The Effect of Voice Communications Latency in High Density, Communications-Intensive Airspace Phase II: Flight Deck Perspective and Comparison of Analog and Digital Systems

Carolina M. Zingale, Ph.D., Titan Corporation
D. Michael McAnulty, Ph.D., ACB-220
Karol Kerns, Ph.D., AND-300, MITRE/CAASD

November 2003
DOT/FAA/CT-TN04/02

U.S. Department of Transportation
Federal Aviation Administration

William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

NOTICE

Technical Report Documentation Page

1. Report No.
DOT/FAA/CT-TN04/02

2. Government Accession No.

3. Recipient's Catalog No.

4. Title and Subtitle
The Effect of Voice Communications Latency in High Density, Communications-Intensive Airspace. Phase II: Flight Deck Perspective and Comparison of Analog and Digital Systems

5. Report Date
November 2003

6. Performing Organization Code

7. Author(s) Carolina M. Zingale, Ph.D., Titan Corporation, D. Michael McAnulty, Ph.D., ACB-220, and Karol Kerns, Ph.D., AND-300; MITRE/CAASD

Performing Organization Report No..
DOT/FAA/CT-TN04/02

Performing Organization Name and Address
Federal Aviation Administration
NAS Human Factors Branch, ACB-220
William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

12. Sponsoring Agency Name and Address
Federal Aviation Administration
Air/Ground Communications Product Team
800 Independence Avenue S.W.
Washington, D.C. 20591

13. Type of Report and Period Covered
Technical Note

14. Sponsoring Agency Code
AND-360

15. Supplementary Notes

16. Abstract

The FAA Next Generation Air-Ground Communications (NEXCOM) program plans to replace the aging analog air traffic communications system with a Very High Frequency (VHF) Digital Link Mode 3 (VDL3) system. VDL3 provides increased channel capacity, can transmit both voice and data, and compensates for known limitations in the analog system through features such as controller override, antiblocking, and a transmit status indicator. The system also virtually eliminates "step-ons." However, VDL3 will have a longer voice throughput delay (up to 350 ms) than the analog system (approximately 70 ms). This report describes the second high fidelity, human-in-the-loop simulation study of VDL3 system performance and operational acceptability. The first study evaluated VDL3 with air traffic controllers. We conducted the current study to validate the findings of the first study with pilots, to compare analog communications to those using VDL3, and to assess analog and VDL3 communications under routine conditions and under adverse weather conditions that further increased demand for the channel. Fourteen airline pilots participated in the study using two realistic flight deck simulators. VDL3 allowed more successful transmissions to be made. Most other communications characteristics did not differ between the radio systems. The effects of adverse weather were similar for both systems. The participants were highly positive in their evaluations of VDL3. They rated the operational acceptability of VDL3 higher than analog, and they nearly always rated the digital system as equal to or better than analog for completing communication tasks. Overall, the results indicate that VDL3 is an acceptable communications system for pilots.

Air-Ground Communications, Digital Voice, Aviation Human Factors, Human-In-The-Loop Simulation, Throughput Delay, Pilot Workload and Acceptance

18. Distribution Statement
This document is available to the public through the National Technical Information Service, Springfield, Virginia, 22161

19. Security Classif. (of this report)
Unclassified

20. Security Classif. (of this page)
Unclassified

21. No. of Pages
62

22. Price

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

Acknowledgements

This Technical Note describes the methodology and results of a highly complex, human-in-the-loop simulation conducted to evaluate the Very High Frequency (VHF) Digital Link Mode 3 (VDL3) communication system for the Next Generation Air/Ground Communications (NEXCOM) program. It reflects the contributions of numerous individuals in diverse disciplines from multiple organizations, and the authors wish to acknowledge and thank all of them for their efforts.

Beth Lyall, Sharon Hecht, and Jay Martin served as human factors observers in the two cockpits we simulated, and they, along with Victor Riley, Steven Rogers, and Robert Davis, contributed to test-plan development and the development of the cockpit data collection instruments and procedures. Other members of the NEXCOM Human Factors Working Group also contributed to the development of the test plan: Mark Cato, Madison Walton, Mark DeBack, Keith Dutch, Harry Eberlin, and Craig Morris.

The following members of the Research, Development and Human Factors Laboratory (RDHFL) technical staff developed, tested, and ran the simulated air traffic and communication systems: Albert Macias, Mary Delemarre, Otto Smith, Wallace Daczkowski, Bob Oliver, Dan Johnson, and Dario Rossilli.

We also wish to thank Scott Ducette and Samantha Fullerton from the Target Generation Facility, Joe McCall and George Bollenbach from the General Aviation Trainer facility, and Al Adkins, Pocholo Bravo, Dave Carty, and Al Rehmann from the Engineering Research Simulator, who developed, tested, and ran their parts of the simulation system.

John Goldman and Joe Toland provided Air Traffic Control subject matter expertise in developing the simulation scenarios. Doug Kelley (Indianapolis Center) and Dave Conley (Jacksonville Center) participated as controllers during shakedown and testing. Mike Cullum served as the pilot flying in the B-747 cockpit. Human factors specialists at the RDHFL helped to design and run the simulation and to collect and analyze data: Randy Sollenberger, Todd Truitt, Mark Hale, Kenneth Allendoerfer, Ben Willems, and Kevin Hallman.

We also thank the simulation pilots who managed the other traffic and communications in the scenarios and the airline pilots who served as participants in the cockpit simulators. The simulation could not have been run successfully without the outstanding efforts of all those individuals we named.

Table of Contents

List of Illustrations

Tables Page

Executive Summary

The FAA Next Generation Air-Ground Communications program plans to replace the aging analog air traffic communications system with a Very High Frequency (VHF) Digital Link Mode 3 (VDL3) system. VDL3 is a time division, multiple access system that provides increased channel capacity and will be capable of transmitting both voice and data. In addition, the VDL3 system compensates for known limitations in the analog system by implementing features such as controller override, antiblocking, and a transmit status indicator. The system virtually eliminates "step-ons" that result when more than one user tries to transmit simultaneously and in which none of the messages is sent clearly and completely. However, the proposed system will also have a longer voice throughput delay (up to 350 ms) than in the analog system (approximately 70 ms). The study described in this report is the second high fidelity, human-in-the-loop simulation study of VDL3 system performance and operational acceptability. The first simulation evaluated the effects of the VDL3 system on air traffic control specialist performance and workload. The objectives of this study were to validate the findings of the first study of VDL3 system effects with pilots as participants, to compare data obtained under analog communications to those obtained using VDL3, and to assess analog and VDL3 communications under routine conditions and under adverse weather conditions that further increased user demand for access to the channel. Fourteen airline pilots participated in the study using two realistic flight deck simulators. One simulator was an air transport category cockpit with a two pilot crew, and the other was a twin engine general aviation cockpit flown by a single pilot. The results showed that the digital system allowed more successful transmissions to be made although users tended to access the channel similarly with both radio systems, demonstrating comparable rates of overlapping transmissions. Communications did increase under the weather conditions, but did so similarly for both radio systems. The participants were highly positive in their evaluations of VDL3, rating the operational acceptability of the digital system significantly higher than that of the analog system regardless of the weather. The digital system was almost always rated as being equal to or better than the analog system for completing communications tasks. Overall, the results indicate that the VDL3 system is an acceptable communications system for pilots.

1. Introduction

1.1 Background

The Federal Aviation Administration (FAA) uses an extensive radio system (individual radios, remote control communications facilities, radio communications outlet facilities, remote transmitter/receiver facilities, and backup emergency communications facilities) to provide air-ground communications between Air Traffic Control Specialists (ATCSs) and airplane pilots. The communications system allows pilots to convey requests for departure or landing and route, altitude, and speed changes; to report position information or weather conditions; and to notify controllers of urgent or emergency situations. Controllers use the system to manage the airspace, ensuring aircraft separation for planes flying under instrument flight rules, maintaining efficient traffic flows, and offering additional services on a time-available basis. Most of the civilian communications occur using frequencies with a bandwidth of 25 KHz in the Very High Frequency (VHF) spectrum allocated for aeronautical communications (117.975 – 137 MHz). The current system is a double sideband, audio-modulated radio, also called an analog radio.

Although it has worked well for years, there are problems with the current analog communications system (FAA, 2000). First, as the volume of air traffic grows, there is a shortage of assignable frequencies in the available VHF radio band to fill the need for new frequency assignments for additional facilities and sectors. Second, the current equipment is also old and becoming increasingly difficult to maintain. A third problem is the system lacks the capability to transmit data in addition to voice and to implement adequate security protections against unauthorized use. Finally, the current system cannot manage difficulties that result when more than one user tries to transmit simultaneously. When these "step-on" situations occur, neither user transmits a message clearly and completely, and neither may be aware that the transmission attempts have failed. Also, when a user has a stuck microphone key, other transmission attempts are unsuccessful. In response to these problems, the FAA Next Generation Air-Ground Communications (NEXCOM) program is developing the VHF Digital Link Mode 3 (VDL3) communications system to provide air traffic services communications.

1.1.1 The VDL3 System

The VDL3 system uses a time division, multiple access process to convert each analog frequency into four digital communication channels. The channels can be set up to carry either voice or data communications. VDL3 was selected from a number of other possible alternatives (e.g., further subdividing current analog frequency bandwidths, satellite communications) because it provides for increased channel capacity and highly reliable transmission capabilities without being cost prohibitive and without compromising safety (Williams, Eck, & Eckstein, 2001). From an operational perspective, the FAA goal for VDL3 voice communications is to remain as consistent as possible with the current voice radio system. For example, VDL3 will maintain the current party-line operation for voice messages, allowing pilots and controllers to overhear information on a common channel. In addition, the established procedures for communication will be retained. However, there are fundamental differences between the analog and the VDL3 systems that may affect the perceived quality and effectiveness of the digital voice service.

First, the VDL3 system uses voice coding (vocoder) techniques to encode and decode voice messages transmitted to and from controllers and pilots. The vocoded speech results in a different voice quality and different throughput latency than the current system. Throughput latency has multiple components. In analog radio, there is a variable system set-up time plus a propagation delay of approximately 70 ms (Nadler, DiSario, Mengert, Sussman, & Spanier, 1990). Once a transmitter is set up in analog radio, it takes approximately 70 ms before everyone on the frequency hears the voice message, regardless of whether the speaker is a controller or pilot. In VDL3 ground-to-air communications, there is channel acquisition time, time to encode the message (initially specified as a maximum of 250 ms; RTCA, 1994), plus an additional time of about 40 ms representing the propagation time and processing the signal through the cockpit avionics. Once a controller acquires the VDL3 channel, the specified delay before the pilots hear the transmission is approximately 290 ms. Air-to-ground communications will have a similar but slightly shorter delay of 260 ms.

A further difference between the current system and VDL3 is the way access to the communications channel will function. In analog radio, communications are partially or totally blocked if more than one person tries to transmit at once or if a pilot or controller has a stuck microphone key. In VDL3, channel access is managed with a ground-based control station establishing and maintaining a communication link between the controller and pilot and managing users' access to the channel. To compensate for the known limitations in the analog system, new features have been proposed for the basic VDL3 voice service. First, the VDL3 system will minimize the possibility of one user stepping on another who is already sending a message. An antiblocking capability will detect when the channel is busy and will block any conflicting transmissions initiated by other pilots by placing the other radios on the channel in "receive only" mode. Second, the VDL3 system will allow the controller to obtain access to a busy channel through an override function. The controller override or priority access feature enables the controller to preempt other users and also allows the controller to free a channel blocked by an active transmitter in a stuck microphone situation. Third, a transmit status indicator cue will provide notification to users that a channel is unavailable. This feature is expected to alert a pilot who attempts to transmit on a busy channel and will also alert a pilot whose transmission is preempted by the controller.

1.1.2 Previous Research on VDL3 Issues

Since the late 1980s, a considerable body of research has been conducted to define and validate the required voice quality of voice coder technology for the VDL Mode 3 system (Child, Cleve, & Grable, 1989; Dehel, Grable, & Child, 1989; Farncome & MacBride, 2000; Fujimori & Ueno, 1999; LaDue, Sollenberger, Bellanger, & Heinze, 1997; Sollenberger, LaDue, Carver, & Heinze, 1997; Renaud, Fistas, Brugere, & Garcia, 1999). The research indicates that the voice quality is acceptable for air traffic communications, although some vocoders are significantly more intelligible and preferred than others (LaDue et al.; Sollenberger et al.).

During radio communications, pause durations provide cues to the listener about when it is time to speak. When pauses are longer than anticipated, a listener may interpret the delay as a signal that a response is expected. Voice throughput delays that are longer than expected can disrupt turn taking, increase conflicting transmissions, and affect the quality, efficiency, and integrity of the communications process. Previous studies of air traffic control (ATC) communication delays in an analog system context indicate that communication delays may affect several operational parameters, including the variability or efficiency in the flight paths of arrival traffic, the workload experienced by the controller, the number of transmissions, the duration of transmissions, and the rate or frequency of step-ons. The research also indicates that these delay effects are context specific and will be observable in a communications-saturated, high workload environment.

Nadler, Mengert, DiSario, Sussman, Grossberg, and Spanier (1993) conducted a high-fidelity, controller-in-the-loop simulation to evaluate four delay conditions representing potential combinations of ground-to-air and air-to ground delays induced by voice switching and satellite transmission equipment. Throughput delays above 400 ms significantly increased the frequency of step-ons and retransmissions. However, the increased incidence of step-ons was observed only when communications workload was high.

Suganuma (1997) investigated the effects of delays while pairs of controllers read 3-digit numerals, ATC terms, and ATC instructions to one another in a soundproof studio. One of the controllers read while the other read back the information. During the read back, a delay ranging from 50 to 500 ms was added. The results showed that about one quarter of the participants noticed the delay when it reached 250 ms and about one-half noticed the delay when it reached 450 ms. The participants reported that they did not expect the 250 ms delay to affect ATC operations. However, the study also found that participants completed only about 95% of their communications tasks under the 250 ms delay condition compared to the 50 ms delay condition.

Farncombe (1997) conducted a series of high fidelity, controller-in-the-loop simulations to determine the effect of four delay conditions (130, 280, 400, and 550 ms) and found that the higher delay conditions reduced the number of transmissions. Though controllers were found to have modified their speech strategies to compensate for the delays (e.g., shorter messages in some instances, longer in others), the higher delays were nevertheless associated with higher variability of the aircraft flight paths. The author attributed this variability to a lack of timely delivery of ATC instructions. Overall, he concluded that delays of about 280 ms would not adversely affect ATC operations, but that delays of 400 ms or more would be unsuitable.

In a very simplified control task, Rantanen, McCarley, Xu, and Yeakel (2001) assessed the sensitivity of ATC operations to different magnitudes of audio throughput delays and analyzed the relative contributions of the audio throughput delay and delays in pilot response to operational effects. There were statistically significant effects of increasing audio delay and random pilot delay on total communications duration and on lateral separation, but the differences between the minimum and maximum delay intervals were very small (3.5 sec and 0.14 nm, respectively). Controller performance was generally worse and some elements of workload were slightly higher in the highest delay conditions (1000 ms).

A recent study by Sollenberger, McAnulty, and Kerns (2003) examined the effect of communications delays in a digital system context. They conducted a high fidelity, controller-in-the loop simulation to evaluate the effects of three different voice throughput delays representing VDL3 ground system delays of 250, 350, and 750 ms. They added an additional 40 ms to the ground delay for signal propagation and airborne system processing. The simulation of the VDL3 system included the antiblocking, transmit status indicator, and controller override features. The results indicated no differences between the 250 ms and 350 ms delay conditions. However, the 750 ms delay condition resulted in differences in communications measures and in controller ratings on the impact of communications. The controllers made fewer transmissions and occupied the communication channel for a shorter duration during the 750 ms delay compared to the 250 ms and 350 ms delays. The controllers reported using compensatory strategies for coping with the 750 ms delay, indicating that they provided fewer optional ATC services (e.g., calling traffic). They also reported that this delay reduced their speech clarity. Controller overrides of pilot transmissions occurred much more frequently in the 750 ms delay condition, although most of them appeared to be unintentional, occurring during the communications system delay before it was possible for the controller to have heard the pilot speaking. The authors concluded that a VDL3 communications system with 350 ms ground system delay and controller override, antiblocking, and a transmit status indicator would be effective and acceptable to controllers.

Sollenberger et al. (2003) also recommended that additional research be conducted to examine the impact of the VDL3 communication system from the flight deck perspective. The initial study was not as realistic from the pilot perspective as it was for the controllers. The simulation facility was configured for three pseudopilot positions to support each controller position. This meant that each pseudopilot was responsible for maneuvering 4 to 8 aircraft at a time, a configuration that reduced the potential for communications conflicts among aircraft. Finally, no data were collected to determine whether the observed operational effects represent losses or gains in performance relative to the current analog system, and, if so, whether they may be counterbalanced by changes elsewhere that affect safety or workload. These recommendations form the basis for the present study.

1.2 Study Objectives and Hypotheses

A human factors plan for the NEXCOM system defined a program of human factors research and engineering tasks needed to facilitate a successful operational deployment of a VDL3 system (FAA, 2001). The human factors evaluation of the VDL3 system is being conducted incrementally based on the current FAA development risks and deployment strategy. The first study under the plan, NEXCOM I, investigated the effects of different delay parameters and basic system features on ATCS performance and their subjective assessment of the system (Sollenberger et al., 2003). The design and affordability of the VDL3 system had been constrained by the initial 250 ms voice delay requirement. Their study demonstrated that the more technologically feasible delay of 350 ms is acceptable for ATCS communications.

In the context of the NEXCOM human factors plan, the research described in this report is the second high fidelity, human-in-the-loop simulation study of VDL3 system performance and operational acceptability, NEXCOM II. This study had three objectives. The first was to validate the findings of the previous study of VDL3 system effectiveness and efficiency by utilizing a similar methodology and collecting comparable measurements for pilots. The second objective was to compare data obtained under analog communications to those obtained using the VDL3 system. The third was to assess analog and VDL3 communications under routine but busy conditions and under severe weather conditions that further increased user demand for access to the communications channel.

We replicated the method and procedures from the previous study, NEXCOM I, with some modifications to increase operational fidelity and to implement the weather conditions. We generated additional scenarios and adapted the test measures to more appropriately address factors relevant to pilots. We collected objective performance measures and subjective measures of system acceptability. We also modified the data analysis programs to correctly analyze data from the analog radio test conditions.

Overall, we anticipated similar communications patterns between the two simulations, but with increased contention in the current study for the radio channel because of the larger number of pilots. We also expected that the VDL3 system would be more efficient (no step-ons) than analog radio and would be highly acceptable to pilots. Finally, we expected increased communications and pilot workload along with reduced communication system acceptability in nonroutine conditions compared to routine conditions, effects that may be more pronounced with the analog system.

2. Method

2.1 Participants

Fourteen certified airline pilots participated in the study. They were recruited through the Air Line Pilots Association (ALPA). A description of the participant requirements is included in the recruitment letter in Appendix A. Twelve participants were current for Instrument Flight Rules (IFR); the other two were furloughed. The airlines they currently or most recently flew for were Air MidWest, AmericaWest, Atlantic Coast Airlines, Continental, Delta, Northwest, Pinnacle, United, USAir, and USAir Express.

The participants ranged in age from 23 to 53, with a mean age of 39. Thirteen were male, one was female. They reported an average of nearly 8,000 total flight hours, including military experience. They reported an average of 2,437 flight hours as Captain and 4,587 hours as First Officer. All had previous experience with simulators, with a mean of 255 hours.

The participants rated the complexity of their typical flights between 3 and 10 on a 10-point scale, with an average rating of 7.5. They rated their level of satisfaction with the current communications system between 2 and 8 on the same scale, with an average rating of 5.1. Their level of motivation to participate in the study was high, with a mean of 9.2 and ranging between 7 and 10. For these scales, 1 represented the lowest rating possible, and 10 represented the highest rating possible.

2.2 Test Facility

We connected three simulation facilities to conduct this experiment. The facilities included two flight simulators and an ATC simulator, all located at the William J. Hughes Technical Center (WJHTC) but in different buildings. The Reconfigurable Cockpit Simulator (RCS) is a fixed-base transport aircraft simulator with a computer-generated external visual scene. For this study, the transport aircraft simulator was configured with an airplane model, avionics, and two-member crew flight deck layout that emulated a Boeing 747-400 (B-747) aircraft. The General Aviation Trainer (GAT) is a motion-based Cessna 421 (C-421) aircraft simulator, with atmospheric turbulence capability. For this study, the C-421 was configured for single-pilot flight.

The ATC simulation configuration consisted of the Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE) and the Target Generator Facility (TGF). DESIREE emulates Display System Replacement (DSR) functions and receives input from the TGF to display radar targets. We used a single DSR workstation configured for single controller operation. The controller workstation includes a high-resolution Sony 2K display monitor, en route keyboard and trackball, headsets with microphones, and push-to-talk (PTT) handsets or foot pedals. DESIREE and the TGF used in this study are located in the Research Development and Human Factors Laboratory (RDHFL) of the WJHTC.

We linked the TGF to the two flight simulators and to 12 pseudopilot workstations. During the study, 10 – 12 pseudopilots supported the simulation, with 11 available for the majority of the test scenarios[1]. Each pseudopilot typically handled about one to two aircraft at a time, a configuration intended to increase demand for the channel and maximize the realism of the pilot communications. The pseudopilot workstations were located in the RDHFL and were separated from one another as much as possible. Most of the pseudopilot workstations were located in one experiment room; however, two were located in a second room. We instructed the pseudopilots to pay attention only to information heard on the communications channel.

The TGF controlled aircraft maneuvers based on simulation pilot and pseudopilot entries and on scripted flight plan data. Figure 1 depicts a schematic representation of the relationship between the various facilities and the simulation equipment involved in the study. The bold lines depict the information flow of aircraft data, whereas the narrow lines depict communications flow.

2.3 Voice Communications Systems

We used a Yamaha D5000 Digital Delay System to implement the communications system delays used in the study. To simulate current analog voice communications, the simulated system provided an audio delay of approximately 70 ms ground-to-air and air-to-ground, with no delay between the pseudopilot positions. In the actual analog system, when two or more simultaneous transmissions occur, a high-pitched squeal can be heard by everyone monitoring the frequency. Each speaker attempting to transmit hears only his or her own voice. In our simulated analog system, no one heard a squeal. Instead, others monitoring the system heard the voices of those attempting to transmit. As in the actual system, those attempting to transmit heard only their own voices.

---

[1] During one of the scenarios, 10 pseudopilots began the test session, but only 9 were available to complete the last half.

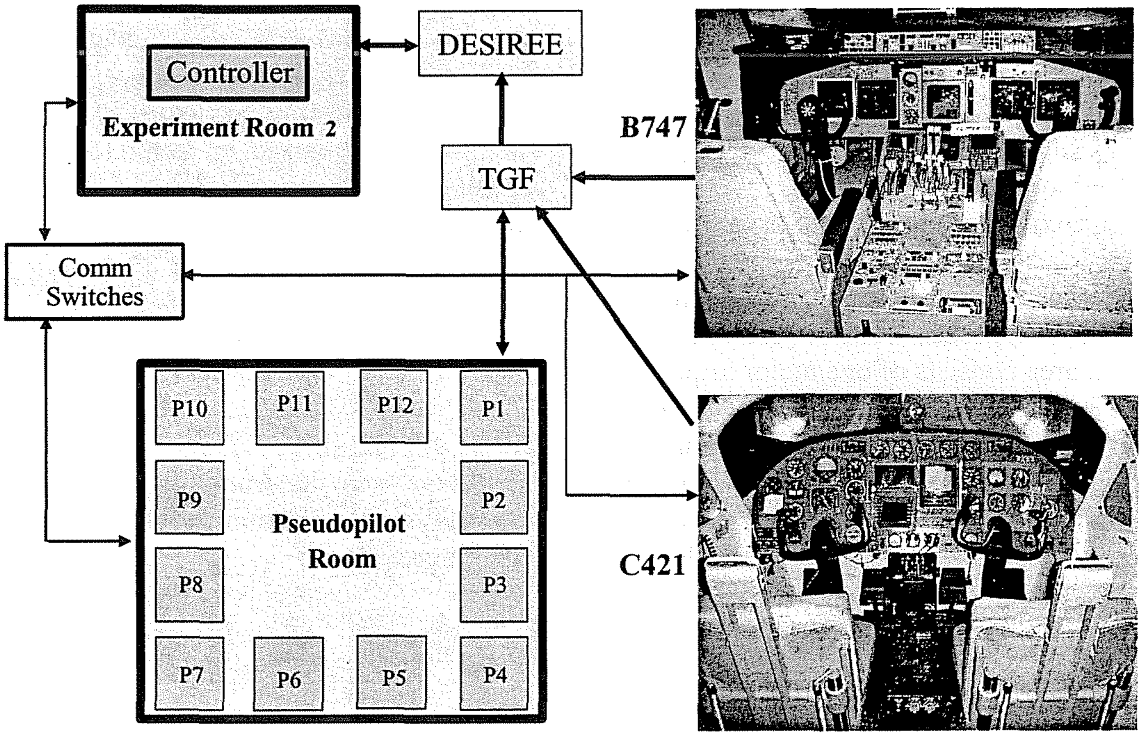


Figure 1. Schematic depiction of aircraft simulators and test facilities.

To simulate the VDL3 system, we implemented a 350 ms ground delay. The total transmission propagation and avionics processing delays resulted in an end-to-end, ground-to-air communication delay of 390 ms and an air-to-ground delay of 360 ms. We further modified the communications system to implement the controller override, antiblocking, and transmit status indicator features exactly as they had been in NEXCOM I. Controller override occurred immediately when the controller pressed the PTT key[2]. The transmit status indicator consisted of an audio signal to the pilot that occurred within 500 ms after the pilot attempted to transmit while the communication channel was occupied[3]. The indicator was a 1 kHz tone that was on for 500 ms and off for 500 ms. The pilots were able to hear communications on the channel when the transmit status indicator was operating. The antiblocking feature prevented any other pilot transmission from occurring when the controller or another pilot was already on the channel.

---

[2] In the actual VDL3 system, the controller override activation will be slightly delayed based on system latency and the system configuration and timing state.

[3] In the actual VDL3 system, the timing of the onset of the transmit status indicator may be slightly different.

2.4 Airspace

The airspace for the simulation was the same as that used in the previous study (Sollenberger et al., 2003) and in other human factors research (e.g., Yuditsky, Sollenberger, Della Rocco, & Friedman-Berg, 2003). Guttman and Stein (1997) created the generic en route airspace to be realistic, yet relatively easy to learn (see Figure 2). The airspace represents an en route, low-altitude transitional sector extending from the surface to 23,000 feet. It is somewhat rectangular in shape, approximately 120 nm north to south and 85 nm east to west. There are five jet routes for arrival, departure, and overflight aircraft. The sector also includes five restricted areas, which were not activated for this study. Arrival aircraft flow generally southbound into GENERA Terminal Radar Approach Control (TRACON), located in the southeast portion of the sector. The terminal area consists of one major airport, Genera, and three satellite airports, Midtown, Downtown & Uptown. Arrival aircraft are handed off to the TRACON between 8,000 ft and 14,000 ft depending on the type of aircraft and destination via the ILL (Illinois) and SGF (Springfield) arrival fixes. The sector also handles north and northwest departures from the TRACON for transition to higher cruise altitudes. Center overflights are normally assigned jet routes; however, point-to-point routes are also used to enhance realism.

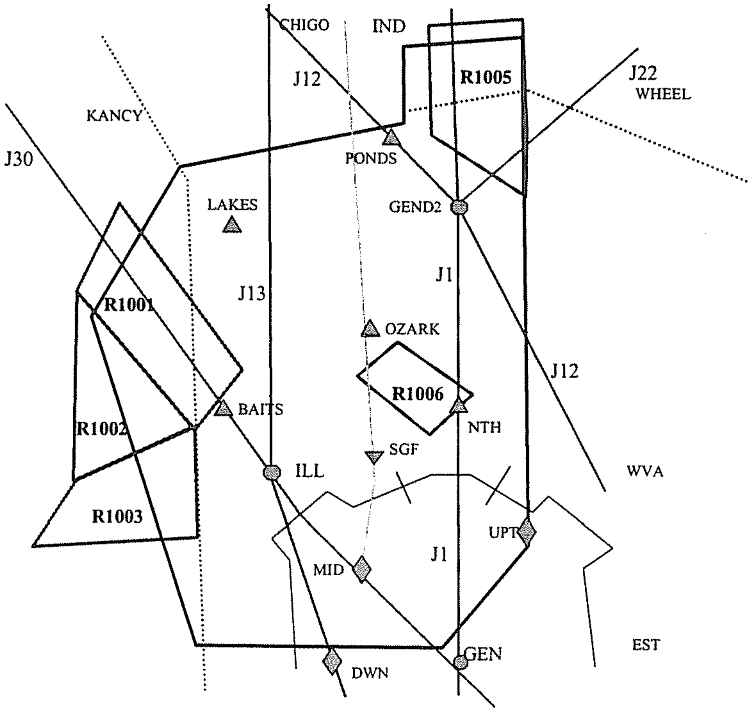


Figure 2. Generic en route sector map.

2.5 Traffic Scenarios

We used the traffic scenarios from NEXCOM I (Sollenberger et al., 2003) to generate scenarios for the practice and test scenarios in the current study. The test scenarios were 60 minutes in duration[4] and represented a communications-intensive environment. Each test scenario consisted of high traffic volume with 94 total aircraft (54 arrivals, 11 departures, and 29 overflights). All test scenarios had the same traffic volume with similar levels of difficulty and complexity; however, aircraft callsigns, spacing, and sequencing were different. For this study, the practice scenario was 30 minutes long, which was enough time for the participants to experience a demonstration of each of the digital system features and become acclimated to the 350 ms voice delay.

During each test scenario, the participants flew 5 of the 94 aircraft in the scenario. The participant in the B-747 flew in three flight segments (one overflight, one arrival, one departure), whereas the participant in the C-421 flew in only two flight segments (one arrival, one departure) because of the slower speed of the C-421 aircraft through the sector.

2.6 Experimental Design

We used a two-factor design with two levels of the communications system (analog and digital) and two levels of the environmental condition (routine and weather). We manipulated both factors within subjects so that each participant experienced all four combinations of communication system and environmental test conditions. To minimize variability in controller performance, one supervisory en route ATCS from a field facility acted as the controller throughout the practice and test scenarios for each group of participants. Two controllers assisted over the course of the study. One controller worked with the first 10 participants, the second worked with the last four.

We assigned half of the participants to the B-747 and half to the C-421. An experienced airline pilot, currently working for the WJHTC, served as the pilot flying (PF) the B-747 throughout all test scenarios. The participant assigned to the B-747 served as the pilot not flying (PNF) and was responsible for handing communications with the controller in support of the PF as needed. The C-421 participant operated the flight controls in the simulator and handled all communications.

We counterbalanced the order of the flight segments across participants. For the B-747, six flight segment orders were possible (departure, arrival, overflight; departure, overflight, arrival; arrival, departure, overflight; arrival, overflight, departure; overflight, departure, arrival; overflight, arrival, departure). For the C-421, two flight segment orders were possible (departure, arrival; arrival, departure). Therefore, we designed 12 scenarios to accommodate all

---

[4] Test scenarios actually ran for 65 minutes, as implemented in NEXCOM I, to ensure that we captured at least 60 minutes of PTT data. We analyzed the communications data starting with the first transmission in the scenario and ended the analysis after 60 minutes.

possible B-747 and C-421 flight segment order combinations. We also counterbalanced the order of the analog and digital conditions across the four test scenarios to reduce any effects of sequencing. The complete experimental design is shown in Table 1.

Table 1. Experimental Design

Radio Condition	Weather	Flight Segment Order	
		C-421	B-747
		Participant 1	**Participant 2**
Analog	No	Arrival; Departure	Arrival; Departure; Overflight
Digital	No	Departure; Arrival	Departure; Arrival; Overflight
Analog	Yes	Departure; Arrival	Overflight; Departure; Arrival
Digital	Yes	Arrival; Departure	Overflight; Arrival; Departure
		Participant 3	**Participant 4**
Digital	No	Departure; Arrival	Arrival; Overflight: Departure
Analog	No	Arrival; Departure	Overflight; Departure; Arrival
Digital	Yes	Arrival; Departure	Departure; Arrival; Overflight
Analog	Yes	Departure; Arrival	Departure; Overflight; Arrival
		Participant 5	**Participant 6**
Analog	No	Arrival; Departure	Departure; Overflight; Arrival
Digital	No	Departure; Arrival	Overflight; Arrival; Departure
Digital	Yes	Arrival; Departure	Arrival, Overflight, Departure
Analog	Yes	Departure; Arrival	Arrival, Departure; Overflight
		Participant 7	**Participant 8**
Digital	No	Arrival; Departure	Overflight; Departure; Arrival
Analog	No	Departure; Arrival	Overflight; Arrival; Departure
Analog	Yes	Arrival; Departure	Departure; Arrival; Overflight
Digital	Yes	Departure; Arrival	Departure; Overflight; Arrival
		Participant 9	**Participant 10**
Analog	No	Departure; Arrival	Arrival; Overflight; Departure
Analog	Yes	Arrival; Departure	Overflight; Arrival; Departure
Digital	No	Arrival; Departure	Departure; Overflight; Arrival
Digital	Yes	Departure; Arrival	Arrival; Departure; Overflight
		Participant 11	**Participant 12**
Digital	No	Arrival; Departure	Arrival; Departure; Overflight
Digital	Yes	Departure; Arrival	Overflight; Departure; Arrival
Analog	No	Departure; Arrival	Departure; Arrival; Overflight
Analog	Yes	Arrival; Departure	Arrival; Overflight; Departure
		Participant 13	**Participant 14**
Digital	No	Departure; Arrival	Arrival; Overflight: Departure
Analog	No	Arrival; Departure	Overflight; Departure; Arrival
Digital	Yes	Arrival; Departure	Departure; Arrival; Overflight
Analog	Yes	Departure; Arrival	Departure; Overflight; Arrival

The participants always worked with a communications system under the routine condition before working with it under the weather condition. For the weather scenarios, we provided the pseudopilots with scripted prompts to indicate when the weather events would occur and what communications needed to be made from the designated aircraft. We instructed the pseudopilots to call in their scripted reports of weather or requests for deviations as indicated by the time on their script sheets. As soon as the weather was scheduled to occur, the B-747 PF prompted the PNF participant to request a deviation from the controller. In the C-421, we instructed the participant to request a deviation immediately after the onset of turbulence. During the introductory briefing, we instructed the participants to assume high severity in these situations and to make the request for a new clearance as soon as possible. We scheduled the weather events to occur simultaneously for the B-747 and C-421.

2.7 Procedure

The participants signed an informed consent form that described the study and their rights and responsibilities as participants (see Appendix B). The participants also completed a background questionnaire to collect information about their piloting experience that may relate to their performance and assessment of the communications systems (see Appendix C).

Each pair of participants spent approximately 1½ days at the WJHTC. Two participants, one in the C-421 and one in the B-747, completed the scenarios over this period. On the first day, the participants attended the introductory briefing session in which one of the researchers provided an overview of the study and a description of the participants' roles and the experimental procedures. We escorted the participants to their respective simulators and gave them a familiarization overview on their cockpit. Then, we demonstrated the VDL3 350 ms delay and basic system features during the practice session. The participants listened to activity on the channel and communicated with the controller to become acclimated to this experimental condition. We asked them to report when they were ready to conclude practice and begin the first test session. They completed all practice sessions within 20 minutes. The participants completed at least one test scenario on the first day of participation. On the second day, they completed the remaining test scenarios and attended a final debriefing session.

2.7.1 Training

We thoroughly briefed the participants on the properties of the airspace, flight routes, and test procedures upon arriving at their respective simulators. A technical expert from the C-421 laboratory trained the C-421 participants to fly the simulator. To further increase workload for those pilots in the C-421, we instructed them not to use the autopilot feature in that simulator.

2.7.2 Data Collection

Each participant assumed responsibility of the first designated aircraft at the beginning of a scenario and completed all necessary procedures until the aircraft exited the airspace. A Human Factors Specialist (HFS) observed each participant in the C-421 and B-747 and recorded the time and type of communications attempted throughout the scenarios, any blocks or overrides experienced, the participants' comments, and other observations. The HFS observers recorded all data on a form specially designed for the study (see Appendix D).

The observers included HFSs with experience pertinent to the environment of each simulator. For the C-421 simulator, the observer was a general aviation pilot and, for the B-747 simulator, the observer had extensive experience in transport line operations. An additional HFS observer was also present in the C-421.

Immediately after completing each flight segment, the HFS asked the participant to provide a measure of workload using a 1 to 10-point rating scale (1= Extremely Low; 10=Extremely High). After a break of approximately 7 to 10 minutes, the participant took responsibility for the next designated aircraft in the scenario. The participants took a 15 to 20-minute break between scenarios. We video and audio recorded all scenarios.

After each test scenario, the participants rated the extent to which the communications system affected various tasks. They also assessed the operational acceptability of the system using a modified version of the Controller Acceptability Rating Scale (Lee, Kerns, & Bone, 2001). After the final test scenario, the participants provided an overall assessment of the scenarios, system realism, and the VDL3 basic system features. They also discussed these and other issues in a final debriefing session with the HFS observers.

2.8 Schedule

We conducted the study during January and February, 2003. Five pairs of participants completed testing between January 29 and February 6. Because of inclement weather, we cancelled testing for the last 2 of the 12 pilots originally scheduled for this simulation. When we rescheduled the sixth pair, we also recruited a seventh pair to increase our sample size to 14. We conducted the additional test sessions on February 24, 25, and 26.

Table 2 shows the schedule for the first week of testing . While one pair of participants was completing their final test scenario and debriefing session, the next pair of participants arrived at the WJHTC for the introductory briefing. We continued the schedule for Week 2, except that we cancelled the last pair. The final two pairs of participants used the same schedule.

Table 2. Week 1 Schedule

Time	Day 1 Wednesday, Jan. 29	Day 2 Thursday, Jan. 30		Day 3 Friday, Jan. 31
8:00 – 9:30	Test Run Equipment check	Group 1: Test Session 2	Group 2 Participants 3 & 4 Travel to Tech Ctr.	Group 2: Test Session 2
9:30 – 9:45	Group 1 Participants 1 & 2 Travel to Tech Ctr.	Break		Break
9:45– 11:15		Group 1: Test Session 3		Group 2: Test Session 3
11:15–12:30	Lunch	Lunch	Lunch	Lunch
12:30–2:00	Inbriefing	Group 1: Test Session 4	Inbriefing	Group 2: Test Session 4
2:00 – 2:15	Break	Group 1: Exit Questionnaire Final Debriefing	Break	Group 2: Exit Questionnaire Debriefing
2:15 – 2:45	Group 1: Familiarization with Equipment		Group 2: Familiarization with Equipment	
2:45 – 3:15	Group 1: Practice Session	Group 1: Travel home	Group 2: Practice Session	Group 2: Travel home
3:15 – 5:00	Group 1: Test Session 1		Group 2: Test Session 1	

2.9 Dependent Measures

We collected a battery of objective and subjective measures to assess the impact of the alternative communications systems and environmental conditions.

2.9.1 Communication Measures

The communications system recorded the time of each transmission for the participants, pseudopilots, and controller. For purposes of the analyses, we defined a transmission as any PTT action from key press to key release, regardless of whether there was any communication made. Therefore, we included actual transmissions as well as transmission attempts. We analyzed the data for each scenario to obtain the following measures.

2.9.1.1 Total Number of Transmissions

We calculated the total number of transmissions made per scenario, as well as the total number made by the pilots and the total number made by the controller.

2.9.1.2 Proportion of Overlapping or Conflicting Transmissions

We defined overlapping transmissions as those that occurred when at least one other user keyed the microphone while another was already keying. For example, if a pilot or the controller keyed the microphone and another pilot keyed before the first released, that resulted in two overlapping transmissions. We calculated the proportion of overlapping transmissions by dividing the number of overlaps by the total number of transmissions in the scenario. We defined this measure the same way for the analog and digital conditions.

2.9.1.3 Proportion of Unsuccessful Transmissions (Step-ons, Blocks, and Overrides)

We defined unsuccessful transmissions as those that that did not result in complete access to the channel from PTT key press to key release. We calculated unsuccessful transmissions differently for the digital and analog radio conditions because of the way that each system allows users to access the channel.

In the digital condition, only pilot transmissions could be unsuccessful. These resulted because a pilot was either blocked by the controller or by another pilot already occupying the channel, or because a pilot was overridden by the controller. A controller override occurred when the controller initiated a transmission while a pilot was already on the channel.

For controller overrides, we further distinguished between those that were unintentional and those that were intentional. We defined unintentional overrides as those that occurred when the controller pushed to talk during the delay time (within 360 ms) before it was possible for the controller to begin hearing the pilot speak. We defined intentional overrides as those that occurred when the controller initiated a transmission more than 360 ms after the pilot had keyed.

In the analog condition, both the pilots and the controller could be unsuccessful because of step-ons. If one user keyed before another released, both attempts were considered unsuccessful. This meant that any transmission involved in an overlap was also unsuccessful. We measured both controller-pilot step-ons and pilot-pilot step-ons.

2.9.1.4 Channel Occupancy Rate

We defined the channel occupancy rate as the total amount of time that the user occupied the channel during each 60-minute scenario. We measured the channel occupancy rate differently for the analog and digital conditions because of the way that each system allows users to access the channel.

In the analog condition, the channel is occupied as long as at least one user is keying the microphone. If one or more users keyed before another released, we defined channel occupancy for that interval as the time from the first key press until the last key release.

In the digital condition, we calculated channel occupancy by summing the durations of each PTT action from key press to key release but excluded the durations of any blocked transmissions because these did not result in channel access. When a pilot was on the channel and subsequently overridden by the controller, we added the duration of the unobstructed portion of the pilot transmission to the duration of the controller transmission to measure channel occupancy for that interval because the initial portion of the pilot transmission did have access to the channel.

2.9.1.5 Average Duration of Pilot Transmissions.

We determined the average length of successful pilot transmissions for each scenario by summing the durations of any transmissions that were not blocked, overridden, or involved in a step-on and dividing by the total number of successful pilot transmissions.

2.9.1.6 Urgent Pilot Request Intervals

We examined the time taken for pilots to receive a clearance around the weather in the non-routine analog and digital conditions. We defined this interval as the time until the pilot request for a reroute was granted, with the start of that interval defined as the time the PF told the PNF to request a reroute in the B-747, or the time the turbulence started in the C-421.

2.9.1.7 Average Duration of Controller Transmissions and Average Number of Controller Transmissions Made per Aircraft

We calculated the average duration of successful controller transmissions in each scenario as well as the average number of transmissions made per aircraft. The latter was defined by taking the total number of controller transmission and dividing by the number of aircraft in each scenario.

2.9.2 Workload Measures

The pilots provided workload ratings using two different techniques. First, they provided unidimensional workload ratings (Stein, 1985; Stein, 1991) after each flight segment. As part of the Post-Scenario Questionnaire (see Appendix E), they also completed the NASA Taskload Index (TLX), a multidimensional rating for which we used five of the six subscales (Hart & Staveland, 1987). Both techniques used 10-point scales requiring participants to make ratings from extremely low (1) to extremely high (10).

2.9.3 Pilot Subjective Ratings

After each scenario, the participants rated the perceived effect of the communications system on various tasks as part of the Post-Scenario Questionnaire. In addition, they rated their awareness of the position and status of other aircraft and the weather, and assessed the controller's performance and the difficulty of the scenario. The participants also provided a rating as to how well they perceived the system to have performed by assessing the degree of problems or deficiencies experienced using the Pilot Acceptance Rating Scale (see Appendix F). Each of the forms used 10-point rating scales in which 1 represented the lowest rating and 10 represented the highest.

At the conclusion of the study, the participants compared the analog and digital systems directly as part of the Exit Questionnaire (see Appendix G). They provided ratings as to which system, if either, they perceived as better at handling communications tasks using a 5-point rating scale. A 1 indicated that the analog system was "much better," 5 indicated that the digital system was "much better," and 3 indicated that there was no difference between them. They also rated the usefulness of each VDL3 basic system feature and the extent to which each feature interfered with their ability to communicate using a 10-point scale.

2.9.4 Air Traffic Measures

The TGF recorded the position and status of all aircraft every second during the test scenarios. These data were processed using the Data Reduction and Analysis Tool (DRAT), allowing measurements to be made on the total flight time and distance of each aircraft; the number of altitude, heading, and airspeed changes; and the number of events representing a loss of standard separation (see Buckley, DeBaryshe, Hitchner, & Kohn, 1983; or Stein & Buckley, 1992, for a description of how the measures are computed).

3. Results

We defined all communications data and aircraft measures per scenario. We chose an alpha level of .1 as our criterion for determining whether there were statistically significant differences between these variables as a function of test condition because of the low statistical power that resulted from our relatively small sample size[5]. The alpha level indicates the rate at which the results would be expected to occur by chance, rather than to real differences between the test conditions. For the subjective measures (e.g., workload, system suitability), in which data from each of the 14 individual participants was analyzed, alpha was set to .05.

We used both multivariate and univariate statistics to analyze the data. We conducted multivariate analyses when we determined two or more variables to be components of a higher order metric. For example, we combined all of the workload measures from the NASA TLX and post-flight segment workload ratings and analyzed these using a multivariate analysis of variance (MANOVA). This analysis addresses concerns about inflated alpha levels when conducting multiple individual tests. If the multivariate results were significant, we then conducted univariate tests. We present the univariate results for easier interpretability. In cases where the variables within a metric were highly correlated, we could conduct only univariate tests. For example, we had planned to analyze rates for channel occupancy, overlapping transmissions, and unsuccessful transmissions together, but because of the high correlation between the proportion of overlapping and unsuccessful transmissions, we tested the overlapping transmission rate separately.

We used proportions for many of the communications measures (e.g., proportion of unsuccessful transmissions). Proportions can be problematic in that these data may not form a normal distribution, making it inappropriate to assess differences using statistical tests such as an analysis of variance. To correct for this effect, we used an arcsin transformation to normalize the data. Analyses of both the transformed values and the original proportions yielded equivalent results in terms of whether we found statistically significant differences. We present the proportions to allow for easier interpretability.

3.1 Between-Study Comparison

We compared the digital, routine test condition from the current study to the 350 ms delay condition in the previous study (Sollenberger et al., 2003). We compared the number of transmissions, the rates of overlapping and unsuccessful transmissions, the channel occupancy rates, and the proportions of blocks and overrides.

Figure 3 shows the number of total, controller, and pilot transmissions made in NEXCOM I and NEXCOM II. Neither the total number of transmissions nor the number of pilot transmissions differed significantly between the studies. However, there were significantly more controller transmissions made in NEXCOM I [$F(1,15) = 4.16, p = .059$]. Controllers made about 300 transmissions on average in NEXCOM I compared to 253 in NEXCOM II. In each study, the controllers made over 40% of the transmissions.

---

[5] Data for six participant groups were analyzed because of missing data from the analog radio, no weather condition for one group. Additionally, for one of the six groups, communications data for the C-421 were missing due to an equipment problem during the analog, no weather scenario. We examined the DVD recording from this session to determine how many actual transmissions were made by the C-421. The data were analyzed with and without the additional transmissions, and we did not find any differences in the outcome of the statistical tests on the available measures.

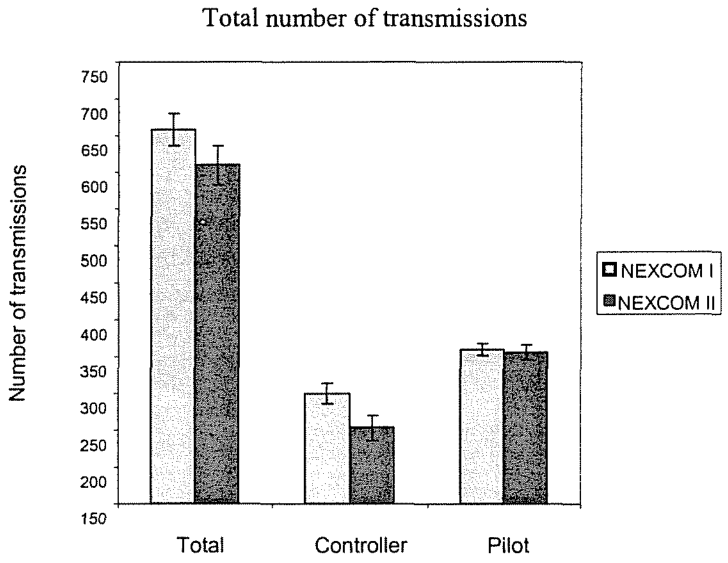


Figure 3. Number of total transmissions, controller transmissions, and pilot transmissions in NEXCOM I and NEXCOM II.

The channel occupancy rate was significantly higher in NEXCOM I than NEXCOM II [$F(1,15) = 8.99, p = .009$]. The channel was occupied for an average of about 34 minutes in NEXCOM I, with a standard deviation (*SD*) of 3.9 minutes, and about 29 minutes (*SD* = 1.6) in NEXCOM II.

The rate of overlapping transmissions was significantly higher in NEXCOM II, $F(1,15) = 6.19, p = .025$, indicating that there was more competition for the channel despite lower channel occupancy in that study. About 15% of the total transmissions overlapped with one another in NEXCOM II (*SD* = 3%), whereas about 10% overlapped in NEXCOM I (*SD* = 4%).

The proportion of unsuccessful transmissions also differed significantly between the two studies, with more unsuccessful attempts made in NEXCOM II [$F(1,15) = 4.18, p = .059$]. Because only pilot transmissions could be unsuccessful, we examined the proportion of unsuccessful transmissions with respect to the total number of pilot transmissions rather than the total number of transmissions made in the scenario. About 14% of pilot transmissions were unsuccessful in NEXCOM II (*SD* = 3%) compared to about 10% of pilot transmissions in NEXCOM I (*SD* = 4%).

We also found differences between the studies with respect to the proportions of blocks and overrides. The left side of Figure 4 shows the total proportion of controller overrides and the proportion that were intentional and unintentional in each study. The right side of Figure 4 shows the total proportion of blocks and the proportion of blocks by the controller and blocks by the pilot in each study. There was a significant difference between the studies on all measures ($p < .1$). There were more controller overrides and more unintentional overrides in NEXCOM I.

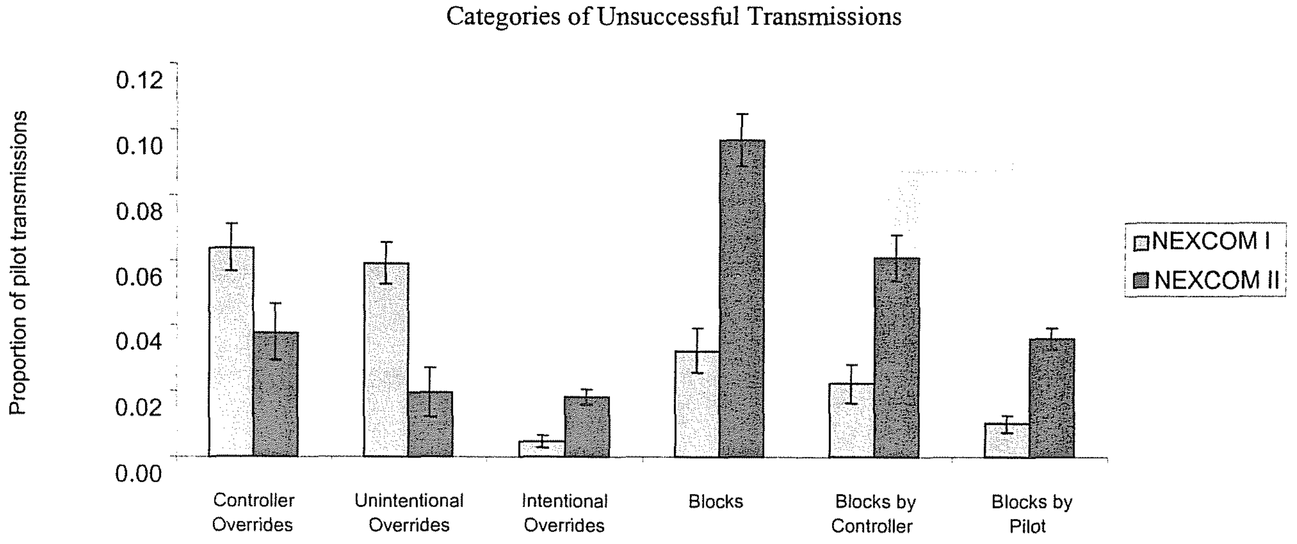


Figure 4. The proportion of overrides (total, unintentional, and intentional) and blocks (total, blocks by controller, and blocks by pilots) in NEXCOM I and NEXCOM II.

A little over 6% of pilot transmissions were overridden in NEXCOM I, whereas about 4% of pilot transmissions were overridden in NEXCOM II. There were more total blocks and more blocks by the controller and pilots as well as more intentional overrides in NEXCOM II. Approximately 10% of pilot transmissions were blocked in NEXCOM II, whereas about 4% were blocked in NEXCOM I.

Overall, we found fairly comparable levels of communications activity, in that the overall number of transmissions did not differ significantly between the two studies. Despite this, there were differences between the studies in terms of the number of controller transmissions, channel occupancy time, unsuccessful transmission rate, and the proportions of blocks and overrides. We attribute these differences to two factors. One is that the controllers in NEXCOM II were highly practiced and familiar with the scenarios as well as the VDL3 system. As a result of their experience, they likely issued fewer clearances, thus reducing the number of transmissions, the channel occupancy rate, and the number of controller overrides. The second factor is that there were more pilots competing for the channel in the second study, which resulted in a greater number of blocks.

3.2 NEXCOM II: Communications Data

We compared the number of transmissions, the rate of overlapping and unsuccessful transmissions, and the channel occupancy rates between the analog and digital conditions, with and without weather. We examined the pilot data further by comparing the rate of unsuccessful pilot transmissions and the average transmission duration between the test conditions, and the time needed for the simulator pilots to receive a clearance around weather in both radio conditions. We examined the controller data by comparing the rate of unsuccessful transmissions (analog condition only), the average number of controller transmissions per aircraft, and the average transmission duration between the test conditions. We also analyzed the characteristics of the overlapping transmissions in each radio condition to determine when users were attempting to access the channel with respect to other transmissions.

3.2.1 Number of Transmissions

There was no difference in the total number of transmissions, the number of controller transmissions, or the number of pilot transmissions made between the radio conditions (see Figure 5). There was an average of 648 transmissions made in the analog condition and an average of 653 transmissions in the digital condition. Controllers made fewer transmissions than pilots. Overall, controllers made an average of 270 transmissions, whereas pilots made an average of 381. Controllers made an average of 276 transmissions in the analog condition and 263 in the digital condition. Pilots made an average of 372 transmissions in the analog condition and 390 in the digital condition.

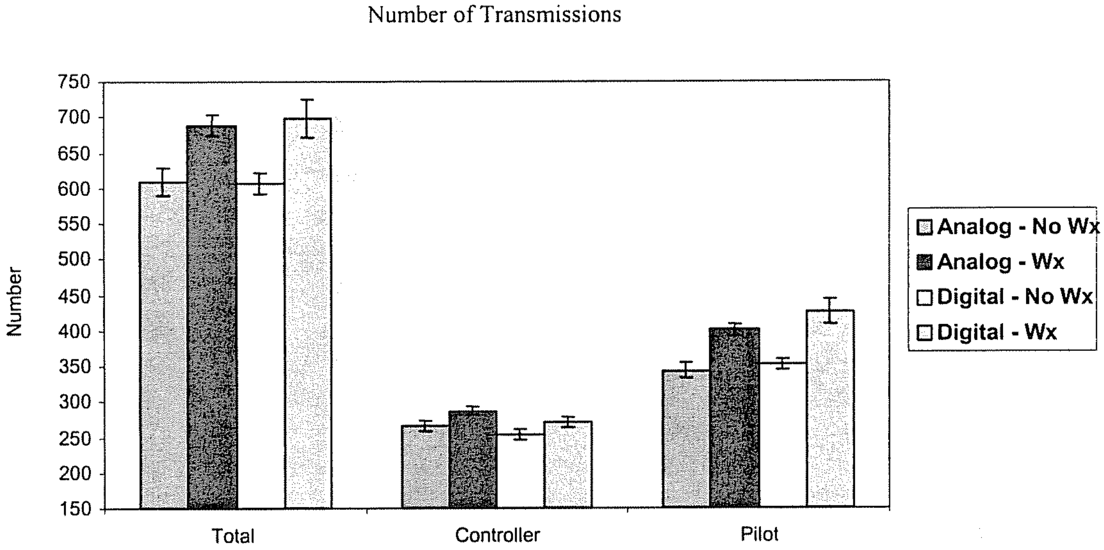


Figure 5. Number of total transmissions, controller transmissions and pilot transmissions as a function of radio and environmental conditions.

The total number of transmissions was significantly higher when weather was involved [$F(1, 5) = 55.64$, $p = .001$], as was the number of controller transmissions [$F(1,5) = 9.14$, $p = .029$], and pilot transmissions [$F(1,5) = 147.78$ $p < .001$]. There was an average of 693 total transmissions made when weather was involved, compared to 609 under routine conditions. Controllers made an average of 279 transmissions in the weather conditions and 260 in routine conditions, whereas pilots made an average of 414 transmissions in the weather conditions and 348 in the routine conditions. The interaction between the radio and environmental conditions was not significant.

3.2.2 Channel Occupancy

The channel occupancy rates did not differ significantly between the analog and digital conditions (see Figure 6). The channel was occupied for about 31 and 32 minutes, respectively, for the digital and analog conditions, about half the duration of the scenario.

Channel occupancy rates were, however, significantly higher under the weather conditions [$F(1,5) = 19.40$, $p < .007$]. The channel was occupied for about 33 minutes in the weather conditions and for about 30 minutes in the routine conditions.

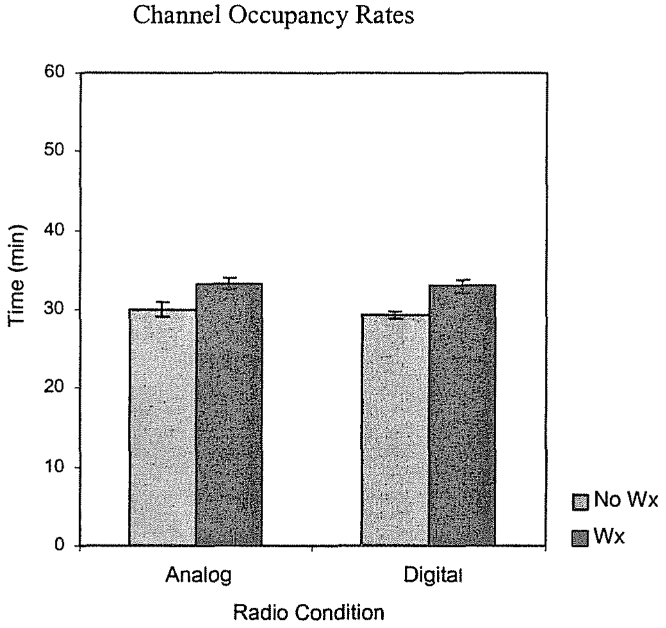


Figure 6. Channel occupancy rate as a function of radio and environmental conditions.

3.2.3 Overlapping and Unsuccessful Transmissions

Figure 7 shows the rates of overlapping and unsuccessful transmissions on the left and right sides, respectively. We based the proportions on the total number of transmissions made in the scenario. The rates of overlapping transmission did not differ between the analog and digital conditions. About 18% of the transmissions overlapped in the analog condition, whereas about 17% overlapped in the digital condition. This result suggests that the general way in which users attempted to access the channel did not differ as a function of which radio they used.

The rate of overlapping transmissions was, however, significantly higher when weather was involved [$F(1,5) = 104.69, p < .001$]. About 21% of the transmissions overlapped in the weather conditions, whereas about 15% overlapped when users communicated under routine conditions.

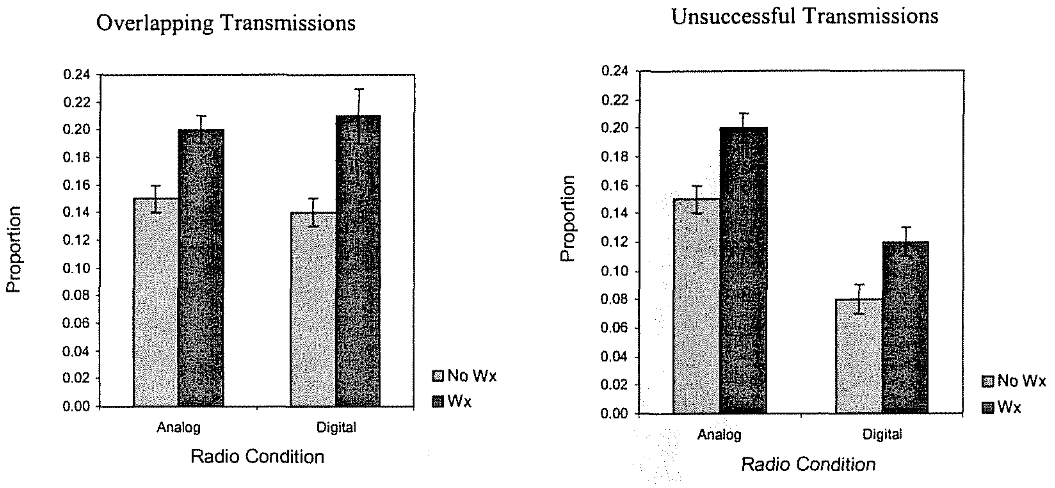


Figure 7. The rate of overlapping and unsuccessful transmissions as a function of radio and environmental conditions.

The rate of unsuccessful transmissions did differ as a function of radio condition, as expected based on the definition of unsuccessful transmissions and the similar rate of overlaps between the two conditions. There were significantly more unsuccessful transmissions in the analog condition [$F(1,5) = 129.97$, $p < .001$]. Approximately 18% of the total number of transmissions made were unsuccessful in the analog condition compared to about 10% in the digital condition. Significantly more unsuccessful transmissions were also made when weather was involved [$F(1,5) = 77.08, p < .001$]. Approximately 16% of the transmissions were unsuccessful in the weather conditions, whereas about 12% were unsuccessful in the routine conditions. The interaction between the radio and environmental conditions was not significant.

We also examined the rate of unsuccessful pilot transmissions separately (see Figure 8) to allow for a more direct comparison of the effects of the test conditions because only pilot transmissions could be unsuccessful with the digital radio. The results were analogous to those presented in Figure 7. The proportion of unsuccessful pilot transmissions was significantly higher in the analog condition [$F(1,5) = 7.87, p = .038$], and significantly higher when there was weather involved [$F(1,5) = 85.92$, $p < .001$]. Approximately 19% of pilot transmissions were unsuccessful in the analog condition compared to 16% in digital. About 21% of pilot transmissions were unsuccessful in the weather conditions, whereas about 15% were unsuccessful under routine conditions. The interaction between the radio and environmental conditions was not significant.

These results indicate that pilots had more unobstructed access to the channel when using the digital radio, despite the antiblocking and controller override features. The addition of weather in the scenarios increased the rate of unsuccessful pilot transmissions, but it did so similarly for each radio condition.

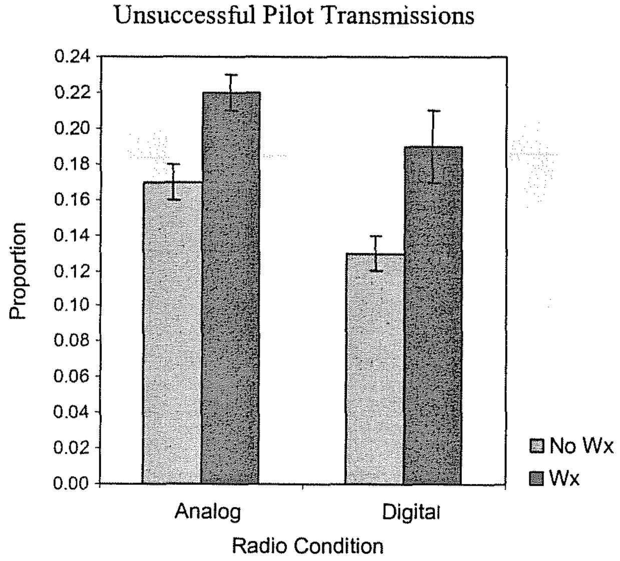


Figure 8. The rate of unsuccessful pilot transmissions as a function of radio and environmental conditions.

The average duration of successful pilot transmissions did not differ significantly between the radio conditions. Transmissions averaged about 2.8 seconds in analog (*SD* = .17) and in digital (*SD* = .12). The durations did differ significantly between the environmental conditions [$F(1,5) = 7.82, p < .05$]. Although statistically significant due to the low variability involved, the difference was small. The average duration of pilot transmissions was 2.7 seconds under routine conditions and 2.8 seconds when weather was involved.

3.2.4 Urgent Pilot Transmissions

We examined the time taken for pilots to get a clearance around weather in the analog and digital conditions. We measured the time between the events that signaled the start of weather for the B-747 and C-421 participants and the time the controller issued them a clearance to deviate. The cockpit observers recorded these data.

There were two separate weather events in each weather scenario[6]. Figure 9 shows the distribution of these individual clearance times for each radio condition. It took about 73 seconds to get a clearance in the digital condition and about 86 seconds in the analog condition, a

---

[6] Three of the individual data points were omitted. In one of the digital scenarios, the controller issued a clearance to the C-421 before the participant had the opportunity to request a deviation. In one of the analog scenarios, the controller's microphone had unplugged resulting in an added delay confounding this measure for both participants.

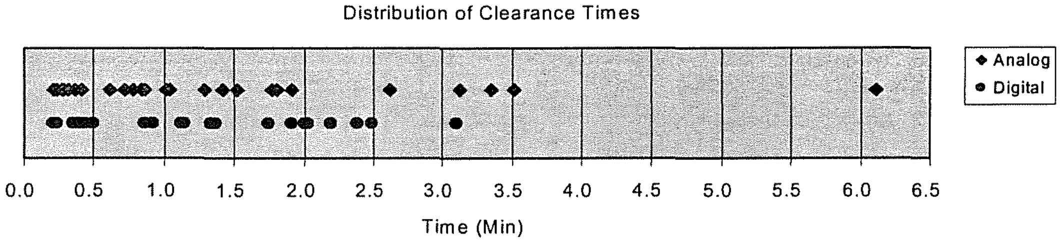


Figure 9. Clearance times for the analog and digital conditions.

difference that was not statistically significant. The time to receive a clearance was somewhat more variable in the analog condition, ranging from 13 seconds to over 6 minutes, with 5 of the 26 clearances taking over 2.5 minutes. In the digital condition, all clearances except one were obtained in less than 2.5 minutes.

We also analyzed these data using a 2x2x2 design, with simulator, radio condition, and weather event within the scenario as factors. This determined whether there was a difference in the time to receive a clearance as a function of simulator. The C-421 participants received a clearance in an average time of 102 s, whereas the B-747 participants received a clearance in about 53 s, a difference that was statistically significant [$F(1,9) = 5.67, p < .05$]. None of the other factors or interactions were significant.

The differences in clearance times between simulators may have been due to workload. Participants in the C-421 flew the aircraft and handled communications, whereas those in the B-747 were responsible for communications only. The C-421 participants may have been occupied by other tasks when the turbulence started, causing them to delay their calls. The B-747 participants may have been more ready to respond to the PF when instructed to request a deviation because no other tasks interfered.

The weather events resulted in congestion on the channel. We found that the channel occupancy rate during these weather events was higher than the rate we observed for the scenarios as a whole. The channel was occupied about 74% of the time during the weather events compared to channel occupancy rates of slightly more than 50% over the full scenario. The channel occupancy rates during the weather events did not differ significantly between the radio conditions.

3.2.5 Controller Transmissions

The average number of controller transmissions did not differ between the radio conditions. There were about 2.9 transmissions made per aircraft in the analog condition ($SD = .18$) and about 2.8 transmissions in digital ($SD = .21$). However, more transmissions were made per aircraft when weather was involved [$F(1,5) = 9.14, p = .029$]. About three transmissions were made per aircraft in the weather conditions ($SD = .20$), whereas about 2.8 were made under routine conditions ($SD = .19$). Although statistically significant due to the low variability involved, the difference was small.

By definition, controller transmissions were always successful in digital conditions. If the controller was already on the channel, any pilot attempting to transmit would be blocked. If the controller keyed the microphone when a pilot was on the channel, the override feature would be activated, enabling the controller to gain access. In analog, both pilot and controller transmissions could be unsuccessful. Overall, 16% of controller transmissions were unsuccessful in the analog condition and significantly more of the transmissions were unsuccessful in the weather condition [$F(1,5) = 4.86$, $p = .079$]. Fourteen percent of controller transmissions were unsuccessful under routine conditions and 18% were unsuccessful when weather was involved.

3.2.6 Characteristics of Overlapping Transmissions

We analyzed overlapping transmissions in more detail to determine at what time during a transmission pilots were blocked, overridden, or stepped on. We analyzed the time at which the controller overrode a pilot in the digital condition, and comparably, at what time the controller stepped on a pilot in the analog condition. Likewise, we analyzed at what point a pilot attempted to get on a channel that was already occupied in both radio conditions. In the digital condition, a pilot attempting to access an already-occupied channel would be blocked, and in the analog condition, that pilot would step on the controller or another pilot.

In the digital condition, 65% of the 237 controller overrides were unintentional. They occurred within the first 360 ms of a pilot key press, the length of the downlink delay, during which time it would not have been possible for the controller to begin hearing the pilot speak (see Figure 10). Of the 35% of controller overrides that we categorized as intentional (i.e., made after the first 360 ms of a pilot key press), 88% were made within the subsequent 360 ms, or 720 ms from the pilot key press. If we assume that the pilot did not speak immediately upon keying the microphone, the effective delay for the controller to begin to hear the pilot speak would be somewhat longer than 360 ms. It is likely, therefore, that some of the overrides we categorized as intentional were probably initiated before the controller actually heard the pilot speak. Of the few overrides made 2 or more seconds into a pilot transmission, all but one was made within the last 230 ms of the pilot transmission. The remaining override occurred 3.8 seconds into the pilot transmission and 550 ms from the end of that transmission. In cases where the controller keyed very near the end of the pilot transmission, it is likely that the pilot had already finished speaking, but had not yet released the microphone.

The comparable situation in the analog condition occurs when the controller steps on a pilot transmission (see Figure 11). For easier comparability, the same scale is used as that for the controller override data. Fourteen percent of the controller step-ons occurred within the first 70 ms of a pilot key press, the length of the delay during which time the controller would not have been able to hear the pilot begin speaking. If we again assume that the pilot did not begin to speak immediately upon keying the microphone, then the effective delay for the controller to begin to hear the pilot speak would be somewhat longer.

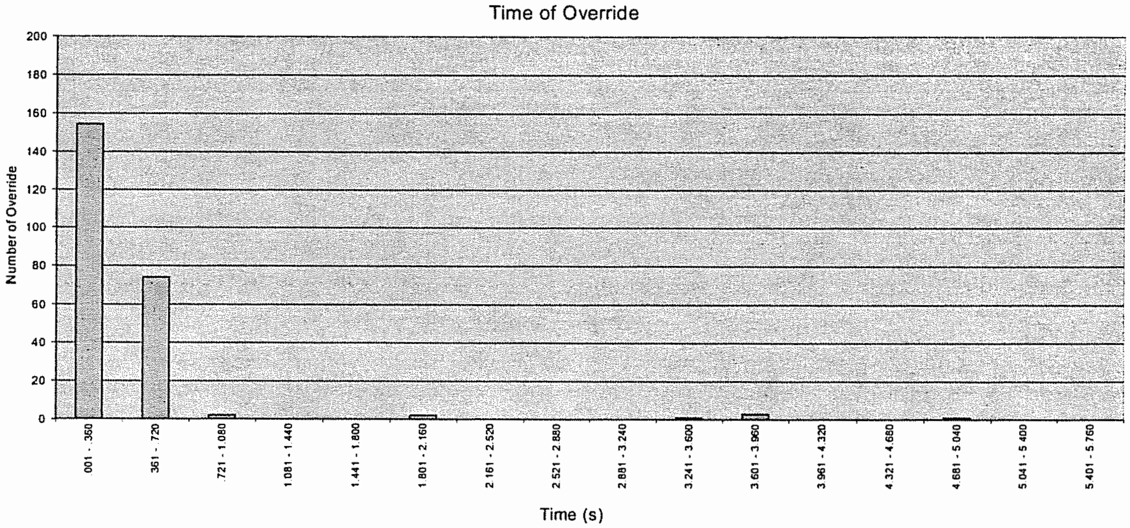


Figure 10. Time between pilot key press and controller override.

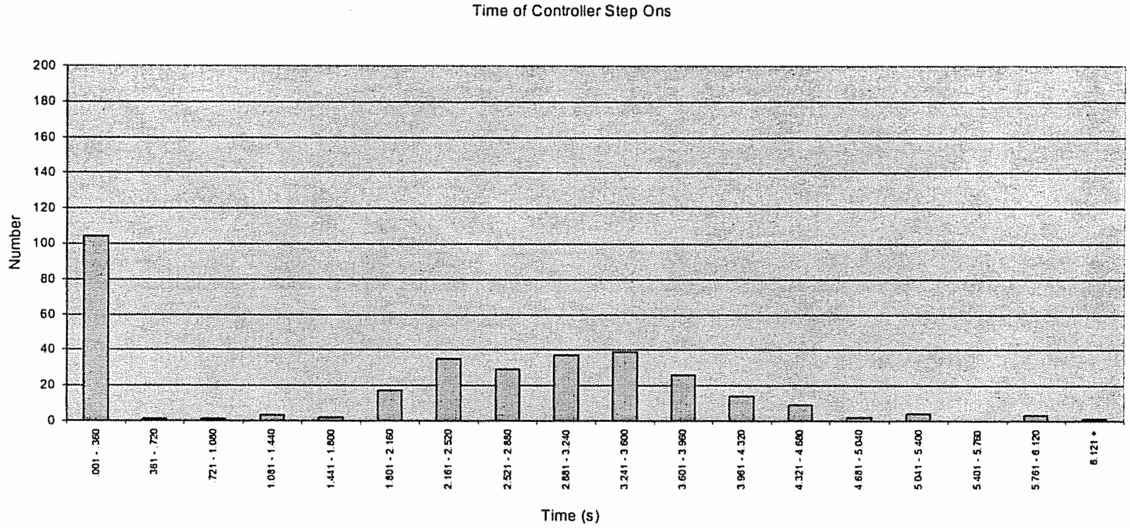


Figure 11. Time between pilot key press and controller step-on.

Overall, 32% of the 327 controller step-ons occurred within the first 360 ms of the pilot key press, relatively early into a pilot transmission. Unlike the digital condition, however, about half of the step-ons occurred more than 2 seconds after the pilot keyed the microphone. We further analyzed these data to determine if these step-ons occurred towards the middle or end of the pilot transmission. We found that about 64% of these step-ons occurred within the last 360 ms of a pilot transmission. This result is not surprising. We would not expect to find controllers deliberately stepping on a pilot transmission because it would not allow them unobstructed access to the channel. It is possible in these situations that the pilot had finished talking but had not yet released the microphone key.

We also analyzed the time between the start of a pilot key press with respect to the start of a transmission already occupying the channel (see Figure 12). In the digital condition, this resulted in the pilot transmission being blocked. In the analog condition, it meant that the pilot stepped on another pilot or the controller. For easier comparability, we present these data using the same scale as that used previously for overrides and step-ons. However, because these data include blocks by pilots and step-ons between pilots, the uplink and downlink delays do not apply to those transmissions. We found that 72% of pilot attempts to access the channel when it was occupied occurred fairly early in either radio condition, within the first 720 ms of those transmissions. On average, these pilot attempts occurred about 2.5 s *(SD* = 1.76) from the end of these transmissions.

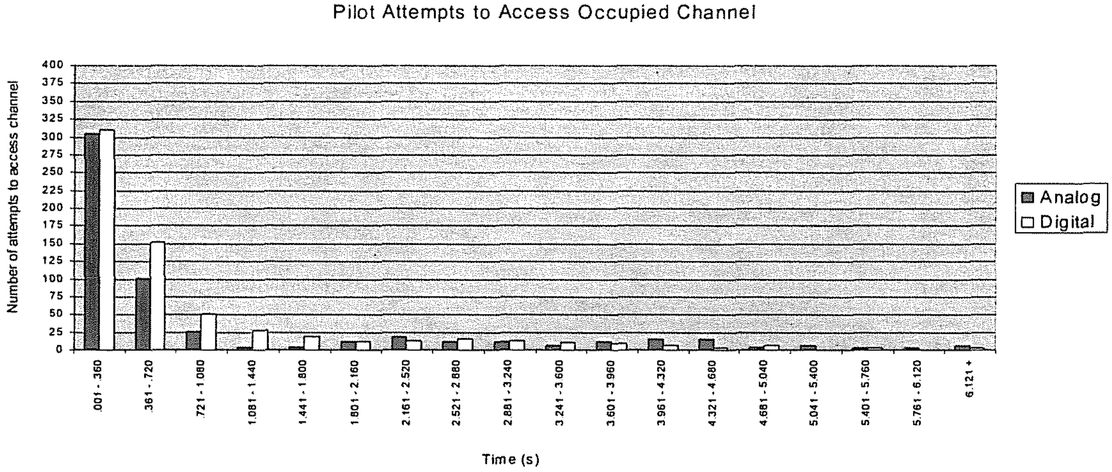


Figure 12. Time between pilot key press and start of a transmission already occupying the channel in the analog and digital conditions.

3.3 Air Traffic Measures

We measured the average flight time and distance of the aircraft and the number of altitude, speed, and heading changes made in each scenario. We analyzed the data using a MANOVA and found that there were no significant differences as a function of the test conditions. Overall, the average flight time per aircraft was about 9 minutes (*SD* = .29) and the average distance flown was about 50 nm (*SD* = 1.92) within the sector[7].

There were more altitude changes made than speed or heading changes. There were an average of 84.5 altitude changes (*SD* = 6.13), 14.8 speed changes (*SD* = 4.03), and 17.3 heading changes (*SD* = 4.84) made overall in each scenario.

---

[7] We analyzed data from 6 groups because altitude, heading, and speed data from an analog, no weather scenario for one group was not available.

3.4 Subjective Measures

3.4.1 Workload Ratings

The participants rated their workload at the end of each flight segment. The B-747 participants provided three flight segment ratings per scenario, and the C-421 participants provided two. We averaged these ratings for each participant to obtain a single measure for each test condition. The participants also provided workload ratings for mental demand, physical demand, temporal demand, effort, and frustration using the NASA-TLX at the end of each scenario. We analyzed these data using a MANOVA and determined that there was no statistically significant difference in the ratings across test conditions. Overall, workload ratings tended to be low, with averages between one and three, with the exception of the "Effort" variable, which they rated moderately at about four (see Figure 13). The participants' comments indicated that they found the communications workload fairly typical for the en route environment that we simulated, but lower than they would expect in a terminal environment.

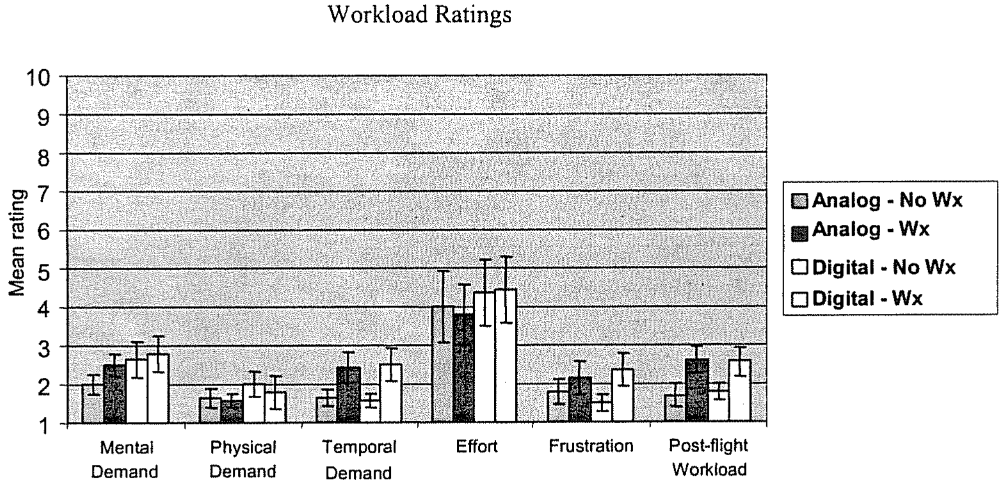


Figure 13. Mean workload ratings as a function of radio and environmental conditions.

3.4.2 Pilot Acceptance Ratings

Mean pilot acceptance ratings of each system were high, indicating that they considered both highly acceptable with minimal deficiencies (see Figure 14). The ratings for the digital system were significantly higher than those for analog [$F(1,12) = 21.13, p = .001$]. The average rating for the digital system was 9.6. The average rating for the analog system was 8.7. The difference in the ratings between the environmental conditions was not significant, nor was the interaction of radio by environmental condition. Ninety percent of the confidence ratings were high, whereas 10% were medium.

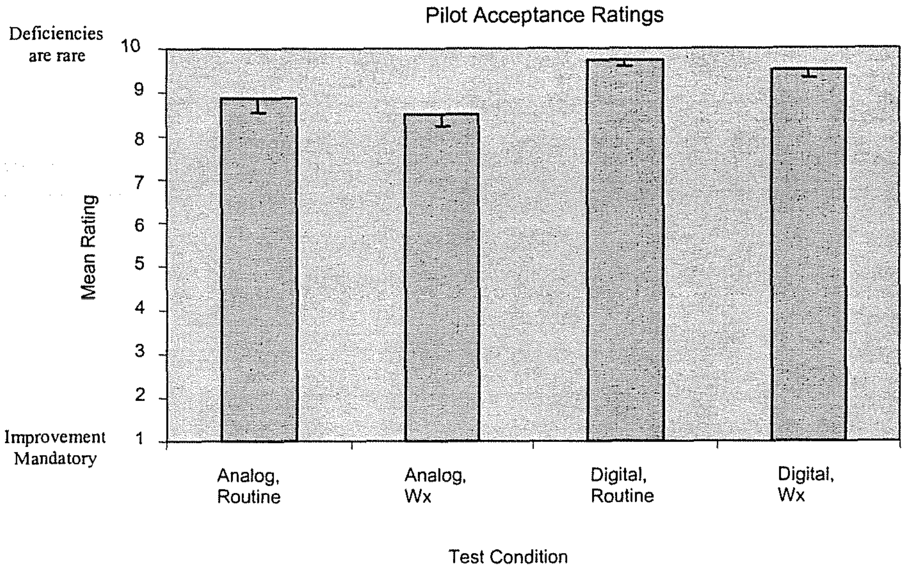


Figure 14. Mean system acceptability ratings as a function of radio and environmental conditions.

3.4.3 Communication Task Ratings

We analyzed the participants' ratings as to the extent that each communications system used in each scenario interfered with various communications tasks using a MANOVA. We found no statistically significant differences between these ratings as a function of test condition. Nearly all ratings averaged less than three (see Figure 15), indicating that participants attributed only a small negative effect to either of the systems.

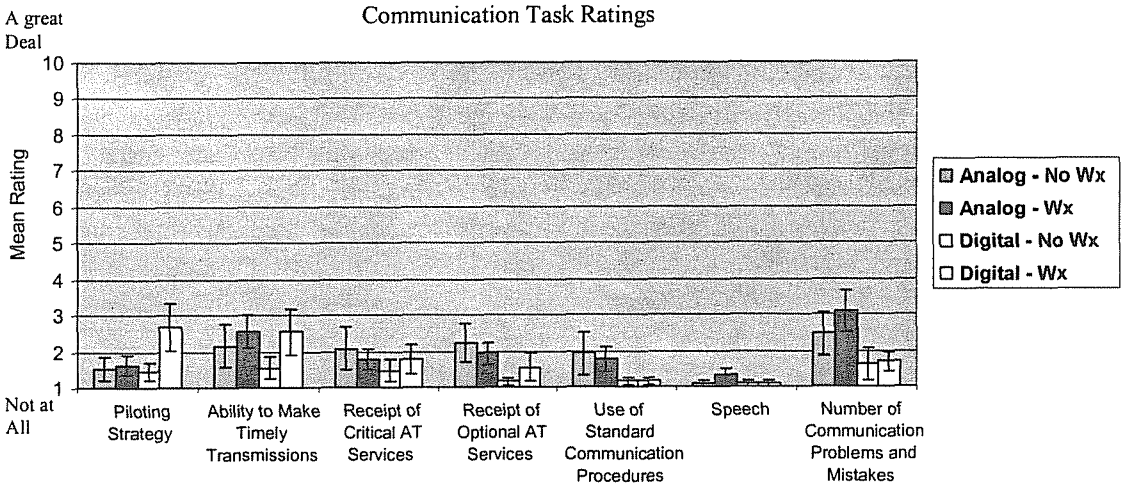


Figure 15. Mean communication task ratings as a function of radio and environmental conditions.

3.4.4 System Comparison Ratings

Table 3 shows the participants' ratings comparing the analog and digital systems on 10 communications functions. The ratings reflect the extent to which participants indicated a system allowed them to perform a function "much better than" or "somewhat better than" the other, or whether there was no difference between the two.

All but two of the 140 individual ratings indicated that the participants perceived the digital system as equivalent to or better than analog. Of the total number of responses, 92 (66%) indicated that the participants performed functions either "somewhat better" or "much better" with the digital system. Those functions that the participants indicated were most effectively accomplished with the digital system were "determining when the channel was busy," "hearing complete messages," and "being confident that the controller received a message." Most participants indicated that the digital system was better overall at "accomplishing all communications tasks." The two individual ratings that favored the analog system came from the same participant. This participant commented that he felt the receipt of transmissions was faster with analog.

Table 3. Distribution of Responses (and Means) for System Comparison Ratings

	Much Better with Analog	**Somewhat Better with Analog**	**No Difference between Analog & Digital**	**Somewhat Better with Digital**	**Much Better with Digital**	**Means (St. Dev.)**
	①	②	③	④	⑤	
Completing routine pilot-initialed radio calls	0	0	5	6	3	3.86 (0.77)
Completing time-critical communications	0	1	5	4	4	3.79 (0.97)
Responding to air traffic controller calls	0	0	8	4	2	3.57 (0.76)
Determining when the channel was busy	0	0	2	2	10	4.57 (0.76)
Determining when the channel was available	0	0	7	4	3	3.71 (0.83)
Being confident controller received message	0	0	3	4	7	4.29 (0.83)
Detecting communications problems	0	0	6	4	4	3.86 (0.86)
Hearing complete messages	0	0	1	5	8	3.79 (1.05)
Receiving timely responses	0	1	6	2	5	3.79 (1.05)
Accomplishing all communications tasks	0	0	3	5	6	4.21 (0.82)
Total responses	**0**	**2**	**46**	**40**	**52**	

3.4.5 Digital System Feature Ratings

The participant ratings about the usefulness of the VDL3 basic features were very high (see Figure 16). Ninety percent of the ratings were eight or higher, and all but one were seven or higher.

The extent to which the antiblocking and controller override features negatively affected communications was rated fairly low on the 10-point scale, with average ratings of 3.5 (*SD* = 2.7) and 2.4 (*SD* = 1.5), respectively. Individual responses varied, however. Nine participants rated the negative impact of the antiblocking feature at three or less, whereas four participants

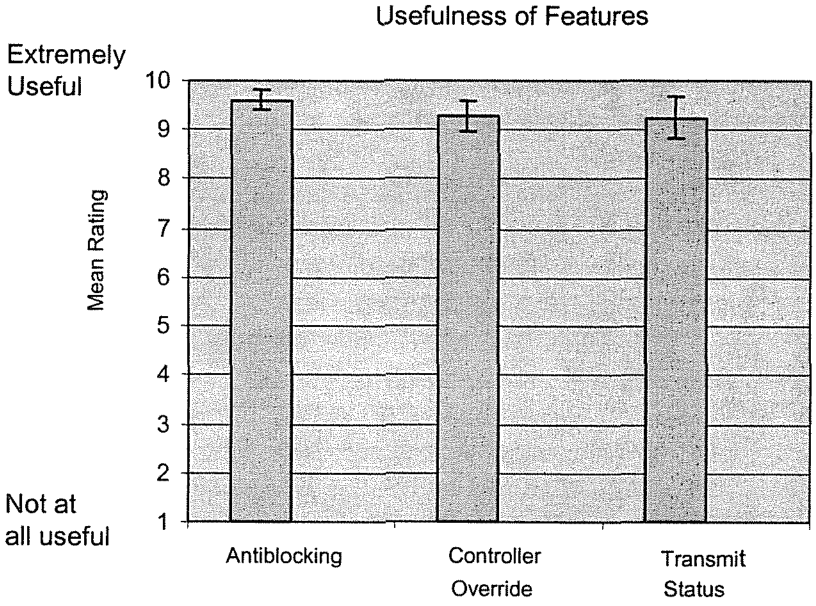


Figure 16. Mean usefulness ratings of basic VDL3 system features.

rated the impact as moderate, with ratings between five and seven. The remaining participant rated the negative impact of antiblocking a 10; however, we believe this may have been a misinterpretation of the scale because the same participant also gave the feature a usefulness rating of 10.

Ten participants rated the negative effect of controller override as very low, with ratings of one or two. The remaining four participants rated it low to moderate, with ratings of four and five. The participants' comments indicated that the override feature would be especially helpful for resolving "stuck mic" situations and for organizing the communications traffic on the channel.

The participants' ratings of the acceptability of the volume, pitch, and on-off cycle rate of the transmit status indicator each averaged about 7. However, these ratings were highly variable, ranging from 3 to 10. There were a few specific comments made about the indicator that suggested that the tone should be louder and made more distinctive or attention-getting. The tone was harder to hear in the C-421 because of the loud engine noise in that simulator. The ratings on the effect of the digital system on safety averaged 4.7 ($SD = .5$) on a 5-point scale, indicating that participants felt the system would positively affect safety.

For these variables, we also examined whether participants in the B-747 and C-421 provided ratings that were significantly different from one another. We analyzed the data using a MANOVA and found a significant effect of simulator [$F(2,9) = 21.80, p < .05$]. The univariate analyses revealed that the B-747 participants gave the override feature more positive ratings than those in the C-421 [$F(1,12) = 5.08, p < .05$], and also rated the digital system as having a more positive effect on safety than those in the C-421 [$F(1,12) = 5.56, p < .05$].

3.4.6 Additional Participant Comments

During the debriefing session, the participants had the opportunity to discuss with the HFS observers their experiences in the test sessions and to offer their thoughts on the usefulness of the system in actual operations. The participants' comments were for the most part favorable with regard to VDL3 and the basic system features. The comments included that the system would "be a positive change" and "improve pilot situation awareness." Other comments were that "communication will be clearer with this design" and that, with the digital system, "it was almost like getting a receipt that your transmission went through; you know the message was sent."

Other comments, however, indicated some areas of concern. For example, one participant responded that the system "will require lots of training for the pilots," whereas another felt that he "had to wait too long to communicate with the digital system." Others focused on advanced features or other implementations that they believed would enhance the current design. For example, one participant responded that "it would be helpful if there was a light when you keyed the mic. Pilots need to know when the channel is free." Others responded that "it would be nice to have a way to cue the controller that you want to speak" and the "the downside to controller override is that pilots do not have a similar way to get to the controllers."

We also asked the participants about one of the advanced features, the urgent downlink request, planned for later implementation. This feature would allow pilots to have a way to contact the controller if a serious, yet non-emergency, situation arose and the channel was too busy to allow timely access. The participants in this study indicated that this feature would be highly desirable in these instances.

4. Conclusions

Fourteen certified airline pilots participated in a high-fidelity, human-in-the-loop simulation involving realistic flight deck simulators. These participants completed scenarios using simulated digital and analog radio communications systems under routine conditions and conditions that involved weather to further increase demand for the channel. The results of the study indicate that the VDL3 communications system with 350 ms voice throughput delay and basic features (antiblocking, controller override, and transmit status indicator) is an acceptable system for pilots.

Overall, the digital system allowed more successful transmissions to be made, though users accessed the channel similarly with both radio systems. We found that the number of transmissions, the amount of time the channel was occupied, and the number of overlapping transmissions did not differ between the analog and digital systems. Additionally, when weather was involved, these communications measures were affected similarly in both systems. Other measures, including the duration of average controller transmissions and pilot transmissions did not differ between the analog and digital radios, nor were there any differences found in the way the aircraft were handled between these conditions.

The participants were generally quite positive in their reactions to VDL3. System acceptability ratings, although high for both systems, were significantly higher for the digital system. In addition, almost all individual communications task ratings indicated that the digital system was at least as good as the analog system and usually was rated as better. The basic system features were also rated very highly and any negative effects of these features were rated fairly low. With respect to the controller override feature, we observed that the controller tended to initiate overrides quite early in the pilot transmission. In many cases, these occurred before it would have been possible for the controller to have begun to hear the pilot speak given the system downlink delay.

The results obtained in this study differed from those obtained in NEXCOM I in several respects. In NEXCOM II, we observed fewer controller transmissions, lower total channel occupancy time, more unsuccessful transmissions, and more blocks but fewer controller overrides. We attributed these differences to the controllers who were highly experienced with the scenarios and VDL3 system in the second study and to more pilots competing for the channel. On the basis of these findings, we recommend that future studies of flight deck issues include a large number of pilot participants to enhance competition for the channel.

References

Buckley, E.P., DeBaryshe, B.D., Hitchner, N. & Kohn, P. (1983). *Methods and measurements in real-time air traffic control system simulation* (DOT/FAA/CT-83/26). Atlantic City International Airport, NJ: DOT/FAA Technical Center.

Child, J., Cleve, R., & Grable, M. (1989). *Evaluation of low data rate voice CODECS for air traffic control applications* (DOT/FAA/CT-TN89/13). Atlantic City International Airport, NJ: DOT/FAA Technical Center.

Dehel, T., Grable, M., & Child, J. (1989). *Phase II testing and evaluation of low data rate voice CODEC equipment* (DOT/FAA/CT-TN89/49). Atlantic City International Airport, NJ: DOT/FAA Technical Center.

Federal Aviation Administration. (2000). *Requirements document for next generation air/ground communications system (NEXCOM), Segment 1.* Washington, DC: Federal Aviation Administration.

Federal Aviation Administration. (2001, July). *NEXCOM human factors plan*, Unpublished report. Washington, DC: Federal Aviation Administration.

Farncombe, D. (1997). *Investigations into the operational effects of the VDL Mode 3 voice throughput delay* (AMCP WG D/7 WP 53). Madrid, Spain: Aeronautical Mobile Communications Panel Working Group D.

Farncombe, D. & MacBride, J. J. (2000). *VDL Mode 3 flight tests using the AMBE-ATC-10B vocoder* (AMCP/7-WP/UK4). Montreal, Canada: Aeronautical Mobile Communications Panel.

Fujimori, T., & Ueno, M.(1999). *Results of vocoder evaluation in Japan* (AMCP/WGD10-WP9). Hawaii, USA: Aeronautical Mobile Communications Panel Working Group D.

Guttman, J., & Stein, E. S. (1997). *En route generic airspace evaluation* (DOT/FAA/CT-TN97/7). Atlantic City International Airport, NJ: DOT/FAA William J. Hughes Technical Center.

Hart, S. G., & Staveland, L. E. (1987). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock and N. Meshkati (Eds.), *Human Mental Workload*. Amsterdam: North-Holland.

LaDue, J., Sollenberger, R. L., Belanger, B., & Heinze, A. (1997). *Human factors evaluation of vocoders for air traffic control environments, Phase I: Field Evaluation* (DOT/FAA/CT-TN97/11). Atlantic City International Airport, NJ: DOT/FAA William J. Hughes Technical Center.

Lee, K.K., Kerns, K., & Bone, R. (2001). Development and validation of the controller acceptance rating scale (CARS): Evidence from empirical research. *Proceedings of the 4th USA/Europe Air Traffic Management R&D Seminar,* Santa Fe, NM

Nadler, E, DiSario, R., Mengert, P., Sussman, E.D., & Spanier, G. (1990). *A simulation study of the effects of communication delay on air traffic control* (DOT-TSC-FAA-90-5). Cambridge, MA: DOT Volpe National Transportation systems Center.

Nadler, E., Mengert, P., DiSario, R., Sussman, D. D., Grossberg, M., & Spanier, G. (1993). Effects of satellite- and voice-switching equipment transmission delays on air traffic control communications. *The International Journal of Aviation Psychology, 3*(4), 315-325.

Rantanen, E., McCarley, J. Xu, X., & Yeakel, S. (2001). *The effect of audio throughput delay of the NEXCOM system on air traffic controller performance and workload (*Technical Report ARL-01-19/FAA-01-5). Savoy, IL: University of Illinois at Urbana-Champaign.

Renaud, P., Fistas, N., Brugere, C., & Garcia, A. (1999). *Results of vocoder intelligibility analysis for European languages.* (AMCP/WG D/11 WP), Toulouse, France: Aeronautical Mobile Communications Panel Working Group D.

RTCA DO-225. (1994). *VHF air-ground communications system improvements alternatives study and selection of proposals for future action.* Washington, DC: RTCA, Inc.

Sollenberger, R. L., LaDue, J., Carver, B., & Heinze, A. (1997). *Human factors evaluation of vocoders for air traffic control environments, Phase II: ATC simulation* (DOT/FAA/CT-TN97/25). Atlantic City International Airport, NJ: DOT/FAA William J. Hughes Technical Center.

Sollenberger, R. L., McAnulty, D. M., & Kerns, K. (2003). *The Effect of voice communications latency in high density, communications-intensive airspace: Final report.* Atlantic City International Airport, NJ: DOT/FAA William J. Hughes Technical Center.

Stein, E. S. (1985). *Graphic simulation and the automated en route air traffic control concept: An FAA Technical Center preliminary study* (DOT/FAA/CT-TN85/29). Atlantic City International Airport, NJ: Federal Aviation Administration Technical Center.

Stein, E. S. (1991). *Air traffic controller memory -- a field survey* (DOT/FAA/CT-TN90/60). Atlantic City International Airport, NJ: Federal Aviation Administration Technical Center.

Stein, E.S., & Buckley, E.P. (1992). Unpublished Simulation Variables.

Suganuma, M. (1997). *Investigation of influences on ATC communications by the delay time generated in voice communications using VDL* (AMCP WG D7 WP 59). Madrid, Spain: Aeronautical Mobile Communications Panel Working Group D.

Williams, J., Eck, J. & Eckstein, B. (2001). *Why VDL-3? The rationale behind the FAA's technology choice for NEXCOM.* Washington, DC: Federal Aviation Administration.

Yuditsky, T., Sollenberger, R. L., Della Rocco, P. S., & Friedman-Berg, F. (2003). *Application of color to reduce air traffic complexity* (DOT/FAA/CT-TN03/01). Atlantic City International Airport, NJ: DOT/FAA William J. Hughes Technical Center.

Acronyms

ALPA	Air Line Pilots Association
ATC	Air Traffic Control
ATCS	Air Traffic Control Specialist
DESIREE	Distributed Environment for Simulation, Rapid Engineering, and Experimentation
DRAT	Data Reduction and Analysis Tool
FAA	Federal Aviation Administration
GAT	General Aviation Trainer
HFS	Human Factors Specialist
IFR	Instrument Flight Rules
MANOVA	Multivariate Analysis of Variance
NEXCOM	Next Generation Air-Ground Communications
PF	Pilot Flying
PNF	Pilot Not Flying
PTT	Push-To-Talk
RCS	Reconfigurable Cockpit Simulator
RDHFL	Development and Human Factors Laboratory
SD	Standard Deviation
TGF	Target Generator Facility
TLX	Taskload Index
TRACON	Terminal Radar Approach Control
VDL3	VHF Digital Link Mode 3
VHF	Very High Frequency
WJHTC	William J. Hughes Technical Center

Appendix A

PARTICIPANT RECRUITMENT LETTER

The Federal Aviation Administration is planning a simulation in January-February 2003 to evaluate a new digital air-ground radio system in comparison to the current analog radio system. The simulation will be conducted at the William J. Hughes Technical Center located at the Atlantic City (NJ) International Airport. We are recruiting ALPA pilots to participate in the simulation in either our fixed-base Boeing 747 simulator or our Cessna 421 simulator. The simulator pilots will be communicating with an air traffic controller in a low altitude, en route generic airspace sector. Other aircraft tracks will be generated by the Target Generation Facility (TGF) and controlled by TGF pilots to increase the amount of activity on the communications channel. Pilots in the 747 simulator will be flying as First Officers and will be responsible for handling the air traffic control communications. Pilots in the 421 simulator will be single pilots responsible for flying and communicating. All runs will originate or terminate at 6,000 ft or above. The focus will be on evaluating the new radio system rather than piloting skills.

Pilots interested in participating will need to spend two consecutive days on the study. On the first day, they will travel to the Technical Center in the morning and participate in an inbriefing beginning at 12 noon. After the briefing, they will receive simulator familiarization and a practice flight, followed by a one-hour flight scenario. The second day, they will participate in three additional one-hour flight scenarios. In each scenario, the 421 pilot will fly two runs through the sector and the 747 pilot will fly three runs. Half the flights will use analog radios and half will use digital radios. After each scenario, the pilots will fill out a short questionnaire about their experience in those runs. After the last scenario, they will also fill out an exit questionnaire about their overall experiences and their evaluation of the two radios. A final debriefing should be completed by 3 PM so the pilots can travel home. All data collected will be confidential, and the results will be reported only as group averages. We do not anticipate any risk to the participants, but they may withdraw at any time if they choose.

Pilots interested in participating should be commercial pilots, but need not be qualified in these specific aircraft models or current for IFR operations. Pilots may volunteer for either the 747 or 421 simulator or both, but will only be able to participate in one or the other. Pilots volunteering for the 421 role should have some light twin engine flying experience. They may also volunteer for specific dates.

Pilots will receive a stipend of $___ per day under a consulting agreement for their travel costs and participation in the study. Lodging costs in the Atlantic City area are reasonable during the "tourist off season" time period for this study.

Appendix B

INFORMED CONSENT

I, ______________________________, understand that this study, entitled "The Effect of Voice Communications Latency in High Density, Communications-Intensive Airspace. Phase II: Flight Deck Perspective and Comparison of Analog and Digital Systems" is sponsored by the Federal Aviation Administration and conducted at the William J. Hughes Technical Center (WJHTC).

Nature and Purpose:

I have been recruited to volunteer as a participant in this project. The purpose of the study is to determine the effects of the Very High Frequency Digital Link Mode 3 (VDL3) communications delay and additional features in a high-fidelity simulation. The results of the study will be used to determine whether and to what extent the delay and features impact pilot communications, workload, and system acceptance.

Experimental Procedures:

All participants will be briefed on the airspace and complete a 30-minute practice session prior to the start of the experiment. They will complete a background questionnaire to provide information on years of experience, etc., which may be useful in interpreting other aspects of the data. The participants will be encouraged to ask questions at any point during the practice and test sessions.

The participants will complete four test sessions in either the General Aviation Trainer (GAT) simulator or Reconfigurable Cockpit Simulator (RCS). Each of the four sessions is expected to take approximately 1 hour to complete. For those in the RCS, each session will include 3 passes through the airspace, one as an overflight, one as an arrival, and one as a departure. For those in the GAT, each session will include 2 passes through the airspace, one as an arrival and one as a departure. The order of these flight types will be varied across sessions. Participants will complete all practice and test sessions over a one and one-half day period, not including travel to and from the WJHTC.

The participants will complete questionnaires at the end of each test session to evaluate the impact of the delay and features on their perceived workload and acceptance. Finally, an automated data collection system will record communications so that data regarding missed transmissions, step-ons, overrides, etc. can be obtained.

Discomfort and Risks:

I understand that I will not be exposed to any foreseeable risks or intrusive measurement techniques.

Confidentiality:

My participation is strictly confidential, and no individual names or identities will be recorded or released in any reports.

Benefits:

I understand that the only benefit to me is that I will be able to provide the researchers with valuable feedback and insight into the effects of the VDL3 throughput delay and features on pilot communications, workload and system acceptability. My data will help the FAA to determine whether the VDL3 system is acceptable.

Participant Responsibilities:

I am aware that to participate in this study I must be a certified pilot and hold a current medical certificate. I will fly in one of the simulators indicated above and answer any questions asked during the study to the best of my abilities. I will not discuss the content of the experiment with anyone until the study is completed.

Participant Assurances:

I understand that my participation in this study is completely voluntary and I can withdraw at any time without penalty. I also understand that the researchers in this study may terminate my participation if they feel this to be in my best interest. I understand that if new findings develop during the course of this research that may relate to my decision to continue my participation, I will be informed.

I have not given up any of my legal rights or released any individual or institution from liability for negligence.

If I have questions about this study or need to report any adverse effects from the research procedures, I will contact Dr. Carolina Zingale at (609) 485-8629.

Compensation and Injury:

I agree to immediately report any injury or suspected adverse effects to Dr. Carolina Zingale at (609) 485-8629. Local clinics and hospitals will provide any treatment, if necessary. I agree to provide, if requested, copies of all insurance and medical records arising from any such care for injuries/medical problems.

Signature Lines:

I have read this informed consent form. I understand its contents, and I freely consent to participate in this study under the conditions described. I understand that, if I want to, I may have a copy of this form.

Research Participant:____________________________________ Date:__________

Investigator:____ Date:__________

Witness:____ Date:__________

Appendix C

BACKGROUND QUESTIONNAIRE

Participant number: ______ ....: Date:________________

Instructions:

This questionnaire is designed to obtain information about your background and experience as a pilot. The information will be used to describe the participants in this study as a group. You will not be identified by name.

Indicate responses by placing a check mark or "X" next to a selection, filling in a blank, filling in a circle, or circling one of the numbers provided on a rating scale.

Demographic Information and Experience

1. What is your **age**?	_____ years _____ months	
2. What is your **gender**?	○ Male	○ Female
3. Are you **current for IFR operations**?	○ Yes	○ No
4. Date (approximate) of **most recent** flight operation.	Date: ____________________	

5. What **aircraft** do you currently fly? (list all applicable)

____

____

____

____

6. Indicate the total **amount of experience** you have as a pilot.	_____ hours

7. If applicable, indicate the **amount of experience** you have as a	
a) captain	_____ hours
b) first officer / co-pilot	_____ hours

8. If applicable, indicate the **amount of experience** you have working in a simulator.	_____ total hours

9. Rate the **complexity** of your typical flights	Low Complexity	①②③④⑤⑥⑦⑧⑨⑩	High Complexity

10. Rate your level of satisfaction with the **current air-ground communications system.**	Completely Dissatisfied	①②③④⑤⑥⑦⑧⑨⑩	Completely Satisfied

General Ratings

11. Rate your **level of stress** about participating in this study.	Not Stressed	①②③④⑤⑥⑦⑧⑨⑩	Extremely Stressed

12. Rate your **level of motivation** to participate in this study.	Not Motivated	①②③④⑤⑥⑦⑧⑨⑩	Extremely Motivated

13. Please include any additional information below that you feel will better enable us to interpret your responses.

Appendix D

HUMAN FACTORS OBSERVER FORM

Directions:

- **Start the video recorder.**
- **Start the timer/stopwatch immediately when instructed by the experimenter at the start of each scenario.**
- **Record** ALL **pilot communications and comments as well as your own observations about events occurring during the scenario. For each, record the** TIME, DESCRIPTION, CODE **(see below), and** NUMBER of attempts **(if relevant).**

Example:

TIME	DESCRIPTION & COMMENTS	CODE	No. attempts
4:30	**Requests altitude change**	**S**	
4:55	**Repeats request for altitude change (controller did not respond to earlier attempt)**	**R**	**2**

- **Ask the participant for a workload evaluation after each flight segment using the scale provided on the form.**
- **At the end of each scenario**
    - **stop the video recording.**
    - **have the participant complete a Post-Scenario Questionnaire and Pilot Acceptance Rating Scale.**
- **At the conclusion of the last test scenario, have the participant complete the Exit Questionnaire.**

EVENT CODES:

- **Pilot request** *I.*
- **Pilot response** *(A).*
- **Time of successful initial contact with controller** *(I).*
- **Time of successful log off** *(L).*
- **Unexpected/other events** *(U).*

- **For** WEATHER **scenarios, record the start time of the weather** (SW) **event and the time at which the pilot request for a reroute around the weather is granted** *(W).*
    - **For the GAT the start time = onset of turbulence**
    - **For the RCS the start time = time the PF tells the participant to request a reroute.**

- **For** ANALOG **test scenarios, record inferred step-ons** *(S)***.**

Use the following cues as indicators that a step-on has occurred:

- **The other participant informs the pilot that the transmission has been stepped on.**
- **The controller does not respond to the pilot request.**
- **The pilot repeats a transmission.**

Record the number of times the participant needed to get the message to the controller (repeat transmissions) and/or the number of controller requests for repeats.

- **For** DIGITAL **test scenarios, record blocked transmissions** *(B)* **and controller overrides** *(O)***.**
    - **Record the attempted transmission and whether the participant was blocked or overridden just after the controller finished speaking.**
    - **Record the number of times the participant needed to get the message to the controller (repeat transmissions) and/or the number of controller requests for repeats.**
    - **Record whether the participant did/did not respond to the controller's message following a controller override of another pilot.**

Participant #:
Date:
Equipment: (circle) **GAT** **RCS**
Radio Condition: (circle) **Digital** **Analog**
Weather Condition: (circle) **Severe/Turbulent** **None**
Scenario:

Time	Description and Comments	Code	No. Attempts

End-of-flight workload assessment.

"Rate your workload due to communications during this flight:"

Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Appendix E

POST SCENARIO QUESTIONNAIRE

Participant #: ____________

Date:____________________

Scenario: ______________

Weather OR no weather

Instructions:
Please answer the following questions based upon your experience in the scenario just completed. Fill in one circle to indicate your level of response.

1. Rate your **level of awareness of the position and status of other aircraft** in this scenario.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

2. Rate your **level of awareness of the weather in the airspace** in this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

3. Rate the **performance of the air traffic controller** in terms of responding to your requests, etc. during this scenario.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

4. Rate the **difficulty** of this scenario.	Extremely Easy	①②③④⑤⑥⑦⑧⑨⑩	Extremely Difficult

5. Please list and comment on any major differences between your experiences this session in comparison to typical operations:

____

____

____

Specific Communications System Ratings

To what extent did the communications system used in this scenario…

6. … affect your typical piloting strategy or style, if applicable?	Not At All	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

7. … interfere with your ability to make timely transmissions?	Not At All	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

8. … interfere with your timely receipt of critical air traffic services?	Not At All	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

9. …affect your receipt of optional air traffic services?	Not At All	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

10. …interfere with your use of standard communication procedures (e.g., phraseology, readbacks)?	Not At All	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

11. …affect your speech (e.g., clarity, rate)?	Not At All	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

12. …impact the number of communications problems and mistakes?	Not At All	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

13. Please describe the type of communication problems and mistakes encountered during this scenario.

NASA-TLX Ratings

Definitions

Mental Demand – how much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Were your tasks easy or demanding, simple or complex, exacting or forgiving?

Physical Demand – how much physical activity was required (e.g., talking, pointing, etc.)? Were your tasks easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Temporal Demand – how much time pressure did you feel due to the rate or pace at which your tasks occurred? Was the pace slow and leisurely or rapid and frantic?

Performance – how successful do you think you were in accomplishing the goals of your tasks? How satisfied were you with your performance in accomplishing these goals?

Effort – how hard did you have to work (mentally and physically) to accomplish this level of performance?

Frustration – how insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel in performing your tasks?

14. Rate your **mental demand** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

15. Rate your **physical demand** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

16. Rate your **temporal demand** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

17. Rate your **performance** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

18. Rate your **effort** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

19. Rate your **frustration** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

20. Do you have any comments or clarifications about these NASA-TLX questions?

Appendix F

PILOT ACCEPTANCE RATING SCALE

Guidelines For Numerical Rating

Procedure:

1. Start at the top left-hand corner of the page.
2. Answer each yes/no question according to the test scenario that you just flew.
3. Use the definitions below to make the judgments.
4. Circle **one** number from 1 to 10 that **best** reflects your experience in the test scenario you just flew.
5. Provide a Confidence Rating according to the definitions below.
6. Please add comments to explain your rating.

Definitions

System:

The system is taken to mean **everything** being rated:

- The pilot's performance,
- The performance of the communications system, and
- The controller's performance

Confidence:

The **Confidence Rating** should describe confidence in the numerical acceptability rating itself. **It is <u>not</u> a rating of how confident one is <u>about</u> the system or operating environment.** The **Confidence Rating does** answer the question, **"How confident am I that the rating I just made is an accurate one, reflecting the overall system performance, based on the amount of information I had available to me?"**

The **Confidence Rating should reflect the amount of information you think you had available to you in making your overall rating.** It should also reflect problems that you encountered that are not necessarily an indication of how the system performed. As in the example above, a controller that is especially unresponsive and uncooperative which results in a difficult traffic situation could mean that any problems encountered in the traffic situation could be due to more than just system performance, the controller response is also a factor. How much a factor is reflected in the confidence rating?

There are 3 levels of Confidence rating:

A. High Confidence

You were able to account for the events that occurred. You are very certain what problems or benefits could be due to air-ground communications, the traffic situation, the controller, etc., and can therefore provide a rating that really reflects how well the communications system performed.

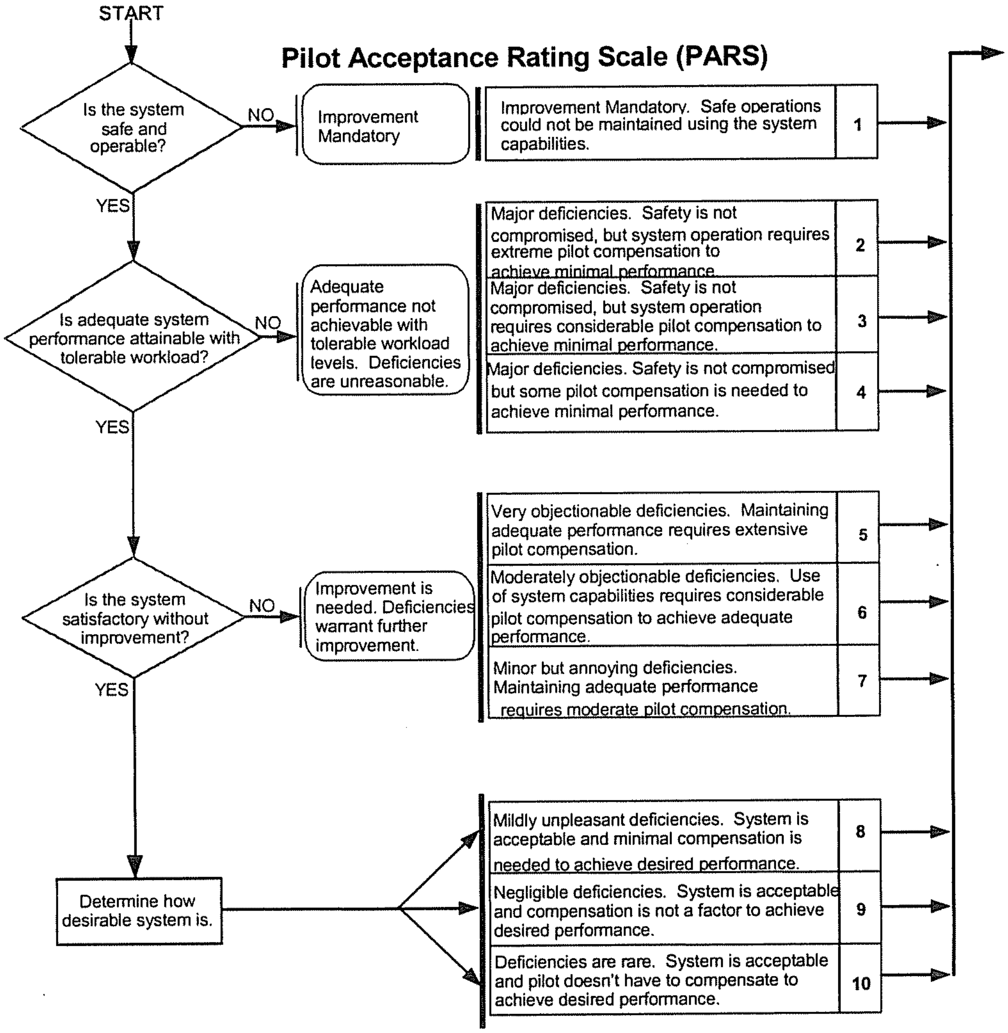
B. Moderate Confidence

You were able to account for some of the events. You are somewhat certain what problems or benefits could be due to air-ground communications, the traffic situation, the controller, etc. There is some uncertainty about how well the communications system performed, given the overall situation. You have some reservations about the accuracy of your numerical rating.

C. Low Confidence

It was difficult to account for the events. There is a great deal of uncertainty about the performance of the communications system, and how you were able to work within the whole system. You have many reservations about the accuracy of your numerical rating because of external factors that you cannot adequately account for.

Pilot Acceptance Rating Scale (PARS)



Improvement Mandatory. Safe operations could not be maintained using the system capabilities.	**1**
Major deficiencies. Safety is not compromised, but system operation requires extreme pilot compensation to achieve minimal performance.	**2**
Major deficiencies. Safety is not compromised, but system operation requires considerable pilot compensation to achieve minimal performance.	**3**
Major deficiencies. Safety is not compromised but some pilot compensation is needed to achieve minimal performance.	**4**
Very objectionable deficiencies. Maintaining adequate performance requires extensive pilot compensation.	**5**
Moderately objectionable deficiencies. Use of system capabilities requires considerable pilot compensation to achieve adequate performance.	**6**
Minor but annoying deficiencies. Maintaining adequate performance requires moderate pilot compensation.	**7**
Mildly unpleasant deficiencies. System is acceptable and minimal compensation is needed to achieve desired performance.	**8**
Negligible deficiencies. System is acceptable and compensation is not a factor to achieve desired performance.	**9**
Deficiencies are rare. System is acceptable and pilot doesn't have to compensate to achieve desired performance.	**10**

Provide Confidence Rating:

A B C

Comments:

____
____
____
____
____
____
____
____
____
____
____
____
____
____

Circle appropriate information below:
Subject ID:
Test Run:
Simulator RCS GAT

Appendix G

EXIT QUESTIONNAIRE

Participant #: _____
Date: ____________

Instructions:
Please answer the following questions based upon your overall experience in the simulation. Fill in one circle with a pen or pencil (or mark with an X) to indicate your level of response.

Simulation Realism and Research Apparatus Ratings

1. Rate the **realism of the overall simulation experience** compared to actual flight operations.	Extremely Unrealistic	①②③④⑤⑥⑦⑧⑨⑩	Extremely Realistic

2. Describe any **differences** you noticed between the types of flights you worked in these scenarios that affect your comments on the communications system design:

____
____
____
____

3. Do you have any comments or suggestions for improving the simulation?

____
____
____

Final Communications Questions

Instructions:

Please respond to each of the following items based upon your overall experience in the simulation. For each statement, fill in one circle (or mark with an X) to indicate your response. Leave an item blank if you have no information on which to base a decision.

	Much Better with the Analog System	Somewhat Better with the Analog System	No Difference Between the Analog and Digital Systems	Somewhat Better with the Digital System	Much Better with the Digital System
4. Completing routine pilot-initiated radio calls.	①	②	③	④	⑤
5. Completing time-critical communications.	①	②	③	④	⑤
6. Responding to air traffic controller calls.	①	②	③	④	⑤
7. Determining when the channel was busy.	①	②	③	④	⑤
8. Determining when the channel was available for use.	①	②	③	④	⑤
9. Being confident that a message was received by the controller.	①	②	③	④	⑤
10. Detecting communications problems and mistakes (e.g., readback errors, stolen clearances).	①	②	③	④	⑤
11. Hearing complete messages (e.g., not clipped, stepped on) from users of the system.	①	②	③	④	⑤
12. Receiving timely responses from the controller.	①	②	③	④	⑤
13. Accomplishing all communications tasks.	①	②	③	④	⑤

Please rate the following aspect of the Analog System:

14. The extent to which step-ons interfered with your ability to communicate with the controller.	Not at all	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

Please rate the following aspects of the Digital System features.

A brief description of each feature is provided below. Please review the descriptions, then think back over the scenarios and try to recall specific instances of antiblocking and controller override before completing the items that follow. Note that the end points of the rating scales vary. Read each carefully before making your choice.

Transmit Status Indicator

- The busy tone, presented as an on and off signal

Antiblocking

- You tried to transmit while another user was already transmitting on the channel. You heard the busy tone.
- You keyed the microphone and tried to transmit when the channel initially sounded quiet. You heard the busy tone.

Controller Override

- You were in the middle of a transmission and were interrupted by the controller. You heard the busy tone.
- You heard another pilot talking and heard the pilot get interrupted by the controller.

Answer #15 and #16 only if you are sure you recall specific instances of each.

15. The usefulness of the **antiblocking** feature for managing communications on the sector frequency.	Not at all Useful	①②③④⑤⑥⑦⑧⑨⑩	Extremely Useful

17. The extent to which the **antiblocking** feature interfered with your ability to communicate with the controller.	Not at all	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

16. The usefulness of the **controller override** feature for managing communications on the sector frequency.	Not at all Useful	①②③④⑤⑥⑦⑧⑨⑩	Extremely Useful

18. The extent to which **controller override** inappropriately interrupted your communications.	Not at all	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

19. The overall usefulness of the **transmit status indicator**.	Not at all Useful	①②③④⑤⑥⑦⑧⑨⑩	Extremely Useful

20. The acceptability of the transmit status indicator **volume level**	Not at all Acceptable	①②③④⑤⑥⑦⑧⑨⑩	Highly Acceptable

21. The acceptability of the transmit status indicator **pitch**.	Not at all Acceptable	①②③④⑤⑥⑦⑧⑨⑩	Highly Acceptable

22. The acceptability of the transmit status indicator **on-off cycle**.	Not at all Acceptable	①②③④⑤⑥⑦⑧⑨⑩	Highly Acceptable

Please note the following item uses a 1 – 5 rating scale:

	Very Negative	Somewhat Negative	Neither Negative nor Positive	Somewhat Positive	Very Positive
23. The impact of the Digital System on safety.	①	②	③	④	⑤

24. Please provide any additional comments that you feel will be helpful in allowing us to understand your experiences working in these scenarios, your reactions to the communications systems, and your responses on this questionnaire.

____

____

____

____

____