include the message identifier when MT and TI messages are sent in a manner similar to that used for the full data block display.

Subsequent consideration of the individual services and procedures resulted in agreement among the subjects that the failure messages (TIMEOUT, NAK, UNABLE) should flash in the data block in order to capture the controller's attention. In addition, significant changes to the design were recommended with respect to the handling of the "UNABLE" and "TIMEOUT" status messages. The controllers indicated that the software should prevent the controller from attempting to resend a message to which the pilot has responded by indicating that he is UNABLE to comply. Likewise, because the timeout message is intended to act only as a cue to the controller that an extended period of time has elapsed since a message was sent, the controllers argued that the timeout should not lock out a subsequent response from the flight deck. According to this recommendation, the design should permit the controller to continue to wait for a WILCO or UNABLE response, if warranted by the ATC situation.

3.1.2 History List.

The history list is a feature that was added for the present study based on controller concerns regarding problems that may arise when attempting to recall Data Link instructions that had been issued to an aircraft. Displayed by controller request, the list provided a record of the last five messages that had been acknowledged by an aircraft.

Five of the six controllers indicated that the history list will be a necessary feature of the design for operational implementation. These subjects felt that the list provided a necessary review capability to ensure that past messages had been sent and successfully acknowledged. It was also noted that the list would be a significant aid when briefing a relief controller.

Four subjects felt that five messages was a sufficient history record. However, two of these, as well as the remaining two subjects, felt that the past three transactions also would be sufficient and acceptable. Five subjects indicated that the existing history display format was acceptable. The tested display presented the messages in reverse order of receipt, with the last transaction completed at the bottom of the list and the oldest item at the top.

For the present study, when requested by a slew action, the history list overlaid the position occupied by the TI and MT lists and was displayed for 8 seconds. Three subjects indicated that the display persistence was adequate. However, all subjects indicated that display time would be dependent on the number of items in the list. The controllers agreed during subsequent discussion that the display time should be adaptable and include an ability to manually terminate the display as in the existing design. The controllers also generally agreed that the list could continue to share its location with the menus, but that a separate location, or sharing positions with the status list, would be preferred.

3.1.3 Initial Contact.

Substantial changes to the initial contact service were recommended by the subjects during this study. In the tested design, the initial contact was accomplished by automatically uplinking an assigned altitude request when the TOC was completed and the receiving controller had obtained Data Link messaging eligibility. Two problems with this procedure were detected by the subjects during the design review and subsequent testing under extended Data Link delays. First, although all six subjects agreed that flashing the display when an altitude request message fails would improve detection of a breakdown in the initial contact process, they also noted that making the receiving controller responsible for the progress of a transaction that had been initiated by the automation would result in excessive monitoring demands. More importantly, as noted by four of the six subjects during the design review, and all subjects after delay testing, the tested design for initial contact requires an additional uplink-downlink transaction during the transition between sectors which can significantly increase the time that the receiving controller must wait before beginning to issue necessary messages and control instructions.

In the improved design recommended by the subjects, the assigned altitude will be downlinked to the sending controller from the aircraft as part of the WILCO response to the TOC message. The altitude message will then be sent to the receiving controller over ground channels and displayed as a flashing "IC altitude" in the third line of the data block. Under this design, the receiving controller will not be required to monitor an altitude request message, and will not be forced to wait for a second transaction to be completed after the TOC before sending ATC instructions.

3.1.4 TI Service.

TI messages were sent in the tested system by selecting a message from a menu list of alternatives. A default message located at the top of the list could be sent using simplified inputs, and any TI message could be combined with a MT message in a single uplink.

All of the subjects found the method used for changing the message designated as the default message acceptable. However, two controllers felt that the system for combining TI messages and MT messages should be enhanced by permitting these combined messages to be entered in any order and by allowing the uplink of more than one MT message with a TI message.

Additional TI service design results were associated with the allowable content of the messages and with the general issue of menu item selection. These findings are presented in section 3.1.7.

3.1.5 TOC.

For the present study, the TOC design was enhanced to provide two operational default modes along with the ability to override either default and change to the alternate mode for a single transaction. The default options were an automatic send mode in which the acceptance of a handoff by the receiving controller automatically sends the TOC message, and a "hold mode" in which the handoff acceptance creates the message, but transmission is withheld until the controller makes a manual input. The active default mode was displayed on the first line of the status list.

All six controllers indicated that each of the four TOC options would be required in an operational implementation in order to maximize the use of automation while maintaining the ability to disassociate the transfer of control and the TOC events. All controllers also agreed that the design should be modified by placing the display of the active default mode in some other location than the status list, since this list is expected to receive minimal use and may be suppressed by a majority of controllers. The system data area was suggested as an alternate display location. In addition, the subjects strongly recommended that when the ATC computer is activated, the TOC mode should initialize in the "hold" default rather than the automatic default to insure safety.

Because the frequency change message is automatically created when the handoff is accepted, the controllers indicated that the design should be changed to permit manual deletion of the held message. This change will accommodate the ability to delete the message and send the aircraft to an alternate radio frequency using voice. Finally, the subjects noted that it will be imperative to change the design so that the controller will be able to offer a handoff to a receiving sector while a Data Link transaction is in progress. The modification will give the controller the ability to initiate required handoff actions without waiting for other clearances to be acknowledged. This could be accomplished without requiring two open Data Link transactions with the same aircraft by permitting the handoff to be completed, but delaying the automatic or held TOC uplink until the prior transaction has been closed.

3.1.6 MT (Speed, Heading, Altitude Clearances).

Both the TI service and MT functions rely heavily on the use of predefined lists of message items which can be selected for uplink by designating an item identifier. The results of this study associated with usage of this menu feature are discussed in section 3.1.7. Findings which are particular to the MT function are presented here.

All six controllers agreed that the MT function must retain the three methods for sending multiple clearances in a single uplink. These included the ability to select up to three item identifiers for uplink in a single message, reserving the use of one predefined item for a combined altitude, speed, and heading clearance, and the ability to compose a "shorthand" multiple clearance using the menu by-pass function. While this flexibility was considered desirable to the controllers, all six subjects recommended at least one change which would further increase the availability of options in sending ATC clearances. These ranged from suggestions to add more multiple clearance line items, to completely eliminating a structured menu approach in which various items are reserved for specific clearance categories. In general, the controller comments indicated that the menu system should be designed to be as flexible as possible so that items can be tailored to the needs of specific sectors and ATC situations.

Other recommendations agreed upon by the subjects were to permit the inclusion of more than one menu item in combination with a TI message, and to change the prefix letter input for a speed clearance in the menu by-pass feature from S to V (velocity). This modification would confine all by-pass prefixes A, H, V) to a single column of the ARTS keyboard, potentially enhancing the ease of key selection.

<u>3.1.7 MT and TI Lists</u>.

Possibly the most important findings of the design review and the subsequent debriefings associated with the controller-Data Link interface were needs for refining the list design and improving the speed and accuracy with which controllers will be able to select items from the TI and MT menus. As a result of problems experienced during the high fidelity simulation trials, the controllers agreed upon several changes in menu design and usage procedures.

In the present study, the TI service list was used to include the final approach clearance and other control messages that were not suited to the structured design of the MT list. The controllers who participated in the study found that this lack of functional separation detracted from the distinctiveness of the two menus and increased the workload associated with selecting the desired message. Because of this, it was recommended that items on the TI service list include only those messages that are informational in nature. All maneuvering clearances should be restricted to the MT list.

A second source of problems noted by the controllers was that the two lists used common sets of digits as item identifiers. This was seen as a potential source of error because it was possible to send an erroneous message by typing the wrong single letter prefix prior to the message identifier digit (e.g., typing M3 when T3 was the desired message). For this reason the subjects suggested using exclusive digit sets to denote items on the two lists or using alphabetic characters for one list and digits for the other.

A third human factors design problem identified by the subjects was a spatial display/control incompatibility between key locations and list positions. As noted previously, all menu messages were selected by typing a prefix letter denoting the appropriate list, followed by an identifier digit. In the tested design, the TI list was displayed above the MT list. However, the M is located above the T in the same column on the ARTS keyboard. The recommended solution was to reverse the positions of the two lists on the display to make them spatially compatible with the key positions.

A final controller suggestion for simplifying menu usage and improving message selection accuracy was to modify the lists so that unneeded message lines could be suppressed. This change would reduce visual clutter in menus individually tailored by the controller to require fewer than the total number of lines available.

The controllers indicated that the use of spaces or labels to separate message categories in the lists also should be tested, but felt that extended practice and operational use would help to overcome many of the item selection problems that were encountered during the study.

3.1.8 Data Link Symbology and Key Location.

The final issues considered during the design review were the graphic symbol set used to denote the Data Link status of an aircraft and the location of a key used to preface some Data Link keyboard inputs. As currently designed, the Data Link system uses a graphic character in the first position of the first line of the full data block to indicate the Data Link status of an aircraft to the controller. No symbology in this position indicates that the aircraft is not equipped with a functioning Data Link system. A graphic symbol is used to indicate that the aircraft is Data Link equipped, but is not eligible to communicate with the viewing controller. A second symbol replaces the first to indicate that the aircraft is equipped and is eligible to receive messages from the viewing controller.

The number of unreserved, nonconfusible graphic symbols in the ARTS IIIA system that are available for displaying this information is extremely limited. Controller evaluation of the possible symbols yielded two options that were compared in this study. The first was a plus (+) sign to indicate equipage and an asterisk (*) to indicate eligibility. The second was a diamond (♦) to indicate equipage and a square (■) to indicate eligibility. During the design review, two subjects felt that either set would be acceptable, while the remaining four preferred the +/* set. This symbol set was used during the remainder of the study. However, comments recorded during the full scale simulation phase of the study indicated that some controllers found that identifying the aircraft that were Data Link equipped became a significant task as workload increased. Consequently, it appears that improved and more salient equipage/eligibility display methods should be considered as Data Link is transitioned to the next generation of more capable terminal controller display systems.

A unique keystroke entry is required by the current design to delete Data Link transactions, suppress or retrieve Data Link lists, change the content of menu message items, temporarily change the TOC between the automatic and hold modes, and some other functions. The function keys located at the top of the ARTS keyboard are candidates for this input action. In the present study, the F9 key located to the far left of the keyboard and the F16 key on the right were compared. All six subjects preferred the F9 key.

3.2 FULL SCALE SIMULATION AND DELAY TESTING.

The full scale simulation phase of this study was conducted to meet two objectives. First, these test runs were used as a basis for expert controller judgments regarding the impact of total transaction time on usability, effectiveness, and workload factors. Second, the full scale testing was used as an opportunity to collect and analyze data on a variety of candidate system and controller performance measures which will be considered for application to future operational testing of the mature terminal Data Link service designs. The findings of the controller evaluations are presented in the following subsections, while the results obtained from the candidate performance metrics are presented in section 4 of this report.

3.2.1 Controller Workload.

3.2.1.1 SWAT Scale Generation.

Each subject completed the SWAT scale development task prior to the start of the 12 full scale simulation test runs. The ordered card sorts generated by the individual subjects were used as input to the SWAT scaling algorithms in order to generate the undimensional, interval level workload scale values for each subject that would be used to interpret the ordinal 1 to 3 ratings on the time, effort, and stress rating scales.

The overall level of agreement among subject orderings of the 27 possible combinations of the time, stress, and effort ratings was assessed by computing Kendall's Coefficient of Concordance. The W statistic attained a value of .7718, indicating that only moderate agreement existed among the subjects' internal models of the way in which the three factors contribute to workload. Because of this, the SWAT prototype analysis was performed to identify any common subgroups within the subject set. This analysis indicated that four of the subjects produced sorts consistent with a workload model which has primary weighting on the stress factor, while the remaining two subjects' data were more consistent with a time weighted model. The average Spearman's correlation between the stress prototype subject sorts and ideal stress prototype models was .93, while the average correlation between the time prototype groups scores and the ideal time prototype models was .95, indicating a strong level of agreement within the two subgroups.

Accordingly, two SWAT analyses were performed to generate independent 0 to 100 workload scales for the two subgroups. Axiom tests performed on the average rankings for the subgroups produced a minimum number of violations under the additive model. The final scale values that were generated for transforming the ordinal ratings for the stress group reflected approximate levels of importance of 23.1 percent for time, 16.8 percent for effort, and 60.1 percent for stress. The scale values for the time group reflected approximate levels of importance of 58.9 percent for time, 18.7 percent for effort, and 22.4 percent for stress.

The raw workload ratings collected after the 12 simulation runs were transformed to SWAT scores by referring to the scale appropriate to the prototype subgroup membership of each subject. All statistical analyses were performed on these scores.

3.2.1.2 Effects of Transaction Delay and Sector Type.

The 72 scores generated by the six subjects after each of the 12 test runs were analyzed in a 3 x 4 (sector type communication condition) multivariate analysis of variance with repeated measures on both factors. The results of the analysis revealed significant effects of both communication condition ($F(3,15)= 9.283$, $p=.002$) and of airspace sector ($F(2,10)= 5.103$, $p=.029$). The interaction term was not significant ($F(6,30)=.700$, $p=.55$).

As shown in figure 3, the results indicate that controller workload increased as a function of increasing delay in the Data Link conditions, and that the effect was additive with the inherently different workloads of the arrival, departure, and final approach sectors. Post-hoc comparisons were performed on the overall cell means for the communication condition factor using the Newman-Keuls procedure. The results indicated that test runs conducted under the shortest Data Link turnaround delay range (12-20 seconds) did not produce a significantly higher level of controller workload than those in which voice radio was used exclusively ($p=.094$). However, both the longer 22-30 and 32-40 second delay ranges produced significantly higher workload than that experienced under the voice conditions ($p=.006$; $p=.001$).

3.2.1.3 Effects of Repeated Testing.

As shown in figure 4, the effects of delay on controller workload may have moderated as a function of repeated testing. Since subjects participated at a given sector only once under each communication condition, meaningful statistical tests of this apparent reduction in workload with practice at varying delay times could not be performed. However, except for the longest (32-40 seconds) transaction delay, examination of the SWAT scores produced under the three repetitions of the three delay conditions indicates a general downward trend in workload as the subjects received more experience with the airspace, the Data Link services, and the potential delays associated with Data Link communications.

3.2.1.4 Controller Comments.

Written comments solicited from the subjects by instructions on the SWAT rating forms offered some additional insights to the factors which influenced perceived workload. A total of 29 statements regarding the sources of rated workload were obtained. Of these, 2 were associated with the three voice-only test runs, 10 with the short delay Data Link runs, 9 with the moderate delay runs, and 8 with the long delay runs. Four of the comments were related to workload experienced at an arrival sector, 12 at a departure sector, and 13 at a final approach sector.

The single factor cited as influencing rated workload most often in the comments was the Data Link turn around delay. Of the 11 comments noting a negative effect of delay, 6 were associated with the final approach sectors, 4 with the departure sectors, and 2 with the arrival sectors. Five of these were recorded after runs at the longest turn around delay, 2 at the moderate delay, and 1 at the shortest delay. Delay effects reported for the final approach sectors emphasized the reduction in time available to maneuver aircraft onto the final approach course when the sequence of TOC and initial contact transactions initiated by the arrival sector were excessively delayed.

Other defined sources of workload increases included low levels of skill and errors committed when using the Data Link procedures (5 comments), unfamiliarity with the airspace (1), and problems with the traffic in the test scenario or with the ARTS input devices (3). Finally, 9 citings of simulator pilot errors were noted as workload problems. Six of these occurred in final approach sectors where the pilots failed to carry out a successfully received and acknowledged Data Link clearance or were not intercepting the localizer. In all cases, subjects reported having to revert to voice procedures to deal with the errors under the time critical conditions of final approach.

Overall, 16 of the comments were relevant to the testing conditions, while the remaining negative effects on workload appeared to be attributable to factors unassociated with Data Link usage or Data Link delays.

3.2.2 Controller Perceptions of Delay Effects.

In addition to the SWAT workload rating, subjects completed a second data form following each of the nine Data Link test trials. The purpose of the evaluation instrument was to elicit impressions of the nature and extent of effects produced by the transaction delays experienced during the preceding trial. The five questions that were presented are summarized below and a sample data form is contained in appendix C.

1. To what extent did the Data Link turn-around delays that you experienced during this run impair your ability to control traffic?

Scale:

"1" - (Delays had no effect on ATC abilities)
to
"7" - (Extremely impaired ATC abilities)

2. How did the Data Link turn-around delays that you experienced during this run affect your use of the Data Link services?

Scale:

"1" - (No delay effect -- Used Data Link at every opportunity)
to
"7" - (Delays prevented any use of Data Link)

3. How often did you use voice instead of Data Link during this run in order to avoid turn-around delays?

Scale:

"1" - (Never used voice when I could use Data Link)
to
"7" - (Had to switch to voice all the time)

4. In what ways did turn-around delays affect you or the way in which you controlled traffic during this run?

Check all that apply:

Data Link delays did not affect me in any way.

I compensated by sending Data Link messages and clearances sooner than I would normally.

I used voice instead of Data Link.

I "got behind the power curve" and found it difficult to keep up with the air traffic.

I increased aircraft separation.

I made minor control errors.

I made major control errors.

I sometimes forgot (temporarily) about an ongoing transaction or its status.

Because of the length of delays, I didn't use Data Link for (when): (Describe services and situations.)

Others effects of the delays (Describe):

5. Overall, if average Data Link turn-around delays in my facility were as long as those I experienced during this test run (and in this type of scenario and control position):

Scale:

"1" - (I would not use Data Link at all)
to
"4" - (I would use Data Link at every possible opportunity)

Questions 1, 2, 3, and 5 utilized rating scales which provided an ordinal scale of measurement. Because of this, statistical tests on differences between the three Data Link delay conditions were performed using distribution-free methods. Question 4 was analyzed by compiling frequency counts of reported occurrences of the listed effects of transaction delays.

3.2.2.1 Impairment of ATC Abilities, Ability to Use Data Link, and Use of Voice to Avoid Delays.

The first three questions produced a common pattern of findings. As shown in figures 5, 6,and 7, median ratings indicate that, in comparison to the longer transaction time conditions, when operating in a shorter delay environment, traffic control abilities were perceived as less impaired, subjects felt they were more able to use the Data Link services, and voice was used less often to avoid the consequences of delay. The ratings obtained for each of the three questions were similarly affected by the type of airspace sector controlled during the test run, with the arrival sectors producing lower (less "negative") ratings than the final sectors. The Friedman Two-Way Analysis of Variance (ANOVA) by ranks was used to test for the significance of the delay effect across sectors for each question. This analysis identified a significant impact of transaction time for all three questions.

Question 1: Chi2=10.33, df=2, p<.005
Question 2: Chi2=6.583, p<.037
Question 3: Chi2=7.75, p<.020)

Subsequent Wilcoxon Matched Pairs tests were used to perform individual comparisons among the delay conditions. In all three cases, these tests indicated that the 12-20 second delay ratings were significantly lower than either the 22-30 second or the 32-40 second test run ratings, and that the two longer delay conditions did not differ.

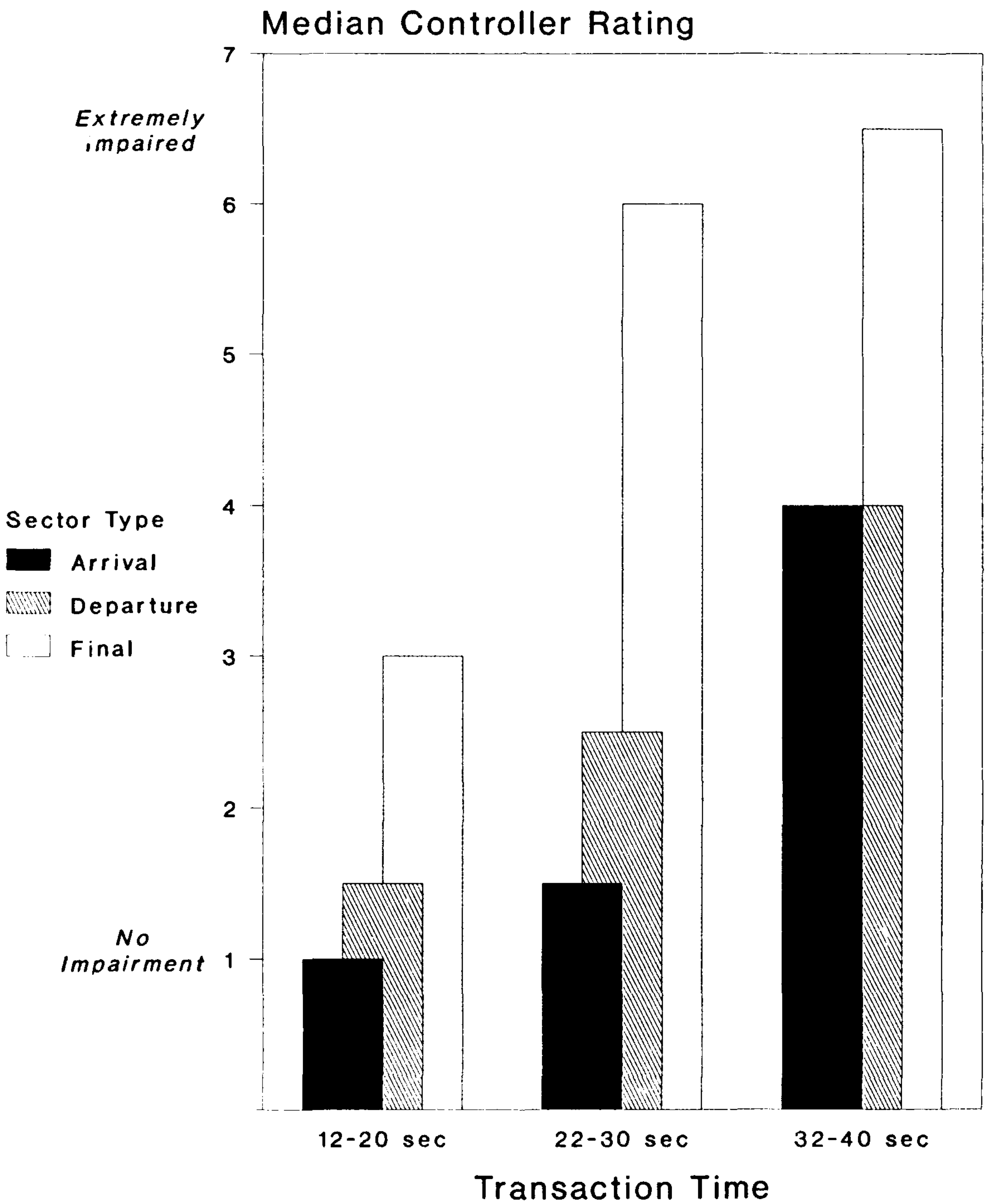


FIGURE 5. RATED IMPAIRMENT OF TRAFFIC CONTROL ABILITY ATTRIBUTED TO DATA LINK DELAYS

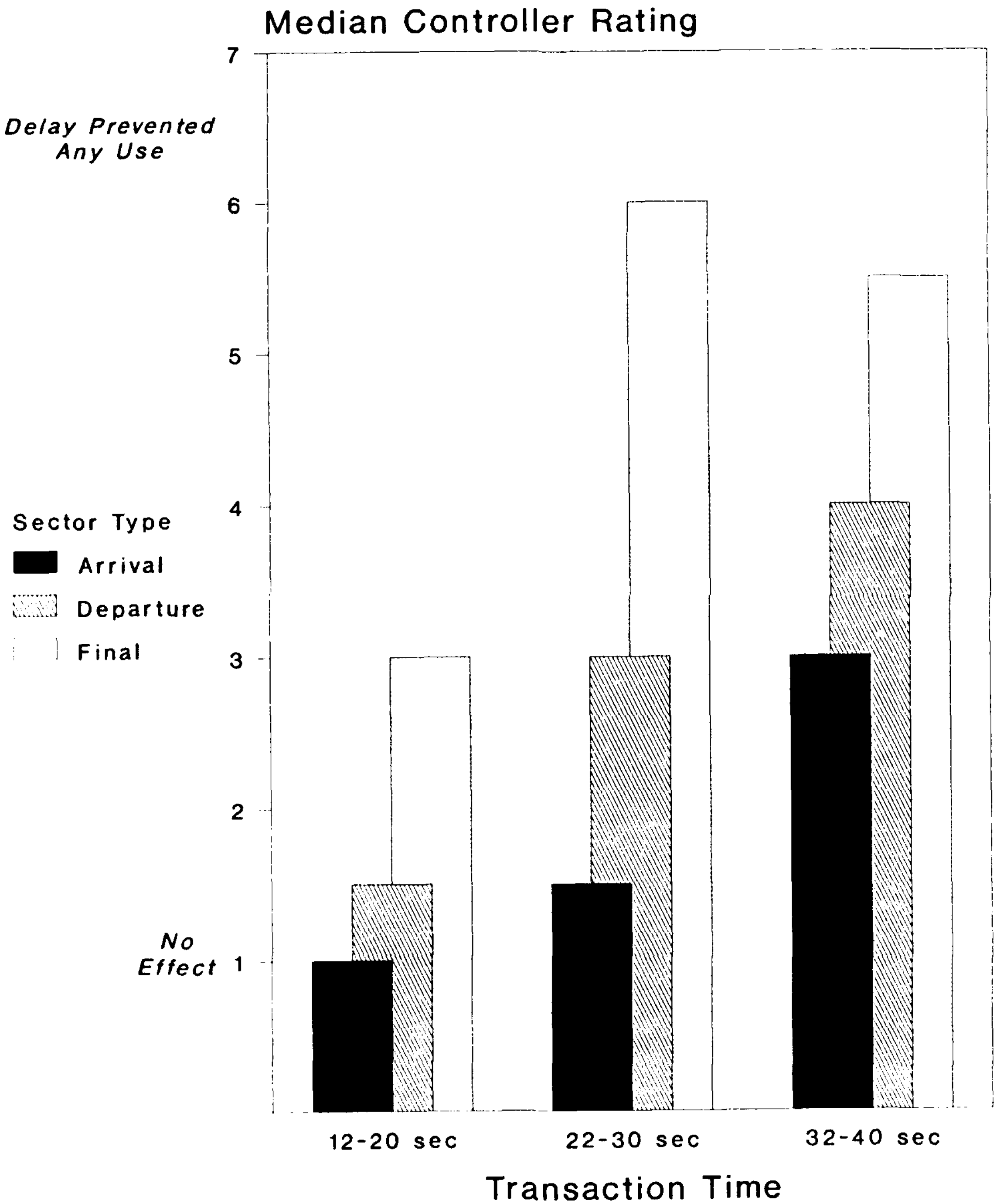


FIGURE 6. RATED EFFECT OF DELAYS ON ABILITY TO USE DATA LINK SERVICES

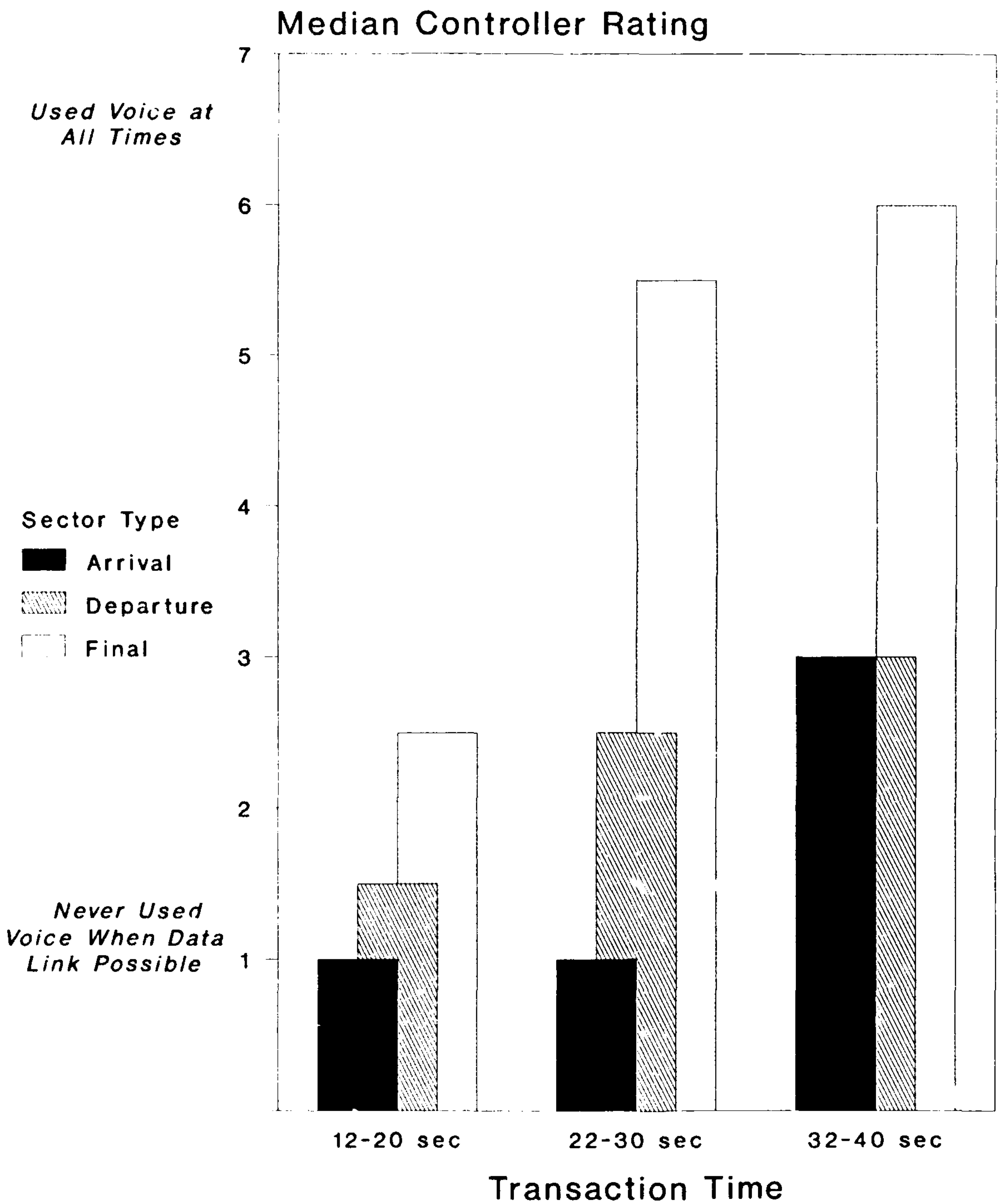


FIGURE 7. RATED FREQUENCY OF USE OF VOICE TO AVOID DATA LINK DELAYS

Wilcoxon Tests

Question 1	z value	p
12-20 vs 22-30	2.201	.027
12-20 vs 32-40	2.201	.027
22-30 vs 32-40	1.467	.142
Question 2		
12-20 vs 22-30	2.201	.027
12-20 vs 32-40	1.991	.046
22-30 vs 32-40	.943	.345
Question 3		
12-20 vs 22-30	2.201	.027
12-20 vs 32-40	2.020	.043
22-30 vs 32-40	1.048	.294

3.2.2.2 Effects of Delay and Adaptive Strategies.

Frequency counts for the eight potential effects of delay reported by subjects under question 4 are shown in table 2. These are presented both as a function of Data Link delay time and airspace sector. In the following discussion of these results it should be noted that, for any of the eight categories of effects, each of the six subjects had nine opportunities to indicate whether it had been experienced during the test trials. Thus, cited percentages are based on the observed frequencies out of a potential total of 56 reports for each listed effect.

The most commonly reported effect of delay was "compensating by sending messages earlier." Of the total 54 opportunities, 67 percent (36) indicated that this strategy had been adopted. The strategy was most common at the two longest delay times (13 and 15 vs 8), and in the arrival and departure sectors rather than the final sectors (13 and 14 vs 9).

The use of voice radio as a means to cope with delays was reported on 38 percent (21) of the opportunities. This was most commonly reported at the two longest delay times (9 and 8 vs 4), and most often in the departure sectors.

Fewer reports of increasing aircraft separation as a response to delay were received. This strategy was reported on only 13 percent

TABLE 2. QUESTION 4 CHECKLIST DATA

Question 4: In what ways did Data Link delays affect you or the way in which you controlled traffic?

Delays had no effect:

Sector Type		Mean Delay 16 sec	26 sec	36 sec	
	Arr.	3	1	1	5
	Dep.	2	0	0	2
	Fin.	3	0	0	3
		8	1	1	

I compensated by sending messages earlier:

Sector Type		Mean Delay 16 sec	26 sec	36 sec	
	Arr.	3	5	5	13
	Dep.	4	6	4	14
	Fin.	1	4	4	9
		8	15	13	

TABLE 2. QUESTION 4 CHECKLIST DATA (CONTINUED)

Question 4: In what ways did Data Link delays affect you or the way in which you controlled traffic?

I used voice radio:

Sector Type		Mean Delay 16 sec	26 sec	36 sec	
	Arr.	1	2	3	6
	Dep.	3	4	4	11
	Fin.	0	2	2	4
		4	8	9	

I found it difficult to keep up with the traffic:

Sector Type		Mean Delay 16 sec	26 sec	36 sec	
	Arr.	0	0	0	0
	Dep.	0	0	0	0
	Fin.	0	1	0	0
		0	1	0	

TABLE 2. QUESTION 4 CHECKLIST DATA (CONTINUED)

Question 4: In what ways did Data Link delays affect you or the way in which you controlled traffic?

I increased aircraft separation:

Sector Type		Mean Delay 16 sec	26 sec	36 sec	
	Arr.	0	2	1	3
	Dep.	0	1	1	2
	Fin.	0	1	1	2
		0	4	3	

I made minor errors:

Sector Type		Mean Delay 16 sec	26 sec	36 sec	
	Arr.	0	1	2	3
	Dep.	2	1	0	3
	Fin.	0	1	1	2
		2	3	3	

TABLE 2. QUESTION 4 CHECKLIST DATA (CONTINUED)

Question 4: In what ways did Data Link delays affect you or the way in which you controlled traffic?

I made major errors:

Sector Type		Mean Delay 16 sec	26 sec	36 sec	
	Arr.	0	0	0	0
	Dep.	1	1	1	3
	Fin.	0	0	0	0
		1	1	1	

I sometimes forgot about an ongoing transaction:

Sector Type		Mean Delay 16 sec	26 sec	36 sec	
	Arr.	0	0	0	0
	Dep.	0	1	0	1
	Fin.	0	1	2	3
		0	2	2	

(7) of the opportunities, and reports were confined to the two longest delay times.

On 18 percent (10) of the opportunities, the subjects reported that delays had no effect. These were reports were concentrated at the lowest transaction delay (8), and were most common in the arrival sector (5).

On 8 of the reporting opportunities, subjects recognized that they made minor control errors because of delay, while 3 reports of major errors due to delay were received. The three major errors reported were confined to departure sectors, while minor errors were distributed across delays and sectors.

No reports of having difficulty keeping up with the task of controlling the air traffic were received, and only four reports of forgetting about the status of ongoing Data Link transactions were recorded.

Two open-ended categories of effects included in question 4 permitted the subjects to indicate what specific tasks were not accomplished using Data Link because of the delays, and to note any other effects of delay not included in the preceding list. Since a total of only 25 reports were received under these two categories, the results were combined for the following discussion.

When working the arrival sectors, the controllers wrote five reports. These included the decision not to use Data Link for one altitude clearance when a potential separation problem arose for the final descent out of the sector, and when a long duration ongoing transaction prevented the controller from sending a subsequent clearance when needed. One additional report noted that an unlisted effect of Data Link was a tendency not to attend to the flight progress strip marking task. Two of the reports were obtained at the shortest delay, one at the moderate delay, and three at the longest delay.

Six reports were received from controllers when working departure sectors. These included not using Data Link to clear up the consequences of a pilot error, problems in remembering how to use the multiple message MT feature, abandoning Data Link under any condition that was not routine, and two reports of switching to voice when an ongoing transaction prevented the controller from issuing a needed clearance. The sixth report indicated that an unlisted delay effect was a perception that turns and climbs occurred later than anticipated. Each delay condition produced two of the reports.

The final approach sector conditions produced 14 reports. Of these, 8 were indications that because of delays Data Link was not used for some or all turns onto the final approach course, and 2 indicated that delays prevented the use of Data Link for any control instructions. Three reports noted the inability to use Data Link when needed during lengthy transaction delays. These cited the problem of significant flight progress within the sector before the delayed Data Link TOC and initial contact were completed. Three of the reports from the final sectors were obtained at the shortest transaction delay, 7 at the moderate delay, and 4 at the long delay.

3.2.2.3 Projected Use of Data Link in the Operational Environment.

The final item on the delay evaluation form asked the subjects to make an expert projection of the extent to which Data Link would be used at their own facility at sectors and delays similar to that experienced during the prior test run. The results obtained from the ratings on the 4-point scale are shown in figure 8. As the median ratings indicate, predicted Data Link usage would depend both on the type of sector and on average turnaround times. In the best case, all six subjects felt that Data Link would be used at every opportunity or slightly less often than if the delays were shorter in arrival sectors with transaction times in the 12 to 20 second range. At the other extreme, in final approach sectors at a 32 to 40 second delay, four subjects indicated that Data Link would be not be used at all, or much less often than if the delays were shorter.

Statistical testing using the Friedman Two-Way ANOVA confirmed the observation that transaction time significantly affected projections of Data Link usage in the operational environment (Chi2=10.33, $p<.005$). Individual tests on pairs of delay conditions using the Wilcoxon procedure showed that, across sectors, the 12 - 20 second delay produced significantly higher projections for Data Link use than the 22 - 30 or the 32 - 40 second delays ($z=2.201$, $p<.027$ for both tests). The two longest delay conditions did not yield significantly different projections for field use ($z=1.825$, $p<.068$).

3.2.2.4 Composite Controller Problem Index.

A primary purpose of the full scale simulation testing conducted during this study was to obtain a preliminary estimate of the impact of Data Link transaction delays in differing terminal ATC sectors. In order to gain a single overall impression of the way in which the subject controllers perceived this effect, a composite controller problem index was developed based on a summary of the detailed findings discussed above.

The composite problem index was derived from the results of 12 of the post-run controller evaluations discussed above. These are expressed below as 12 problem factors that could have been experienced under each of the nine combinations of transaction delay and sector type:

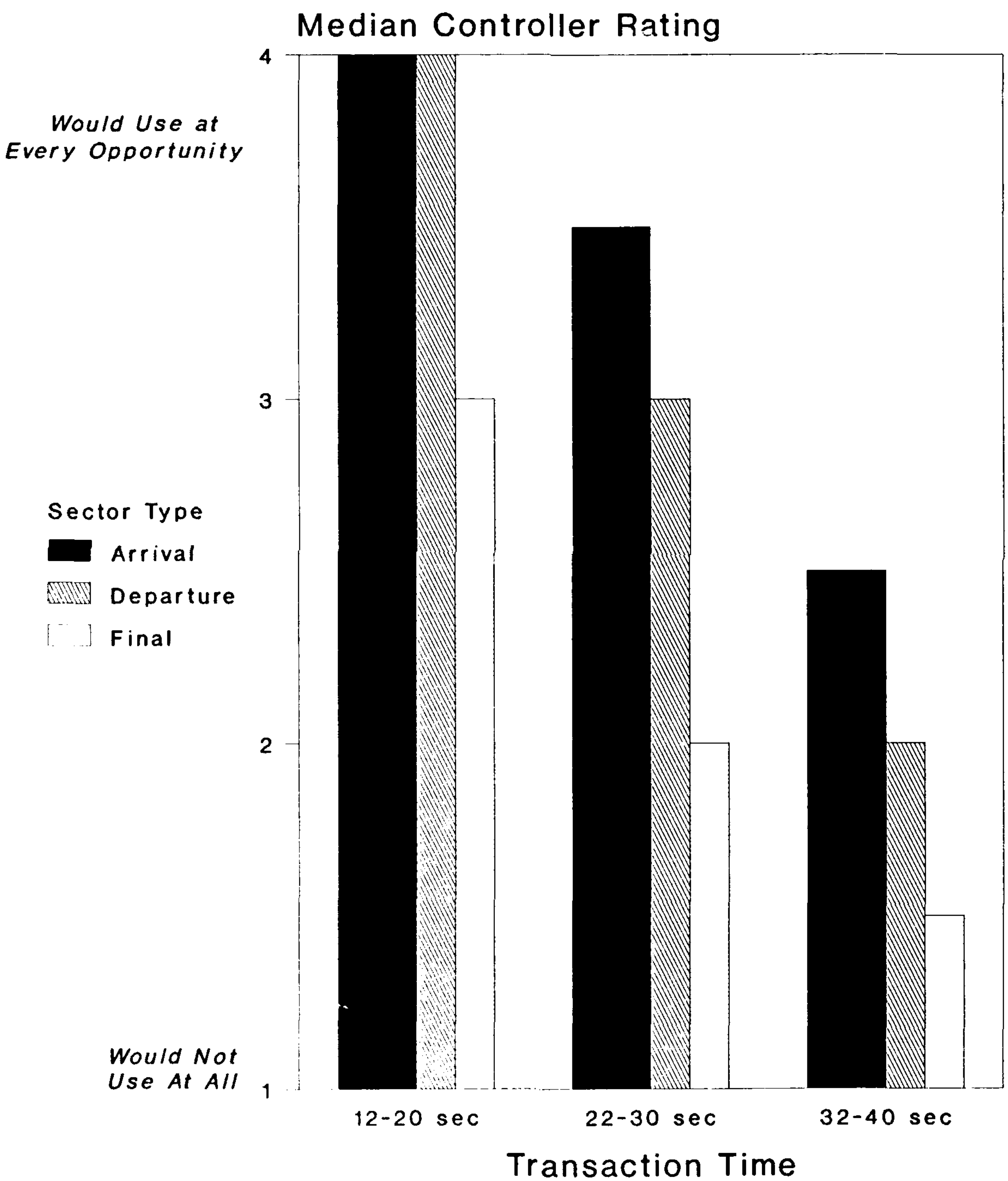


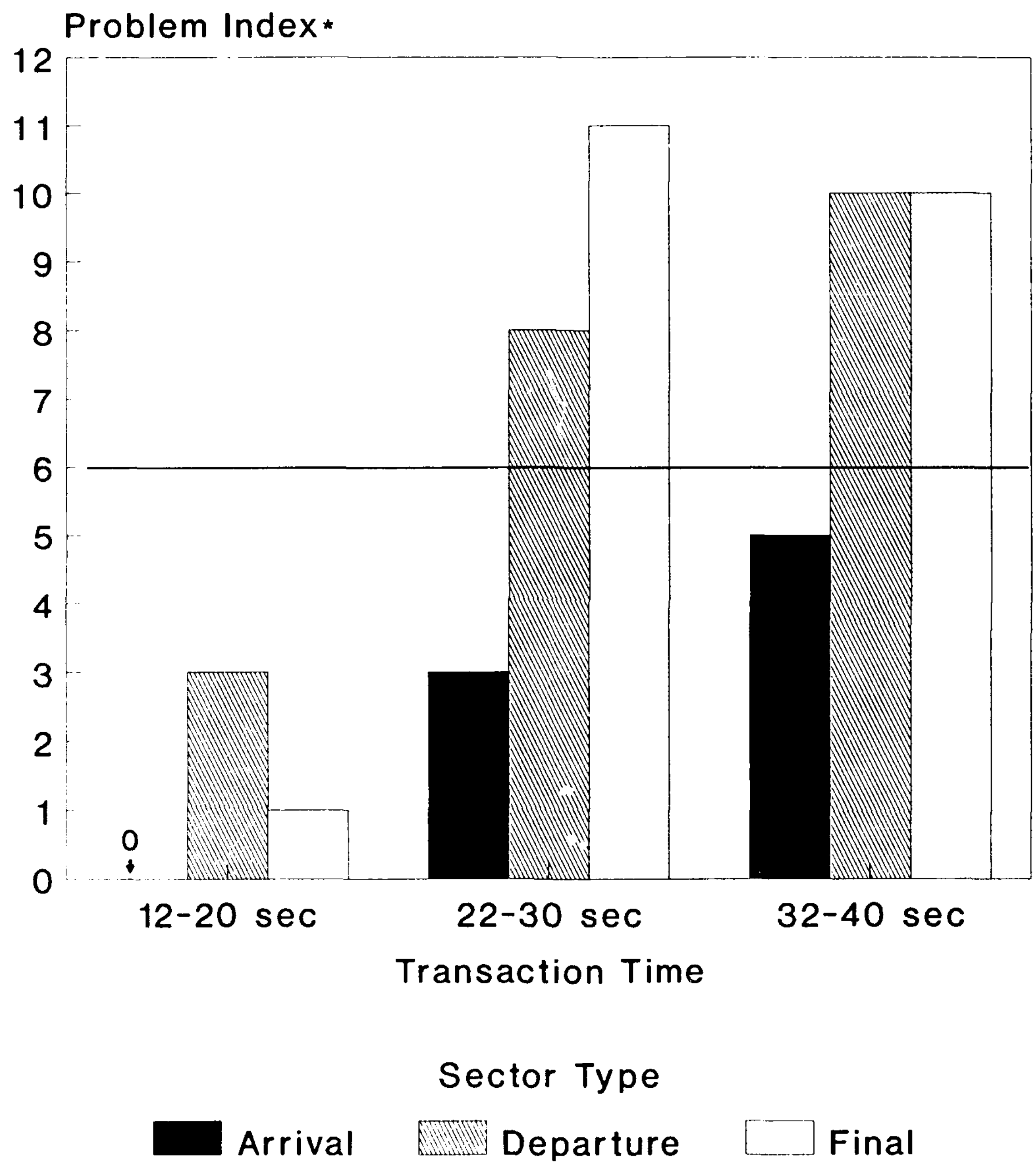
FIGURE 8. PROJECTED USE OF DATA LINK AT HOME FACILITY SECTORS

1. High Workload
2. Impaired ATC Abilities
3. High Voice Radio Usage
4. Low Predicted Data Link Usage in Field Settings
5. Negative Delay Effects
6. Need to Send Messages Early
7. Difficulty in Keeping Up with Air Traffic
8. Need to Increase Separation
9. Minor Control Errors
10. Major Control Errors
11. Memory Problems
12. Inability to Use Data Link

For each factor, the overall scale range used by the subjects was determined. The ranges were then split at the midpoint, and these values were used as judgement criteria. For each factor, the mean, median, or frequency count within each of the nine combinations of sector type and delay duration was compared to the criterion. If the value in a cell exceeded the midpoint criterion, one problem index point was accumulated for the associated sector/delay condition. The composite index score for each condition was computed as a simple sum of the problem index points received. This produced a scale ranging from 0 (no problems) to 12 (high problem frequency).

The composite controller problem index scores derived in this manner are presented in figure 9. When interpreting the data produced by this simplified combinatorial method it should be noted that the problem factors upon which the index was based are probably nonindependent, and that, although they may differ in importance, each factor was given equal weighting. Nevertheless, inspection of the results offers what appears to be a reasonable summary of the accumulated findings of this study derived from expert controller perceptions on multiple dimensions.

If an arbitrary score of greater than six points is assumed as a criterion for significance, figure 9 clearly indicates that more problems with the usability of Data Link and the effectiveness of the controller's performance emerged when total transaction times were equal to or greater than 22 seconds in the departure and final approach sectors. Conversely, long transaction delays appear to



*Count of above scale midpoint means, medians or frequencies on 12 evaluation factors

FIGURE 9. COMPOSITE CONTROLLER PROBLEM INDEX

have been more tolerable in the arrival sector, and all sectors produced low composite problem scores when delays were confined to a range of 12 to 20 seconds.

3.2.3 Projected Utility of Tested Data Link Service Designs.

After completing all 12 simulation runs, the subjects were asked to provide an overall evaluation on two dimensions for each of the Data Link service designs as tested in this study. First, controllers rated the projected operational effectiveness/suitability of each service. This dimension was defined as the extent to which a service design could meet the technical requirements of the ATC function that it was designed to perform. The second dimension evaluated for each service design was controller acceptance and preference. This factor was defined as the extent to which the displays, inputs, and procedures associated with the service design would make the service usable by controllers and compatible with their styles of controlling air traffic. Ratings obtained on both dimensions for each service are presented in figures 10 and 11.

3.2.3.1 Operational Effectiveness and Suitability.

None of the controllers rated any of the four tested service designs as "not operationally suitable." In addition, all ratings of effectiveness for the TI, TOC, and MT services were either at, or above the midpoint of the scale (meets most requirements). A Friedman Two-Way ANOVA revealed no statistically significant differences among the service designs in effectiveness/suitability ratings (Chi2=3.75, $p<.289$).

3.2.3.2 Controller Acceptance and Preference.

As with the effectiveness ratings, none of the controllers judged that any of the four service designs would be "completely unacceptable." However, ratings below the scale midpoint (moderately preferred) and lower median ratings than those assigned on the operational effectiveness scale were obtained for the TI service and MT. The Friedman Two-Way ANOVA corroborated this apparent difference between the ratings on the two dimensions by detecting a significant difference among the services in controller acceptance and preference (Chi2=7.80, $p<.05$). Subsequent Wilcoxon tests performed on service pairs indicated that the tested MT design was significantly less preferred than the designs for the initial contact ($z=2.022$, $p<.043$) and the TOC ($z=2.022$, $p<.043$).

Written comments provided by the subjects, with respect to their effectiveness and preference ratings, are summarized below for each service design.

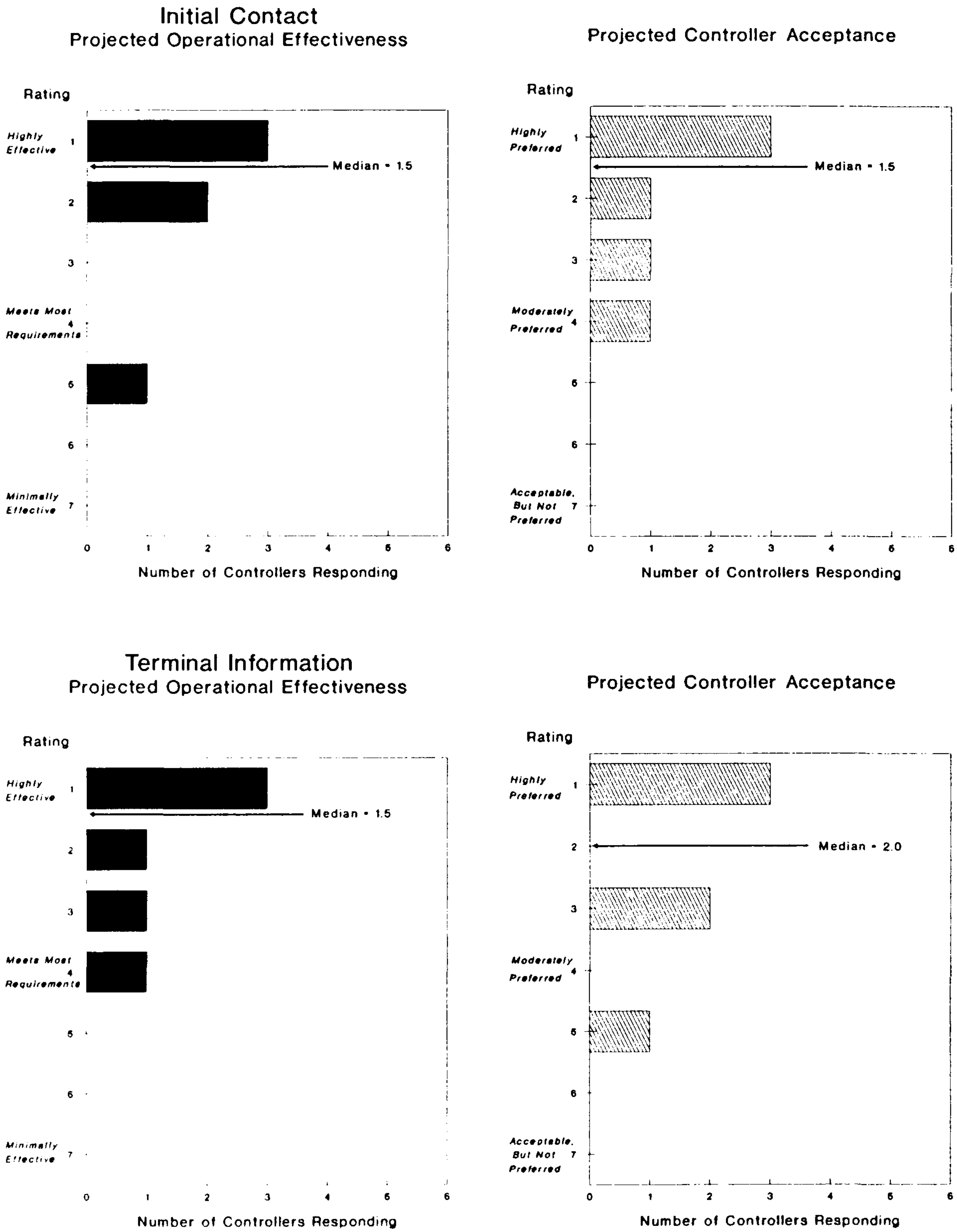


FIGURE 10. EFFECTIVENESS AND ACCEPTANCE - INITIAL CONTACT AND TERMINAL INFORMATION SERVICES

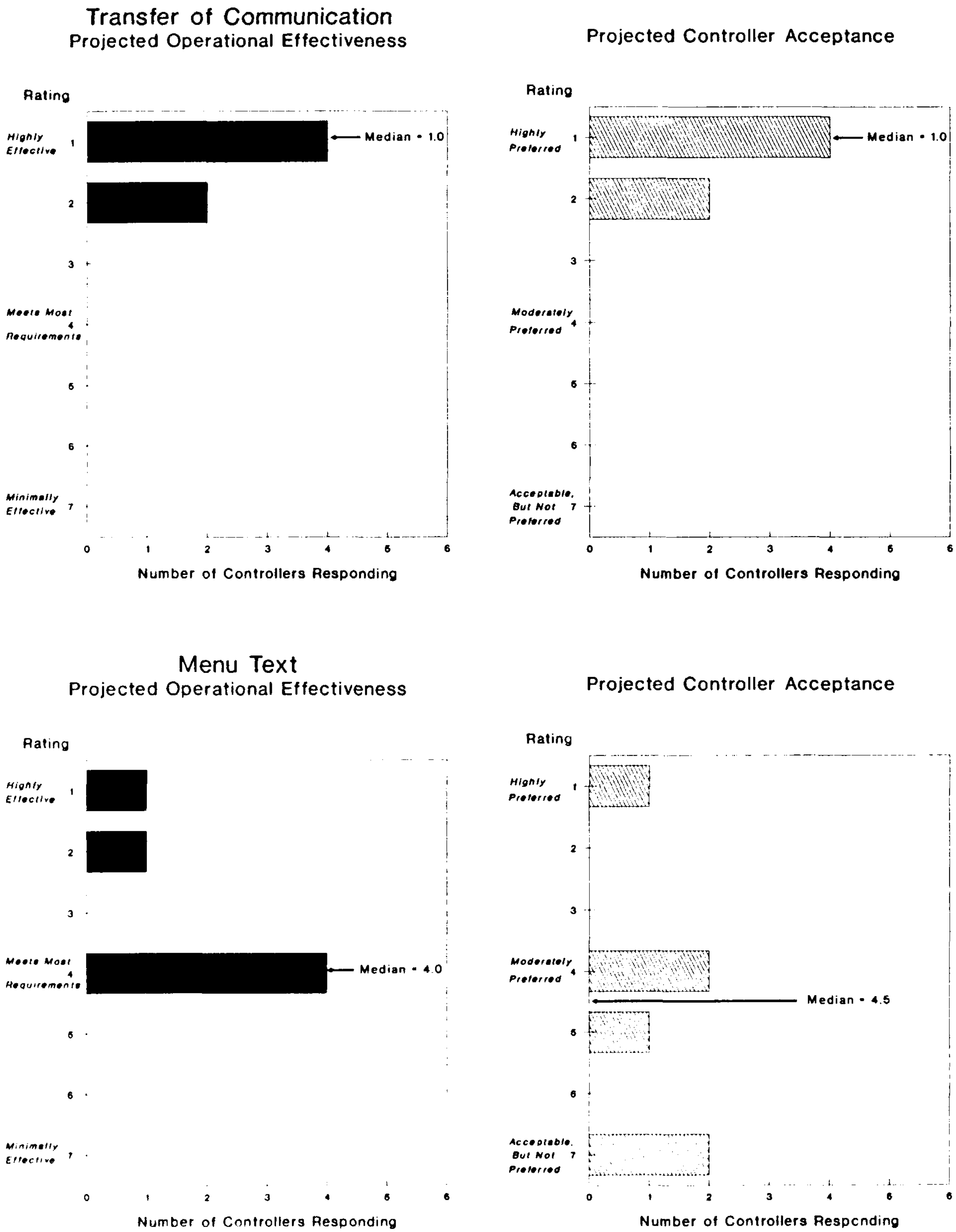


FIGURE 11. EFFECTIVENESS AND ACCEPTANCE - TRANSFER OF COMMUNICATION AND MENU TEXT SERVICES

The single below scale midpoint value (5) for the effectiveness of the initial contact service was explained by the rater as a problem with the time taken to send the altitude request and receive the downlinked initial contact message. Three comments on the effectiveness of the TI service design indicated that the maximum permissible message length was too short to accommodate the lengthy messages for which the service is considered most effective.

TOC, the service design with the best median ratings on effectiveness and acceptance, received one comment indicating that it would be used today if available.

MT elicited five comments from the subjects. Two of these referred to limitations in the design of the service noting that (1) it should be possible to reduce the number of items in the menu if a shorter list is desired, and (2) that the design imposes limits on the order in which combined clearances can be entered and sent. Other comments indicated that the effort required to send control clearances via Data Link was too high to be acceptable; that control clearances should be restricted to one menu; and informational messages to another in order to avoid confusion, and that the value of the service will depend on developing brief and effective menus that can be rapidly memorized by the controller.

3.3 DEBRIEFING.

The final day of this study was devoted to a debriefing and discussion session which covered the topics considered in the following four subsections.

3.3.1 Data Link and Airspace Training.

A secondary objective of this study was to assess the quality of airspace and Data Link training that was provided and to generate improvements in the training protocol that will be needed for future operational evaluation research. Training levels directly affect the types of conclusions that can be drawn from the data that are collected in manned simulation research. Because of this, increasingly thorough amounts of subject training and higher levels of acquired expertise are required as research progresses from preliminary Data Link ATC service design evaluation studies through intermediate design studies such as the one reported here and, finally, to operational evaluation testing. At the operational evaluation stage, it must be possible to demonstrate the effects of Data Link services on system level performance factors such as safety, efficiency, and capacity. For such measurements to be valid, the level of simulation fidelity must be extremely high, and subject controllers should be as familiar with the airspace procedures and with the Data Link services as they might be if they were working in their home facility using communication methods with which they have had the opportunity to develop a high degree of proficiency.

During the present study, several indicators suggested that improvements in training will be required for operational evaluation testing. For example, as testing proceeded over the 12 full scale simulation trials, controller workload decreased and the negative effects of transaction delay on Data Link usage appeared to diminish. Controller comments recorded during the testing also suggested that the practice effect was quite strong. Improvements were apparent in the facility with which the controllers employed the Data Link services and in their familiarity with the air traffic procedures. In addition, the controllers appeared to develop strategies for using Data Link even under extreme transaction delays. These included improved timing in issuing Data Link clearances, and better recognition of situations which called for immediate use of voice radio rather than Data Link. Further evidence of the development of enhanced strategies was provided by numerous observations of the controllers modifying their MT lists to include clearances that were tailored to their improving capabilities to use Data Link effectively.

Taken together, these findings suggest that for operational evaluation studies, the subjects should be given much more extensive practice with both the test airspace and the use of Data Link before data collection begins. While additional classroom, separate airspace, and Data Link procedural training may not be required, practice time under full fidelity simulation conditions probably should be increased to at least twice the amount of time devoted to full scale testing during the present study (16 to 24 hours). Furthermore, formal testing of airspace familiarity and knowledge of Data Link inputs should be included in the protocol.

Beyond the general finding discussed above, the controllers suggested a number of additional changes to improve the training protocol and materials. With respect to the test airspace, it was recommended that the airspace maps and written procedures be provided to the subjects 1 week prior to arriving for training and testing. It was also suggested that the overhead area on the ARTS workstations be used during training and testing to display maps and procedural information in lieu of the booklets provided during this study.

The controllers also indicated that Data Link training would be enhanced by reserving sufficient time prior to training for the subjects to study functional descriptions of each of the Data Link services. Descriptions similar to those used during the design review portion of the present study were recommended (see appendix C). Finally, the subjects felt that a quick reference card summarizing Data Link input procedures would be useful during training and testing at the ARTS workstations.

3.3.2 Data Link Usability in Terminal Airspace Sectors.

A second topic which was discussed during the debriefing was the controllers' perceptions of the degree to which Data Link services would be usable in different sectors of terminal airspace. The consensus opinion of the subjects was that Data Link usage could not be mandated or precluded simply by considering the general type of sector involved. Instead, they indicated that the degree of use would be driven by the time criticality of specific ATC situations and by the duration of Data Link total transaction times. For example, it was noted that while using Data Link to issue tactical speed, heading, and altitude clearances in the final approach sector at a very busy airport may not be feasible, such actions may be common at a low traffic airport or where the approach segment can be simplified to the point that the landing can be directed by a single message sent as the aircraft enters the final approach airspace. Likewise, whereas Data Link was seen as extremely useful for issuing clearances to departing aircraft on particular headings in the present study, other departures which required complex turn and climb instructions immediately after takeoff to achieve a restricted altitude window were nearly impossible with long transaction delays.

In general, the subjects argued that the communication roles played by Data Link will have to be adapted to specific airspace and procedural requirements, and that transaction times will be a major determinant of the level of usage that Data Link receives.

One specific Data Link design feature that was discussed in the context of this issue was the tested method for accomplishing the TOC and initial contact sequence. The controllers indicated that the utility of these services would be compromised under all but the shortest transaction delay conditions because two uplink-downlink sequences were required before the receiving controller could communicate with an aircraft. For this reason, the subjects placed a high priority on integrating the initial contact with the TOC as described in section 3.1.3 of these results.

3.3.3 Potential Causes of Controller Error.

During the debriefing the subjects also were asked to reflect on their experiences under full scale simulation conditions, and to attempt to identify the types of errors that might be made by controllers when employing Data Link. Prior terminal Data Link studies produced anecdotal data which suggested that the controller's short term memory for messages issued via Data Link may be poorer than memory for voiced messages. The controllers who participated in the present study also noted this difference, and suggested that part of the problem may be associated with reduced attention to maintaining flight strips due to the increased keyboard activity associated with Data Link. However, the subjects also indicated that the history list feature added to the design for the present study offered a viable solution to occasional memory lapses, and was used effectively during the simulation tests.

A second type of error discussed was the inadvertent attempt to send Data Link messages to nonequipped aircraft. The controllers indicated that determining the Data Link eligibility of an aircraft was a nontrivial task during full scale simulation and that this additional effort may detract from concurrent visual scanning and planning tasks. As discussed in section 3.1.8, improved equipage/eligibility displays may have to be deferred until the advanced capabilities of future terminal controller displays become available. However, because visual pattern recognition tends to become highly automatic with experience, it is also likely that increased practice with the current symbology may reduce the attentional demands of determining the Data Link status of an aircraft track.

A final error category discussed by the subjects was inappropriate data entries on the ARTS keyboard. The present service designs use abbreviated keystroke sequences to designate messages contained on the MT and TI lists. While this is a desirable feature because it limits the manual task demand associated with sending Data Link messages, it is also problematic. Highly abstracted and abbreviated inputs are inherently subject to error because of low redundancy in the message string and the limited amount of feedback provided to the controller during the data entry task. The subjects made several suggestions for reducing this type of error (see section 3.1.7). In addition, because the results of an erroneous uplink may not be directly apparent until the aircraft which receives the message begins to execute an action not intended by the controller, it was suggested during the discussion that the transaction status information presented in the full data block be improved to provide less abstract feedback regarding the content of the message sent by the controller. For example, the current design presents MT7 in the data block to indicate that an altitude clearance contained in item 7 of the menu had been sent. Confirmation that this was the attended message requires the controller to consult his memory for the contents of item 7, or to check the menu displayed on the screen. Based on the discussions conducted during the debriefing, it may improve self-detection of input errors if the third line presented more direct feedback about clearance content (e.g., DTAM 9000).

3.3.4 Impact of Data Link on Terminal ATC.

Finally, in light of the testing experience, the controllers were asked to express summary opinions concerning the impact that Data Link services can be expected to have on the overall operation of the ATC system. The controllers reiterated several proximal effects of Data Link which they felt would benefit the system. In general, the subjects agreed that the quality and effectiveness of

ATC communications would be improved significantly. Increases in system safety would be achieved through fewer communications errors and decreased frequency congestion. The controllers also indicated that the ability to automate lengthy and repetitive advisory messages and to reduce the demand associated with TOC and initial contact would be reflected in improvements in system efficiency. Efficiency was also expected to increase in situations where the controller will be able to use voice and Data Link in combination to accomplish more tasks in a limited period of time. Although they noted that system capacity was ultimately limited by the maximum arrival acceptance rate at an airport, they also indicated that the addition of the Data Link channel is likely to increase the number of opportunities in which current effective capacity limits can be overcome by eliminating delays attributable to the single channel communications capability of voice radio.

4. EVALUATION OF EXPERIMENTAL PERFORMANCE MEASURES.

4.1 PERFORMANCE RESULTS.

The results reported in this document are to be considered as tentative, given the experimental nature of the measures involved. When, after a careful analysis of the results, a chosen measure was not considered reliable, only general results are reported. A number of measures will require further development during future Data Link studies in order to improve their usefulness and accuracy.

In each analysis, the effect of side of airspace was tested and found not to be significant. The Huynh-Feldt correction to the degrees of freedom for the within-subjects Multivariate Analysis of Variance (MANOVA) was applied when violations of the assumption of sphericity were found.

4.1.1 ATC System Safety.

The primary measures of safety in Mini Study 2 were the number and severity of conflicts. Both the NSSF and ARTS systems tracked conflicts using similar, but not identical approaches. Due to the effort required to manipulate and reduce the ARTS data, the NSSF reports were relied upon for the conflict analysis. The ARTS conflict reports were used as a cross reference.

There were differences in the frequency of conflicts over the four communication conditions (voice only and the three Data Link delays). The number of conflicts while aircraft were exclusively under voice control was nearly identical to the number in the shortest Data Link condition. Number of conflicts increased as Data Link delays became longer. However, further study of the conflict measure is needed to determine that all of the conflicts detected were valid. A qualitative analysis of the most severe conflicts would then be required to determine whether the use of Data Link was a causative factor.

ANOVA's would normally be applied in this situation to assess whether differences existed in conflict severity between the four communication conditions. Measures to test would include the duration of conflict, API, and CPA values calculated by the NSSF computer for each conflict. However, the duration, API, and CPA variables were markedly skewed and this ruled out the application of parametric tests such as ANOVA. Alternately, nonparametric tests could be used to test the results for significance. In any case, the conflict severity variables could only be employed if it was initially determined that the conflicts detected by the NSSF were valid.

4.1.2 ATC System Efficiency and Capacity.

4.1.2.1 Flight Time.

The NSSF system measured the amount of time an aircraft spent in each sector (between controller acceptance and handoff). If Data Link with voice supported more efficient management of aircraft, some benefit might be detected by considering the time aircraft spent in each sector.

Flight times were categorized by communication condition and position type. An MANOVA was used to test the one between-subjects variables and two within-subjects variables design and showed that the main effect of communication type was not significant ($F(3, 12) = 1.66$, $p = .247$), but that the effect of position ($F(2, 8) = 49.48$, $p = .000$) was significant. The interaction of the two factors was also significant ($F(6, 24) = 4.15$, $p = .005$). The cell means are shown in figure 12.

Although figure 12 suggests that communication condition and position jointly affected flight times, the significant interaction of communication type and position in the above statistical test should be ignored. It is most likely an artifact of the test procedure. However, when the departure sectors were tested alone, a significant effect of communication condition emerged. MANOVA results were ($F(3, 12) = 4.73$, $p = .021$). It is legitimate to analyze each type of airspace independently, given that different functions and kinds of activities were performed in each.

4.1.2.2 Distance Flown.

Average distance flown in each sector was also evaluated as a measure of efficiency. Distances were analyzed by communication condition and position type. An MANOVA showed no significant main effect of communication type ($F(3, 12) = 2.89$, $p = .079$) and a significant effect of position ($F(2, 8) = 51.87$, $p = .000$). The

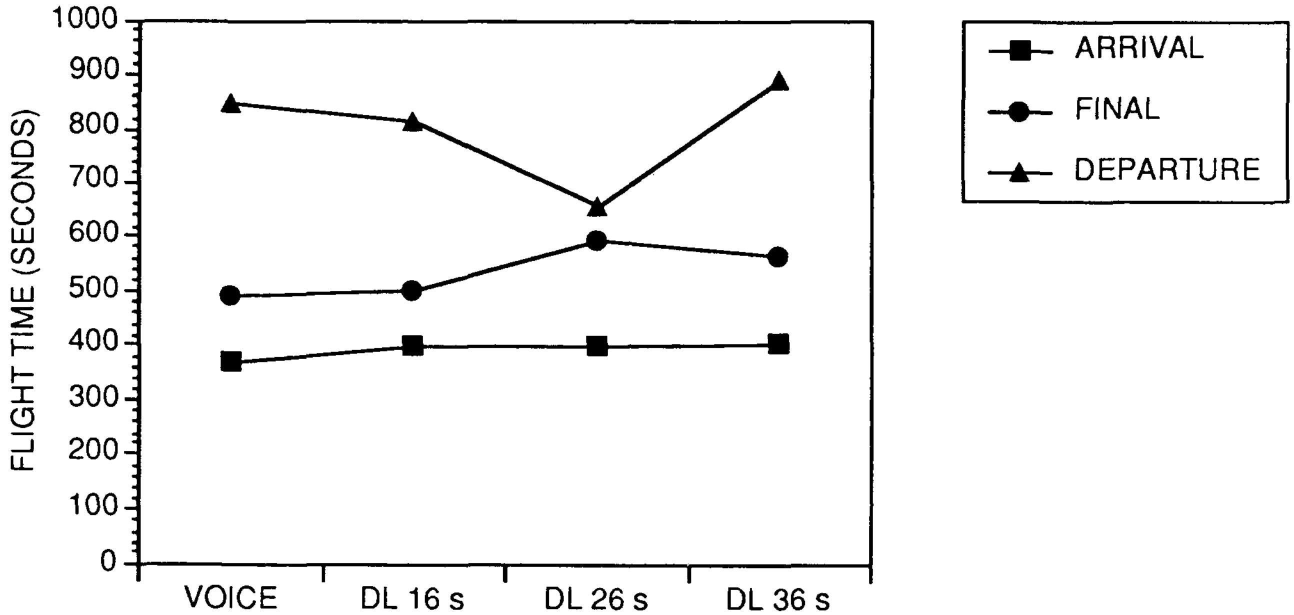


FIGURE 12. AIRCRAFT SECTOR FLIGHT TIME BY COMMUNICATION CONDITION AND POSITION

interaction of the two factors was significant (F(6, 24) = 4.14, p = .018), but again must be interpreted with caution. The cell means are shown in figure 13.

A significant effect of communication condition emerged when the departure sectors were tested alone; MANOVA results were (F(3, 12) = 4.39, p = .026).

4.1.2.3 Path Changes.

The path changes made by each aircraft in each sector were automatically recorded by the NSSF computer system. It was thought that Data Link with voice might allow controllers to be more efficient in issuing ATC requests resulting in a lower number of changes in altitude, speed, or heading made by aircraft while in a sector.

An MANOVA did not show a significant main effect of communication type (F(3, 12) = .95, p = .446), but indicated a significant effect of position (F(2, 8) = 56.45, p = .000). The interaction of the two factors was not significant (F(6, 24) = .80, p = .517). The data are plotted in figure 14.

The significant effect for position is most likely attributable to the larger number of path changes needed in the final sectors.

4.1.2.4 Traffic Load.

In many studies of ATC, the number of aircraft a controller can safely handle is employed as a measure of the effectiveness of the system under test. In this study, traffic load was held constant across all ATC scenarios and, therefore, there was no expectation that differences in traffic load would appear across communication conditions.

A test of the mean number of aircraft handled in each communication condition did not prove to be significant (F(3, 12) = 2.17, p = .184). However, there were significant differences between positions (F(2, 8) = 365.82, p = .000). The interaction of the two factors was not significant (F(6, 24) = 1.12, p = .383). Individual controllers worked an average of 16.5 to 18.4 aircraft.

4.1.3 Voice and Data Link Usage.

4.1.3.1 Voice.

It was expected that the availability of Data Link services would reduce the need to use the radio telephone system to communicate ATC instructions to aircraft. As is demonstrated in figure 15, there was considerably less use of the voice channel (expressed as a proportion of the total run time) in the shortest Data Link condition for all positions. In the longer Data Link delay

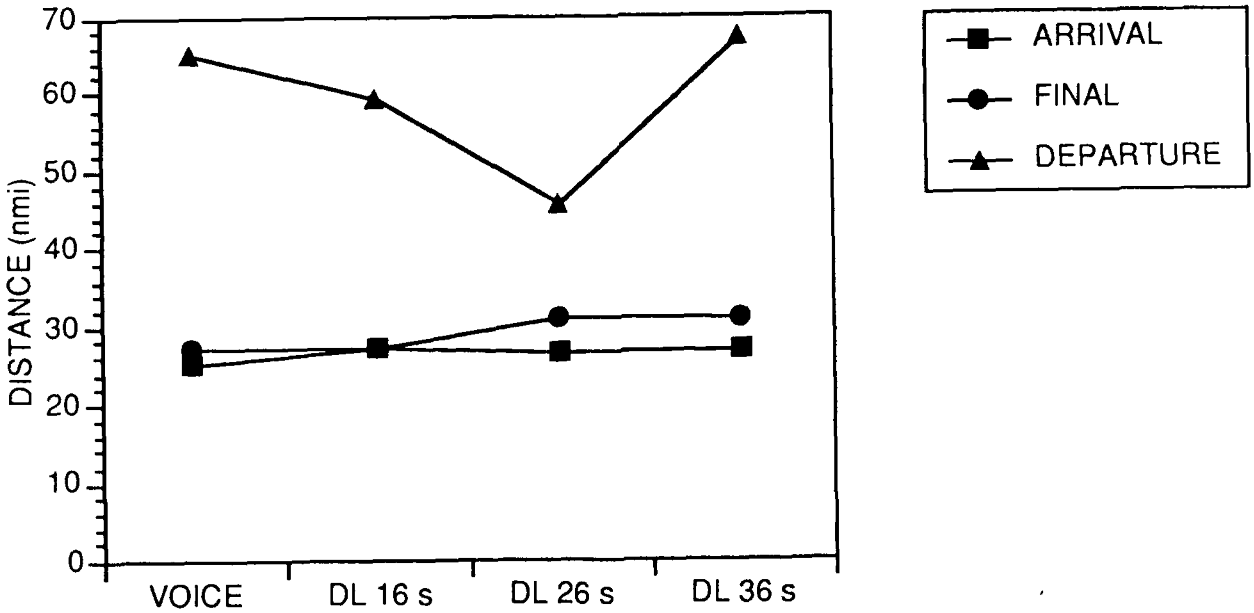


FIGURE 13. AVERAGE AIRCRAFT DISTANCE FLOWN BY COMMUNICATION CONDITION AND POSITION

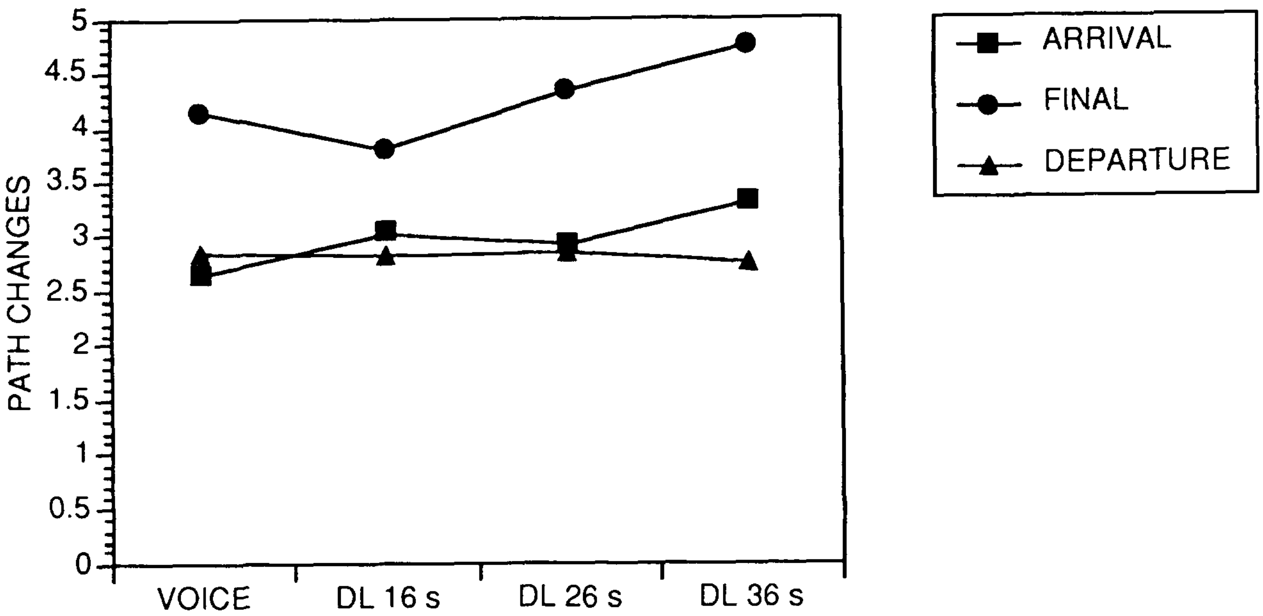


FIGURE 14. PATH CHANGES BY COMMUNICATION CONDITION AND POSITION

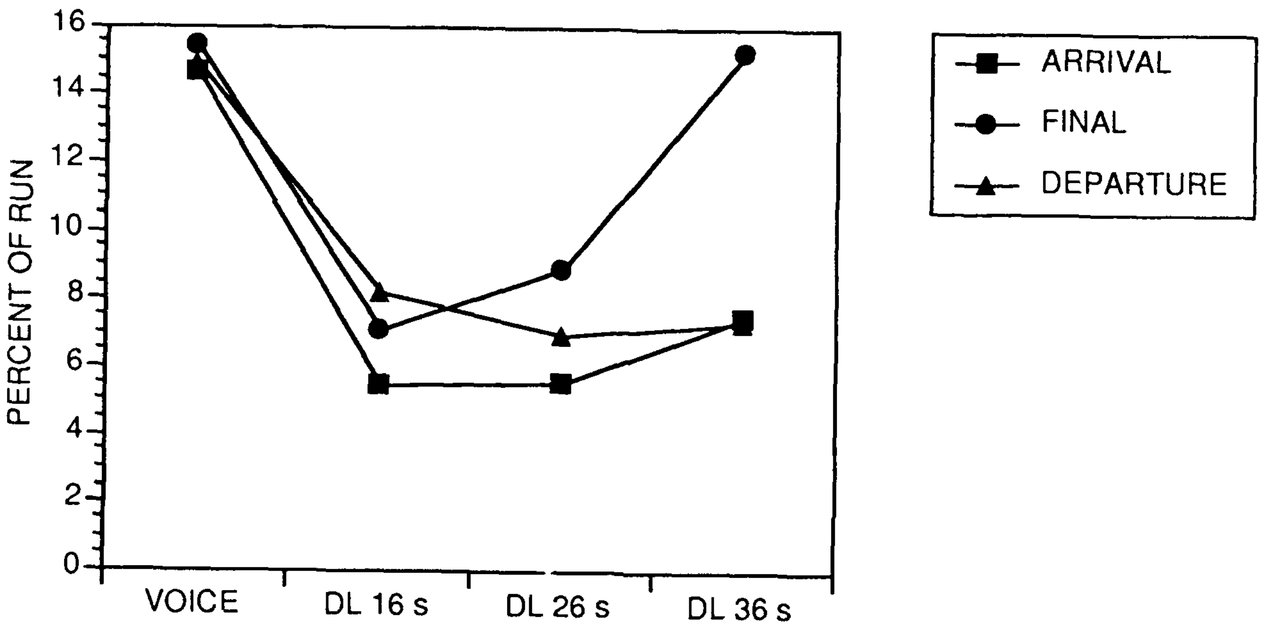


FIGURE 15. PERCENT OF TOTAL SCENARIO RUN TIME CONTROLLERS USED THE VOICE CHANNEL BY COMMUNICATION CONDITION AND POSITION

situations, voice usage appeared to rise in the final sectors but remain low in the arrival and departure positions.

A test of the differences between the groups (using MANOVA) resulted in a significant main effect of communication condition ($F(3, 12) = 43.03$, $p = .000$), but not of position ($F(2, 8) = 3.76$, $p = .071$). Their interaction was not significant ($F(6, 24) = 1.98$, $p = .118$). This means that, although there were significant differences between communication conditions, this was not so for position, and the communication and position factors did not jointly affect voice usage. The appearance of large differences in voice usage between communication conditions as a function of position may not be meaningful due to large variations in how subjects used voice. However, there was a trend in the data toward increased voice usage in the final sectors as Data Link delay increased. This should be studied further.

Voice messages were very consistent in length (about 4 seconds). (Only messages from controller to pilot were considered in this study.)

Similar results emerged for the number of push-to-talk actions. A voice transmission was considered as the communication of one packet of information to the pilot. The message might consist of a query, information useful to the pilot, an ATC instruction (such as a clearance or altitude, speed, or heading change), or a combination of the above. As indicated in figure 16, there was a decrease in the number of voice transmissions per minute across all positions for the shortest Data Link condition.

In parallel to the voice message time data reported above, the rate of message production seemed to remain about the same in all positions except the finals, where the frequency of use of the radio system appeared to increase in the medium and long Data Link delay conditions. MANOVA was used to test the differences between the groups and resulted in a significant effect of communication condition ($F(3, 12) = 27.51$, $p = .000$) but not of position ($F(2, 8) = 3.81$, $p = .069$). Their interaction was not significant ($F(6, 24) = 2.54$, $p = .072$). Similar to the voice time data, there were significant differences for push-to-talk activity between different communication conditions but not for position. The communication type and position factors did not jointly affect the number of push-to-talk actions. Therefore, the appearance of large differences in voice radio activity in the communication conditions as a function of position may not be meaningful due to large variations in how subjects used voice. However, there was again a trend in the data toward increased voice usage in the final sectors as Data Link delay increased which should be studied further.

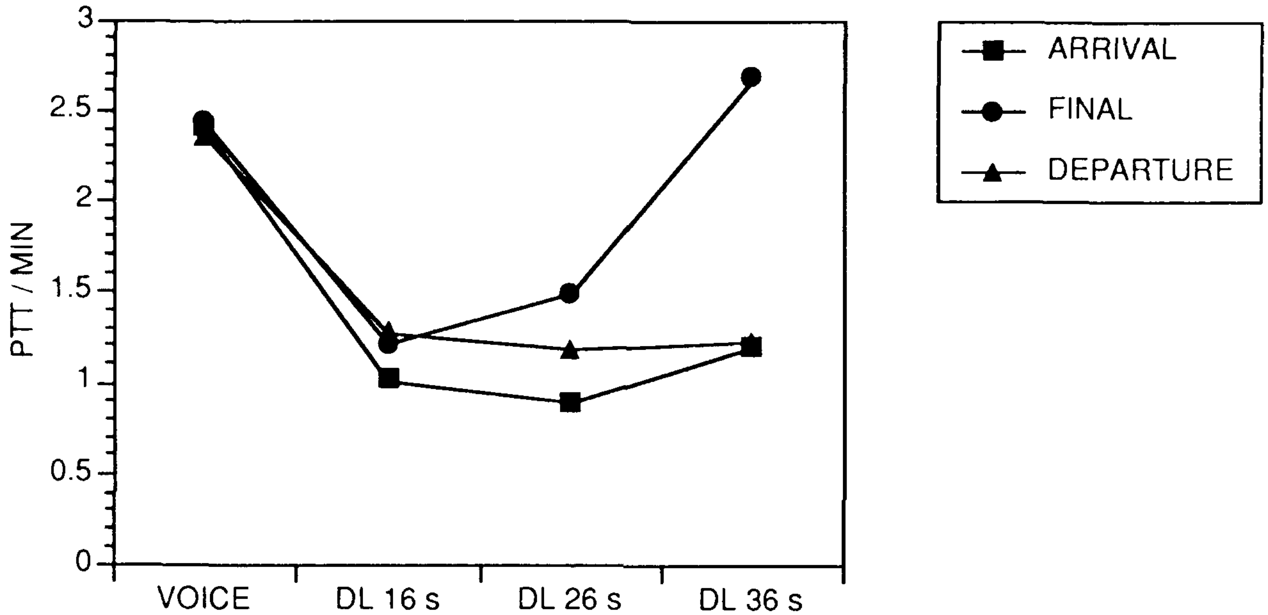


FIGURE 16. PUSH-TO-TALK ACTIONS PER MINUTE FOR EACH COMMUNICATION CONDITION

4.1.3.2 Data Link.

Sources of information regarding Data Link messages included the VAX and NSSF computer systems and videotapes of the ARTS IIIA radar screens during the simulation experiment.

The VAX computer continuously recorded all Data Link transactions for each run with no labels for originating ATC position, but with time stamps for message and receipt of WILCO. Due to the format of the VAX Data Link records, it proved to be very time consuming to assign each message to a controller position. Therefore, only a sample of the VAX data were analyzed to determine Data Link delay times and to correlate with other records.

Calculations of the Data Link automatic delay times from the VAX records (the interval between pressing the ENTER key and receiving a WILCO) yielded the following times:

Short:	average = 17 seconds; standard deviation = 2 seconds
Medium:	average = 27 seconds; standard deviation = 2 seconds
Long:	average = 37 seconds; standard deviation = 2 seconds

The NSSF computer recorded all pilot entries. Pilots were asked to press a special Data Link key whenever a Data Link communication was received from a controller. This allowed a count to be made of Data Link transmissions received per run.

The NSSF path change records proved to be less accurate than anticipated with regard to the assignment of a message to voice or Data Link origination. Correlation of one of the NSSF path change lists with VAX Data Link records showed an 8 percent error rate. In most of these cases, pilots appear to have omitted to press the special Data Link key and in a few instances it seemed that the key was pressed inadvertently.

In order to determine the amount of time required to compose and send Data Link messages, videotape records of Data Link entries initiated from the keyboard were made from the ARTS IIIA screen preview area. (Recordings were made of only half of the Data Link runs (27 sessions.))

The time subjects spent composing and sending a Data Link message was assumed to be the equivalent of keying and speaking into the microphone to send a voice radio transmission. The actions timed from the videotape included key presses, slewing the trackball to the aircraft, and pressing the ENTER key.

There was a good correlation of the Data Link entries observed on the ARTS IIIA screen with those recorded by the VAX computer system. However, during each run a number of Data Link entries were found on videotape of the preview area which were not found in the VAX printouts (and were not sent to pilots). It was decided to include these Data Link message attempts in the analysis because they reflected a realistic use of Data Link and required controller attention.

Observation of the controllers during the experiment and inspection of the message composition durations revealed differences in Data Link usage style. Most of the time subjects composed, slewed, and entered each Data Link message in less than 5 seconds. However, on some occasions the controllers delayed sending the message for up to 30 seconds, presumably to wait for the right time to contact an aircraft. Because such delays represented an ATC strategy rather than the actual time need to send a Data Link message, they were excluded from the data set. Any time over 8 seconds was replaced by the mean of the times under 8 seconds.

Figure 17 shows the average total times (expressed as a proportion of run time) required to compose Data Link messages for each communication condition and position. A statistical test (MANOVA) of the differences between group means was not significant for communication condition ($F(2, 4) = 1.55$, $p = .317$) and was also not significant for ATC position ($F(2, 4) = 8.19$, $p = .097$). The interaction of the two factors was not significant ($F(4, 8) = 1.89$, $p = .206$). The test lacked power because of the small number of subjects in each group. However, there were trends in the results indicating reduced use of Data Link in the final sector as Data Link delay increased and more use of Data Link in the departure sector in all communication conditions. This should be studied further.

The data for the rate of sending Data Link messages appeared similar to those for message composition time. Figure 18 shows the average number of Data Link messages sent per minute for the positions which were videotaped. An MANOVA was completed to test the differences between group means. The results were not significant for communication condition ($F(2, 4) = .43$, $p = .679$), but were significant for ATC control position ($F(2, 4) = 11.42$, $p = .022$). The interaction of the two factors was not significant ($F(4, 8) = .38$, $p = .820$). The test again lacked power because of the small number of subjects in each group. However, there was a trend in the results indicating reduced use of Data Link in the final sector as Data Link delay increased. This should be studied further.

The average time to compose Data Link messages was 3.7 seconds. However, there was considerable variation between composition times. This average time is similar to the that required to send voice messages.

<u>4.1.3.3 Voice and Data Link</u>.

One goal of this study was to compare voice only communication with Data Link with voice. The results reviewed in the preceding two

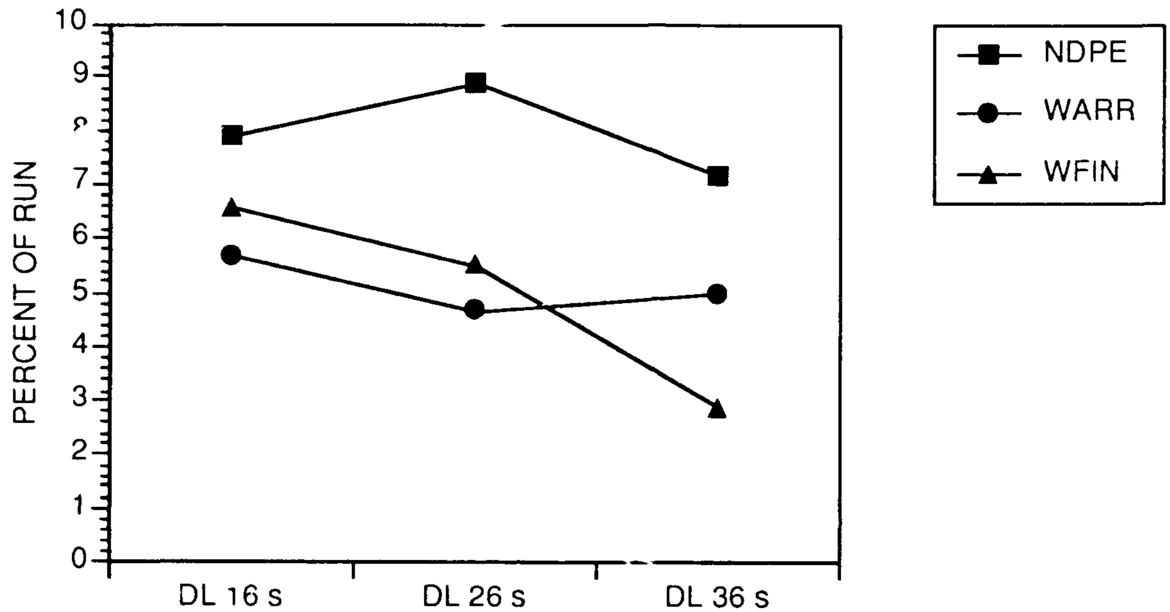


FIGURE 17. PERCENT OF SIMULATION RUN TIME REQUIRED TO COMPOSE AND SEND DATA LINK MESSAGES FOR THREE OF THE SIX POSITIONS

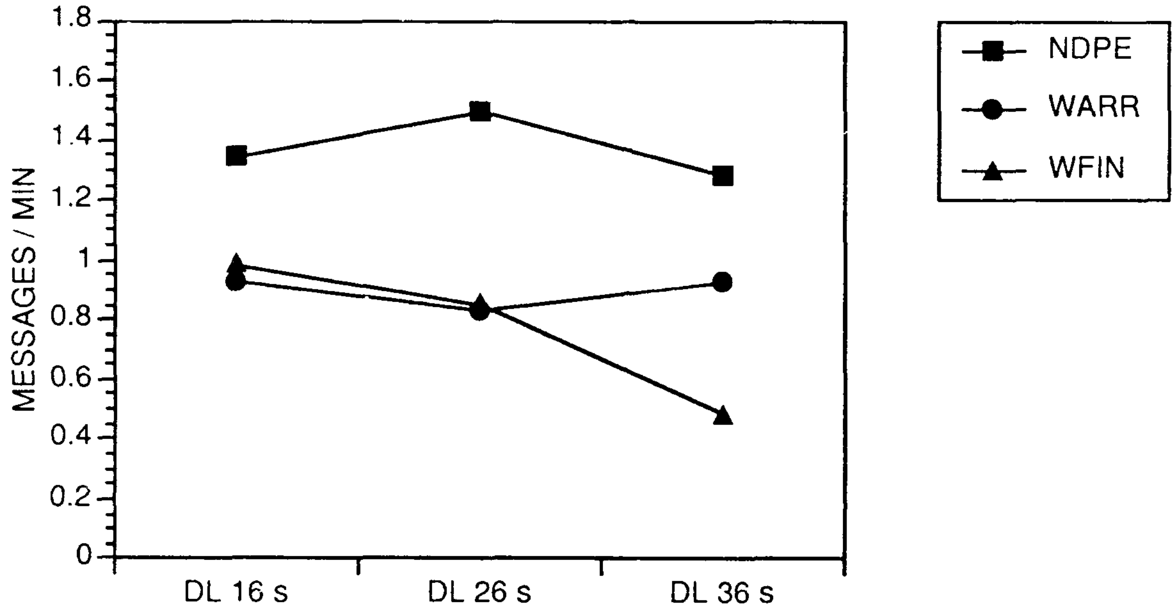


FIGURE 18. DATA LINK MESSAGES SENT PER MINUTE FOR THREE OF THE SIX POSITIONS

sections allow some comparisons of the efficiency of voice and Data Link in the terminal environment.

Figure 19 combines NSSF data for length of voice transmissions with the time required to compose and send Data Link messages observed on the videotapes. In the three Data Link delay conditions, total time required to send communications to aircraft was a combination of the voice and Data Link times. It should be noted that the Data Link communication times included errors and history list requests as well as successfully sent messages. It is probable that miscommunications and other types of discussions unrelated to ATC instructions were also included in the voice times.

A within-subjects ANOVA indicated no significant differences in message time for communication condition ($F(3,6) = 2.84$, $p = .1279$). Therefore, these results must be interpreted with caution pending further research.

Figure 20 again combined NSSF and videotape data to show the rate of voice and Data Link message composition for each communication condition. An ANOVA yielded no significant differences in messages per minute for communication condition ($F(3, 6) = 3.84$, $p = .073$), although the results approached a significance level of $p < .05$.

4.1.4 Cross Communications.

It was considered important to record voice communications with Data Link equipped aircraft and attempts to send Data Link messages to voice-only aircraft. Use of voice radio with Data Link aircraft was not discouraged; controllers were asked to use whichever system seemed most effective. However, such actions could be considered as information about the usefulness or usability of Data Link and were counted by the experiment facilitators. These individuals also counted how many times subjects inadvertently tried to send a Data Link message to a voice only aircraft.

Figure 21 shows the frequency of voice communication with Data Link aircraft for the three Data Link communication conditions. There was a general trend (as shown in figure 22) toward fewer of these actions in each communication condition as the experiment progressed.

Figures 23 and 24 show the frequency of attempts to send Data Link messages to voice-only aircraft. The number of these messages increased as Data Link delay time lengthened. There was some indication that this activity decreased later in the study (as shown in figure 24).

4.1.5 Pilot Callbacks.

Pilot requests for a repeat of a voice or Data Link message might serve as an indication of the efficiency of each medium.

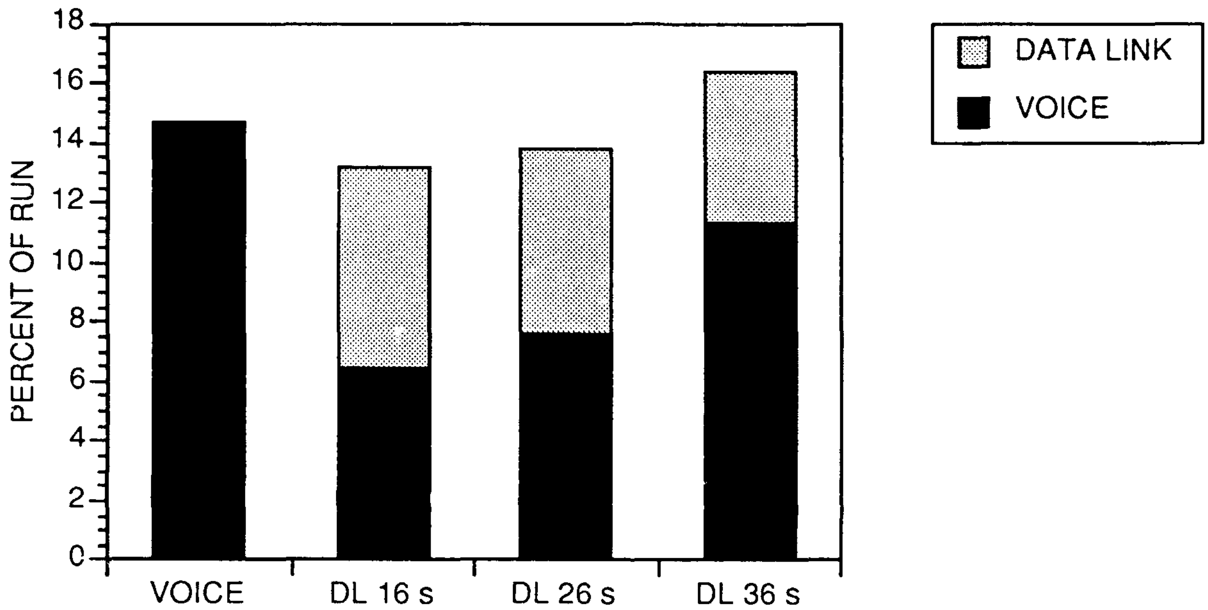


FIGURE 19. AVERAGE TIME TO COMMUNICATE IN EACH COMMUNICATION CONDITION

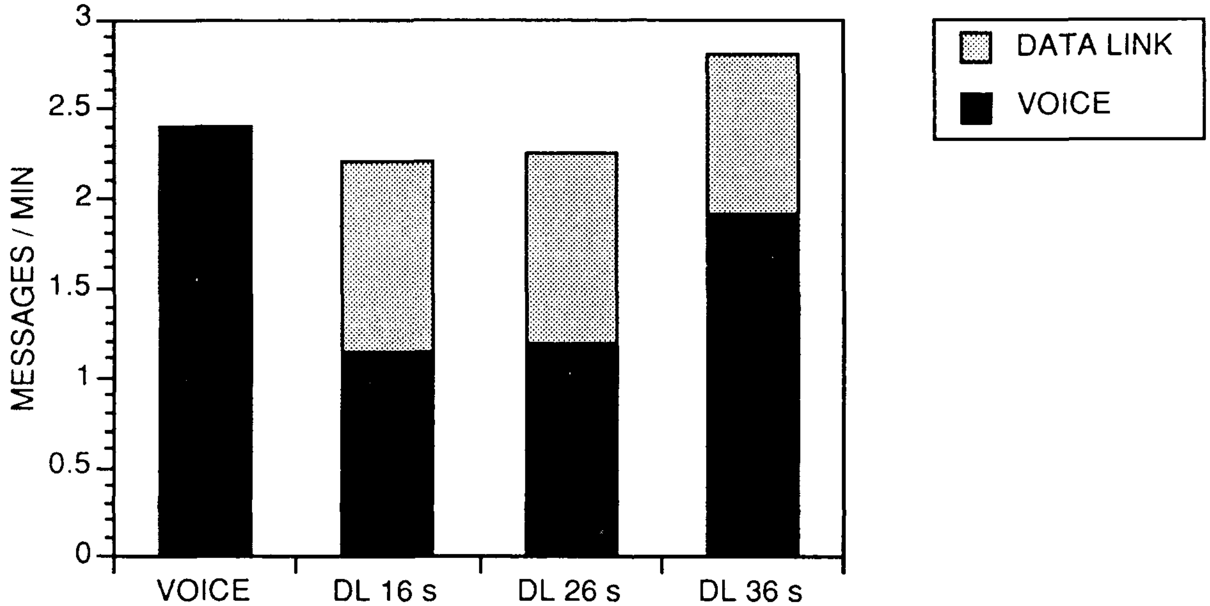


FIGURE 20. AVERAGE VOICE AND DATA LINK MESSAGE RATES FOR EACH COMMUNICATION CONDITION

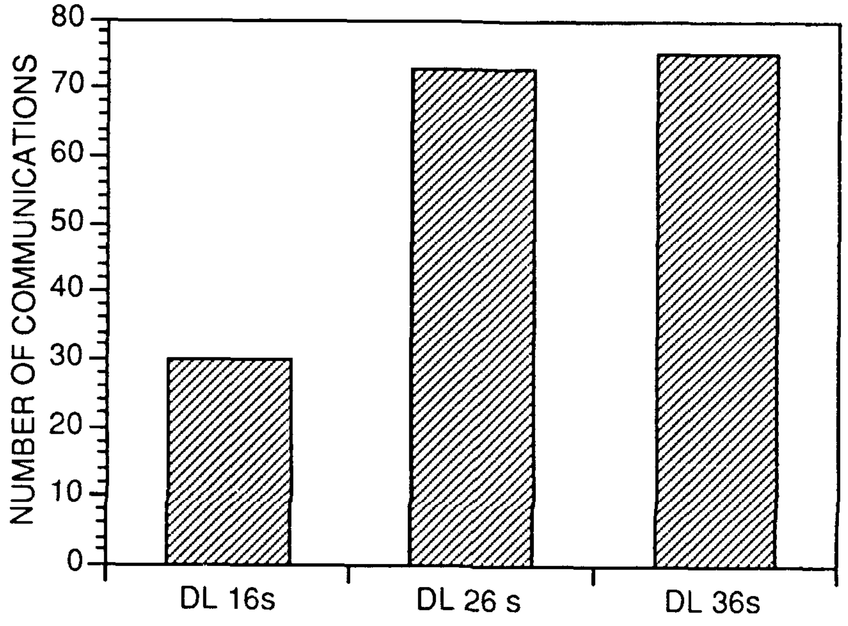


FIGURE 21. AVERAGE NUMBER OF VOICE COMMUNICATIONS TO DATA LINK AIRCRAFT PER RUN FOR THE THREE DATA LINK DELAY CONDITIONS

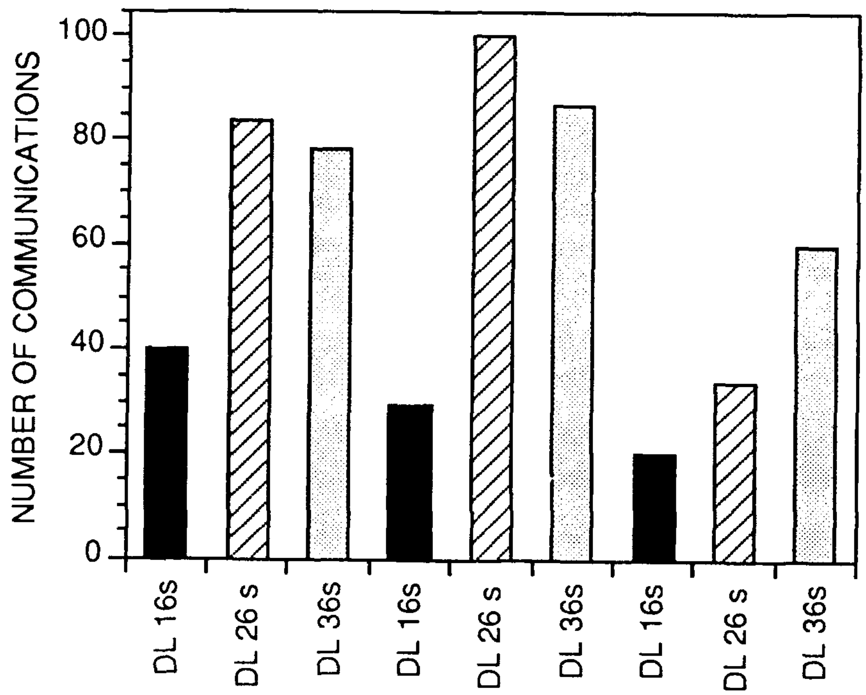


FIGURE 22. FREQUENCY OF VOICE COMMUNICATIONS WITH DATA LINK EQUIPPED AIRCRAFT OVER THE COURSE OF THE EXPERIMENT

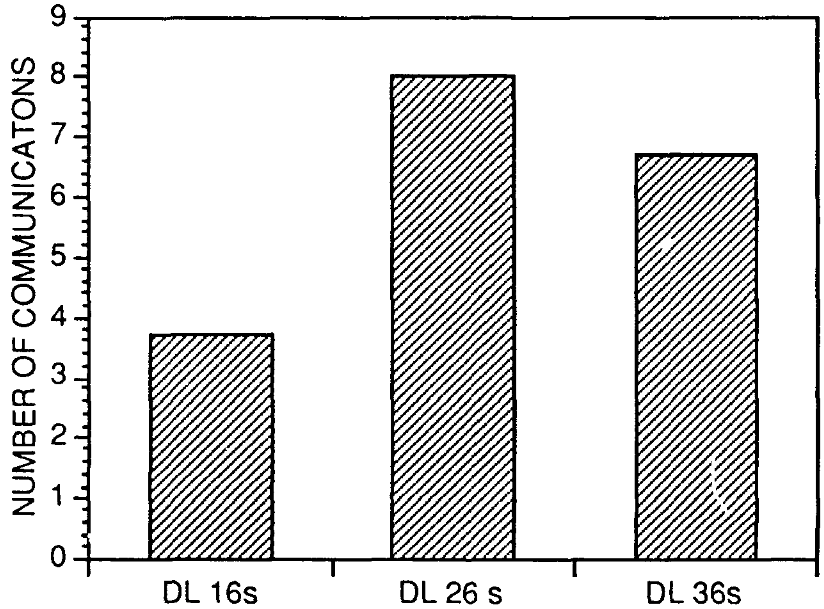


FIGURE 23. AVERAGE NUMBER OF DATA LINK COMMUNICATIONS ATTEMPTED WITH VOICE AIRCRAFT PER RUN

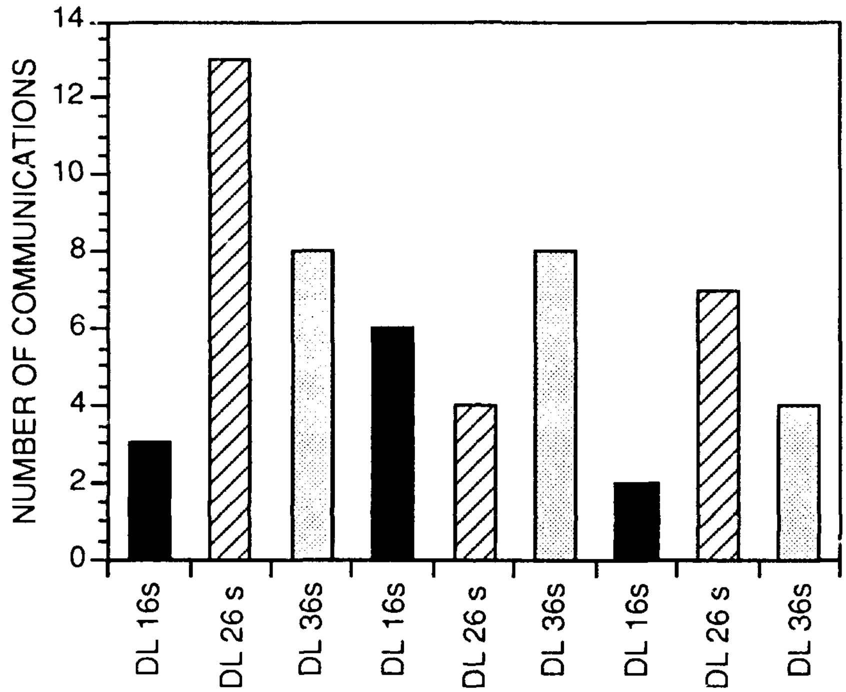


FIGURE 24. FREQUENCY OF ATTEMPTED DATA LINK COMMUNICATIONS WITH VOICE AIRCRAFT OVER THE COURSE OF THE EXPERIMENT

Figure 25 shows the number of pilot callbacks noted by the facilitators for each communication condition.

4.2 DISCUSSION OF PERFORMANCE MEASURE RESULTS.

The preceding results provide useful information about the safety, efficiency, capacity, and communications activities in the terminal environment using voice radio and Data Link with voice. At this point a number of implications can be drawn from these data.

4.2.1 ATC System Safety.

The NSSF conflict file initially contained a large number of incidents. After inspecting these, it was discovered that the list included numerous false alarms. Rules were developed to screen these errors but a sizeable number of conflicts remained. Comparison of the number of conflicts between the four communication conditions suggested that use of the shortest Data Link delay condition resulted in about the same number of conflicts as with voice. However, further study of the conflict measure is needed before it can be considered as a reliable indicator of safety.

4.2.2 ATC System Efficiency and Capacity.

4.2.2.1 Time and Distance.

It is not surprising that differences in aircraft flight time and distance travelled were evident between ATC positions given that the sectors in the experiment were of different sizes. However, overall differences in efficiency also existed as a function of communication medium (when the departure sectors were considered separately).

While there was some evidence that the shortest Data Link condition supported a reduction in flight time and distance for departure sectors, it seemed that the medium Data Link condition created a larger effect. This is not what would be expected if it is assumed that the shortest Data Link condition should offer the most advantages to the controller.

A possible reason for this finding is that the controllers, faced with increasing, but not unworkable, delays in system response time to Data Link messages, changed their communication strategies. A count was made of Data Link initiated path change messages in departure sectors which contained two or more instructions (such as for a speed and altitude change). This showed that controllers in these sectors sent about 55 percent more of this type of message in the medium and long Data Link delay conditions as compared to the short delay condition.

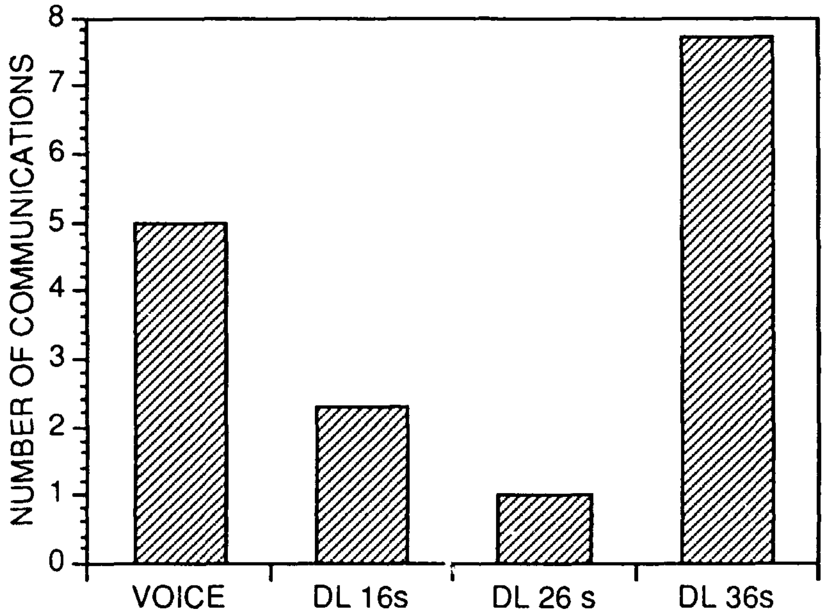


FIGURE 25. NUMBER OF PILOT CALLBACKS FOR EACH COMMUNICATION CONDITION

The increase in Data Link response times may have encouraged the controllers to send more flightpath change requests in each Data Link transmission and this could have had a positive effect on aircraft flight time and distance flown. This improvement was most likely to occur in the departure sectors given that the departing aircraft flew nearly twice as far as aircraft in the arrival and final sectors; there was a longer window of opportunity for a more efficient strategy to have an effect.

Although this explanation might hold true for the medium Data Link delay condition, flight time and distance were not similarly affected in the long delay condition. This is in spite of the fact that the incidence of more complex path change inputs were about the same in the long delay situation as in the medium delay. The reason for this could be that the longer delays made the strategy less effective.

4.2.2.2 Path Changes.

The results for number of path changes did not show significant differences between communication conditions. The requirements for positioning aircraft for landing probably accounted for the increase in path changes in the final sectors.

4.2.2.3 Traffic Load.

It was no surprise that differences in traffic load did not emerge under different Data Link response time conditions. Traffic activity was kept constant for all scenarios in the experiment. Had this been manipulated, it might have been possible to observe differences in capacity as a function of communication type.

4.2.3 Voice and Data Link Usage.

4.2.3.1 Voice.

There was considerable evidence that the use of the voice radio channel decreased with the availability of Data Link services. Both the proportion of time spent using voice per ATC scenario and the frequency of push-to-talk actions dropped dramatically in the short Data Link delay condition and appeared to rise in the final sectors as Data Link delays increased. It is likely that controllers reverted to the voice channel as it became more difficult to use Data Link. This suggests that the final sectors were most sensitive to longer Data Link delays.

Figures 15 and 16 seem to show a difference in voice use across ATC positions in the medium and especially in the long Data Link condition. However, the effect of position and the interaction of communication condition and position were not statistically significant in either case. Further study of the observed trends is needed.

4.2.3.2 Data Link.

Figure 17 suggests (although without statistical support) that more time was spent composing Data Link messages in the departure sectors as compared to the arrival and final sectors. The rate of Data Link message composition (figure 18) followed the same trend. It is reasonable that more time was employed sending messages given that aircraft flew farther in the departure sectors. However, this does not explain why the rate of sending Data Link messages was higher.

In spite of an apparent reduction in Data Link message composition times as delay increased in the final sector, there was no significant effect of Data Link delay condition on Data Link message composition time. This may be due to the limited amount of data available upon which to base the statistical tests.

Although no statistically significant findings emerged as a function of communication type in the Data Link results in figures 17 and 18, it is, nevertheless, interesting to compare the trends to the voice results. Figures 15 and 16 seem to indicate that voice use increased in the final sectors with longer Data Link delays. This was paralleled by a trend in figures 17 and 18 toward decreased Data Link activity in the final sectors as Data Link delay lengthened. Longer Data Link delay times appeared to have the effect of reducing Data Link usage in the final sectors as the controllers increasingly reverted to voice communications. However, further study is needed to confirm these observations.

4.2.3.3 Voice and Data Link.

When the times required to communicate using voice radio and the times needed to compose Data Link messages in the different delay conditions were combined, some interesting results emerged. Figures 19 and 20 suggest that when voice and Data Link activities were considered together, some time savings occurred for the short and medium Data Link delay conditions. However, statistical tests did not indicate that the differences were meaningful. A different result might have emerged if the errors in composing Data Link messages had been removed from the data.

4.2.4 Cross Communications.

Figure 21 graphically demonstrates the effects of long Data Link delays. Voice communications with Data Link aircraft increased over twofold in the medium and long delay conditions. However, figure 22 suggests that this pattern changed over the course of time. There was an apparent decrease in voice to Data Link communications during the experiment; this was especially pronounced for the final runs of the medium and long delay conditions. This implies a learning effect. Controllers may have been able to improve their ability to cope with the long Data Link delays, thus, reducing their need to send voice transmissions to Data Link equipped aircraft.

The number of Data Link messages attempted with voice-only aircraft also appeared to increase with Data Link delay (see figure 23). While figure 24 suggests that these errors decreased over time, the trend is inconsistent. This measure might be an indirect indicator of controller workload. Attention was required to determine from the ARTS IIIA display which aircraft were Data Link equipped. A symbol to denote communication equipage was displayed in the aircraft's data block. If it can be assumed that lorg Data Link delays increased workload, it would be reasonable to expect that less time could be devoted to checking aircraft equipage resulting in Data Link messages being sent to the wrong targets. The issuance of such messages could lead to conflict situations (if it were thought the message had been delivered).

4.2.5 Pilot Callbacks.

The number of times pilots had to contact controllers to ask for clarification of an ATC request is another measure of the efficiency of the communication media. Figure 25 suggests that the short and medium Data Link delay conditions resulted in fewer callbacks from the simulator pilots. On the other hand, the long delay situation created the need for more callbacks. This was another indication of the problems created by long Data Link delays.

5. CONCLUSIONS.

The results of the study presented in this report warrant the following conclusions regarding the initial terminal air traffic control (ATC) services, required changes to the service designs, potential effects of extended transaction times on the utility of Data Link, and requirements for future terminal Data Link research.

a. The air traffic controller members of the Air Traffic Data Link Validation Team (ATDLVT) who participated in this study expressed a continued belief that implementation of Data Link services in the terminal environment will have the potential to significantly improve the quality and effectiveness of air-ground ATC communications. Predicted beneficial effects include reductions in voice frequency congestion and communication errors, and improvements in efficiency made possible by shifting lengthy and repetitive advisory messages and clearances from the voice channel to the Data Link communications system.

b. The results of the full scale simulation phase of this study confirmed the hypothesis that the viability of Data Link ATC services in the terminal environment will be strongly dependent on the length of total transaction delays between message dispatch by the controller and receipt of an acknowledgement from the flight crew. While this preliminary study was conducted with minimal controller training using Data Link service designs which are still under development, the results indicated that total transaction delays exceeding 20 seconds may significantly increase controller workload, impair performance, and impede the use of Data Link. The severity of the effects of transaction delays were dependent on the type of airspace sector under control and the time criticality of the message. In addition, the findings suggested that the negative effects may be partially ameliorated by practice as controllers develop effective strategies for overcoming time delays.

c. The design review, and debriefings which followed the full scale simulation phase, provided resolutions for all open design issues and generated several modifications and enhancements to the Data Link service designs, which the subject controllers indicated would enhance performance and help to reduce transaction delays. Among the most important of the modifications agreed upon by the controllers were: (1) a change in the design of the initial contact service to reduce the number of Data Link transactions and amount of transaction time required to perform this function, and (2) extensive modifications to the controller interface of the MT and TI services to improve the speed and accuracy of message selection.

d. Observed reductions in controller workload and improvements in performance that occurred over the course of testing, as well as inputs from the subjects, suggested that modifications to the Data Link and airspace training protocol will be required for future operational evaluation research. In particular, it is concluded that the ability to make valid measurements of the operational effectiveness of Data Link services will depend upon increasing the duration of the practice phase of training in order to permit the development of proficient strategies for employing the system.

e. Several of the experimental measures of system and controller performance examined during this study effectively reflected and complemented the controller evaluations, and are, therefore, expected to be essential for future operational evaluation research. This research will be performed using the fully mature Data Link service designs and extensively trained controllers, and will provide the opportunity to derive valid conclusions regarding the impact of Data Link services on the safety, efficiency, and capacity of the ATC system. The results of this preliminary evaluation indicated that additional development of these measures must be conducted in studies prior to operational evaluation to: (1) improve the ease with which performance data can be extracted from the simulation computer records, (2) insure the validity of measures of aircraft separation conflicts, and (3) explore the ability of these performance metrics to reflect the potential operational benefits of Data Link.

6. RECOMMENDATIONS.

The following recommendations for future actions under the Federal Aviation Administration (FAA) terminal air traffic control (ATC) Data Link program are derived from the results and conclusions of the present research:

a. The results of this second design development study indicate that the initial Data Link ATC services continue to have significant potential for enhancing terminal ATC operations. Therefore, it is recommended that additional design, development, and evaluation efforts be pursued to implement these services at the earliest possible date.

b. Because of the significant effects of total transaction times on the effectiveness and usability of Data Link that are predicted by the findings of this preliminary study, it is recommended that technical analyses and flight deck Data Link research be initiated to define the distribution of transaction times that can be expected in operational implementation. When available, these time parameters should be used in future ATC service evaluation research, and should serve as a baseline for evaluating improvements in aircrew interfaces that are intended to minimize message acknowledgement response latencies.

c. It is also recommended that, prior to operational evaluation of the initial terminal Data Link ATC services, additional research be conducted to further refine the service designs and to optimize test and evaluation procedures. As part of this research, the service design modifications identified during the present study should be implemented in the ARTS IIIA section of the Data Link Test Bed for review by Air Traffic Data Link Validation Team (ATDLVT) controllers. Furthermore, in preparation for operational evaluation, this research should test more extensive Data Link training protocols, exercise improved performance data collection procedures, and attempt to identify new and more sensitive measures which will reliably reflect the impact of Data Link on ATC system and controller performance.

d. Finally, the results of the present study suggest that, with extended practice, controllers may spontaneously develop strategies for employing Data Link which can mitigate the impact of transaction delay and improve the effectiveness of ATC in a combined voice radio, and Data Link communications environment. Because of this, it is recommended that future research efforts include the development and application of human factors metrics which can define and document these strategies. Information derived from these measures should be used to develop optimal Data Link procedures and training materials that will be needed to complement functional specifications for the Data Link service designs and controller/system interfaces.

7. BIBLIOGRAPHY.

Billings, C. E. and Reynard, W. D., Dimensions of the Information Transfer Problem, in C. Billings and E. Cheaney (Eds.), Information Transfer Problems in the Aviation System, NASA Technical Paper 1875, National Aeronautics and Space Administration, 1981.

Computers to Reshape Air Control Job in 90's, New York Times, July 29, 1988.

Data Link Development Team, Controller Evaluation of Initial Data Link Terminal Air Traffic Control Services: Mini Study 1, DOT/FAA/CT-90/29, January, 1991.

Shingledecker, C. A., Human Factors Research in the Development of Data Link Air Traffic Control Services, Proceedings of the Aeronautical Telecommunications Symposium on Data Link Integration, ARINC/FAA, 1990.