DOT/FAA/TC-08/04

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

Human-in-the-Loop Evaluation of an Integrated Arrival/Departure Air Traffic Control Service for Major Metropolitan Airspaces

Carolina M. Zingale, Ph.D., Human Factors Team – Atlantic City, ATO-P
Todd R. Truitt, Ph.D., Human Factors Team – Atlantic City, ATO-P
D. M. McAnulty, Ph.D., Human Factors Team – Atlantic City, ATO-P

March 2008

Technical Report

U.S. Department of Transportation
Federal Aviation Administration

Technical Report Documentation Page

1. Report No. DOT/FAA/TC-08/04	**2. Government Accession No.**	**3. Recipient's Catalog No.**

4. Title and Subtitle

Human-in-the-Loop Evaluation of an Integrated Arrival/Departure Air Traffic Control Service for Major Metropolitan Airspaces

5. Report Date

March 2008

6. Performing Organization Code

AJP-6110

7. Author(s)

Carolina M. Zingale, Ph.D., Human Factors Team – Atlantic City, ATO-P
Todd R. Truitt, Ph.D., Human Factors Team – Atlantic City, ATO-P
D. M. McAnulty, Ph.D., Human Factors Team – Atlantic City, ATO-P

8. Performing Organization Report No.

DOT/FAA/TC-08/04

9. Performing Organization Name and Address

Federal Aviation Administration
Human Factors Team – Atlantic City, ATO-P
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
Human Factors Research and Engineering Group
800 Independence Avenue, S.W.
Washington, DC 20591

13. Type of Report and Period Covered

Technical Report

14. Sponsoring Agency Code

ATO-P

15. Supplementary Notes

16. Abstract

The increasing number of U.S. air flights has placed a severe strain on the National Airspace System (NAS), especially airspace surrounding major metropolitan areas. In a recent study, Truitt, McAnulty, and Willems (2004) tested and found benefits in procedures designed to address some of the pressures around New York airspace. The procedures included extending terminal lateral separation standards (3 nm) and procedures (diverging courses) into en route airspace as well as collocating terminal and en route facilities to promote more effective communication and coordination. The Integrated Arrival/Departure Air Traffic Control Service, termed the Big Airspace (BA) concept, applies these procedures to other busy areas and includes the use of Area Navigation (RNAV) routes as well as dynamic resectorization to make airspace boundaries more flexible. Twenty-four controllers from en route and terminal facilities participated in a simulation that compared a baseline (BL) condition using current airspace standards and procedures to two BA conditions. In one condition, en route and terminal participants managed traffic in the same control room, and in the other condition they worked as if in separate facilities. Overall, the results provided support for the BA concept. The aircraft moved through the arrival corridor more efficiently in the BA conditions than in the BL condition, and participants made fewer ground-ground transmissions and issued fewer altitude and heading clearances. Subjective ratings of performance, situation awareness, and the ability to move traffic through the sector were also higher in the BA conditions.

17. Key Words

Area Navigation Routes
Dynamic Resectorization
Integrated Arrival/Departure Air Traffic Control Service
Reduced Separation
Transitional Airspace

18. Distribution Statement

This document is available to the public through the National Technical Information Service, Springfield, Virginia, 22161. A copy is retained for reference at the William J. Hughes Technical Center Library.

19. Security Classif. (of this report)	**20. Security Classif. (of this page)**	**21. No. of Pages**	**22. Price**
Unclassified	Unclassified	85	

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

Table of Contents

List of Illustrations

Tables Page

Acknowledgments

We thank our sponsor, the Air Traffic System Concept Development Group, led by Michele Merkle, part of the Research and Technology Development Office within the Operations Planning Service Unit of the Air Traffic Organization (ATO-P). We also thank Brian Bagstad, who recruited and scheduled participants, co-led airspace and procedures training, and served as a Subject Matter Expert (SME) during the experiment. We thank Phil Bassett, Traffic Management Unit supervisor at Jacksonville Air Route Traffic Control Center (ZJX), who helped adapt the airspace used in testing, co-led the airspace and procedures training, and served as an SME. We also thank Tom Dury, Support Manager for Operations at Indianapolis Air Route Traffic Control Center, who was on detail at the Research Development and Human Factors Laboratory (RDHFL) as an SME and provided invaluable assistance in recruiting participants and conducting shakedown activities.

We thank the following individuals from the programming, simulation, and engineering support team at the RDHFL for their outstanding efforts to ensure that the simulator, communication system, and data collection tools were functioning optimally: Nick Marzelli, Joel Grochowski, Dallas Shiles, Ed Little, and Gary Mueller developed new features and evaluated and resolved problems with the air traffic control simulator during shakedown and testing; Valentina Velez and AJ Mieskolainen for their efforts in developing the weather depicted on the en route controller displays; Otto Smith, Wallace Daczkowski, and John Dilks for ensuring that the communication system and audio and video recording equipment were operating as needed; and Bob Oliver for his efforts in gathering and formatting the communications data for analysis.

Special thanks are extended to Nick Marzelli for his additional efforts in creating a new database tool for analyzing aircraft performance data. This tool integrated key components of the aircraft data and allowed us to consolidate information for the data sets we wished to examine. It also provided a way for us to visually track individual aircraft and correlate aircraft position with the commands entered by controllers and simulation pilots at points throughout the scenarios. This allowed us to have an enriched understanding of aircraft performance issues and helped extensively with our analysis efforts.

We thank members and support staff of Rick Ozmore's Simulation and Analysis Group: Sherri Magyarits, who coordinated the efforts of the fast-time and real-time simulation subteams and managed simulation pilot staffing needs during shakedown and testing; Al Schwartz and Doug Baart, who created the initial fast-time scenarios; Evan Darby (now at ZJX), who helped adapt the airspace, develop procedures, and conduct shakedown. We thank Nicole Racine, who assisted during shakedown and provided expertise and guidance on structuring the aircraft data analysis files; Kim Bender and Scott Terrace, who continued that effort and analyzed a large portion of the aircraft performance data; Gabrielle Jerdan (now at Atlanta Air Route Traffic Control Center), who served as one of our controllers during shakedown, collected face-to-face communication data during the experiment, and participated in aircraft data analysis efforts. We also thank Tim Holth and Jack Graph, retired controllers and SMEs, who participated in shakedown, helped develop procedures for the ghost-controller position, and staffed that position during the experiment.

We thank Joe Toland, Jim McGhee, and Rich Ridgway, RDHFL air traffic SMEs and retired controllers. Joe and Jim adapted the original fast-time scenarios and airspace maps for use in the experiment. They and Rich Ridgway were extensively involved in shakedown efforts needed to troubleshoot and resolve problems, and they also served as observers during testing. We thank Ben Willems, RDHFL Engineering Research Psychologist, for his expertise in adapting the User Request Evaluation Tool to the airspace and for his advice and assistance in conducting the experiment and troubleshooting issues with the simulator. Our thanks also go to Mark Hale, Kevin Hallman, and Tina DiProspero, RDHFL, for their work in organizing, collating, and coding test materials as well as assisting with data recording, data entry, and data analysis.

We thank Al Rehmann and the simulation pilot staff for their efforts throughout shakedown and testing: Mary Rozier-Wilks, Tru Hall, Dale Laudenslager, Bill Hickman, Bob Engiles, Lee Dixon, Bruce Slack, Mary Lou Hagan, Audrey Wilmer, Tom Vispisiano, Joe Flood, Ross Spath, Dorothy LaSure, J. C. Lester, Steve Allen, George Bigge, Tom Callahan, and Ed Fitzpatrick.

We also thank members of the Target Generation Facility (TGF) staff, especially Scott Doucett, who provided assistance during simulation preparation; Dana Whicker and Mike Konyak, who ensured that required system commands and aircraft functionality were implemented and operational; Samantha Fullerton and Jim Swantek, who provided assistance during shakedown and testing; and George Chachis for his efforts in meeting project personnel requirements for the TGF and contract simulation pilots.

Executive Summary

The increasing number of U.S. air flights has placed a severe strain on the efficiency of the National Airspace System. These problems are especially apparent in the airspace surrounding major metropolitan areas. In a recent study, Truitt, McAnulty, and Willems (2004) tested and found benefits in procedures designed to address some of these system pressures. They evaluated a New York Integrated Control Complex concept that extended terminal separation standards (i.e., 3 nm instead of 5 nm lateral separation) and other terminal procedures (i.e., diverging courses) to airspace farther away from airports to ease the traffic flow in and out of those areas and collocated terminal and en route facilities to promote more effective communication and coordination. The Integrated Arrival/Departure Air Traffic Control Service, termed the Big Airspace (BA) concept, was designed to evaluate those procedures in other busy areas outside of New York airspace. In addition, the BA concept included the use of Area Navigation (RNAV) routes as well as dynamic resectorization capabilities to make airspace boundaries more flexible so that traffic can be more easily rerouted when weather, equipment outages, or active special use airspace disrupt normal flows.

The real-time human-in-the-loop experiment summarized in this report was part of a broader effort to evaluate the concept. Other components of the evaluation included fast-time simulations, human performance modeling, and cost-benefit analyses. For this experiment, we examined controller performance in a high-fidelity simulation that compared a baseline (BL) condition to two alternative operating conditions, one that simulated a common en route and terminal control room environment and another in which the control rooms were not combined. The experiment included an arrival and a departure sector for en route and terminal airspace and weather that impacted arrival traffic and drove the need for dynamic resectorization. We collected and evaluated aircraft performance data to evaluate efficiency and safety for each experimental condition. We also collected communications data and subjective measures of performance, including participant workload, situation awareness, and evaluations by Subject Matter Experts (SMEs).

Our results indicated support for the BA concept. We found that aircraft moved through the busy arrival sectors more efficiently in both of the BA conditions than in the BL condition. Aircraft spent less time and traveled less distance through the BA airspace, and the participants working those sectors made fewer ground-ground transmissions and issued fewer altitude and heading clearances. The participants also needed less assistance from the ghost controllers managing traffic outside the en route arrival sector in the BA conditions. The number of operational errors did not differ across conditions.

Many of the subjective measures also supported the concept. The en route participant on-line workload ratings were lower in the second half of the scenarios in the BA conditions than in the BL condition, indicating that the participants found it easier to manage traffic after the dynamic resectorization took place. The SMEs rated most of the en route performance measures higher and noted fewer problems in the BA conditions. Participant ratings of performance, situation awareness, and ability to move traffic through the sector were among the other measures that were also higher in the BA conditions. There were very few meaningful differences found between the two BA conditions.

1. INTRODUCTION

1.1 Background

The National Airspace System (NAS) in the U.S. is one of the busiest, most complex, and safest in the world. The NAS includes numerous facilities and thousands of pieces of equipment to support surveillance, navigation, and communication functions. The facilities include the Air Traffic Control System Command Center (ATCSCC), Air Route Traffic Control Centers (ARTCCs), Terminal Radar Approach Controls (TRACONs), and Airport Traffic Control Towers (ATCTs). These facilities are staffed by Certified Professional Controllers (CPCs), technical operations support staff, and other related personnel to ensure system efficiency, functionality, and safety.

ATCTs are responsible for the airspace within 5 miles of the airport, and aircraft are primarily controlled through visual sighting. TRACON facilities are responsible for the airspace that extends approximately 40 miles from the primary airport (the area may also include secondary airports) and from 3,000 ft above the airport to approximately 10,000 ft above mean sea level (MSL). ARTCCs control aircraft operating above the TRACONs. The TRACONs and ARTCCs use radar surveillance to control aircraft, but the TRACONs and ARTCCs have different separation standards. In the TRACON, aircraft flying under Instrument Flight Rules (IFR) must be separated from other aircraft by 3 nm horizontally and 1,000 ft vertically unless the aircraft are on diverging courses. Aircraft flying under Visual Flight Rules (VFR) are responsible for maintaining visual separation from other aircraft, but they may receive traffic and other advisories from CPCs if requested and time is available. In ARTCCs, aircraft flying higher than 18,000 ft above MSL must be on IFR and must be separated from other aircraft by 5 nm horizontally and 1,000 ft vertically. Aircraft below that flight level may be on either an IFR or VFR flight plan, depending on aircraft capabilities and intent and on weather conditions.

Within each facility, the airspace is divided into sectors of responsibility that are separated by horizontal and vertical boundaries. Depending on how complex and busy it is, each sector is staffed with one to three CPCs to provide separation assurance, aircraft sequencing, and advisory information. Finally, the ATCSCC takes a system-wide view of the NAS. The ATCSCC monitors traffic flows and weather conditions and also coordinates with the airlines, the military, and Air Traffic Control (ATC) to maintain an optimum flow of traffic across the nation. Traffic Flow Management (TFM) personnel at each ARTCC and major large TRACONS coordinate with the ATCSCC to determine the appropriate flow of aircraft through the airspace, including what constraints may need to be implemented if weather or other problems arise.

The increasing number of U.S. air flights has placed a severe strain on the efficiency of the NAS. These problems are especially apparent in the airspace surrounding major metropolitan areas. The arrival and departure airspace surrounding major metropolitan areas is complex, which impacts operational efficiencies. Route structures in these areas are fairly inflexible, which can cause traffic flow disruptions far from the existing terminal boundaries. In addition, existing boundary structures limit the types of procedures available to controllers and limit the controllers' ability to optimize airspace, forcing controllers to spend much of their time communicating and coordinating with surrounding control facilities. Increasing congestion and decreasing efficiency increases costs for the Federal Aviation Administration (FAA) as well as for the airlines and consumers. To alleviate some of the stress resulting from an increasingly crowded NAS, the

FAA is attempting to develop and implement changes in ATC procedures, airspace boundaries, and routing structures to improve NAS performance and increase system efficiency. These changes should also result in an increase in controller productivity and a decrease in controller workload.

In a recent study, Truitt, McAnulty, and Willems (2004) tested procedures designed to address some of these system pressures. They evaluated a New York Integrated Control Complex (NYICC) concept designed to deal with congestion in the Northeast corridor around the New York airspace. The NYICC concept proposed two primary adjustments to address congestion issues. First, it proposed collocating terminal and en route facilities to facilitate communications and coordination related to ATC operations (collocation). Second, it proposed extending terminal separation standards (i.e., 3 nm instead of 5 nm lateral separation) and other terminal procedures (i.e., diverging courses) to airspace farther away from the terminal area to ease traffic flow to and from the major airports (i.e., terminalization). Truitt et al. found that both of the proposed changes facilitated CPC performance, and those performance enhancements occurred in both the arrival and the departure sectors. Additionally, Truitt et al. found that although there were benefits in the collocation condition, the benefits were greatest in the terminal separation with collocation condition. Some of the benefits realized in the terminal separation with collocation condition included an increase in the number of arrivals and departures, a reduction in the number and duration of holds, and a reduction in the number and duration of departure stops.

Although these findings are intriguing, they are limited in their scope to the New York airspace as it is currently designed. To generalize these findings to other crowded airspace, it is important to replicate them in an experiment using more generic airspace with CPCs recruited from a broader range of facilities. Recently, the FAA developed a concept of operations for an integrated arrival/departure control service termed the Big Airspace (BA) concept. This concept proposes to extend the changes from the NYICC experiment to other congested airspace. The BA concept also includes an increase in the number of Area Navigation (RNAV) routes so that more Standard Instrument Departures (SIDs) and Standard Terminal Arrival Routes (STARs) are available. The concept also incorporates dynamic resectorization (Hadley, Sollenberger, D'Arcy, & Bassett, 2000; Stein, Della Rocco, & Sollenberger, 2006), a procedure that makes airspace boundaries more flexible so that traffic can be more easily rerouted when weather, equipment outages, or active special use airspace disrupt normal flows. The goal is that by moving the artificial barriers separating en route and terminal airspace to a point farther from congested airport airspace, the FAA will reduce procedural and airspace inefficiencies, thereby achieving smoother, more efficient air traffic flows into and out of major metropolitan airports. To introduce such major changes into the Standard Operating Procedure (SOP) for the NAS, it is first necessary to test the proposed changes to see how they might impact CPC and system performance. To this end, researchers initially visited six major TRACONs and four ARTCCs, and conducted a cognitive walkthrough with Subject Matter Experts (SMEs). These efforts focused on identifying major issues and helped to identify airspace for subsequent validation efforts. Other groups in the FAA performed validation efforts via fast-time modeling and human performance modeling (Air Midas: see Corker & Smith, 1993) to evaluate the impact of the proposed changes on the NAS (FAA, 2007). In this experiment, we used a high-fidelity human-in-the-loop simulation to test the impact of the proposed changes on controller efficiency and safety.

1.2 Purpose

This experiment examined the effects of extending 3 nm lateral separation and aircraft divergence procedures to approximately 100 nm from the airport for both arrival and departure sectors, dynamic resectorization, and control room configuration on aircraft and controller performance. We conducted a high-fidelity human-in-the-loop experiment that compared a baseline (BL) condition with existing airspace and procedures to two alternative BA conditions, one that simulated a common en route and terminal control room environment (BAC) and another in which the control rooms were not combined (BANC).

2. METHOD

2.1 Participants

Twenty-four controllers (21 men and 3 women) from five en route (Levels 10-12) and seven terminal (Levels 3-5) facilities participated in the experiment. Twelve participants were current in en route, and 9 of them also had experience in the terminal domain. Twelve participants were current in terminal, and 4 of them also had experience in the en route domain. One of the terminal participants was currently working as a supervisor and another was currently working in the Traffic Management Unit (TMU) at their respective facilities.

Table 1 presents summary information obtained from the participants' Biographical Questionnaires. On average, the participants had over 21 years experience controlling air traffic, almost 18 of which were as CPCs for the FAA. They rated (1 = *lowest*, 10 = *highest*) their current skill level as high and indicated that they were very motivated to participate in the experiment.

Table 1. Means and Standard Deviations for Biographical Questionnaire Items

Questionnaire Item	**Mean *(SD)***
Age of participant	43.9 (4.06)
Years as an Air Traffic Controller (including FAA and military experience)	21.8 (4.81)
Years as a CPC for the FAA	17.9 (4.77)
Years actively controlling traffic in en route domain (12 en route participants)	14.0 (5.53)
Years actively controlling traffic in the en route domain (4 terminal participants)	10.5 (12.33)
Years actively controlling traffic in the terminal domain (12 terminal participants)	17.1 (8.06)
Years actively controlling traffic in the terminal domain (9 en route participants)	7.6 (3.91)
Number of months in past year actively controlling traffic	11.9 (0.41)
Current skill level as a CPC	7.8 (1.49)
Level of motivation to participate in this study	9.3 (0.90)

2.2 Research Personnel

A Principal Investigator and co-Principal Investigator conducted the experiment. They supervised the preparation and operation of the simulator equipment and administered the instructions and questionnaires to the participants. Two research assistants prepared data collection instruments, helped collect data, and entered information into spreadsheets for analysis. Three other research assistants helped to reformat the data for analysis and ran some of the statistical analyses.

In preparation for the experiment, two air traffic SMEs modified the basic scenarios developed for the fast-time modeling analysis. They added additional aircraft to keep pressure on the primary airport (GENERA). Hardware and software engineers prepared all the experiment tools, including the display configurations, workstation operation, and communication system used in the experiment. The engineers were on standby to assist during the experiment.

Two controllers from Orlando TRACON assisted in verifying the scenarios and corresponding procedures during early shakedown efforts. Four additional controllers (two with en route experience and two with terminal experience) from other facilities assisted later in shakedown to evaluate the training procedures on unfamiliar airspace.

Twelve simulation pilots managed the aircraft during shakedown and testing. During testing, two SMEs served as over-the-shoulder observers, two other SMEs collected data on participant coordination, and two confederates operated as ghost-sector controllers. One of the experimenters acted as a supervisor to indicate to participants when the weather would require the dynamic resectorization of airspace.

2.3 Equipment

We conducted the experiment at the FAA William J. Hughes Technical Center (WJHTC) Research Development and Human Factors Laboratory (RDHFL). The simulation configuration consisted of the Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE) and the Target Generator Facility (TGF). DESIREE emulates ATC display functions and receives input from the TGF to display radar targets.

2.3.1 Hardware

The CPC workstations and associated equipment were located at the RDHFL in Experiment Room (ER) 1 and ER 2 (see Figure 1). The equipment for the ghost sector was located in an adjacent room. The simulation pilot workstations were located in the simulation pilot workstation room.

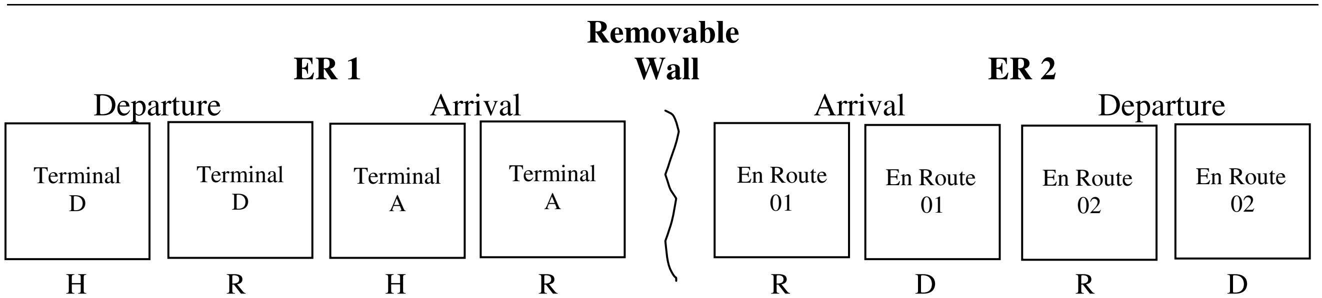

Figure 1. A depiction of the en route and terminal workstation console configuration. (*Note*. ER = Experiment room, R = R-side position, D = D-side position, and H = Handoff position.)

2.3.1.1 Air Traffic Control En Route Workstation Consoles

The experiment used four en route workstation consoles. The Radar (R)-side console contained the Display System Replacement (DSR) radar display with a Computer Readout Display (CRD), whereas the Data (D)-side console contained the User Request Evaluation Tool (URET) display and CRD. Each en route R-side and D-side console had communication equipment, a keyboard, and a trackball. The en route controllers did not have use of the Traffic Management Advisor (TMA) tool to modulate spacing, but the SMEs structured the scenarios to reflect TMA sequencing when aircraft entered the arrival sector. We did not include Datalink in the experiment because this tool was not planned to be available in the en route environment in the BA timeframe, and its concept of use in the terminal environment was not yet determined. Therefore, all air-ground communications were voice communications. We also limited the D-side use of URET to reflect the way en route controllers currently use the tool in the field – to update the NAS with flight plan changes but not for conflict probes.

2.3.1.2 Air Traffic Control Terminal Workstation Consoles

The experiment required four terminal Standard Terminal Automation Replacement System (STARS) workstation consoles (as shown in Figure 1), of which two were operational. The handoff (H) controllers sat at a non-operational console and observed and interacted with the R-side radarscopes. Each of the four terminal consoles contained a set of communication equipment, a keyboard, and a trackball.

2.3.1.3 Simulation Pilot Workstations

The experiment utilized 12 simulation pilot workstations. Each workstation consisted of a computer, keyboard, monitor, and communication equipment. Each simulation pilot also had a plan view display of traffic and a list of assigned aircraft. For each assigned aircraft, the simulation pilots had information regarding the aircraft's current state and corresponding flight plan data. The simulation pilots also had weather displayed on their workstations and were instructed to request deviations –not greater than 20 degrees– because of weather for affected aircraft.

2.3.1.4 Communications

Each console had communication panels and headsets. The R-side CPCs had two-way voice communication via headsets with their respective simulation pilots. All CPCs had two-way voice communication via headsets with the other sectors involved in the simulation, including the ghost sectors.

2.3.1.5 Workload Assessment Keypad

Each R-side position had a Workload Assessment Keypad (WAK) positioned near the console. The WAK consists of a touch panel display with 10 numbered buttons. The WAK prompts the participants to press a button to provide their subjective workload ratings by using auditory and visual signals. In this experiment, we set the WAK to prompt the participants for a rating every 4 minutes. During the prompt, the numbered buttons on each device illuminated and the device emitted a brief tone. The participants indicated their current level of workload by pressing one of the numbered buttons in which 1 indicated a *very low* workload and 10 indicated a *very high* workload. The buttons remained illuminated for the duration of the response period (20 s) or until a participant made a response, whichever occurred first. The participants received complete WAK instructions at the beginning of the experiment and at the daily in-briefing. They also received brief reminders before each practice scenario and before the actual scenarios to refresh their memories and to increase the likelihood that they would use the same rating criteria every time.

2.3.2 Software

The experimenters used the TGF and DESIREE ATC simulator to present air traffic scenarios. Software engineers at the FAA WJHTC developed both of these systems.

The TGF uses preset flight plans to generate radar track and data block information on the controller and simulation pilot displays. The TGF also provides an interface that allows the simulation pilots to enter flight plan changes. The TGF algorithms can control aircraft maneuvers so that they appear to the controllers to represent realistic aircraft climb, descent, and turn rates. Finally, the TGF allows researchers to capture information about aircraft trajectories, aircraft proximity, and other relevant data for subsequent analyses.

DESIREE emulates both en route and terminal controller functions. Its purpose is to enable researchers to modify or add information and functionality to current ATC workstations and to evaluate new concepts and procedures. DESIREE receives input from the TGF that allows it to display information on the radarscope, including radar tracks, data blocks, and sector maps. It also allows controllers to perform their typical functions in an operational environment (e.g., perform handoffs; enter data into the HOST computer). Like TGF, DESIREE has data collection capabilities and can collect information on all controller entries made during a scenario.

2.4 Materials

We based our materials on the questionnaires, rating forms, and instructions used by Truitt et al. (2004). Therefore, our materials contain some differences in terminology regarding the airspace and control room configurations from that used in the body of this report.

2.4.1 Informed Consent Form

Before the experiment, each participant read and signed an informed consent statement. This form summarized the objectives and discussed participant rights and responsibilities (see Appendix A).

2.4.2 Biographical Questionnaire

Each participant completed a Biographical Questionnaire before the experiment. This questionnaire allowed participants to provide background information (e.g., years of ATC experience) that could be useful in interpreting the results (see Appendix B).

2.4.3 Post-Scenario Questionnaires

After completing each test scenario, participants provided subjective ratings of their own performance, situation awareness, and workload on the Post-Scenario Questionnaire (PSQ-1) (see Appendix C). Using a 10-point Likert scale, performance and situation awareness ratings ranged from 1 (*poor*) to 10 (*excellent*), and workload ratings ranged from 1 (*extremely low*) to 10 (*extremely high*). The participants also had the opportunity to provide open-ended responses so that they could include any information about the scenario that they considered relevant (see

Appendix C). In the BA conditions, the participants also provided ratings on a second Post-Scenario Questionnaire (PSQ-2) to indicate how their performance was affected by the BA procedures. These questions used a 9-point rating scale in which a rating of 1 indicated a *negative effect*, a rating of 5 indicated *no effect*, and a rating of 9 indicated a *positive effect*. The PSQ-2 also allowed the participants to add additional comments to explain their ratings (see Appendix D).

2.4.4 Post-Experiment Questionnaire

The participants completed a Post-Experiment Questionnaire (PEQ) after completing the entire experiment (see Appendix E). On the PEQ, the participants provided ratings to compare the effect of the BA conditions to the BL condition on their control strategies. These questions used a 9-point Likert rating scale in which a rating of 1 indicated a *negative effect*, a rating of 5 indicated *no effect,* and a rating of 9 indicated a *positive effect*. The participants also indicated the extent to which their communication strategies changed between the test conditions and the extent to which the WAK device interfered with their performance. These questions used rating scales that ranged from 1 (*not at all*) to 10 (*a great deal*). Finally, the participants provided responses about simulation realism using scales that ranged from 1 (*extremely unrealistic*) to 10 (*extremely realistic*). Like the PSQ-1 and PSQ-2, the PEQ also posed open-ended questions.

2.4.5 Communication Score Sheet

The experimenters used the Communication Score Sheet during the BAC condition to record verbal and nonverbal communication behavior (Peterson, Bailey, & Willems, 2001; Truitt et al., 2004; see Appendix F).

2.4.6 Observer Rating Form

The SMEs used a modified version of the Observer Rating Form (ORF) to make ratings about the participants (Sollenberger, Stein, & Gromelski, 1997; Vardaman & Stein, 1998; see Appendix G). The SMEs rated the terminal and en route participants separately. Performance ratings were made using scales that ranged from 1 (*least effective*) to 8 (*most effective*). Additional questions pertained to the frequency of occurrence of problematic events, such as issuing clearances earlier or later than appropriate. These questions used 5-point rating scales in which a rating of 1 indicated that an event *never occurred*, a rating of 5 indicated that an event *occurred unacceptably often*, and a rating of 3 indicated that an event *occurred, but within normal limits of operational acceptability.* Each SME filled out two ORF forms, one for the Arrival sector and one for the Departure sector.

2.4.7 Standard Operating Procedures and Letters of Agreement

The participants adhered to the SOPs and the Letters of Agreement (LOAs) for either the en route or the terminal environment. The SOPs and LOAs varied with the sector and the current experimental condition.

2.4.8 Airspace

The experiment required the development of two different airspace designs: one for the BL condition and one for the BA conditions. Both were modifications to current airspace in central Florida. Between the BL and BA conditions, the sector boundaries and route structures along with the fixes and waypoints changed. The BL airspace had two en route sectors, BAASS (Arrival Sector 01) and GRUPR (Departure Sector 02), and two terminal sectors, Arrival and Departure (see Figure 2), which used their respective separation procedures. The airspace for the BA conditions also contained two transitional airspace sectors, BAASS (Arrival Transition Sector 01) and GRUPR (Departure Transition Sector 02), and two near-airport sectors, Feeder and Airport Departure (see Figure 3). In the BA conditions, controllers used 3 nm lateral separation and diverging courses in all four sectors. The volume of the airspace in the BL and BA conditions was the same, but the boundaries of the near-airport sectors in the BA conditions were closer to the airport than the terminal sector boundaries in the BL condition. Thus, the transition sectors were bigger, which allowed more room for maneuvering so that aircraft could be aligned and spaced more effectively.

All conditions contained RNAV SIDs and STARs, but there were more of them in the BA conditions because of the reduced separation requirements. We treated all aircraft as RNAV equipped in all conditions. Along with these capabilities, the airspace in the BA conditions had dynamic airspace boundaries that would allow us to change arrival-to-departure and departure-to-arrival routes. In our experiment, we shifted airspace during the BA conditions from GRUPR to BAASS (refer to Dynamic Airspace 2 in Figure 3). The airspace in the BL condition had dedicated arrival and departure routes and did not have dynamic sector boundaries.

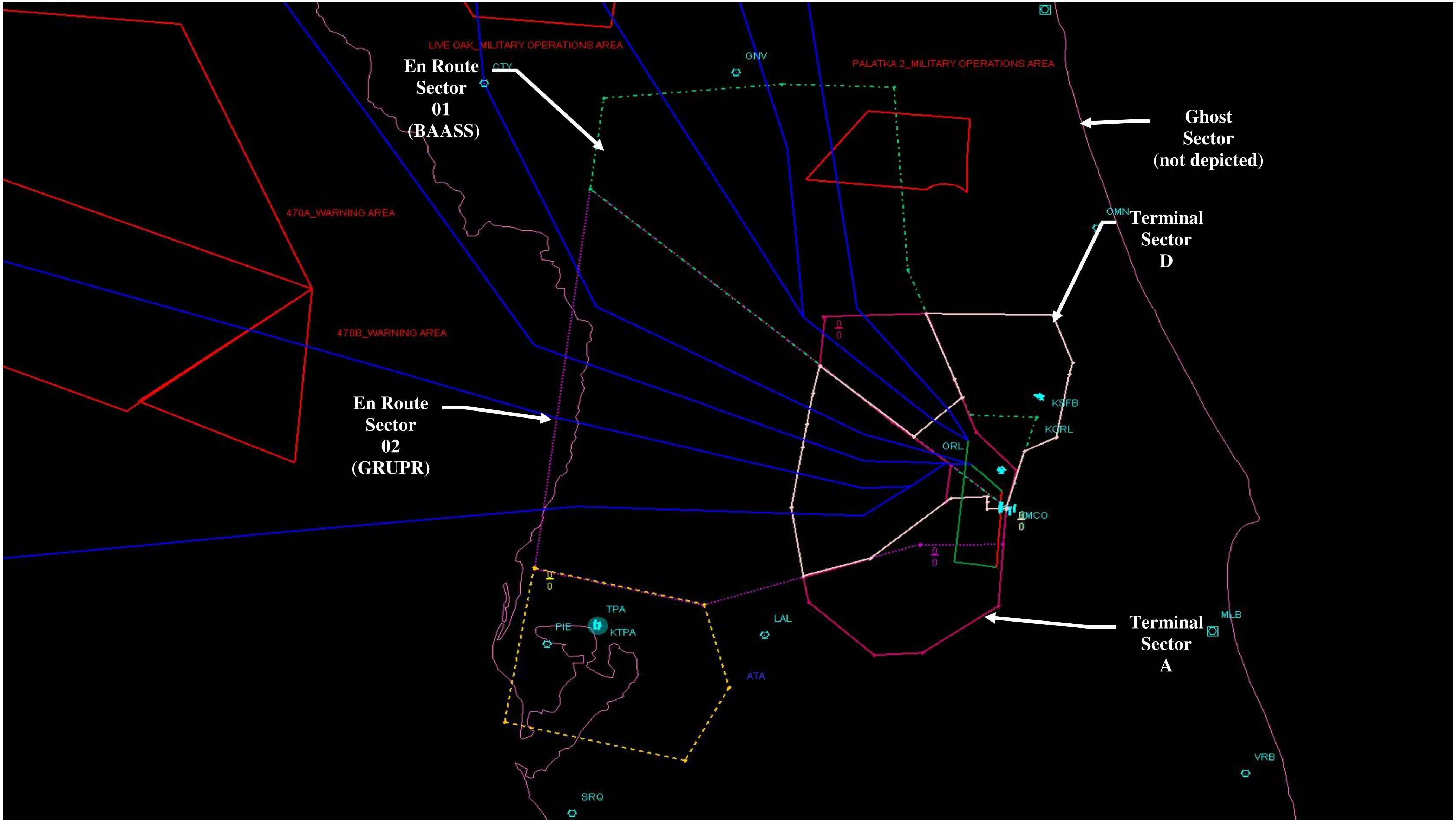

Figure 2. A chart of the airspace for the Baseline condition.

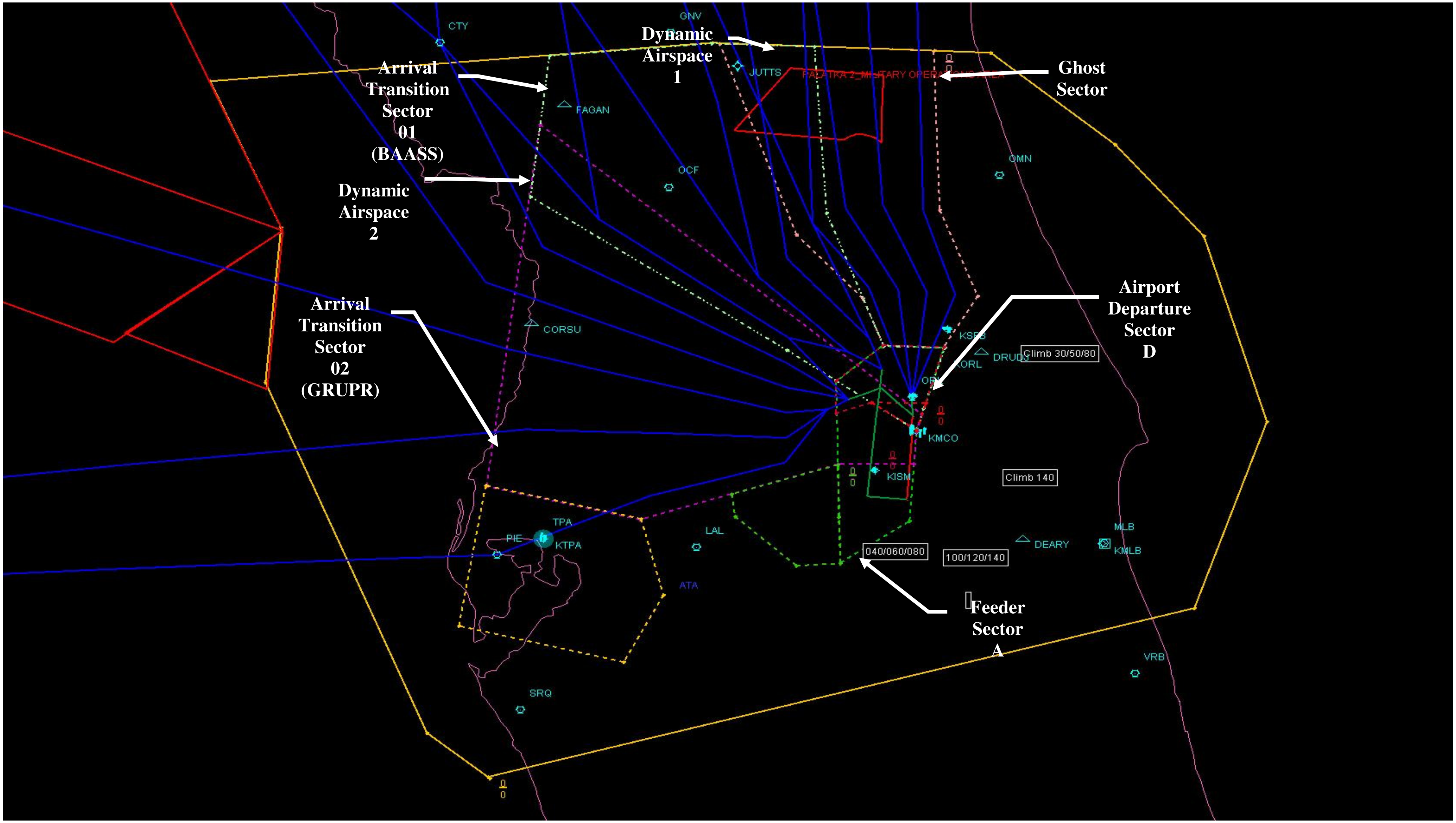

Figure 3. A chart of the airspace for the Big Airspace conditions.

2.4.9 Traffic Scenarios

The experiment required the development of traffic scenarios that contained the same number and type of aircraft for both the BL and BA conditions and that differed only with respect to the routes flown. We developed one basic test scenario for the BA conditions and one for the BL condition. Then, we created four variations of each basic scenario for each test condition that differed only in the aircraft callsigns. Each test scenario began with full traffic and was 50 minutes in length. We also developed separate scenarios for use in the practice sessions. We developed one basic practice scenario for the BL condition and one for the BA conditions with four variations each that differed only in the aircraft callsigns. The practice scenarios began with traffic levels that were comparable to the traffic levels in the test scenarios but were designed to run for 30-40 minutes. We also developed two warm-up scenarios, one for each test condition. Each of the warm-up scenarios began with about half the volume of traffic of the practice and test scenarios and built to about three quarters of the full traffic volume by the latter part of the scenario. We used the warm-up scenarios to introduce the airspace and procedures to the participants and designed those scenarios to run for 30-40 minutes.

2.4.10 Weather Scenarios

The experiment included weather in all of the practice and experimental sessions that impacted routes in the north. The weather updated every 2 minutes on the DSR displays. It contained convective weather cells and was present from the beginning of the scenario. Through about minute 15, the weather began to impact the northern ghost departure sector to the east of BAASS, shutting down those departure routes. At this point, affected aircraft were sent out on a ghosted eastern departure route. From minute 15 to minute 26, the convective weather grew and impacted the BAASS sector (sector 01), shutting down arrival routes. In the BA conditions, at minute 26, the experimenter acting as the area supervisor resectorized the airspace between BAASS and GRUPR so that the northernmost departure route in GRUPR became available to BAASS as an arrival route (see Figure 3). In doing this, it was assumed that the TMU had already directed the TRACON not to send any more departing aircraft out on that route and the ARTCC to send arrivals in on that route. During the interval from minute 26 through approximately minute 37, the participants worked with the dynamically resectorized airspace and cleared remaining departure aircraft out of the sector and began to accept arrival aircraft entering on the new arrival route. The participants worked the remainder of the scenarios, from minute 37 to minute 50, using the resectorized airspace. In the BL condition, no resectorization took place and weather required controllers to use only the available routes.

2.5 Design

2.5.1 Experimental Design

The en route/transitional airspace consisted of 2 two-person sectors, BAASS (01) and GRUPR (02), with an R-side and a D-side. The terminal/near-airport airspace also consisted of 2 two-person sectors, Arrival/Feeder (A) and Departure/Airport Departure (D), with an R-side and an H position. Each participant controlled traffic at each of the positions in his or her airspace. While at each position, each participant ran one scenario in each of the three conditions in the experiment: BL, BAC, and BANC (see Table 2).

Table 2. Counterbalancing Order of Test Conditions

		En Route/Transitional Sectors				Terminal/Near-Airport Sectors					
Group	**Experiment Run**	**01**	**01-D**	**02**	**02-D**	**A**	**A-H**	**D**	**D-H**	**Condition**	**Day**
1	1	E1	E2	E3	E4	T3	T2	T1	T4	BL	M
	2	E1	E2	E3	E4	T3	T2	T1	T4	BAC	M
	3	E1	E2	E3	E4	T3	T2	T1	T4	BANC	M
	4	E2	E3	E4	E1	T4	T3	T2	T1	BAC	M
	5	E2	E3	E4	E1	T4	T3	T2	T1	BANC	T
	6	E2	E3	E4	E1	T4	T3	T2	T1	BL	T
	7	E3	E4	E1	E2	T1	T4	T3	T2	BANC	T
	8	E3	E4	E1	E2	T1	T4	T3	T2	BL	T
	9	E3	E4	E1	E2	T1	T4	T3	T2	BAC	T
	10	E4	E1	E2	E3	T2	T1	T4	T3	BAC	W
	11	E4	E1	E2	E3	T2	T1	T4	T3	BL	W
	12	E4	E1	E2	E3	T2	T1	T4	T3	BANC	W
2	13	E5	E7	E6	E8	T7	T8	T5	T6	BANC	M
	14	E5	E7	E6	E8	T7	T8	T5	T6	BAC	M
	15	E5	E7	E6	E8	T7	T8	T5	T6	BL	M
	16	E7	E6	E8	E5	T6	T7	T8	T5	BL	M
	17	E7	E6	E8	E5	T6	T7	T8	T5	BANC	T
	18	E7	E6	E8	E5	T6	T7	T8	T5	BAC	T
	19	E6	E8	E5	E7	T5	T6	T7	T8	BAC	T
	20	E6	E8	E5	E7	T5	T6	T7	T8	BL	T
	21	E6	E8	E5	E7	T5	T6	T7	T8	BANC	T
	22	E8	E5	E7	E6	T8	T5	T6	T7	BANC	W
	23	E8	E5	E7	E6	T8	T5	T6	T7	BL	W
	24	E8	E5	E7	E6	T8	T5	T6	T7	BAC	W
3	25	E9	E12	E11	E10	T9	T11	T12	T10	BAC	M
	26	E9	E12	E11	E10	T9	T11	T12	T10	BANC	M
	27	E9	E12	E11	E10	T9	T11	T12	T10	BL	M
	28	E10	E9	E12	E11	T11	T12	T10	T9	BL	M
	29	E10	E9	E12	E11	T11	T12	T10	T9	BAC	T
	30	E10	E9	E12	E11	T11	T12	T10	T9	BANC	T
	31	E11	E10	E9	E12	T12	T10	T9	T11	BANC	T
	32	E11	E10	E9	E12	T12	T10	T9	T11	BAC	T
	33	E11	E10	E9	E12	T12	T10	T9	T11	BL	T
	34	E12	E11	E10	E9	T10	T9	T11	T12	BL	W
	35	E12	E11	E10	E9	T10	T9	T11	T12	BANC	W
	36	E12	E11	E10	E9	T10	T9	T11	T12	BAC	W

Note. E1 = Participant 1 for en route; T1 = Participant 1 for terminal.

During the BL condition, the participants controlled traffic as they normally would in the field, and a wall physically separated the en route/transitional sectors from the terminal/near-airport sectors. During the BA conditions, the lateral separation standards for the transitional sectors were reduced from 5 nm to 3 nm, and the participants were also able to use diverging course procedures. Visual separation was not used because the simulation pilot configuration prevented pilots from having the capability to conduct this procedure. For the BANC condition, the wall remained in place. During the BAC condition, the participants were in the same room, and face-to-face communication between them was possible. In both of the BA conditions, the en route controllers continued to use their en route consoles, but the radar display updated at the terminal rate of 5 s rather than the en route rate of 12 s.

2.5.2 Dependent Variables

For each condition, we collected system and CPC measures of efficiency, performance, and communication. We also collected subjective measures of performance and workload.

2.5.2.1 System Performance Measures

We collected many system performance measures for each sector and for the overall experiment to provide information regarding efficiency and safety for each experimental condition. These measures included the number of flights completed; number of departures; number of altitude, heading, and airspeed commands issued by controllers; time and distance flown (in nautical miles) for all aircraft on the controllers' frequency; time and distance on RNAV routes; number of handoffs; number and duration of time (in seconds) of airborne holds; number and duration of ground stops; number and duration of departure delays; and losses of separation and operational errors.

2.5.2.2 Subjective Measures

The SMEs used the ORF (see Appendix G) to collect over-the-shoulder performance ratings for the terminal and en route participants. The SMEs provided an assessment of the participants' performance in maintaining a safe and efficient traffic flow, sequencing aircraft efficiently, and providing control information. They also rated the frequency of occurrence of improper task performance, if any. The participants made subjective ratings of their workload, situation awareness, and control performance on the PSQs and PEQ.

2.5.2.3 On-line Workload Measures

We recorded all WAK ratings made by each R-side controller every 4 minutes during the scenarios. If a controller made no response within 20 s following a prompt, a missing data code was recorded.

2.5.2.4 Communications

We automatically recorded Push-To-Talk (PTT) communications, including both ground-ground and ground-air transmissions. We recorded the number of times each participant transmitted a message and whether that transmission was from a controller to a controller or a controller to a pilot. An observer recorded the frequency and categorized the general content of face-to-face communication between the Arrival Transition and Feeder sectors, and between the Departure Transition and Airport Departure sectors in the BAC condition. The observer also recorded nonverbal gestures, such as pointing to a display. This enabled us to better evaluate what types of correspondence occurred in the event that PTT communications between the en route and terminal participants decreased during the BAC condition compared to the BL and BANC conditions.

2.6 Procedure

2.6.1 General Schedule of Events

The en route and terminal participants were involved in the experiment for 6 days. They traveled to the RDHFL on a Tuesday and left on Thursday of the following week. Table 3 shows the daily schedule of events.

Table 3. Daily Event Schedule

Week 1					
Time	**Wednesday**	**Time**	**Thursday**	**Time**	**Friday**
8:30	Introduction, Forms, Baseline Airspace & LOA/SOP Familiarization	8:30	Daily In-Briefing & Big Airspace Review	8:30	Daily In-Briefing & Baseline Review
10:00	Break	9:00	Practice 6 & 7	9:00	Practice 13 &14
10:15	Practice 1 & 2	10:30	Break	10:30	Break
11:45	Lunch	10:45	Practice 8	10:45	Practice 15
12:45	Review Baseline Rules	11:45	Lunch	11:30	Lunch
1:00	Practice 3 & 4	12:45	Collocation Instructions	12:30	Practice 16
2:30	Break	1:00	Practice 9 & 10	1:30	Break
2:45	Big Airspace & LOA/SOP Familiarization	2:30	Break	1:45	Review Questionnaires, Issues, and Schedule
3:30	Practice 5	2:45	Practice 11 & 12		
4:15	Caucus	4:00	Break		
		4:15	Caucus		
Week 2					
Time	**Monday**	**Time**	**Tuesday**		**Wednesday**
8:30	Daily In Briefing	8:30	Daily In Briefing	8:30	Daily In Briefing
10:00	Break	9:00	Experiment 5	9:00	Experiment 10
10:15	Experiment 1	10:00	Break	10:00	Break
11:15	Break	10:15	Experiment 6	10:15	Experiment 11
11:30	Experiment 2	11:15	Lunch	11:15	Lunch
12:30	Lunch	12:15	Experiment 7	12:15	Experiment 12
1:30	Experiment 3	1:15	Break	1:15	Break
2:30	Break	1:30	Experiment 8	1:30	Questionnaires & Final Caucus
2:45	Experiment 4	2:30	Break		
3:45	Break	2:45	Experiment 9		
4:00	Caucus	3:45	Break		
		4:00	Caucus		

2.6.2 In Briefing

The experimenter reviewed the schedule of events and explained the general procedures for the experiment, including the dependent measures that would be collected. The experimenter also reviewed the participants' rights and responsibilities as summarized in the informed consent statement (see Appendix A). Next, two SMEs briefed the participants on the hardware and software used in the experiment and presented the SOPs and LOAs for each experimental condition. The SMEs instructed the participants that Instrument Meteorological Conditions would be in effect and also informed them that the ghost controllers would be available to handle requests for aircraft outside of the participant-controlled airspace. The ghost controllers were primarily responsible for managing the arrival aircraft into BAASS. The SMEs also instructed the participants to communicate with the ghost controllers if they wanted to hold or regulate the traffic (e.g., reduce speeds) entering BAASS.

After listening to all of the in-briefing information and asking questions, the participants read and signed the informed consent statement and completed the Biographical Questionnaire (see Appendix B). The experimenter and a witness also signed the informed consent statement. The experimenter gave copies of the briefing slides to the participants so that they could take notes on the maps and refer to them, as needed, when they worked the scenarios.

2.6.3 Practice Scenarios

The participants completed a minimum of 16 practice scenarios (including warm-up scenarios). Each practice or warm-up scenario ran for approximately 30 to 40 minutes and was intended to familiarize the participants with the different sectors in the generic airspace, the equipment, and the different experimental conditions. The participants received instructions about the scenario they were about to work and instructions about the WAK device and rating scale (see Appendix H). They also used the WAK device during the practice scenarios to become accustomed to it. The participants completed the practice scenarios starting with the BL condition followed by the BANC and BAC conditions. The participants worked at each of the positions, under each condition, as illustrated by the sample counterbalancing scheme in Table 4.

Table 4. Sample Sequence of Counterbalancing Order of Practice Conditions

	En Route/Transitional Sector				**Terminal/Near-Airport Sector**					
Practice Run	**01**	**01 D**	**02**	**02 D**	**A**	**A-H**	**D**	**D-H**	**Condition**	**Day**
1	E1	E2	E3	E4	T1	T2	T3	T4	BL – warm up	W
2	E2	E1	E4	E3	T2	T1	T4	T3	BL – warm up	W
3	E4	E3	E2	E1	T4	T3	T2	T1	BL – warm up	W
4	E3	E4	E1	E2	T3	T4	T1	T2	BL – warm up	W
5	E2	E3	E4	E1	T2	T3	T4	T1	BL	W
6	E3	E2	E1	E4	T3	T2	T1	T4	BL	Th
7	E1	E4	E3	E2	T1	T4	T3	T2	BL	Th
8	E4	E1	E2	E3	T4	T1	T2	T3	BL	Th
9	E1	E3	E2	E4	T1	T3	T2	T4	BANC	Th
10	E3	E1	E4	E2	T3	T1	T4	T2	BANC	Th
11	E4	E2	E3	E1	T4	T2	T3	T1	BANC	Th
12	E2	E4	E1	E3	T2	T4	T1	T3	BANC	Th
13	E3	E4	E1	E2	T3	T4	T1	T2	BAC	Fr
14	E4	E3	E2	E1	T4	T3	T2	T1	BAC	Fr
15	E2	E1	E4	E3	T2	T1	T4	T3	BAC	Fr
16	E1	E2	E3	E4	T1	T2	T3	T4	BAC	Fr

Note. E1 = Participant 1 for en route; T1 = Participant 1 for terminal.

2.6.4 Data Collection Procedure

During data collection, the participants completed scenarios as indicated by the counterbalancing scheme shown in Table 2. First, the participants received general instructions about the current experimental condition (see Appendix H). For the BL condition, the experimenters informed the participants that they should control traffic as they normally would in the field. Prior to beginning the BA conditions, the experimenters informed the participants about the airspace boundary changes, that the transitional sector lateral separation minimum would be 3 nm, and that the same separation procedures would be in effect for all sectors. The experimenter also reminded the participants about the dynamic resectorization capability and the additional RNAV routes in the BA conditions. For the BAC condition, the experimenters informed the participants that the transitional and near-airport sectors would be located within one room and that they may use face-to-face communication if they wished.

After the participants received all of the instructions and the experimenters answered all questions relating to the current condition, the participants completed a final radio check and the 50-minute test scenario began. During each scenario, the experimenters and the observers collected the dependent measures, and the participants provided subjective ratings of workload at 4-minute intervals. In addition, video and audio equipment recorded the participants' communications and actions during the experiment in case the researchers needed to review the experiment later.

As soon as the scenario ended, the participants completed PSQ-1 and PSQ-2, if appropriate. The participants then took a break for about 15 minutes before the next scenario began. The participants moved to a new position within their domain (en route/transitional or terminal/near-airport) after every three experiment runs. Before the participants began controlling traffic at a new position, they had time to familiarize themselves with the equipment and adjust their display preferences.

The participants completed the PEQ after completing all of the experimental scenarios. We also held a final debriefing to discuss the experiment and the effects of the BA conditions, as well as any additional requirements (e.g., automation, procedures) needed to support the concept.

3. RESULTS AND DISCUSSION

It takes a little time for participants to acclimate to a scenario. There is also a typical decline in performance at the end of a scenario. Therefore, we included the data from the 4-minute mark to the 48-minute mark in our analyses.

We analyzed the data using a repeated measures analysis of variance (ANOVA) to compare the three airspace conditions (BL, BAC, BANC); see Appendix I for information on repeated measures designs. We analyzed all data from the terminal and en route participants separately (Truitt et al., 2004), except for the system performance measures. We collapsed the data across the two arrival sectors and across the two departure sectors in each airspace condition because the size of the en route/transitional and the terminal/near-airport sectors differed between the BL and BA conditions. Collapsing these data allowed us to evaluate these measures in the same total volume of airspace across all conditions. We conducted 2 (Sector) x 3 (Condition) repeated measures ANOVAs. We also analyzed some measures by the four 11-minute weather intervals across the airspace conditions because we expected that the weather might affect these measures differently.

For all analyses, we determined results as significant when *p* values were less than .05, and we report the *F* values for each relevant analysis. When sphericity was violated, we present the adjusted degree of freedom (*df*) for those tests. When significant interactions were found, we present only the results of the interaction because any significant main effects would not be meaningful. When we found significant effects, we also ran Tukey's Honestly Significant Difference (HSD) post hoc test analyses and report which pairs of differences were significant when *p* values were less than .05.

3.1 System Performance Measures

We analyzed system performance measures to test for hypothesized differences between the conditions. In these analyses, it was important to consider that the size of the en route/transitional and terminal/near-airport sectors changes when going from the BL condition to the BA conditions. Therefore, we did not compare performance in the terminal sectors in the BL condition to the near-airport sectors in the BA conditions or performance in the en route sectors in the BL condition to the transitional sectors in the BA conditions because of the differences in the sizes of the sectors. Consequently, we collapsed the data across the en route and terminal sectors in the BL condition and across the transitional and near-airport sectors in the BAC and BANC conditions to get an overall performance metric for the three airspace conditions. For simplification, these will be referred to as the BAASS + Arrival and GRUPR + Departure sectors for each condition. This issue affected the analysis of all of the system performance metrics, including the total distance flown, the average distance flown per aircraft, the number of aircraft handled, the duration of aircraft handled, the number of holds, the duration of holds, and so on. We also evaluated the number and duration of holds and the number of altitude, speed, and heading changes made by the ghost controller to determine the amount and type of maneuvering of aircraft needed prior to their entry into the BAASS + Arrival sector across the test conditions.

3.1.1 Number of Flights Handled

The number of flights handled was higher in the GRUPR + Departure sector (M = 117.3, SD = 4.23) than in the BAASS + Arrival sector (M = 74.2, SD = 2.55), $F(1, 11)$ = 2,293.6. The GRUPR + Departure sector was geographically larger and handled traffic into and out of satellite airports in one of the ghost sectors in addition to departure aircraft from the primary airport.

We found an average of 94.2 (SD = 3.35) aircraft handled in the BL condition, 96.4 (SD = 4.54) in the BAC condition, and 96.7 (SD = 5.22) in the BANC condition, though these differences were not statistically significant. However, because we only analyzed data for a 44-minute interval (and weather only affected the BAASS sector for approximately 33 minutes), these small differences may be operationally significant if they are extrapolated to longer periods of time.

3.1.2 Time and Distance Flown in Sectors and RNAV Routes

We found that the average time aircraft were in the airspace in the BAASS + Arrival and GRUPR + Departure sectors was affected differently by condition, $F(2, 22)$ = 4.45. The aircraft were in the BAASS + Arrival sector for a longer period of time in the BL condition than in either the BAC or BANC conditions (see Figure 4). The average time in the airspace did not differ across conditions in the GRUPR + Departure sector.

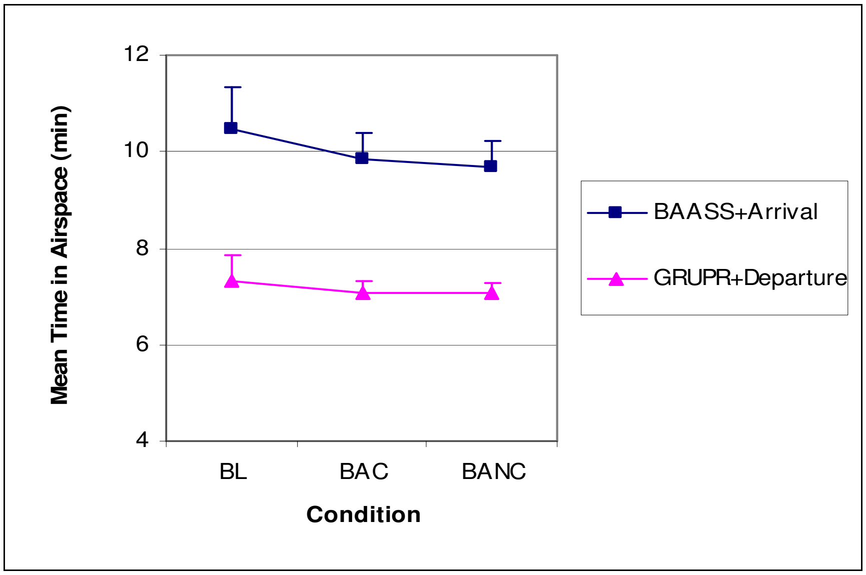

Figure 4. Mean time in airspace by Sector and Condition.

The average distance that the aircraft traveled through the airspace was higher in the BL condition (M = 46.4 nm, SD = 4.07) than in either BAC (M = 43.5, SD = 1.13) or BANC conditions (M = 43.2, SD = 1.46), $F(2, 22) = 13.24$.

We also examined the proportion of time that aircraft were on the RNAV routes (within .25 nm lateral), but we did not find any significant differences across conditions. The aircraft did spend proportionately more time on the RNAV routes in the GRUPR + Departure sector (M = .89, SD = .05) than in the BAASS + Arrival sector (M = .56, SD = .13), $F(1, 11) = 361.87$. This result is not surprising because the weather affected only the BAASS + Arrival sector.

3.1.3 Number of Completed Flights and Number of Departures

We defined the number of completed flights as those that participants handed off to approach control and were below an altitude of 1,200 ft above MSL. We did this to eliminate any instances in which an aircraft would not have landed at the airport because of a technical problem or an error not attributable to the participant. We found an average of 28.8 (SD = 2.12) flights completed in the BANC condition, 27.4 (SD = 4.03) in the BAC condition, and 26.7 (SD = 3.47) in the BL condition. These differences were not significant. However, because of the 44-minute analysis window, these differences may be operationally significant if examined over a longer period.

We evaluated the number of departures for each scenario and did not find any differences across conditions. A total of 45 departures were recorded for each scenario, 29 of which were handed off from the departure sector to GRUPR and 16 of which were handed off from the departure sector to a ghost sector to the east. These numbers did not vary because none of the participants initiated ground stops in the test scenarios.

3.1.4 Losses of Separation

We examined losses of separation differently in the BA and BL conditions for the en route/ transitional sectors because of the different procedures used in those conditions. In the BL condition, en route losses of separation occurred when aircraft were separated by less than 5 nm horizontally and 1,000 ft vertically. Terminal losses of separation occurred when aircraft were separated by less than 3 nm horizontally and 1,000 ft vertically. The terminal separation standards were also used in the transitional sectors in the BA conditions.

We eliminated any losses of separation that occurred only in the ghost sectors, including those that occurred below an altitude of 2,000 ft because these would have been the responsibility of the ghost approach control sector. We also eliminated any losses of separation that were shorter than the duration of one sweep of the radar (12 s in BL for en route, 5 s in the BA conditions and in terminal). We also eliminated other aircraft pairs that were separated by 900 ft to 1,000 ft vertically because the controller does not have information available to indicate separations of less than 100 ft.

The SMEs evaluated the remaining separation violations to determine whether other circumstances warranted that other aircraft pairs should be excluded. For example, if the participants used diverging courses in the terminal environment or in the transitional sectors in the BA conditions, we eliminated these losses of separation as well as any that were determined to have been caused by a pilot's error.

We found eight operational errors in one of the BL scenarios, whereas the other scenarios had from zero to three operational errors. The SME observer's notes from the eight-error scenario indicated that the participants working the BAASS + Arrival sector were experiencing *more than normal* difficulty. The observer also noted that the en route participants took handoffs late and that compression was an issue for traffic downstream when aircraft were handed off to Arrival. The observer also noted that these participants probably would have stopped taking aircraft in this scenario in the real world. As a result, we eliminated this outlier from our analyses and did not find a significant difference in the number of operational errors across conditions. The mean number of operational errors was .72 (SD = .97) in the BL condition, .36 (SD = .82) in the BAC condition, and .23 (SD = .51) in the BANC condition.

Most of the operational errors occurred either in the terminal/near-airport airspace in both the BL condition and BA conditions or close to the boundary between these sectors and the en route/transitional sectors. Several occurred between an aircraft that was arriving or departing and an overflight that was traveling east to west through the terminal/near-airport sectors from one of the satellite airports to another.

3.1.5 Altitude Clearances

The participants issued more altitude clearances in the BL condition (M = 173, SD = 14.12) than in either the BAC condition (M = 154.7, SD = 21.25) or the BANC condition (M = 155.4, SD = 18.51), $F(2, 22) = 11.5$. They also issued more altitude clearances in the BAASS + Arrival sector (M = 172.3, SD = 26.84) than in the GRUPR + Departure sector (M = 149.8, SD = 14.34), $F(1, 11) = 7.3$.

When we examined these data by weather interval, we found differences in the pattern of results obtained across conditions and intervals, $F(6, 66) = 2.9$ (see Figure 5). Overall, the participants issued fewer altitude clearances during the first interval than in each of the others. However, in the BL condition, the participants issued more clearances in both the second and third intervals than the first interval, although in the BA conditions, the participants issued more clearances in only the second interval. The participants issued fewer clearances in the third interval after the airspace resectorization occurred.

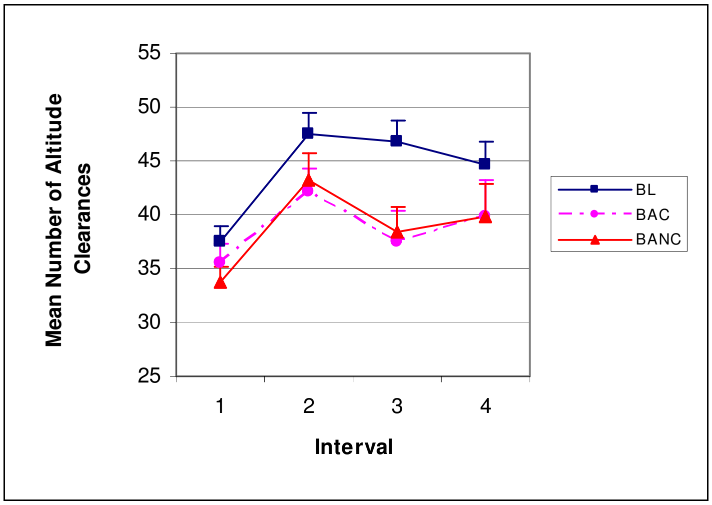

Figure 5. Mean number of altitude clearances issued by Condition and Interval.

The participants issued fewer altitude clearances in the BAASS + Arrival sector in the first interval than in the second and fourth intervals, $F(3, 33) = 11.81$. However, they issued more clearances in the GRUPR + Departure sector in the second interval than in any of the others (see Figure 6).

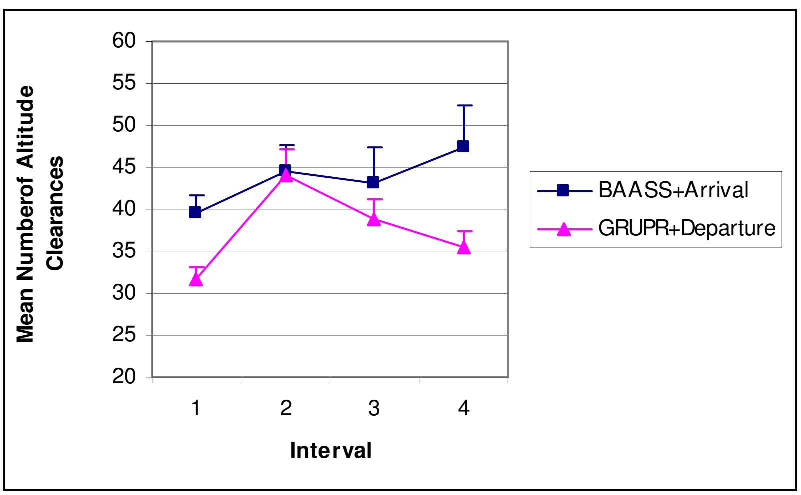

Figure 6. Mean number of altitude clearances issued by Sector and Interval.

In the ghost sector, the ghost controllers issued fewer altitude clearances in the first interval (M = 4.6, SD = 1.66) than in the second (M = 7.9, SD = 3.5) or third (M = 7.1, SD = 3.76), and made more altitude changes during the second and third intervals than the fourth, $F(3, 33) = 14.58$.

3.1.6 Heading Clearances

The participants issued more heading clearances in the BL condition (M = 40.7, SD = 11.97) than in the BANC condition (M = 31.2, SD = 13.34), though neither differed significantly from the BAC condition (M = 35.8, SD = 11.18), $F(2, 22) = 3.95$.

We also found that the participants issued about 10 times more heading clearances in the BAASS + Arrival sector (M = 65.4, SD = 19.61) than in the GRUPR + Departure sector (M = 6.3, SD = 3.11), $F(1, 11) = 111.85$. This finding is not surprising given that more vectoring of aircraft would be expected in the sector affected by weather.

The number of heading changes also differed significantly by condition and weather interval, $F(6, 66) = 2.84$. The participants issued the fewest heading clearances in the first interval for all conditions, but they issued more in the third and fourth intervals in the BL condition than the BANC condition (see Figure 7).

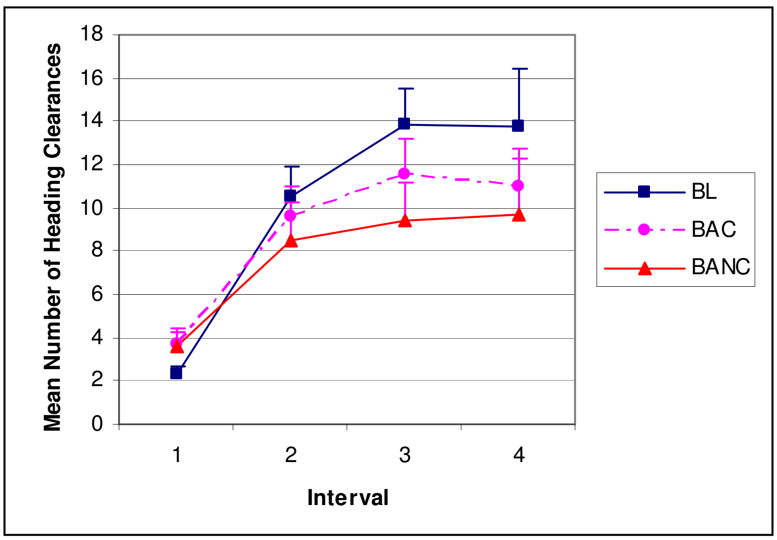

Figure 7. Mean number of heading clearances issued by Condition and Interval.

The number of heading clearances issued in the BAASS + Arrival sector increased across the first three intervals, whereas those in the GRUPR + Departure sector remained the same, $F(1.9, 20.8) = 36.1$ (see Figure 8).

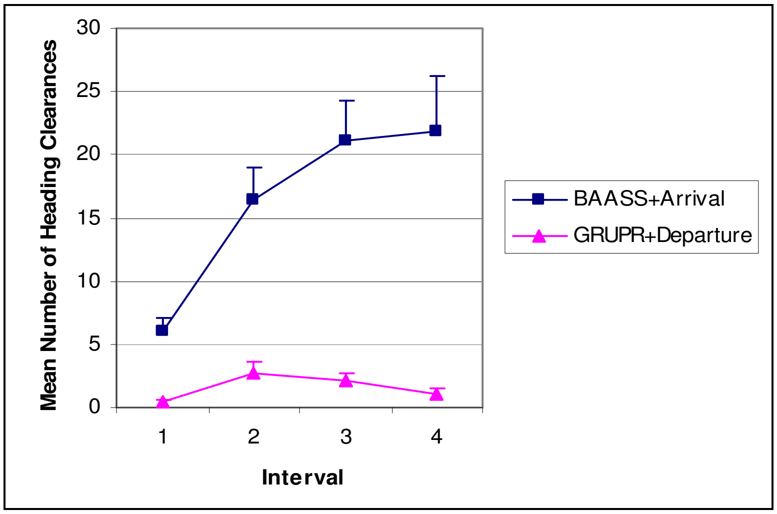

Figure 8. Mean number of heading clearances issued by Sector and Interval.

In the ghost sector, the ghost controller issued more heading clearances in the BL condition (M = 12.8, SD = 4.58) than in either the BAC condition (M = 6.1, SD = 5.66) or BANC condition (M = 4.0, SD = 5.46), $F(2, 22) = 7.97$, indicating that the participants required more assistance in managing traffic before aircraft entered the BAASS + Arrival sector in the BL condition than in the BA conditions.

When we examined these data across weather intervals, and we found that the ghost controller issued more heading clearances during the third and fourth intervals in the BL condition than in either of the BA conditions during that timeframe, $F(2, 22.4) = 7.56$ (see Figure 9). This suggests that less assistance from the ghost controller was utilized when the airspace was resectorized.

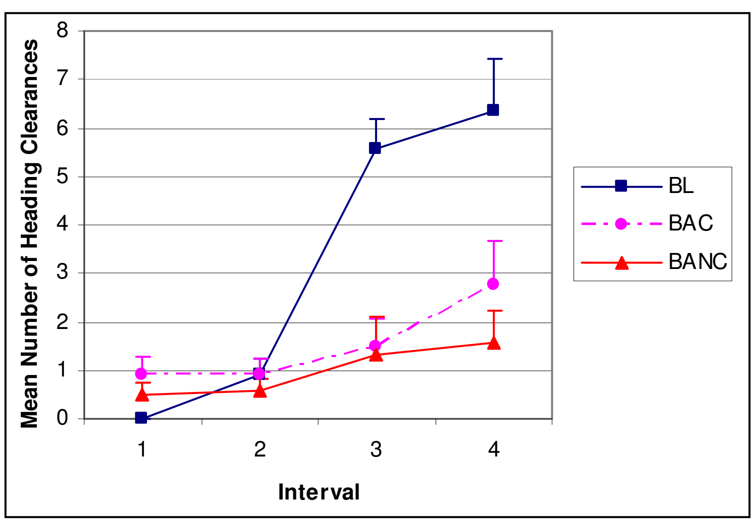

Figure 9. Mean number of heading clearances issued by ghost controller by Condition and Interval.

3.1.7 Speed Clearances

We found that speed clearances in the BAASS + Arrival and GRUPR + Departure sectors were affected differently by condition, $F(2, 22) = 4.31$. The participants issued more speed clearances in the BAC and BANC conditions than in the BL condition in the BAASS + Arrival sector, but they issued the same number of speed clearances across conditions in the GRUPR + Departure sector (see Figure 10). Speed clearances were issued frequently in the BAASS + Arrival sector and rarely in the GRUPR + Departure sector.

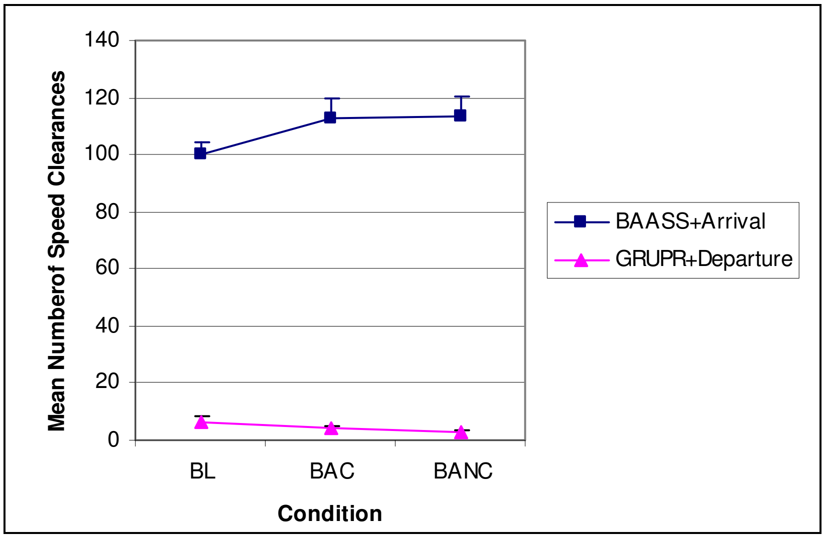

Figure 10. Mean number of speed clearances issued by Sector and Condition.

When we analyzed the data by weather interval, we found that the participants issued fewer speed clearances in the first interval than in the second, third, or fourth, but only for the BAASS + Arrival sector, $F(3, 33) = 29.37$ (see Figure 11).

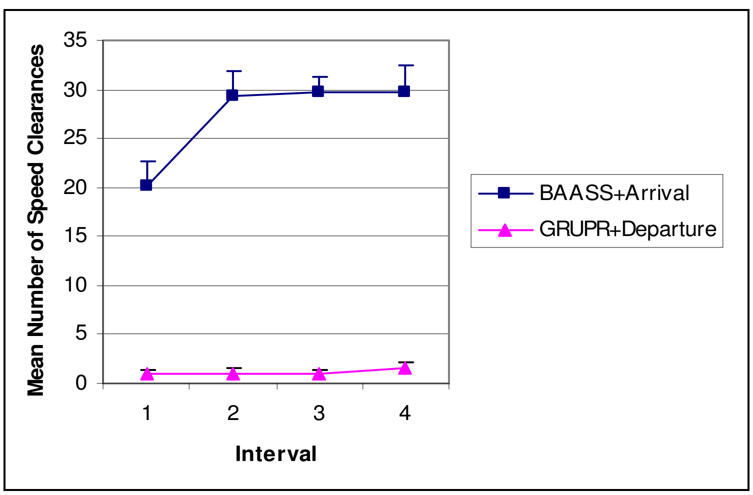

Figure 11. Mean number of speed clearances by Sector and Interval.

The ghost controller issued more speed clearances in the BL condition (M = 34.1, SD = 13.26) than in either the BAC (M = 19.9, SD = 15.47) or BANC conditions (M = 19.5, SD = 15.12), $F(2, 22) = 4.82$. This indicated that the ghost controller provided more assistance in maneuvering aircraft before they entered the BAASS + Arrival sector in the BL condition.

When we analyzed the data by weather interval, we found that the number of speed clearances issued by the ghost controller differed by condition across interval, $F(2.87, 31.51) = 4.76$. The ghost controller issued an increasing number of clearances from the first through the third interval, but they issued more clearances in the third and fourth intervals in the BL condition than in either the BAC or BANC conditions (see Figure 12). This suggested that the participants sought more assistance from the ghost controller when the resectorization was unavailable.

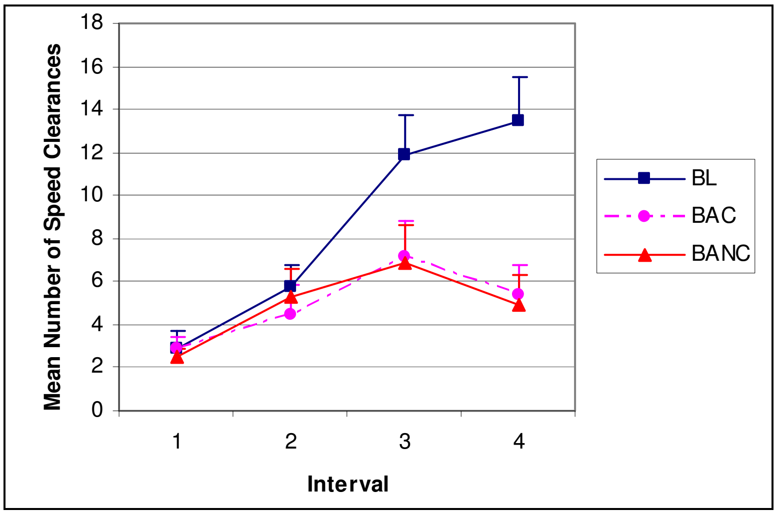

Figure 12. Mean number of speed clearances issued by the ghost controller by Condition and Interval.

3.1.8 Number and Duration of Holds

The participants working the BAASS + Arrival sector could hold aircraft at two fixes within that sector or coordinate with the ghost controller to hold aircraft outside the sector. Overall, the participants did not hold many aircraft in their sectors (see Table 5).

Table 5. Mean Number and Standard Deviation of Hold Commands Issued

	BL	BAC	BANC
BAASS + Arrival	1.42 (2.90)	0	.08 (.29)
Ghost	5.00 (6.47)	.08 (.29)	.08 (.29)

The effect of condition was significant for the number of aircraft held by the ghost controller outside of the BAASS + Arrival sector, $F(1.01, 11.07) = 6.97$. More holding was done in the BL condition than in either the BAC or BANC conditions. The duration of holds outside the sector also differed significantly between the BL and BAC conditions, $F(2, 22) = 4.29$. Mean holding duration was 3.2 minutes (*SD* = 3.71) in the BL condition, but it was less than a minute in the BAC (*M* = .25, *SD* = .86) and BANC conditions (*M* = .72, *SD* = 2.49). Within the BAASS + Arrival sector, the difference in the number and duration of holds did not differ significantly across conditions.

Due to the limited amount of holding data, we did not perform any statistical analyses across the weather intervals. We did find that *no holding* occurred in the first interval either within the BAASS + Arrival sector or within the ghost sector for any condition. However, in the BL condition, the number of holds increased from the second through fourth intervals in both of those sectors, whereas the few holds that occurred in the BA conditions were scattered across those intervals.

3.2 Communications

We measured the mean number of ground-ground and ground-air PTT transmissions for the en route and terminal participants separately. We eliminated any transmissions that were 250 msec in duration or less. It would not have been possible for participants to transmit a meaningful verbal message within this timeframe. We also evaluated the number and type of communications made during the BAC condition when participants had the opportunity to talk face-to-face.

3.2.1 En Route Push-To-Talk Communications

The ground-ground communications included all transmissions made from one participant to another or to the ghost controller. The en route participants made more ground-ground transmissions in the BL condition in the BAASS sector than in either the BAC or BANC conditions, $F(1.3, 14.5) = 9.85$; however, the number did not differ significantly across conditions in the GRUPR sector (see Figure 13).

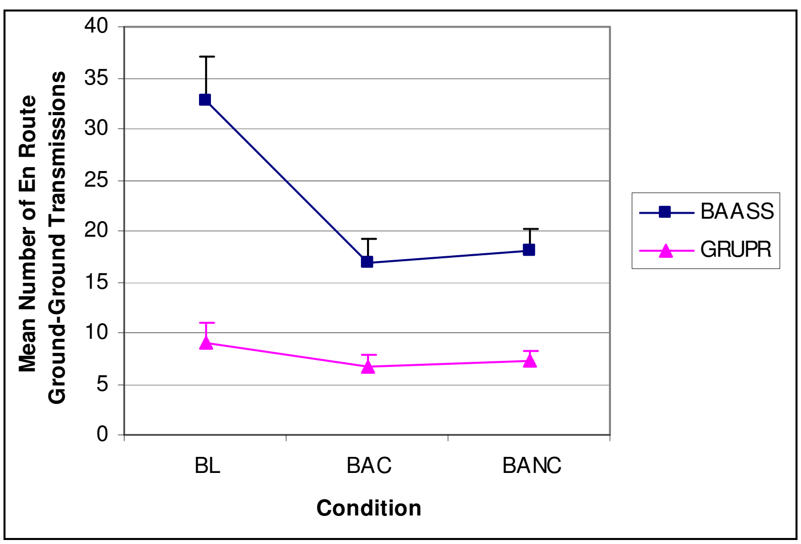

Figure 13. Mean number of en route ground-ground transmissions by Sector and Condition.

To test the effects of a combined control room more closely, we also analyzed the data after eliminating the transmissions that the participants made to the ghost controller because the ghost position was not in the same room as the participants. Though the participants made somewhat more transmissions in the BANC condition (M = 9.8, SD = 6.15) than in the BAC condition (M = 7.4, SD = 4.44), this difference was not significant. The other analyses we ran on these data indicated the same overall effects of condition, sector, and interval reported elsewhere in this section.

We also evaluated the number of ground-ground transmissions across the four weather intervals and found that the number of transmissions increased between the first and second intervals in all conditions and continued to increase in the third interval in the BL condition, $F(6, 66) = 9.47$ (see Figure 14). However, the number of transmissions decreased in the BAC and BANC conditions in the intervals after the airspace resectorization.

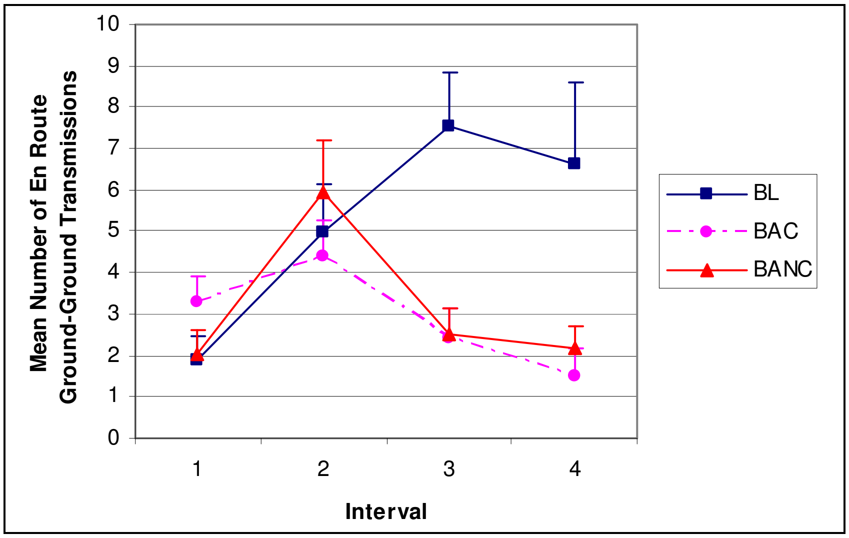

Figure 14. Mean number of en route ground-ground transmissions by Condition and Interval.

We analyzed the ground-air communications similarly to the ground-ground communications. En route participants made more transmissions in the BAC and BANC conditions than in the BL condition in the BAASS sector, $F(2, 22) = 20.06$ (see Figure 15). However, the opposite was found for the GRUPR sector. More transmissions were made in the BL condition than in either the BAC or BANC conditions.

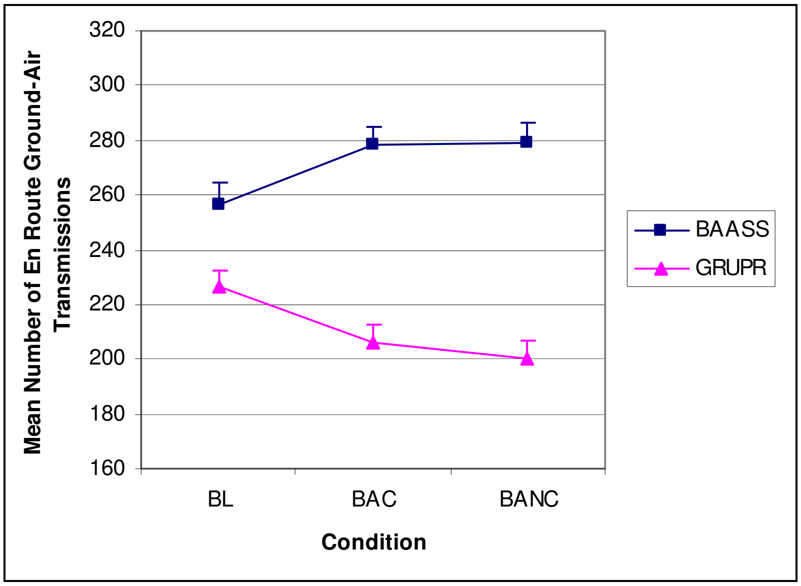

Figure 15. Mean number of en route ground-air transmissions by Sector and Condition.

The most likely reasons for this result are the relative size of the en route/transitional airspace sectors, the effects of resectorization, and the impact of weather in the BAASS sector. In the BL condition, the BAASS sector had a smaller volume of airspace, and the aircraft were handed off earlier and higher to the Arrival sector, so fewer transmissions were needed than in the BA conditions. Aircraft speeds needed to be reduced more in the BAASS sector in the BA conditions to hand off to the smaller Feeder sector. Following resectorization in the BA conditions, more room was also available to maneuver aircraft in the BAASS sector. In the BL condition, the GRUPR sector had a smaller volume of airspace and fewer RNAV routes than in the BA conditions. In the BA conditions, the availability of additional RNAV routes required fewer transmissions.

We also examined the number of ground-air transmissions across the four scenario intervals. We found a significant interaction of Condition x Interval, $F(6, 66) = 5.38$. For all conditions, the participants made more transmissions in the second and third intervals than in the first (see Figure 16). However, participants made more transmissions in the fourth interval than the first in the BL and BAC conditions. They made fewer transmissions in the last interval in the BANC condition than in the BL condition.

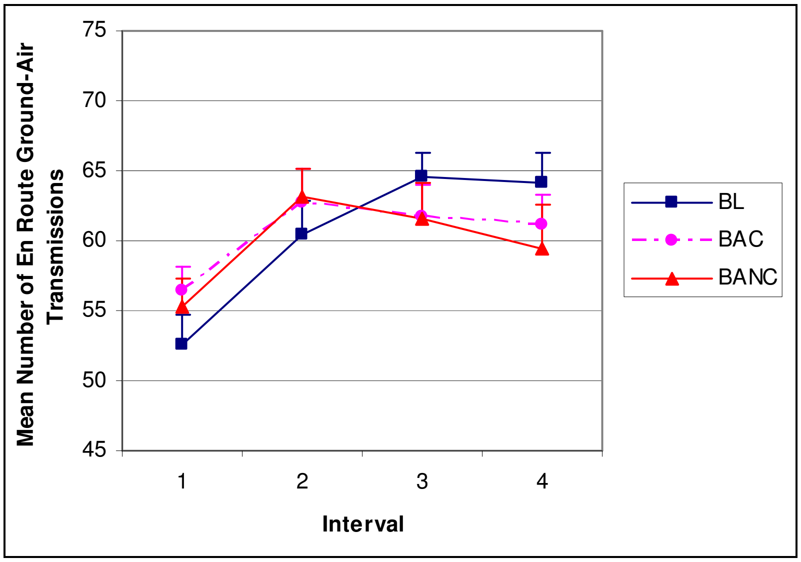

Figure 16. Mean number of en route ground-air transmissions by Condition and Interval.

We also found that the number of transmissions increased similarly in both sectors over the first three intervals, but it increased in the fourth interval in the BAASS sector and decreased in the GRUPR sector to the level observed in the first interval, $F(3, 33) = 29.52$ (see Figure 17). It is possible that this difference was related to the increase in the size of the airspace in the BAASS sector and the decrease in the size of the GRUPR sector following resectorization. More airspace was available to maneuver aircraft around weather in the BAASS sector, whereas the aircraft in the GRUPR sector were well established on the RNAV routes as the scenarios progressed.

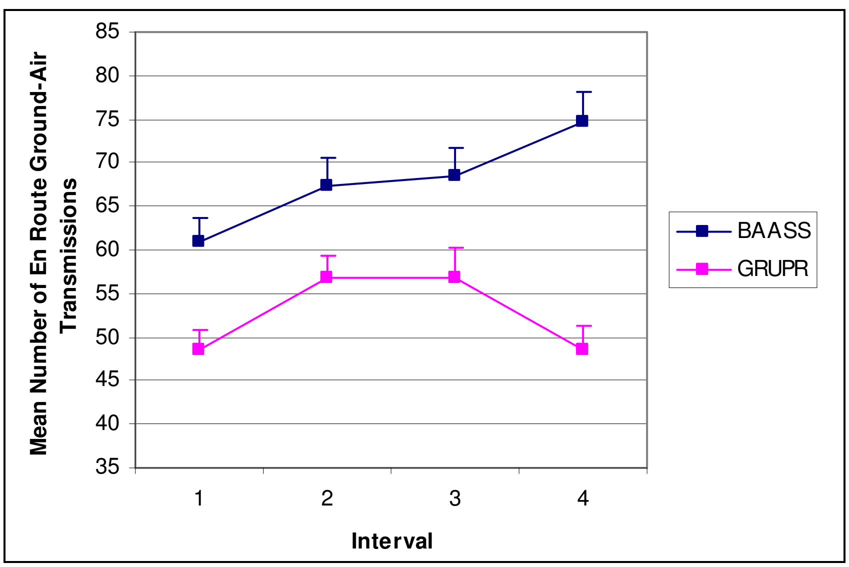

Figure 17. Mean number of en route ground-air transmissions by Sector and Interval.

3.2.2 Terminal Push-To-Talk Communications

The number of ground-ground communications made by terminal participants varied widely. We found that the terminal participants made more transmissions in the Arrival sector in the BL condition than in the Feeder sector in either the BAC or BANC conditions, $F(2, 22) = 3.64$ (see Figure 18). This may have been associated with the relative decrease in the sector size in the BA conditions. The participants made very few transmissions in the Departure/Airport Departure sector and the number of transmissions did not differ significantly across conditions.

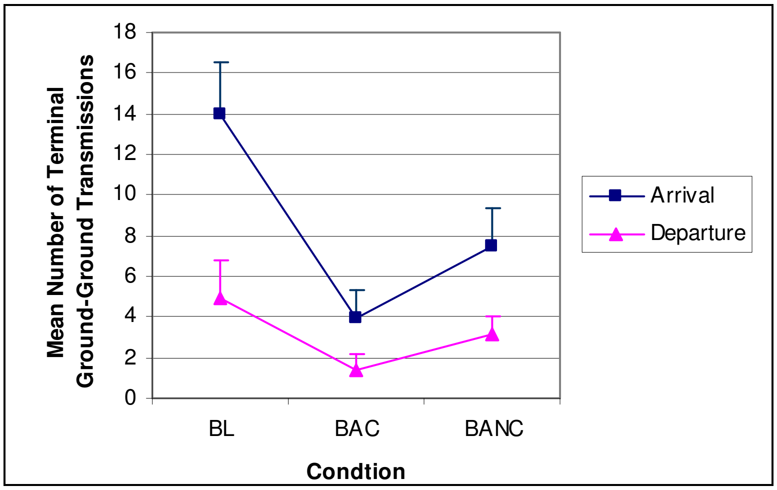

Figure 18. Mean number of terminal ground-ground transmissions by Sector and Condition.

We also found a significant effect of interval, $F(3, 33) = 7.38$. The participants made the fewest transmissions in the first interval ($M = .6$, $SD = 1.55$). The means for the second, third, and fourth intervals did not differ significantly. In order, the means for intervals 2 through 4 were 2.0, 1.5, and 1.75 (*SDs* = 4.08, 2.63, and 2.9, respectively.

The participants made more ground-air transmissions in the BL condition than in either the BAC or BANC conditions in the Arrival/Feeder sector, but the number of transmissions did not differ across conditions in the Departure/Airport Departure sector, $F(2, 22) = 6.17$ (see Figure 19). The number of transmissions made in the Arrival/Feeder sector was almost double the number made in the Departure/Airport Departure sector.

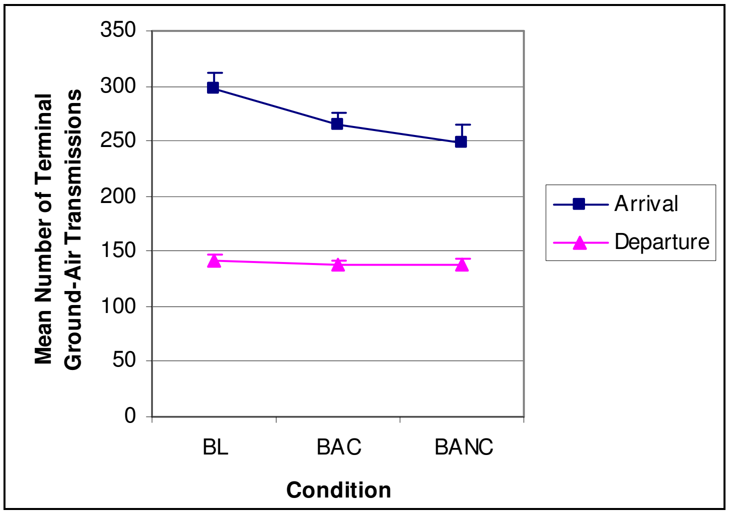

Figure 19. Mean number of terminal ground-air transmissions by Sector and Condition.

The participants made fewer ground-air transmissions in the first interval (M = 39, SD = 12.45) than in the second (M = 58.6, SD = 16.15), third (M = 55.4, SD = 20.22), or fourth (M = 51.4, SD = 16.9) intervals, $F(3, 33) = 88.03$.

3.2.3 Face-to-Face Communication

We examined the number and type of communications made between the en route and terminal participants in the BAC condition when the opportunity to directly interact with one another was possible. We categorized any viewing behavior by a participant as a glance regardless of how long or short the duration, and we did observe instances in which a participant spent a few minutes viewing another display.

We observed relatively few interactions, but this varied widely by participant. On average, the participants looked at one another's displays 2.2 times (SD = 3.36) per scenario and communicated verbally 2.7 times (SD = 3.64) per scenario. The greatest number of verbal communications pertained to speeds (30 total observations across all scenarios) followed by communications regarding frequencies, traffic flow, approvals, altitudes, handoffs, routes, traffic, point outs, and equipment.

Overall, the terminal participants initiated more glances and verbal communications than the en route participants. The terminal participants viewed the en route displays an average of 2.8 times (SD = 4.0) per scenario, whereas the en route participants viewed the terminal displays an average of 1.6 times (SD = 2.48). The terminal participants initiated an average of 3.3 (SD = 4.2) communications, and the en route participants initiated an average of 2.1 (SD = 2.9) communications per scenario.

The participants working the Arrival Transition and Feeder sectors initiated more glances and verbal communications than those working the Departure Transition and Airport Departure sectors. The participants working the Feeder sector initiated an average of 4.9 (SD = 4.81) glances and 6.2 (SD = 4.28) communications. The participants working the BAASS sector made an average of 2.6 (*SD* = 3.15) glances and 3.6 (*SD* = 3.50) communications. Those working the GRUPR sector and the Airport Departure sector made an average of less than one glance and one verbal communication per scenario.

It is likely that some of the differences observed in the number of participant interactions were due to differences in the perceived difficulty between the Arrival Transition and Feeder sectors and the Departure Transition and Airport Departure sectors. However, the laboratory configuration was also likely to have influenced the interactions. The BAASS and Feeder sectors were located side-by-side but, due to room constraints, the GRUPR and Airport Departure sectors were not located adjacently. Therefore, position layout may have influenced the way in which the participants interacted. In the final debriefing, the participants commented that the related sectors should be placed in close proximity to maximize benefits.

3.3 ATC Observer Ratings

The SMEs evaluated the performance of the participants in each of the scenarios using rating scales that ranged from 1 (*least effective*) to 8 (*most effective*) on the ORF. Additional questions pertained to the frequency of occurrence of problematic events and used 5-point rating scales in which a rating of 1 indicated that an event *never occurred*, a rating of 5 indicated that an event *occurred unacceptably often*, and a rating of 3 indicated that an event *occurred, but within normal limits of operational acceptability.* One SME provided ratings for the en route participants and another provided ratings for the terminal participants. We analyzed the ratings separately for each group and evaluated whether ratings differed significantly by sector and condition.

3.3.1 En Route Observer Ratings

Overall, the observer rated the en route participants' task performance as very effective, with mean ratings in each category over 6.5 (and *SDs* close to 1). The ratings of participant ability to maintain separation and resolve potential conflicts differed by condition, $F(2, 22) = 21.32$. The ratings were higher in the BAC ($M = 7.1$, $SD = .66$) and BANC conditions ($M = 7.0$, $SD = .56$) than in the BL condition ($M = 6.0$, $SD = .82$). We also found that ratings in the GRUPR sector ($M = 7.0$, $SD = .58$) were higher than ratings in the BAASS sector ($M = 6.3$, $SD = 1.27$), $F(2, 22) = 6.91$.

The observer ratings for sequencing aircraft efficiently were higher in the BAC ($M = 7.3$, $SD = .63$) and BANC conditions ($M = 7.2$, $SD = .77$) than in the BL condition ($M = 6.0$, $SD = 1.04$), $F(1, 11) = 21.52$. The effect of sector was also significant, with higher ratings in the GRUPR sector ($M = 7.2$, $SD = .76$) than the BAASS sector ($M = 6.5$, $SD = .90$), $F(1, 11) = 13.89$.

The ratings for using control instructions effectively/efficiently were also higher in the BAC ($M = 7.3$, $SD = .56$) and BANC conditions ($M = 7.2$, $SD = .70$) than in the BL condition ($M = 6.0$, $SD = .82$), $F(2, 22) = 28.28$. The ratings of overall safe and efficient traffic flow were lowest in the BL condition in the BAASS sector, $F(1.2, 12.6) = 6.67$ (see Figure 20).

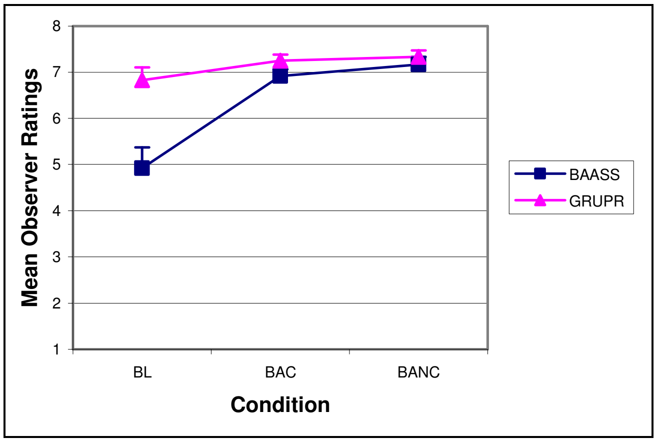

Figure 20. Mean observer ratings for safe and efficient traffic flow by Sector and Condition.

We did not find significant differences for the ratings of D-side communication and coordination. However, the observer rated D-side flight plan amendments and management of data blocks higher in the BAC condition (M = 7.2, SD = .55) and BANC condition (M = 7.0, SD = .77) than in the BL condition (M = 6.7, SD = .63), $F(2, 22)$ = 5.30. The ratings in the GRUPR sector (M = 7.1, SD = .58) were higher than ratings in the BAASS sector (M = 6.8, SD = .83), $F(1, 11)$ = 6.76.

Most of the ratings about the frequency of occurrence of problematic events were low, with mean ratings typically at or below 2. Only a small number (less than 5%) of the individual ratings were either 4 or 5. Most (80%) of those high ratings were given in the BL condition. We found significant differences in these ratings on four of the eight tasks.

The observer noted that more clearances were issued inappropriately in the BAASS sector in the BL condition than in the other conditions, $F(2, 22)$ = 6.95 (see Figure 21).

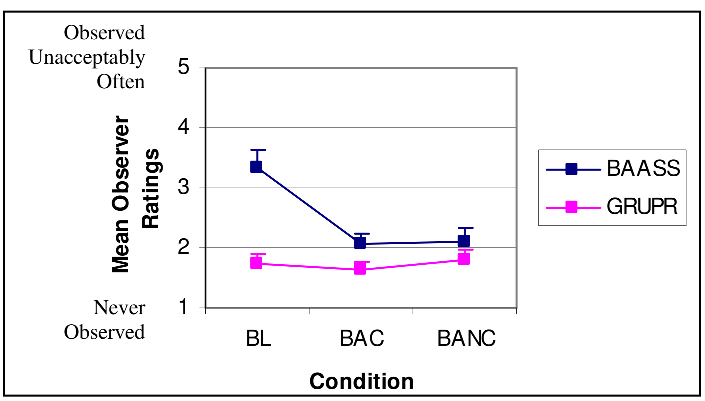

Figure 21. Mean observer ratings for issuing clearances earlier or later than appropriate by Sector and Condition.

The observer also found that more handoffs were offered later than appropriate in the BAASS sector in the BL condition, $F(2, 22) = 3.71$ (see Figure 22). The observer also rated that handoffs were accepted later than appropriate more often in the BL condition in the BAASS sector than in the BAC condition, $F(2, 22) = 3.894$ (see Figure 23). The ratings in GRUPR did not differ significantly across conditions.

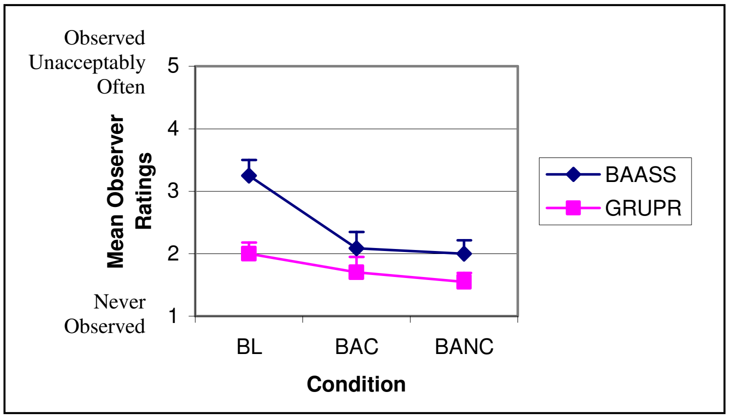

Figure 22. Mean observer ratings for offering handoffs later than appropriate by Sector and Condition.

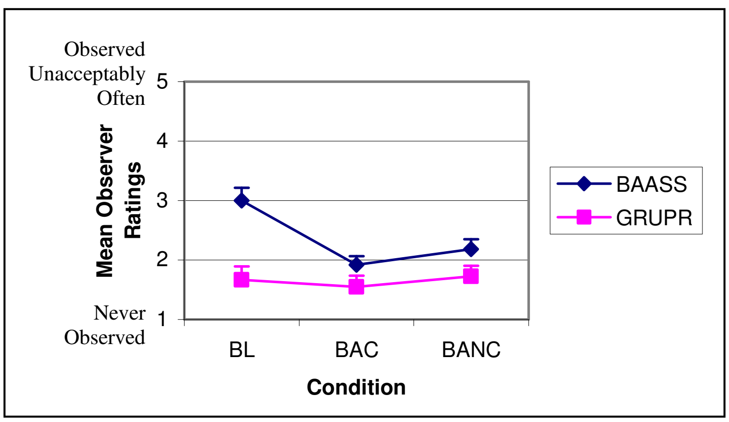

Figure 23. Mean observer ratings for accepting handoffs later than appropriate by Sector and Condition.

Finally, the observer rated that communications were transferred later than appropriate more often in the BAASS sector (M = 2.5, SD = .86) than in the GRUPR sector (M = 1.6, SD = .62), $F(1, 11) = 16.60$.

Comments from the observer elaborated on these ratings. In the BL condition, the observer noted difficulties sequencing aircraft, having room to vector aircraft, and keeping up with the pace of traffic once the weather impacted the BAASS sector. The observer also noted the use of some holding and some late descents and missed handoffs for this condition and sector. Fewer problems were noted for GRUPR and fewer negative comments were made about performance in the BA conditions. Comments on those conditions indicated that the use of speeds was effective, and that there was generally a smoother flow of traffic than in the BL condition. Only three negative comments were noted in the BA conditions in which participants were observed to have done considerable vectoring in the BAASS sector or allowed traffic to become compressed into the Feeder sector.

3.3.2 Terminal Observer Ratings

Overall, the observer rated the terminal participants' performance very highly. There was little to no variability across the test conditions, which made it impossible to analyze these data statistically. The mean rating was 7.8 (SD = .57) for maintaining separation and resolving potential conflicts, 7.9 (SD = .42) for sequencing aircraft efficiently, 7.8 (SD = .43) for using control instructions effectively/efficiently, and 7.8 (SD = .43) for overall safe and efficient traffic flow. The D-side handoff communication and coordination ratings averaged 7.9 (SD = .26).

We also found very little variability for the frequency of occurrence ratings. Most of these ratings were very low, indicating that problematic instances were rarely observed. Mean ratings were between 1.0 and 1.2 (and *SDs* between .17 and .4) for each of these variables.

3.4 WAK Ratings

We analyzed the WAK ratings separately for the en route and terminal participants. We coded instances in which participants did not respond as missing data and included the mean rating obtained for an interval in that cell so that we would not have to drop data from the analysis.[1] We chose to do this rather than to assign the highest workload rating of 10 to instances in which participants did not respond because we could not be certain why a participant did not respond. The participant may have been very busy, but he may simply have been occupied with another task (e.g., making a call) that diverted his attention from the WAK prompt. We prompted each R-side participant for a response every 4 minutes throughout the 50-minute scenarios (a total of 12 prompts), and we took an average of those responses to obtain an overall WAK rating for each individual.

[1] In a repeated measures design, all data for a participant are omitted from the analysis when one or more cells contain missing data. Because the participants had 144 opportunities to respond to the WAK prompt across all of the test scenarios, it was likely that there would be at least one missed response. To enable us to conduct an analysis of these data, we employed the mean substitution procedure (e.g., Tabachnick & Fidell, 1989).

3.4.1 En Route Participant WAK Ratings

For the en route participants, WAK ratings were in the low-moderate range but were highly variable. Mean ratings were 3.8 (*SD* = 2.66), 3.5 (*SD* = 3.08), and 3.6 (*SD* = 3.06) for the BL, BAC, and BANC conditions, respectively, which did not differ significantly. However, the participants reported higher workload levels when working the BAASS sector (M = 4.1, SD = 3.85) than when working the GRUPR sector (M = 3.12, SD = 3.52), $F(1, 11) = 17.51$.

To examine workload across the weather intervals, we averaged the three individual ratings in each interval to obtain an overall interval workload rating. We found that average workload ratings were higher in the last two intervals in the BL condition than in the BAC and BANC conditions, $F(2.7, 29.5) = 5.52$ (see Figure 24). The last two intervals included the workload ratings made after the dynamic resectorization in the BA conditions.

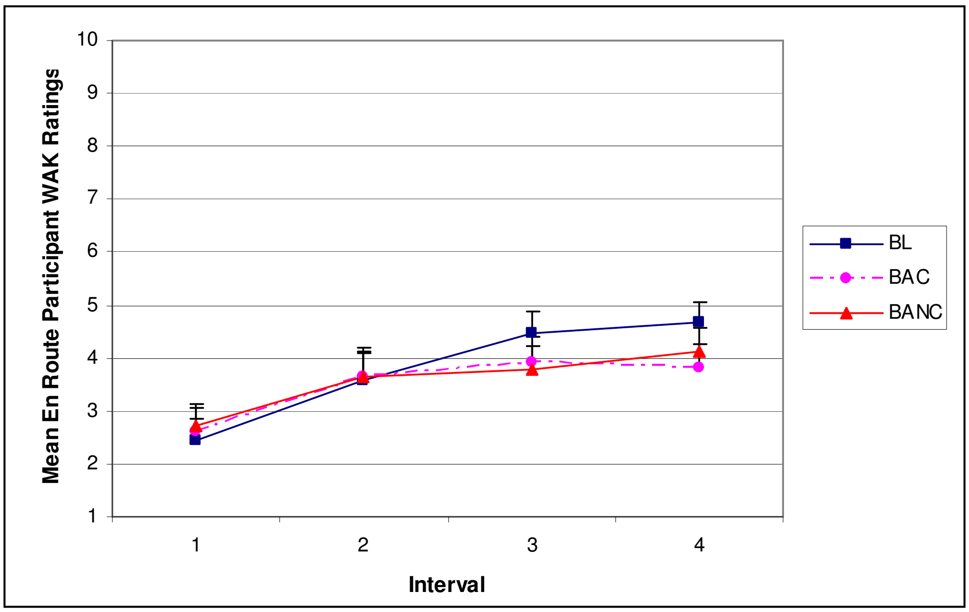

Figure 24. Mean en route participant WAK ratings by Condition and Interval.

We also found that average workload levels increased across intervals in the BAASS sector but did not increase similarly in the GRUPR sector, $F(3, 33) = 20.29$ (see Figure 25). Each of the successive means was significantly higher than the previous one in the BAASS sector, but only the first interval rating was significantly lower than each of the other ratings in the GRUPR sector.

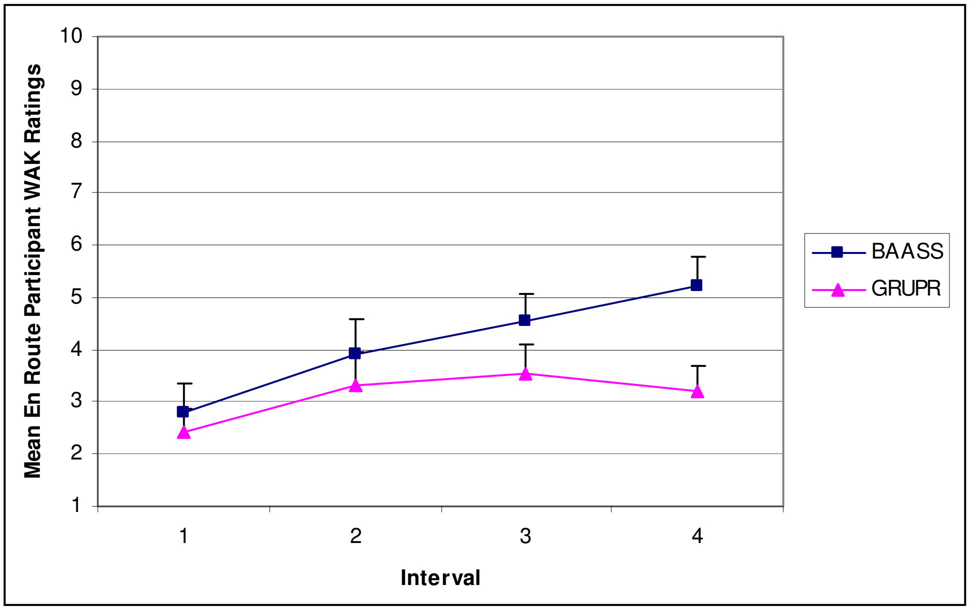

Figure 25. Mean en route participant WAK ratings by Sector and Interval.

3.4.2 Terminal Participant WAK Ratings

Terminal participant WAK ratings were low overall, with average ratings about three or less. WAK ratings were higher for the BL condition (*M* = 2.7, *SD* = 3.13) than for either the BAC (*M* = 2.4, *SD* = 2.36) or BANC conditions (*M* = 2.3, *SD* = 2.29), $F(1.3, 14.5) = 7.38$. WAK ratings were also higher in the Arrival/Feeder sector (*M* = 3.3, *SD* = 4.76) than in the Departure/ Airport Departure sector (*M* = 1.66, *SD* = 1.64), $F(1, 11) = 31.63$.

Terminal participant WAK ratings were lower in the first interval than in any of the others for all conditions, but they were also higher in the third interval than the second in the BAC condition, $F(1.2, 13.6) = 2.79$ (see Figure 26). In the second interval, WAK ratings were higher in the BL condition than in either the BAC or BANC conditions. In the third interval, WAK ratings were higher in the BL condition than in the BANC condition.

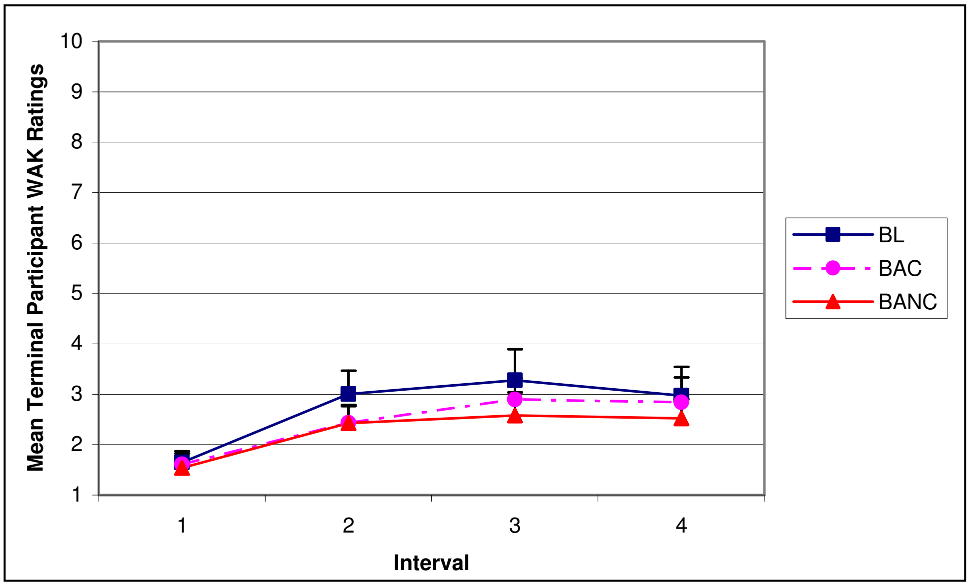

Figure 26. Mean terminal participant WAK ratings by Condition and Interval.

WAK ratings also increased across intervals more in the Arrival/Feeder sector than in the Departure/Airport Departure sector, $F(1.3, 14.2) = 12.6$ (see Figure 27). In the Arrival/Feeder sector, the first interval ratings were lower than ratings in the other intervals, and ratings in the last interval were also higher than those in the second interval. In the Departure/Airport Departure sector, WAK ratings differed significantly only between the first interval and the second interval and between the first interval and the third interval.

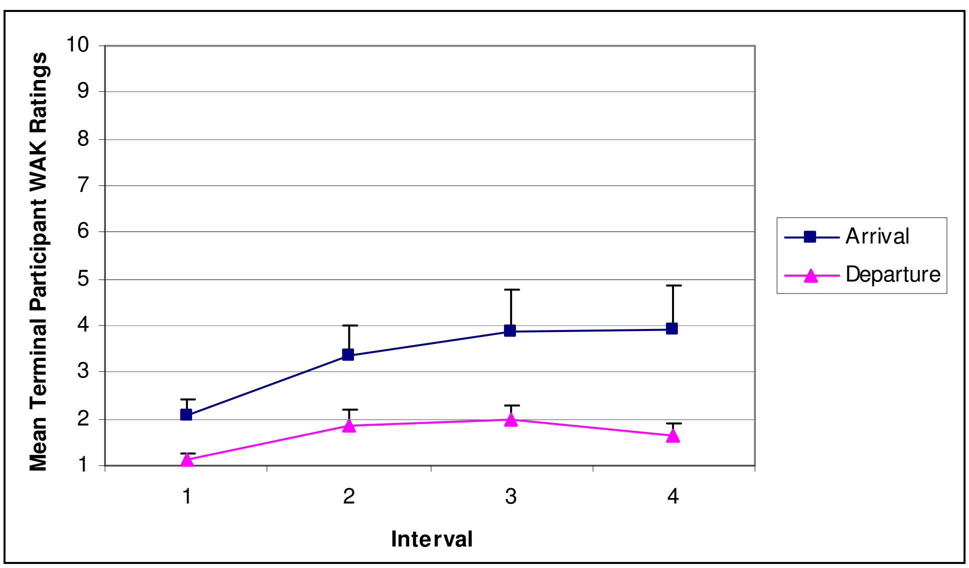

Figure 27. Mean terminal participant WAK ratings by Sector and Interval.

3.5 Post-Scenario Questionnaires

Items on the PSQ-1 utilized a 10-point scale rating (see Appendix C). We analyzed each item on the PSQ-1 separately for the en route and terminal participants, and we analyzed the data for the R-side and D-side (or handoff) positions separately.

3.5.1 En Route Post-Scenario Questionnaire 1

The en route participants rated their ATC performance fairly high, overall, with mean ratings greater than 7. The R-side participants rated their performance higher in the BAC condition (M = 7.9, SD = 1.39) than in the BL condition (M = 7.2, SD = 1.56), $F(2, 20) = 4.26$. They also rated their performance higher in the GRUPR sector (M = 8.1, SD = 1.4) than in the BAASS sector (M = 7.2, SD = 2.04), $F(1, 10) = 7.5$. The D-side participants rated their performance lowest in the BAASS sector in the BL condition, $F(2, 22) = 8.22$ (see Figure 28).

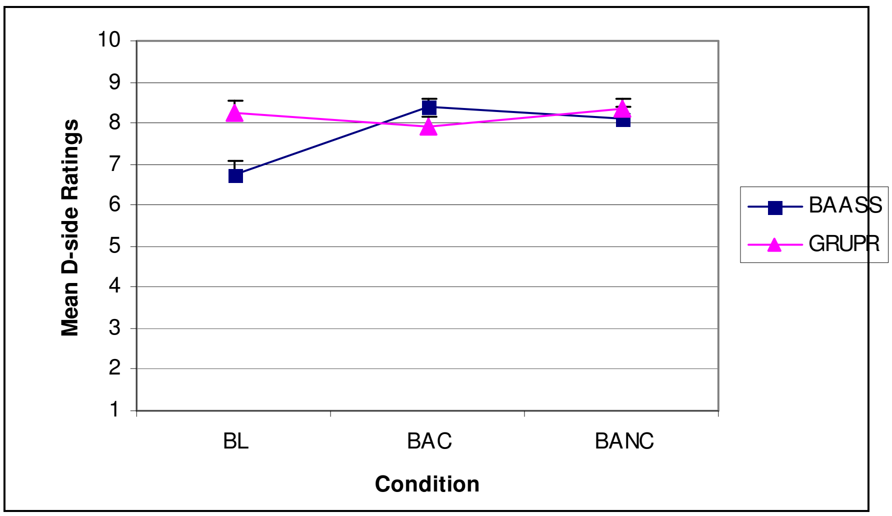

Figure 28. Mean D-side ratings of ATC performance by Sector and Condition.

Four of the PSQ-1 items pertained to situation awareness. These included overall situation awareness, situation awareness for current aircraft locations, situation awareness for projected aircraft locations, and situation awareness for potential loss of separation. In general, when we found significant differences in these ratings, they favored the BA conditions and the GRUPR sector.

The R-side participants rated their overall situation awareness higher in the GRUPR sector (M = 8.3, SD = 1.4) than in the BAASS sector (M = 7.19, SD = 2.13), $F(1, 11) = 9.9$. The D-side participants rated their overall situation awareness higher in the BANC condition (M = 8.5, SD = .93) than in the BL condition (M = 7.9, SD = .8); neither of which differed from the BAC condition (M = 8.4, SD = 1.2), $F(2 ,22) = 4.86$.

The R-side participants rated their situation awareness for current aircraft locations higher in the BAC condition (M = 8.0, SD = 1.22) and the BANC condition (M = 8.0, SD = 1.36) than in the BL condition (M = 7.1, SD = 1.67), $F(2, 22) = 6.62$. They also rated this variable higher in the

GRUPR sector (M = 8.2, SD = 1.44) than in the BAASS sector (M = 7.2, SD = 2.15), $F(2, 22)$ = 6.91. There were no significant differences found for the D-side participants on this variable.

The R-side participant ratings of situation awareness for projected aircraft locations varied by condition and sector, $F(2, 22) = 4.42$. In the BAASS sector, ratings were higher in the BAC condition than in the BL condition. In the GRUPR sector, the ratings did not differ between conditions (see Figure 29). We also found this result for the D-side participants, $F(2, 22) = 4.98$.

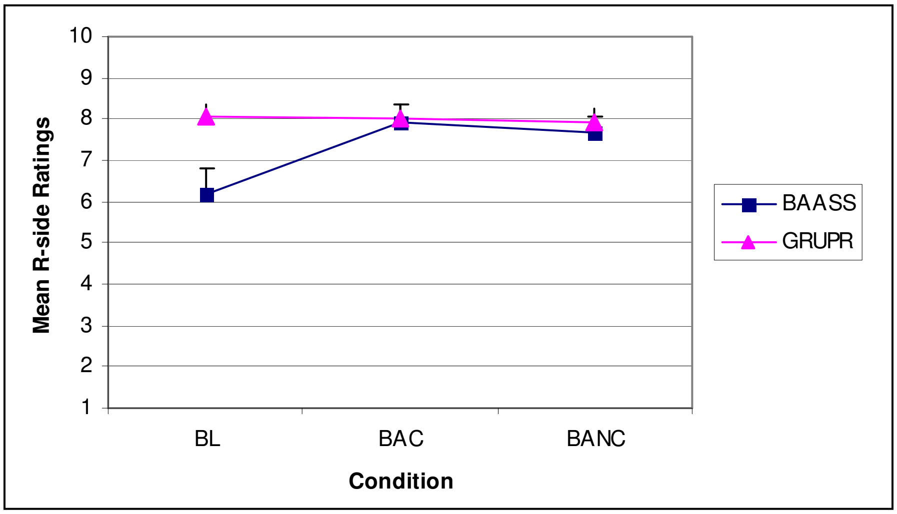

Figure 29. Mean R-side ratings of situation awareness for projected aircraft locations.

Both the R-side and D-side participants rated situation awareness for potential loss of separation higher in the BA conditions than in the BL condition, $F(1.26, 13.84) = 8.88$ and $F(2, 22) = 6.34$, respectively (see Figure 30).

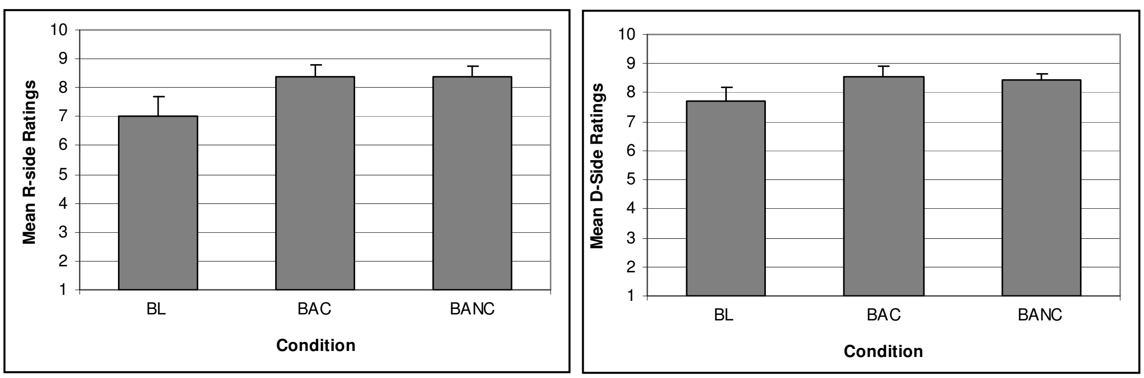

Figure 30. Mean situation awareness ratings for potential loss of separation for R-side (left) and D-side (right) participants by Condition.

Workload ratings due to ground–ground communications for D-side participants were higher in the BL condition (*M* = 5.0, *SD* = 2.6) than in the BAC condition (*M* = 3.7, *SD* = 2.75), but neither differed from the BANC condition (*M* = 4.3, *SD* = 2.56), *F*(2, 22) = 3.89. These workload ratings were also higher in the BAASS sector (*M* = 5.2, *SD* = 3.71) than the GRUPR sector (*M* = 3.5, *SD* = 3.5), *F*(1, 11) = 4.98.

Both the R-side and D-side participants rated their overall workload higher in the BL condition than in the BAC condition, *F*(2, 22) = 8.93 and *F*(2, 22) = 5.04, respectively. For the R-side, we also found higher ratings in the BL condition than in the BANC condition (see Figure 31).

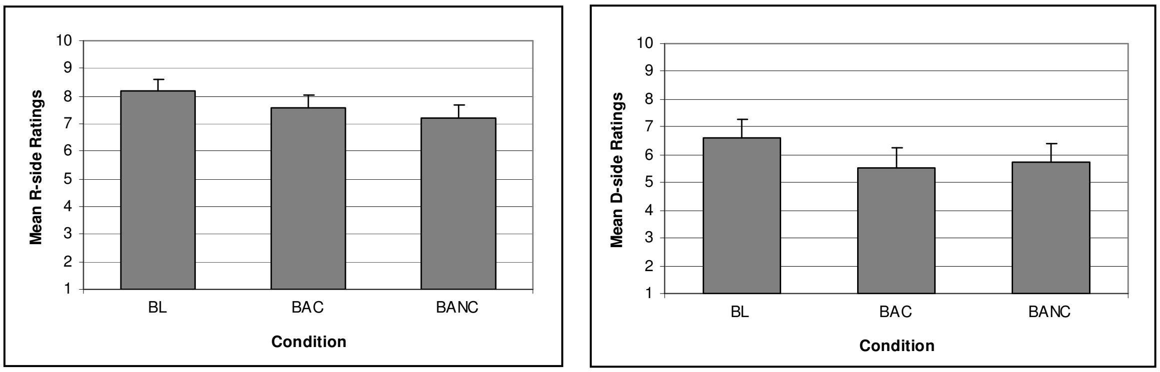

Figure 31. Mean overall workload ratings for R-side (left) and D-side (right) participants by Condition.

The R-side and D-side participants also rated their overall workload higher in the BAASS sector (R-side mean = 8.5, *SD* = 1.83; D-side mean = 6.9, *SD* = 2.75) than in the GRUPR sector (R-side mean = 6.7, *SD* = 2.12; D-side mean = 5.0, *SD* = 3.16), *F*(1, 11) = 27.78 and *F*(1, 11) = 18.27, respectively.

Three other PSQ-1 items were administered only to R-side participants. Participants rated their ability to move aircraft through a sector higher in both the BAC condition (*M* = 8.4, *SD* = 1.76) and the BANC condition (*M* = 8.5, *SD* = 1.26) than the BL condition (*M* = 6.7, *SD* = 1.49), *F*(1.1, 12.3) = 16.43. They also rated workload due to air-ground transmissions higher in the BAASS sector (*M* = 7.8, *SD* = 2.11) than in the GRUPR sector (*M* = 6.1, *SD* = 4.26), *F*(1, 11) = 7.38, and they rated pilot performance higher in the GRUPR sector (*M* = 8.3, *SD* = 1.77) than in the BAASS sector (*M* = 7.8, *SD* = 1.71), *F*(1, 11) = 9.5.

3.5.2 Terminal Post-Scenario Questionnaire 1

We conducted the same analyses for terminal participants that we conducted for en route. The ratings of ATC performance were high, with averages greater than 7.5. The R-side and handoff participants rated their performance higher in the Departure/Airport Departure sector (R-side mean = 8.6, *SD* = 1.27; Handoff mean = 8.7, *SD* = 1.98) than in the Arrival/Feeder sector (R-side mean = 7.6, *SD* = 1.78; Handoff mean = 8.2, *SD* = 1.80), *F*(1, 11) = 8.55 and *F*(1, 11) = 9.16, respectively.

There were no significant differences for overall situation awareness and situation awareness for potential loss of separation for either the R-side or handoff positions. There were also no significant effects for the R-side participants on ratings of situation awareness for current aircraft locations. However, those working the handoff position rated that variable higher in the Departure/Airport Departure sector (M = 8.72, SD = 1.57) than the Arrival/Feeder sector (M = 8.31, SD = 1.36), $F(1, 11) = 12.69$.

The R-side participants rated their situation awareness of projected aircraft locations higher in both the BAC condition (M = 8.3, SD = 1.61) and BANC condition (M = 8.3, SD = 1.26) than in the BL condition (M = 7.8, SD = 1.33), $F(2, 22) = 4.89$. The handoff participants rated their awareness higher in the Departure/Airport Departure sector (M = 8.6, SD = 1.94) than in the Arrival/Feeder sector (M = 8.0, SD = 1.53), $F(1, 11) = 9.78$.

Workload ratings due to ground-ground communications were highly variable. The R-side and Handoff participants rated workload due to ground-ground communications higher in the Arrival/Feeder sector (R-side mean = 4.4, SD = 4.9; Handoff mean = 3.2, SD = 3.51) than in the Departure/Airport Departure (R-side mean = 2.7, SD = 2.79; Handoff mean = 2.1, SD = 1.96), $F(1, 11) = 6.73$ and $F(1, 11) = 5.08$, respectively. Both the R-side and Handoff participants also rated overall workload higher in the Arrival/Feeder sector (R-side mean = 6.2, SD = 3.08; Handoff mean = 4.7, SD = 3.15) than in the Departure/Airport Departure sector (R-side mean = 3.2, SD = 2.49; Handoff mean = 2.4, SD = 2.47), $F(1, 10) = 23.26$ and $F(1, 11) = 14.35$, respectively.

The three other PSQ-1 items pertained only to R-side participants. They rated their ability to move aircraft through the sector higher in the Departure/Airport Departure sector (M = 8.61, SD = 1.51) than in the Arrival/Feeder sector (M = 7.89, SD = 1.75), $F(1, 11) = 5.18$ and their air-ground communication workload higher in the Arrival/Feeder sector (M = 5.86, SD = 2.93) than in the Departure/Airport Departure sector (M = 3.92, SD = 3.39), $F(1, 11) = 12.85$. There was no difference in ratings of simulation pilot performance.

3.5.3 Post-Scenario Questionnaire 2

We administered the PSQ-2 (see Appendix D) following completion of the BAC and BANC scenarios to allow participants to respond more specifically to the effect of the procedures used in these conditions compared to the BL condition. The participants made ratings on a 9-point scale in which a rating of 1 indicated a *negative effect*, a rating of 5 indicated *no effect,* and a rating of 9 indicated a *positive effect*. We conducted a 2 (Sector) x 2 (Condition) repeated measures ANOVA to determine whether there were any significant differences.

Overall, en route R-side participants rated the effect of reduced lateral separation on their ability to control traffic as very positive, with mean ratings of about 8 overall. The ratings were higher in the BAASS sector (M = 8.2, SD = 1.30) than in the GRUPR sector (M = 7.9, SD = 1.59), $F(1, 11) = 6.49$. There were no significant differences for D-side ratings.

The participants also rated the effect of using other terminal procedures on their ability to control traffic somewhat positively, with average ratings over 6.5 (SD = 1.62). There were no significant differences for R-side or D-side participants on this rating.

Both the R-side and D-side participants rated the effect of dynamic sector boundaries on their ability to control traffic higher in the BAASS sector – R-side (M = 7.1, SD = 2.04) and D-side (M = 7.4, SD = 2.0) – than in the GRUPR sector – R-side (M = 5.5, SD = 1.77 and D-side (M = 5.5, SD = 1.69), $F(1, 11) = 8.09$, and $F(1, 11) = 20.47$, respectively. This indicates that the dynamic sector boundary had a positive effect for BAASS (the sector that received the airspace) but did not negatively affect GRUPR.

The R-side participants rated the effect of increasing RNAV routes on their ability to control traffic higher in the GRUPR sector (M = 6.8, SD = 2.0) than in the BAASS sector (M = 5.4, SD = 2.7), $F(1, 11) = 5.05$. For the D-side participants, the interaction of Sector x Condition was significant, $F(1, 10) = 5.4$. Although the Tukey HSD post hoc tests were not significant, the mean ratings were about the same in the BAC condition (M = 5.4, SD = .85) and BANC condition (M = 5.7, SD = 1.4) in the BAASS sector, but the mean ratings were somewhat higher in the BAC condition (M = 6.1, SD = 1.65) than in the BANC condition (M = 5.6, SD = 1.5) in the GRUPR sector.

The terminal participants also completed the PSQ-2. However, their responses on these measures were much less variable (means were between 5.3 and 5.8, and *SDs* were approximately 1.0), and no significant results were obtained. This is not surprising because terminal participants did not directly experience changes to their normal work procedures in the BA conditions.

3.6 Post-Experiment Questionnaire

At the end of the experiment, each group of participants completed the PEQ (see Appendix E) and participated in a final debriefing session to discuss reactions to the BA concept and to provide additional comments about the feasibility of its implementation.

Two of the questions on the PEQ asked participants to rate what effect, if any, the BA conditions had on their control strategies compared to the BL condition. Those questions used a 9-point scale in which a rating of 1 indicated *a highly negative effect*, a rating of 5 indicated *no effect,* and a rating of 9 indicated *a highly positive effect*.

The en route participants indicated that, compared to the BL condition, the BA conditions would have a positive effect on control strategies, with averages of 7.1 (SD = 1.73) and 7.5 (SD = 1.57), respectively. The terminal participants indicated that the BANC condition would have a slightly positive effect (M = 5.7, SD = 1.42) on control strategies, whereas the BAC condition was rated as having a more positive effect (M = 6.9, SD = 1.68). The difference between the participants' ratings is not surprising because participants who worked the en route/transitional airspace experienced the effect of changing separation strategies and procedures between the test conditions. Their comments indicated that using reduced lateral separation and having the ability to dynamically resectorize the airspace in the BA conditions allowed more room to maneuver aircraft and resulted in more efficient flows of aircraft to the airport. Terminal participants, although still positive, included a few negative comments in their responses. Those comments primarily focused on the more limited airspace available in the BA conditions, which caused the final sequence to be essentially set as the aircraft entered the Feeder sector and produced some increase in complexity because there was less room to maneuver the aircraft.

Two other questions on the PEQ asked participants to rate the extent to which communication strategies were affected by the BA conditions compared to the BL condition. These questions utilized a 10-point scale in which a rating of 1 indicated that *communication was not affected at all* and a rating of 10 indicated that *communication was affected a great deal*.

Compared to the BL condition, the participants rated the BANC condition as having only a moderate effect on communication. Average ratings were 4.4 (*SD* = 1.96) and 4.3 (*SD* = 2.57) for the en route and terminal participants, respectively. The ratings were higher for the BAC condition. En route participants rated the BAC condition's effect on communication an average of 5.3 (*SD* = 3.32), and terminal participants rated its effect substantially higher, with an average rating of 6.8 (*SD* = 2.63). The participants' comments identified benefits of face-to-face communication. In general, they commented that the common control room fostered a more cooperative work environment even though individuals differed with respect to how much they took advantage of being located in the same room. Some of the participants got up and moved around the room to coordinate with other participants or to view traffic on other displays. When working the Feeder sector handoff position, one of the participants, sat between the Feeder and the Arrival Transition radar positions and acted much like a multisector planner. The comments typically indicated that being able to see and hear the other participants and view their displays helped in assessing how busy the other participants were. It also enabled them to see the traffic entering a sector so that planning and decision making could be done earlier. As one participant commented, "It allowed me to function better as an integrated team member." Some participants, who did not take as much advantage of the common control room during the experiment, thought it might take time to get accustomed to working that way. One participant commented, "I went over to the other side only once, as a novelty, but still called on the landline for communications." Another participant indicated that, "In time, as controllers grow more accustomed to this condition, they would coordinate better."

The participants were also asked to indicate the most highly positive and negative aspects of the BA concept. Most participants cited benefits, including the increased sector capacity enabled by the reduced separation standards as well as the use of terminal procedures ("Three-mile separation gives you more room to move aircraft"), the enhanced communication, coordination, and cooperation in the common control room environment ("Increased team concept"), the increased use of RNAV routes ("RNAV routes for arrivals are the way to go"), and the dynamic resectorization ("Being able to take control of the airspace you need is much easier and safer than borrowing it").

There were only a few negative comments made about the concept. One comment indicated that having more room to move aircraft could potentially result in sector saturation and lead to an unsafe environment. Another indicated that the workload would simply shift from one sector to another. Other comments indicated that the airspace would have to be worked by highly cooperative and skilled controllers, and that a third person or coordinator would be needed to manage the traffic between the Arrival Transition and Feeder sectors.

Finally, we included questions on the PEQ that pertained to experiment and equipment realism. These questions utilized rating scales that ranged from 1 (*extremely unrealistic*) to 10 (*extremely realistic*). The average responses indicated that these aspects of the experiment were fairly realistic. The mean rating of the overall simulation realism was 6.4 (*SD* = 1.53), the realism of the simulation hardware was 6.1 (*SD* = 1.75), the realism of the simulation software was 5.7 (*SD* = 1.81), and the realism of the traffic scenarios was 6.5 (*SD* = 1.91). The participants also indicated that the WAK online workload rating did not interfere with their control of traffic (*M* = 1.8, *SD* = .85) on a 10-point scale in which 1 indicated that the *WAK did not interfere at all* and 10 indicated that the *WAK interfered a great deal*.

In the final debriefing, we asked the participants to provide additional comments including any display enhancements or procedures that would be necessary to implement the BA concept in the field. Most participants responded that the J-ring or Continuous Range Readout function would provide sufficient spacing guidance regardless of the separation required, although one participant commented that having a clear indication of heavy aircraft would be essential when 3 nm separation is in effect. There were mixed responses regarding the ease of working different sectors that use different separation standards. A few who responded reported that they would find it difficult to work a position with one standard and then move to another position that required a different separation minimum. However, other participants responded that it would be a relatively easy transition. One of the participants reported that while transitioning between 3 nm and 5 nm would not be an issue in itself, knowing when other terminal procedures (e.g., diverging courses) were in use could be problematic without additional cues.

Other comments stressed the benefits of having the high and low altitude sectors located side by side to better enhance coordination. In the experiment, the Arrival Transition and Feeder sectors were located adjacently, but the Departure Transition and Airport Departure sectors were not similarly configured due to constraints in the laboratory. Most participants reported that it would be important for those working the higher and lower altitude sectors to be trained similarly and to use the same equipment.

Most of those who commented during the debriefing indicated that dynamic resectorization would not be problematic because controllers in busy facilities are already familiar with combining and decombining sectors in the field and because the controllers involved in the resectorization are not coming in cold. They are aware of the traffic in the affected sectors, so there is little that needs to be briefed prior to resectorization.

4. CONCLUSION

Overall, the results of this experiment provided support for the BA concept. The aircraft moved through the busy Arrival Transition and Feeder sectors more efficiently in the BA conditions than through the en route and terminal arrival sectors in the BL condition. The participants working the Arrival Transition and Feeder sectors made fewer ground-ground transmissions, issued fewer altitude and heading clearances, and required less assistance holding and maneuvering aircraft outside of the airspace in the BA conditions than in the BL condition. Few differences were found between the BAC and BANC conditions. The participants varied in the extent to which they viewed one another’s displays and communicated face-to-face in the BAC condition. Overall, few operational errors were observed in the experiment and their numbers did not differ across conditions.

Many of the subjective measures also provided support for the concept. The en route participant WAK ratings were lower in the second half of the scenarios in the BA conditions compared to the BL condition, indicating that it was easier to manage traffic after dynamic resectorization of the airspace occurred. The SMEs also rated most of the en route participant performance measures higher and noted fewer problems in the BA conditions than in the BL condition. Participant ratings of performance, situation awareness, and ability to move traffic through the sector were also higher in the BA conditions.

We found that several measures differed between sectors, particularly between the BAASS and GRUPR sectors. For example, the participants made more ground-ground communications and rated their workload higher in the BAASS sector; they issued more heading, altitude, and speed clearances in the BAASS + Arrival sector than in the GRUPR + Departure sector. The test conditions and weather intervals often affected the sectors differently. These results are not surprising because weather affected only the BAASS sector and more vectoring of aircraft was required to manage the arrival traffic, particularly when weather moved in.

During the debriefing sessions, the participants provided feedback about concept feasibility and implementation. The comments indicated that the participants found the BA procedures to be beneficial and that a combined control facility would promote more effective communication and coordination. No special modifications in equipment or automation were cited as necessary for implementation, even though a couple of comments indicated that controllers would need to have an indication as to when other procedures (e.g., diverging courses) are in use and to identify heavy aircraft when 3 nm separation standards are in effect.

References

Algina, J., & Keselman, H. J. (1997). Detecting repeated measures effects with univariate and multivariate statistics. *Psychological Methods*, *2*, 208-218.

Corker, K., & Smith, B. (1993). An architecture model for cognitive engineering simulation analysis: Application to advanced aviation analysis. *AIAA conference on Computing in Aerospace*, San Diego, CA.

Federal Aviation Administration. (2007). *Integrated arrival/departure control service (Big Airspace) concept validation.* Washington, DC: Air Traffic Organization Operations Planning Research & Technology Development Office Air Traffic System Concept Development, AJP-66.

Geisser, S., & Greenhouse, S. W. (1958). An extension of Box's results on the use of the F distribution in multivariate analysis. *Annals of Mathematical Statistics, 29*, 885-891.

Hadley, J., Sollenberger, R., D'Arcy, J., & Bassett, P. (2000). *Interfacility boundary adjustment* (DOT/FAA/CT-TN00/06). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Huynh, H., & Feldt, L. S. (1976). Estimation of the Box correction for degrees of freedom from sample data in randomized block and slit-plot designs. *Journal of Educational Statistics*, *1*, 69-82.

Myers, J. L., & Well, A. D. (2003). *Research design and statistical analysis* (2nd ed.). New Jersey: Lawrence Erlbaum Associates.

O'Brien, R. G., & Kaiser, M. K. (1985). MANOVA method for analyzing repeated measures designs: An extensive primer. *Psychological Bulletin, 97,* 316-333.

Peterson, L. M., Bailey, L. L., & Willems, B. F. (2001). *Controller-to-controller communication and coordination taxonomy (C^4T)* (DOT/FAA/AM-01/19). Washington, DC: Office of Aerospace Medicine.

Sollenberger, R. L., Stein, E. S., & Gromelski, S. (1997). *The development and evaluation of a behaviorally based rating form for assessing air traffic controller performance* (DOT/FAA/CT-TN96/16). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Stein, E. S. (1985). *Air traffic controller workload: An examination of workload probe* (DOT/FAA/CT-TN84/24). Atlantic City International Airport, NJ: Federal Aviation Administration Technical Center.

Stein, E. S., Della Rocco, P. S., & Sollenberger, R. L. (2006). *Dynamic resectorization in air traffic control: A human factors perspective* (DOT/FAA/CT-TN06/19). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Tabachnick, B. G., & Fidell, L. S. (1989). *Using multivariate statistic* (2nd ed.). New York: Harper Collins.

Truitt, T., McAnulty, D. M., & Willems, B. (2004). *Effects of collocation and reduced lateral separation standards in the New York integrated control complex* (DOT/FAA/CT-TN04/08). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Vardaman, J. J., & Stein, E. S. (1998). *The development and evaluation of a behaviorally based rating form for the assessment of en route air traffic controller performance* (DOT/FAA/CT-TN98/5). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Acronyms

ANOVA	Analysis of Variance
ARTCC	Air Route Traffic Control Center
ATC	Air Traffic Control
ATCSCC	Air Traffic Control System Command Center
ATCT	Airport Traffic Control Tower
BA	Big Airspace
BAC	Big Airspace/Collocated
BANC	Big Airspace/Non-Collocated
BL	Baseline
CPC	Certified Professional Controller
CRD	Computer Readout Display
D-Side	Data-Side
DESIREE	Distributed Environment for Simulation, Rapid Engineering, and Experimentation
DSR	Display System Replacement
ER	Experiment Room
FAA	Federal Aviation Administration
GENERA	Generic TRACON airspace
IFR	Instrument Flight Rules
LOA	Letter of Agreement
MSL	Mean Sea Level
NAS	National Airspace System
NYICC	New York Integrated Control Complex
ORF	Observer Rating Form
PEQ	Post-Experiment Questionnaire
PSQ	Post-Scenario Questionnaire
PTT	Push-To-Talk
R-Side	Radar-Side
RDHFL	Research Development and Human Factors Laboratory
RNAV	Area Navigation
SD	Standard Deviation
SID	Standard Instrument Departure

SME	Subject Matter Expert
SOP	Standard Operating Procedure
STAR	Standard Terminal Arrival Route
STARS	Standard Terminal Automation Replacement System
TGF	Target Generation Facility
TMA	Traffic Management Advisor
TMU	Traffic Management Unit
TRACON	Terminal Radar Approach Control
URET	User Request Evaluation Tool
VFR	Visual Flight Rules
WAK	Workload Assessment Keypad
WJHTC	William J. Hughes Technical Center

Appendix A

Informed Consent Form

Informed Consent Statement

I, ______________________________, understand that this study, entitled "Big Airspace: A Human-in-the-Loop Evaluation of an Integrated Arrival/Departure Control Service" is sponsored by the Federal Aviation Administration and is being directed by Dr. Mike McAnulty.

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 alternative air traffic control procedures in a high-fidelity, controller-in-the-loop simulation. The results of the study will be used to establish the feasibility of implementing these alternative or similar air traffic control procedures in an operational environment.

Experimental Procedures:

En route Certified Professional Controllers (CPCs) and Terminal CPCs will arrive at the simulation laboratory in groups of eight and will participate for 6 days over a 2-week simulation session. Each participant will work complex traffic scenarios that involve handoffs with other participants. The first 3 days of the simulation will consist of a project briefing, equipment familiarization, and practice scenarios. During the second week, the CPCs will work twelve 50-minute scenarios. A daily caucus will be scheduled at the end of each test day. On the final day, both en route and terminal CPCs will participate in a 1-hour debriefing session. The participants will work from about 8:30 AM to about 5:00 PM every day with a lunch break and at least two rest breaks.

The participants will control traffic under each of three different experimental procedures. After each scenario, the participants will complete questionnaires to evaluate the impact of the alternative procedures on participant workload and acceptance. In addition, subject-matter experts will make over-the-shoulder observations during the simulation to further assess the procedures. Finally, an automated data collection system will record system operations and generate a set of standard ATC simulation measures, which include safety, capacity, efficiency, and communications measures. The simulation will be audio-video recorded in case researchers need to reexamine any important simulation events.

Confidentiality:

My participation is strictly confidential, and I understand that no individual names or identities will be associated with the data 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 alternative ATC procedures for use in en route and terminal airspace. My data will help the FAA to establish the feasibility of these procedures within such an environment.

Participant Responsibilities:

I am aware that to participate in this study I must be a certified professional controller who is qualified at my facility and holds a current medical certificate. I will control traffic and answer 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 believe 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 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.

The research team has adequately answered all the questions I have asked about this study, my participation, and the procedures involved. I understand that Dr. McAnulty or another member of the research team will be available to answer any questions concerning procedures throughout this study.

If I have questions about this study or need to report any adverse effects from the research procedures, I will contact Dr. McAnulty at (609) 485-5380.

Discomfort and Risks:

I understand that I will not be exposed to any foreseeable risks or intrusive measurement techniques. I agree to immediately report any injury or suspected adverse effect to Dr. Mike McAnulty at (609) 485-5380. 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 B

Biographical Questionnaire

Biographical Questionnaire

Instructions:
This questionnaire is designed to obtain information about your background and experience as a certified professional controller (CPC). Researchers will only use this information to describe the participants in this study as a group. Your identity will remain anonymous.

Demographic Information and Experience

1. What is your **gender**?	○ Male		○ Female	

2. What is your **age**?	_____ years
3. How long have you worked as **an Air Traffic Controller (include both FAA and military experience)**?	_____ years _____ months
4. How long have you worked as a **CPC for the FAA**?	_____ years _____ months
5. How long have you **actively controlled traffic** in the en route environment?	_____ years _____ months
6. How long have you **actively controlled traffic** in the terminal environment?	_____ years _____ months
7. How many of the **past 12 months** have you actively controlled traffic?	_____ months

8. Rate your current **skill as a CPC**.	Not Skilled	①②③④⑤⑥⑦⑧⑨⑩	Extremely Skilled
9. Rate your **level of motivation** to participate in this study.	Not Motivated	①②③④⑤⑥⑦⑧⑨⑩	Extremely Motivated

Appendix C

Post-Scenario Questionnaire 1

P#_____ Date________________ Condition: baseline BA/C BA/N **Scenario:______**
Sector: 10 10D 20 20D A AH D DH

Post-Scenario Questionnaire 1

Instructions:
Please answer the following questions based upon your experience in the scenario just completed. Your identity will remain anonymous.

1. Rate your **overall level of ATC performance** during this scenario.	Poor	①②③④⑤⑥⑦⑧⑨⑩	Excellent
2. Rate your **overall level of situation awareness** during this scenario.	Poor	①②③④⑤⑥⑦⑧⑨⑩	Excellent
3. Rate your **situation awareness for current aircraft locations** during this scenario.	Poor	①②③④⑤⑥⑦⑧⑨⑩	Excellent
4. Rate your **situation awareness for projected aircraft locations** during this scenario.	Poor	①②③④⑤⑥⑦⑧⑨⑩	Excellent
5. Rate your **situation awareness for potential aircraft loss-of-separation** during this scenario.	Poor	①②③④⑤⑥⑦⑧⑨⑩	Excellent
6. Rate your workload due to **ground-to-ground communications** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High
7. Rate your **overall workload** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

For R-side Controllers Only:

8. Rate your **ability to move aircraft through the sector** during this scenario.	Poor	①②③④⑤⑥⑦⑧⑨⑩	Excellent
9. Rate your **workload due to air-to-ground communications** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High
10. Rate the **performance of the simulation pilots** in terms of their responding to your control instructions and providing readbacks.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

Appendix D

Post-Scenario Questionnaire 2

P#_____ **Date**________________ **Condition:** BA/C BA/N **Scenario:**______
Sector: 10 10D 20 20D A AH D DH

Post-Scenario Questionnaire 2

For Big Airspace Conditions Only:

1. What effect, if any, did the reduced lateral separation standard (3 nm) have on your **ability to control traffic?**	Negative Effect	①②③④⑤⑥⑦⑧⑨ | None	Positive Effect

Explain how the reduced lateral separation standards affected your **ability to control traffic**, if at all.

____

____

____

____

____

____

____

____

2. What effect, if any, did the use of other terminal procedures (e.g., green between, diverging courses) have on your **ability to control traffic?**	Negative Effect	①②③④⑤⑥⑦⑧⑨ | None	Positive Effect

Explain how the use of other terminal procedures affected your **ability to control traffic**, if at all.

____

____

____

____

____

____

____

P#_____ Date_______________ Condition: BA/C BA/N **Scenario:______**
Sector: 10 10D 20 20D A AH D DH

3. What effect, if any, did the dynamic sector boundaries have on your **ability to control traffic** (e.g., in terms of timeliness, coordination with other sectors, impact on workload, and traffic flows)?	Negative Effect	①②③④⑤⑥⑦⑧⑨ | None	Positive Effect

Explain how the dynamic sector boundaries affected your **ability to control traffic**, if at all.

____

____

____

____

____

____

____

4. What effect, if any, did the increase in the number of RNAV routes have on your **ability to control traffic?**	Negative Effect	①②③④⑤⑥⑦⑧⑨ | None	Positive Effect

Explain how the increase in the number of RNAV routes affected your **ability to control traffic**, if at all.

____

____

____

____

____

____

____

5. Do you have any additional comments or clarifications about your experience in the simulation?

____

____

____

Appendix E

Post-Experiment Questionnaire

P#_____ Date_______________
Airspace Type: En Route/BA Terminal

Post-Experiment Questionnaire

Instructions:
Please answer the following questions based upon your overall experience in the simulation. Your answers will remain anonymous.

1. **Compared to baseline**, what effect, if any, did the **'Big Airspace'/*non-collocated*** condition have on your **control strategies**?	Negative Effect	①②③④⑤⑥⑦⑧⑨ | None	Positive Effect

Explain how the **'Big Airspace'/*non-collocated*** condition affected your **control strategies**, if at all.

____

____

____

____

____

2. **Compared to baseline**, did your **communication strategies** change during the **'Big Airspace'/*non-collocated*** condition?	Not At All	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

Explain how the **'Big Airspace'/*non-collocated*** condition affected your **communication strategies**, if at all.

____

____

____

____

____

P#_____ Date_______________
Airspace Type: En Route/BA Terminal

3. **Compared to baseline**, what effect, if any, did the **'Big Airspace'/*collocated*** condition have on your **control strategies**?	Negative Effect	①②③④⑤⑥⑦⑧⑨ | None	Positive Effect

Explain how the **'Big Airspace'/*collocated*** condition affected your **control strategies**, if at all.

____

____

____

____

____

4. **Compared to baseline**, did your **communication strategies** change during the **'Big Airspace'/*collocated*** condition?	Not At All	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

Explain how the **'Big Airspace'/*collocated*** condition affected your **communication strategies**, if at all.

____

____

____

____

____

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

6. Rate the **realism of the simulation hardware** compared to actual equipment.	Extremely Unrealistic	①②③④⑤⑥⑦⑧⑨⑩	Extremely Realistic

7. Rate the **realism of the simulation software** compared to actual functionality.	Extremely Unrealistic	①②③④⑤⑥⑦⑧⑨⑩	Extremely Realistic

8. Rate the **realism of the simulation traffic scenarios** compared to actual NAS traffic.	Extremely Unrealistic	①②③④⑤⑥⑦⑧⑨⑩	Extremely Realistic

P#_____ Date_______________
Airspace Type: En Route/BA Terminal

9. To what extent did the **WAK online workload rating technique interfere** with your ATC performance?	None At All	①②③④⑤⑥⑦⑧⑨⑩	A Great Deal

10. Are there any additional requirements (e.g., for communications, automation, surveillance) you feel are necessary for controllers to implement the Big Airspace concept in an operational setting?

____

____

____

____

____

11. Describe any aspects of the Big Airspace concept that are highly positive.

____

____

____

____

____

12. Describe any aspects of the Big Airspace concept that are highly negative.

____

____

____

____

____

P#_____ Date_______________
Airspace Type: En Route/BA Terminal

13. Do you have any comments or suggestions for improvement about our simulation capability?

14. Is there anything about the study that we should have asked or that you would like to comment about?

Appendix F

Communication Score Sheet

Date________________

Communication Score Sheet

(Arrival)

Communication Type	A >>>> 01	01 >>>> A
Glance		
Approval		
Handoff		
Point Out		
Traffic		
Altitude		
Route		
Speed		
Weather		
Frequency		
Flow Messages		
Equipment		
ACID		
Non-verbal (pointing)		
Non-ATC		
Other		
Could Not Code		

Appendix G
Observer Rating Form

Date_______________ **Condition:** baseline BA/C BA/NC **Scenario:______**

Observer Rating Form (ORF)

This form is designed to be used by Subject Matter Experts (SMEs) to evaluate the effectiveness of controllers working in simulations. You will observe and rate the controllers' performance on several different performance dimensions using a rating scale of 1 to 8, with 1 indicating the least effective performance and 8 indicating the most effective performance. Most controller performance is at or above the minimum standards regarding safety and efficiency. The goal of the rating system is to differentiate performance above this minimum. The lowest rating should be assigned for meeting minimum standards and also for anything below the minimum since this should be a rare event. It is important for the observer/rater to feel comfortable using the entire scale and to understand that all ratings should be based on behavior that is actually observed. The rating scale is provided at the top of the ORF, so you can refer to it as you make your ratings.

- **Use the entire scale range.**

- **Write down your observations.**
  Space is provided on the second page of the ORF for comments. Wait until the scenario is finished before making your final ratings. Remain flexible until the end of the scenario so you have an opportunity to see all the available behavior.

- **At all times, focus on what you actually see and hear.**
  This includes what the controller does and what you might reasonably infer from the actions of the pilots. If you do not observe relevant behavior or the results of that behavior, you may leave a specific rating blank.

- **Remember to rate the arrival controllers and the departure controllers on separate forms.**

- **Do not write your name on the form.**
  Enter only the observer code assigned to you.

- **The observations you make may include other areas that you think are important.**

Date_______________ **Condition:** baseline BA/C BA/NC **Scenario:**______

Rating Scale Descriptors

Least Effective	①②③④⑤⑥⑦⑧	Most Effective

I - MAINTAINING SAFE AND EFFICIENT TRAFFIC FLOW								
1. Maintaining Separation and Resolving Potential Conflicts................................ • using control instructions that maintain appropriate aircraft and airspace separation • detecting and resolving impending conflicts early • recognizing the need for speed restrictions and wake turbulence separation	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
2. Sequencing Aircraft Efficiently .... • using efficient and orderly spacing techniques for arrival, departure, and en route aircraft • maintaining safe arrival and departure intervals that minimize delays	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
3. Using Control Instructions Effectively/Efficiently .... • providing accurate navigational assistance to pilots • issuing economical clearances that result in need for few additional instructions to handle aircraft completely • ensuring clearances require minimum necessary flight path changes	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
4. Overall Safe and Efficient Traffic Flow Rating.............................	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**

Questions 5 & 6: Handoff position/D-side only

5. Handoff position/D-side – Communication and Coordination...........................	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**

6. D-side – Entering Flight Plan Amendments..	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**

Questions 7 through 13: Frequency of Occurrence Ratings

Task	Never Occurred	Rarely Occurred	Occurred, but within Normal Limits of Operational Acceptability	Occurred More Than Normal	Occurred Unacceptably Often
7. Gave arriving aircraft descent too early	①	②	③	④	⑤
8. Gave departing aircraft climb too late	①	②	③	④	⑤
9. Issued clearances earlier or later than appropriate	①	②	③	④	⑤
10. Offered handoffs earlier than appropriate	①	②	③	④	⑤
11. Offered handoffs later than appropriate	①	②	③	④	⑤
12. Accepted handoffs later than appropriate	①	②	③	④	⑤
13.Transfered communications later than appropriate	①	②	③	④	⑤

Date_______________ **Condition:** baseline BA/C BA/NC **Scenario:**______

Notes about observations:

Explanatory comments supporting the ratings:

Differences in performance between sectors or positions:

Appendix H

Instructions for Participants

Instructions for Participants

Practice Scenario Instructions

During this brief practice scenario, please take the opportunity to familiarize yourself with your position. Familiarize yourself with the landlines and the Workload Assessment Keypads, or WAKs as we call them. This practice scenario is for your benefit and you should use this time to prepare for the scenarios that will follow. I will now read the WAK instructions to you.

Baseline Condition Instructions (Practice and Experiment)

During this scenario, please control traffic as you normally would in the field. As in every scenario, you will be making workload ratings using the WAK. I will now read the WAK instructions to you.

Big Airspace Condition(s) Instructions (Practice and Experiment)

During this scenario, we will simulate the Big Airspace concept. Sector 10 and Sector 20 will use terminal separation and procedures. The lateral separation standard will be reduced from 5 nautical miles to 3 nautical miles. You may also use other terminal separation procedures such as diverging courses and green between separation criteria. The halos and conflict alert algorithm will be adjusted accordingly. The radar sweep will also be 5 s. You will also have dynamic sector boundaries, which will allow you to swap routes with the adjacent sectors, and additional RNAV routes. You will be informed by one of the experimenters (who will be acting as a supervisor) when the dynamic resectorization occurs. ***For Collocation, add the following italicized statement*****:** *During this scenario, we will simulate the collocation of terminal and en route facilities. Because there is no physical barrier between the terminal and en route sectors, face-to-face communication is possible and you may use it at your discretion.* As in every scenario, you will be making workload ratings using the WAK. I will now read the WAK instructions to you.

WAK Instructions

(The full set of instructions will be read at the beginning of each test day). An abbreviated set of instructions will be read prior to each experimental run. The abbreviated instructions will omit the first paragraph below.)

One purpose of this research is to obtain an accurate evaluation of controller workload. By workload, we mean all the physical and mental effort that you must exert to do your job. This includes maintaining the “picture,” planning, coordinating, decision making, communicating, and whatever else is required to maintain a safe and expeditious traffic flow. Workload is your perception of how hard you must work to perform all of the tasks necessary to meet these demands, not necessarily a measure of how much traffic you are working. Workload levels fluctuate. All controllers, no matter how proficient, will experience all levels of workload at one time or another. It does not detract from a controller’s professionalism when he indicates that he is working very hard at certain times or that he is hardly working at other times.

Every 4 minutes the WAK device, located at your position, will emit a brief tone and the 10 buttons will illuminate. The buttons will remain lit for 20 s. Please tell us what your workload is at that moment by pushing one of the buttons numbered from 1 to 10.

At the low end of the scale (1 or 2), your workload is low - you can accomplish everything easily. As the numbers increase, your workload is getting higher. The numbers 3, 4, and 5 represent increasing levels of moderate workload where the chance of making a mistake (e.g., leaving a task unfinished) is still low but steadily increasing. The numbers 6, 7, and 8 reflect relatively high workload where there is some chance of making a mistake. At the high end of the scale are the numbers 9 and 10, which represent a very high workload, where it is likely that you will have to leave some tasks unfinished. Feel free to use the entire rating scale and tell us honestly how hard you are working at the instant that you are prompted. Do not sacrifice the safe and expeditious flow of traffic in order to respond to the WAK device.

Does anyone have any questions? *(After answering questions, if any, instruct participants to do comm check with pilots and adjacent sectors and centers.)*

Appendix I

Comments on the Repeated Measures Experimental Design

Comments on the Repeated Measures Experimental Design

The experiment uses a repeated measures design in which each participant is tested under each experimental condition. Experimenters often use a repeated measures (or within-subjects) design to control variability due to differences between participants. Too much variability related to participant differences may prevent the researcher from detecting significant effects that are due to the experimental conditions. However, there are certain assumptions that must be met when analyzing data from a repeated measures design. The data must be evaluated and determined not to violate sphericity, the assumption that the variances of the difference scores between the conditions are homogeneous. To address instances when there are violations of sphericity, some researchers employ multivariate analysis of variance (MANOVA) when analyzing repeated-measures data (Myers & Well, 2003; O'Brien & Kaiser, 1985). This analysis helps to avoid potentially inflated Type I error rates (incorrectly rejecting a true null hypothesis). Other researchers perform unfocused significance tests (i.e., omnibus ANOVA tests) to screen for differences in the data. When the omnibus ANOVA is significant but sphericity is violated, then a conservative F test is completed by adjusting the degrees of freedom (e.g., Geisser & Greenhouse, 1958; Huynh & Feldt, 1976) used to calculate the F statistic.[2] If the conservative F test is significant, the data are then analyzed using conservative post hoc procedures for pair-wise comparisons (e.g., Tukey's Honest Significant Difference (HSD) test). This second procedure is the procedure we will employ.

[2] Although the MANOVA avoids sphericity problems and inflated Type I error, it suffers even more severely from inflation of Type II error rates. Because there are methods that correct degrees of freedom (*df*) that reduce the risk of a Type I error, we recommend using these corrections in all but the most severe cases (see Algina & Keselman, 1997, for specific recommendations).