DOT/FAA/TC-12/64

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

High Altitude Airspace Study

Erin J. Higgins, PhD, TASC, Inc.
Randy L. Sollenberger, PhD, FAA Human Factors Branch
Mark Hale, CSSI, Inc.
Anton Koros, FAA Concept Development & Validation Branch
Phil Bassett, FAA Advanced Operational Concepts Division
John DiRico, TASC, Inc.
Andy Andrelczyk, Hi-Tec Systems, Inc.

October 2012

Technical Report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No. DOT/FAA/TC-12/64	**2. Government Accession No.**	**3. Recipient's Catalog No.**
4. Title and Subtitle High Altitude Airspace Study		**5. Report Date** October 2012
		6. Performing Organization Code ANG-E25
7. Author(s) Erin J. Higgins, PhD, TASC, Inc. Randy L. Sollenberger, PhD, FAA Human Factors Branch Mark Hale, CSSI, Inc. Anton Koros, FAA Concept Development & Validation Branch Phil Bassett, FAA Advanced Operational Concepts Division John DiRico, TASC, Inc. Andy Andrelczyk, Hi-Tec Systems, Inc.		**8. Performing Organization Report No.** DOT/FAA/TC-12/64
9. Performing Organization Name and Address Federal Aviation Administration Human Factors Branch 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 Office of Advanced Concepts & Technology Development 800 Independence Avenue, S.W. Washington, DC 20591		**13. Type of Report and Period Covered** Technical Report
		14. Sponsoring Agency Code ANG-C1
15. Supplementary Notes		

16. Abstract

Objective: The High Altitude (HA) Airspace concept is part of the Next Generation Air Transportation System (NextGen) initiative and is designed to address current issues in HA Airspace, such as the inability to fly wind optimal routes, flight congestion, flight inefficiencies, and flight delays. Researchers at the Research Development and Human Factors Laboratory (RDHFL) performed a human-in-the-loop simulation investigating two concepts proposed for HA Airspace: the Generic Sector concept and the High-Performance Routes (HPRs) concept. **Background**: The Generic Sector concept is designed to enable staffing flexibility. Controllers will be able to certify for "generic" sectors when they achieve their area rating. The HPRs concept is based on the concepts of Q-routes and Oceanic routes. HPRs will be dynamic, wind-optimized, Area Navigation (RNAV)-2 routes that will allow for greater flexibility and additional routing options. **Method**: Twelve air traffic controllers participated in the simulation, which consisted of three experiments. Two experiments were designed to investigate the HPRs concept, and one experiment was designed to examine the Generic Sector concept. Researchers collected system performance data, subjective workload ratings, over-the-shoulder observer ratings, and questionnaire responses. **Results**: We identified the human performance issues related to the HPRs and Generic Sector concepts using a combination of objective system performance data and subjective participant ratings and responses. For the HPRs concept, results are discussed in terms of differences between HPRs and current jet routes. For the Generic Sector concept, results focus on human performance as a function of experience with the generic sector. **Conclusions**: In Experiments 1 and 2, we found some evidence for the benefits of HPRs compared with Jet Routes, though in Experiment 1 subjective feedback indicated difficulties with the HPRs concept. In Experiment 2, controllers felt more comfortable with the HPRs concept. In Experiment 3, we found some evidence that controllers can operate traffic in the generic sector with minimal training. If the Generic Sector concept will be implemented with the Controller Information Tool (CIT), we recommend future research to address usability issues identified in the simulation.

17. Key Words Controller Information Tool Generic Sector High Altitude Airspace High-Performance Routes		**18. Distribution Statement** This document is available to the public through the National Technical Information Service, Alexandria, Virginia, 22312. A copy is retained for reference at the William J. Hughes Technical Center Library.	
19. Security Classification (of this report) Unclassified	**20. Security Classification (of this page)** Unclassified	**21. No. of Pages** 153	**22. Price**
Form DOT F 1700.7 (8-72)	**Reproduction of completed page authorized**		

THIS PAGE IS BLANK INTENTIONALLY.

Table of Contents

List of Illustrations

Figures **Page**

Tables **Page**

Acknowledgments

We would like to thank Gary Mueller, Joel Grochowski, and the rest of the Concepts and Systems Integration team for developing the simulation software, supporting the simulation during data collection, and managing the data during and after data collection. We thank Daniel R. Johnson and his team for performing complex data processing and reduction so that we could analyze our data efficiently. We thank Jim McGhee (TASC, Inc.) and Richard Ridgeway (formerly of TASC, Inc.) for assisting with scenario development. We thank Steve Kennedy (Metron) for his suggestions and feedback during scenario development. We acknowledge Dr. Richard Mogford and his team from the NASA-Ames Research Center for development of the Controller Information Tool. We thank the FAA Simulation Team pilots and staff for their assistance and support during data collection. Finally, we would like to thank all of the FAA controllers who volunteered to participate in this study.

THIS PAGE IS BLANK INTENTIONALLY.

Executive Summary

The Federal Aviation Administration (FAA) will face challenges in the coming years as increased demand for air travel strains the capacity of our National Airspace System (NAS). In addition, aging air traffic systems will need to be either upgraded or replaced entirely to ensure effective functioning and integration. The FAA is addressing these issues through an initiative known as the Next Generation Air Transportation System (NextGen). NextGen will not only transform Air Traffic Management (ATM), but it will enable important operational improvements to Air Traffic Control (ATC) systems. NextGen will transform the current surveillance, navigation, and communication systems and change the role of pilots and air traffic controllers (Joint Planning and Development Office, 2007). As a result, airline operators and the flying public will reap the benefits of improved safety, increased capacity, and reduced environmental impact.

Some of the key NextGen midterm enhancements to the NAS are envisioned for High Altitude (HA) Airspace. This human-in-the-loop (HITL) simulation investigated two concepts within HA Airspace: the Generic Sector concept and the High-Performance Routes (HPRs) concept. The Generic Sector concept would allow controllers to certify for predetermined "generic" sectors when they achieve their area rating, resulting in increased staffing flexibility. The HPR concept combines the advantages of both Q-routes and Oceanic routes. HPRs will be dynamic, wind-optimized, Area Navigation (RNAV)-2 routes that will allow for greater flexibility and additional routing options. Aircraft on the HPR will receive a higher priority over non-HPR traffic enabling them to continue to operate at their optimal speeds and altitudes. HPRs may also support different uses and procedures to maximize their efficiency depending upon the airspace and sector operations. In this HITL, we explored three different types of HPR lane usage with procedures designed to organize traffic flow by speed, destination, or equipment.

Three experiments were designed to address the HPR and Generic Sector concepts. In Experiment 1, we investigated the transition from HA HPRs to lower altitude airspace. We compared the baseline traffic scenarios with jet routes to the HPR scenarios. In Experiment 2, we investigated different lane usage strategies for HPRs. We compared the baseline traffic scenarios with jet routes to HPRs using lanes for different aircraft destinations, different aircraft speeds, and for aircraft that were either equipped or not equipped to support Optimum Profile Descent (OPD). In both experiments, scenarios were designed to examine the experimental conditions under both medium- and high-traffic levels. In Experiment 3, we investigated whether controllers could safely and efficiently manage traffic in an unfamiliar, generic sector with the Controller Information Tool (CIT) and minimal sector training. The CIT is a Radar (R)-side display tool developed by the National Aeronautics and Space Administration (NASA) to support the Generic Sector concept (Mogford, 2010). It shows necessary sector-specific information for all sectors in the area, including a map with sector boundaries, sector names, traffic routes, altitudes, radio frequencies, and fix names and locations.

Twelve participants spent four days participating in the HA HITL at the FAA William J. Hughes Technical Center (WJHTC), Research Development and Human Factors Laboratory (RDHFL) in Atlantic City, NJ. Four participants at a time traveled to the RDHFL and worked as R-side and Data (D)-side teams as well as R-side only controllers in medium- and high-traffic scenarios. Experiment 1 consisted of six training scenarios and four experimental scenarios, Experiment 2 consisted of four training scenarios and four experimental scenarios, and Experiment 3 consisted of four experimental scenarios. When working in R-side/D-side teams, controllers switched positions and performed the same scenario procedures once at each position.

We conducted the study at the FAA WJHTC RDHFL. The simulation configuration consisted of the Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE) ATC simulator, the Target Generator Facility (TGF), and the Java En Route Development Initiative (JEDI)/User Request Evaluation Tool (URET) prototype. All three systems work together to provide a realistic ATC simulation for controllers.

In Experiments 1 and 2, we used three sectors from Cleveland (ZOB) Air Route Traffic Control Center (ARTCC) to test the HPR concept. We combined the Jamestown sector (ZOB-79) with the adjacent North sector and the adjacent South sector to form a combined sector for the simulation airspace. In Experiment 3, we used two sectors from Kansas City ARTCC (ZKC) to test the Generic Sector concept. We selected ZKC-21 and ZKC-7 to form a combined sector for the simulation airspace. Massachusetts Institute of Technology Research and Engineering (MITRE) Corporation identified these sectors as relatively simple sectors with routine traffic flow that represent good candidates for Generic Sector operations.

For the HPR concept, traffic scenarios were built based on actual traffic samples from Cleveland ARTCC; however, the samples were extensively modified in volume and routes to meet the requirements of the study. All experimental scenarios were 45 minutes in duration. Traffic levels differed between the first and second half of the scenario. The Monitor Alert Parameter (MAP) value for the sector was set at 22 aircraft. In the first half of each scenario, traffic quickly increased to approximately 85% of the sector MAP value. In the second half of each scenario, traffic continued to build to a maximum of 115% of the sector MAP value.

For the Generic Sector concept, traffic scenarios were built based upon actual traffic samples from Kansas City ARTCC in November, 2011. We did not change the scenario traffic in any way from the original samples in order to simulate actual traffic operations as accurately as possible. All experimental scenarios were 30 minutes in duration.

Across all experiments, we collected system and participant performance metrics, subjective workload ratings, over-the-shoulder observer ratings, and questionnaire responses. The primary measure of safety was loss of aircraft separation. The measures of capacity were the number of aircraft accepted, number of handoffs initiated, and number of aircraft under control. The measures for efficiency were the aircraft time and distance flown through each sector as well as the number of control commands (e.g., altitude, heading, and speed commands) issued. We used the data recorded by the communications system to analyze the frequency and duration of controller and pilot communications. In addition, we measured controller interactions with support tools.

In Experiment 1, we observed that aircraft flew shorter times and distances on HPRs compared with jet routes under medium-traffic levels; however, there were no differences between conditions for either flight times or distances under high-traffic levels. We also found that participants issued fewer altitude, speed, and (marginally fewer) heading commands in the HPRs condition compared with the jet routes condition. There were no other notable system performance differences observed between conditions. Despite some evidence that HPRs improved efficiency, participants reported higher workload levels in the HPRs condition. In addition, they rated their performance and situation awareness as lower in the HPRs condition. It is possible that insufficient training and the complexity of the airspace, rather than an inherent difference between jet routes and HPRs, led to these subjective rating differences.

In Experiment 2, we found that aircraft flew shorter times and distances in all HPR lane usage conditions compared with the jet routes condition, but only under high-traffic levels. Aircraft flight times and distances were similar across conditions under medium-traffic levels. These results demonstrate the potential for HPRs by a single sorting procedure to be beneficial when traffic volumes are high, most likely because it allows an aircraft to get around other aircraft that may be in its way without reducing speed, changing altitude, or flying a heading. However, when traffic volumes are lower, moving an aircraft to another lane adds additional time and distance to the route without providing as many benefits (as it is less likely that other aircraft will be in that aircraft's way). There were no other notable system performance differences observed between conditions. Unlike in Experiment 1, there were no statistically significant differences between conditions for subjective ratings of workload, situation awareness, and performance. We suggest that the combination of more experience with HPRs as well as fewer complexities in the airspace resulted in similar levels of subjective workload, performance, and situation awareness.

In Experiment 3, we found no statistically significant differences between scenario runs for any of the system performance metrics. In addition, no losses of separation were reported. Participants' subjective workload was relatively low across all four scenario runs, but there was a noticeable drop in workload between the first and third scenario run. After one hour of controlling traffic in the generic sector, participants had already reduced their workload. Most notably, by the end of the third scenario run, 11 out of 12 participants indicated that they felt they had controlled a safe sector. We also recorded participants' interactions with the CIT. Participants spent most of their time using the CIT function that enabled them to see sector boundaries, numbers, altitudes, and frequencies. They also used the function that displayed information for a specific route. Participants commented that they found the CIT useful for displaying sector information on their scopes. Despite the positive feedback, they did not use the CIT as much as expected. Participants noted on the Exit Questionnaire that they would have been more likely to use the tool if they could have accessed the information more efficiently. In its current form, the CIT is hierarchical, so it takes multiple clicks to display information. Another common suggestion was to include brightness controls. Future research should address these issues and re-examine the design of the CIT.

THIS PAGE IS BLANK INTENTIONALLY.

1. INTRODUCTION

1.1 Background

The Federal Aviation Administration (FAA) will face challenges in the coming years as increased demand for air travel strains the capacity of our National Airspace System (NAS). In addition, aging air traffic systems will need to be either upgraded or replaced entirely to ensure effective functioning and integration. The FAA is addressing these issues through an initiative known as the Next Generation Air Transportation System (NextGen). NextGen will not only transform Air Traffic Management (ATM), but it will enable important operational improvements to Air Traffic Control (ATC) systems. NextGen will transform the current surveillance, navigation, and communication systems and change the role of pilots and air traffic controllers (Joint Planning and Development Office, 2007). As a result, airline operators and the flying public will reap the benefits of improved safety, increased capacity, and reduced environmental impact.

Some of the key NextGen midterm enhancements to the NAS are envisioned for High Altitude (HA) Airspace. This airspace is defined as "Class A" Airspace, extending from Flight Level 290 to Flight Level 600 inclusive (FAA, 2012). Aircraft flying at these altitudes are typically in the cruise phase of flight. The HA Airspace concept is proposed to help overcome current shortfalls in HA Airspace, such as

- inability to fly wind optimal routes, resulting in increased fuel burn and higher costs to the airlines, the FAA, and the flying public;
- flight congestion and bottlenecks in the NAS;
- flight inefficiencies in and out of major airports; and
- flight delays due to weather and other off-nominal conditions.

1.2 Generic Sector Operations

Currently, Certified Professional Controllers (CPC) are certified to work in a specific Area of Specialization (AOS) within their facility. Controllers gain certification after an extensive amount of training and demonstrated proficiency in their area. They must memorize large amounts of information about the sectors in their AOS, such as the airspace structure, traffic flows, and procedures to certify in the area. Not only is this process time-consuming, but it is very expensive. Also, because training is focused in a specific area, it does not generalize to other areas within their facility. This policy limits staffing flexibility and prevents controllers from working outside of their AOS. In the HA Generic Sector concept, controllers will be able to certify for predetermined "generic" sectors when they achieve their area rating.

As part of the HA Airspace concept, some portions of HA Airspace would qualify for operation as a generic sector. In order for airspace to be suitable for generic operations, it must fulfill each of the following qualities:

- aircraft will predominantly be in level cruise flight,
- aircraft will have infrequent climbs and descents,
- aircraft will not have complex crossing patterns, and
- the sector will have a low to moderate traffic complexity.

1.3 High Performance Routes

The HA High Performance Route (HPR) concept evolved from Q-routes and Oceanic route applications. Q-routes are static Area Navigation (RNAV)-2 routes that provide additional direct routing options to save fuel and travel time while increasing airspace capacity and throughput in areas with a high-traffic density. Oceanic routes are dynamic routes designed to provide a wind optimal structure in the oceanic environment. Oceanic routes are published twice a day for east-bound and west-bound traffic several hours in advance based on the forecast for winds.

The HPR concept combines the advantages of both Q-routes and Oceanic routes. HPRs are envisioned to be dynamic RNAV-2 routes that do not rely on ground-based navigational aids, and may be published multiple times a day in order to adapt to the traffic demand, wind, and weather changes. The greatest benefit of HPRs is wind optimized routing that it will result in reduced fuel consumption. Aircraft on the HPR will receive a higher priority over non-HPR traffic enabling them to continue to operate at their optimal speeds and altitudes.

Other than the characteristics mentioned, HPRs may also support different uses and procedures to maximize their efficiency depending upon the airspace and sector operations. For example, in regions where high-altitude cruise operations occur, different HPR lanes may be used to support aircraft flying at different speeds, similar to speed lanes on the highway. In other regions where arrival and departure operations are common, different HPR lanes may be used to organize the traffic flow by destination airport or equipment supporting special arrival routes. In our study, we explored three different types of HPR lane usage with procedures designed to organize traffic flow by speed, destination, or equipment.

1.4 Purpose

The HA Airspace Study is conducted under the FAA's NextGen trajectory based operations portfolio. The HA concept encompasses three key components: HPRs, generic sector operations, and flexible airspace. The purpose of this study is to investigate two of the HA Airspace concepts planned for the midterm NextGen en route environment. The first concept is an implementation of HPRs and the second concept is an implementation of a Generic Sector concept. Specifically, the objectives of the present study are to conduct a series of human-in-the-loop (HITL) simulations with CPCs:

1. to evaluate the operational viability and human performance issues of HPRs in NextGen HA airspace,
2. to investigate the transition from HA HPRs to lower altitude airspace,
3. to investigate alternative lane usage options for HA HPRs, and
4. to demonstrate that controllers can safely and efficiently manage unfamiliar sector traffic in HA airspace with support tools and training.

2. METHOD

2.1 Participants

Twelve CPCs were recruited from Air Route Traffic Control Centers (ARTCC) nationwide to serve as voluntary participants in the study. All participants were nonsupervisory controllers who were qualified at their facility and held a current medical certificate. We excluded controllers from Kansas City ARTCC who were already experienced with the sectors we used in the simulation. Table 1 shows a summary of participants' responses to the Biographical Questionnaire. Controllers gave high ratings on a 10-point rating scale (1 = *lowest*, 10 = *highest*) for their skill level and motivation to participate.

Table 1. Means and Standard Deviations of Responses to the Biographical Questionnaire

Questionnaire Item	Mean	*SD*
Age of participant	41.24	10.38
Years of experience as an air traffic controller, including FAA and military	15.91	10.92
Years of experience as a Certified Professional Controller for the FAA	13.34	10.73
Years of experience controlling traffic in the en route environment	14.53	10.45
Years of experience controlling traffic in the terminal environment	0.00	0.00
Number of the past 12 months actively controlling traffic	12.00	3.69
Skill level as a Certified Professional Controller (from 1-10)	8.50	1.61
Motivation level to participate in this experiment (from 1-10)	9.00	1.10

Note. FAA = Federal Aviation Administration.

The study required six weeks to complete. A new group of four controllers traveled to the FAA William J. Hughes Technical Center (WJHTC) to participate in the study every two weeks. In the first week, each group of controllers participated in the HA Airspace Study. In the second week, the controllers participated in the Conflict Resolution Advisory Study, an unrelated HITL simulation organized by a separate group of researchers. The present report describes details of the HA Airspace Study.

All participating controllers worked some simulation runs in Radar (R)-side/Data (D)-side teams and other simulations runs as an R-side only position, as detailed in the study procedure below. The controllers worked independent traffic scenarios that did not require coordination with each other. The adjacent sector functions were handled by "ghost" controller automation.

The principal investigator informed the controllers of their rights as participants in a research study, and each participant read and signed an Informed Consent Statement. The FAA WJHTC Local Institutional Review Board (IRB) reviewed the routine ethical considerations and approved this study.

2.2 Research Personnel

An Engineering Research Psychologist (ERP) served as the principal investigator and conducted the simulation. The ERP briefed the participants, collected the data, and led the group discussions with controllers. The ERP supervised the operation of the simulation equipment and coordinated the work of the research personnel. A Human Factors Specialist assisted the principal investigator by operating the simulation software. Hardware and Software Engineers prepared the simulator and ensured the equipment was operating properly.

Eight Subject Matter Experts (SMEs) were involved in the study. Two of the SMEs were retired ATC supervisors who are now contract personnel and have experience with our simulator. In preparation for the simulation, these two SMEs helped to develop the practice and test scenarios. In addition, they trained the participants to operate the sectors. The remaining six SMEs were front-line managers who were recruited to be over-the-shoulder (OTS) observers for the study. Like the controllers, every two weeks, a new group of two front-line managers traveled to the FAA WJHTC. During the HITL simulation, the front-line managers along with the contract personnel acted as OTS observers and completed OTS evaluation forms for participants after each traffic scenario.

We required between six and twelve simulation pilots for the study depending upon the configuration of controllers and type of scenario. Three simulation pilots were used to support each R-side/D-side team during training scenarios and six were used during test scenarios. Three simulation pilots were used to support each R-side only position during both training and experimental scenarios. The simulation pilots operated pilot workstations, communicated with controllers using proper ATC phraseology, and maneuvered the simulation aircraft based upon controller instructions.

2.3 Simulation Environment

2.3.1 Research Facility

We conducted the study in the FAA WJHTC Research Development and Human Factors Laboratory (RDHFL). The RDHFL is a state-of-the-art facility with experiment rooms, ATC workstations, and human performance measurement equipment to support aviation human factors research. The simulation configuration consisted of the Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE) ATC simulator, the Target Generator Facility (TGF), and the Java En Route Development Initiative (JEDI)/User Request Evaluation Tool (URET) prototype. All three systems work together to provide a realistic ATC simulation for controllers.

2.3.2 Software

Software engineers at the FAA WJHTC developed the DESIREE ATC simulator and the TGF to support air traffic research, development, and testing and evaluation activities. The DESIREE ATC simulator emulates both en route and terminal controller functions. DESIREE provides a flexible platform for researchers to modify the displayed information and functionality of controller workstations to evaluate new ATC concepts and procedures. In the present study, DESIREE emulated En Route Automation Modernization (ERAM) and received input from the TGF to display aircraft targets and flight information on the controller displays. DESIREE also acted as a ghost controller and automated the aircraft handoff functions for the adjacent sectors in the simulation.

The MITRE Corporation developed the JEDI/URET prototype as a conflict probe and trial planning tool. The JEDI/URET prototype is similar to the URET system that controllers currently use in the field, but it can be implemented without ERAM using DESIREE. JEDI/URET presented the Aircraft List, Plans Display, and the Graphic Plan Display windows on the D-side controller display. JEDI/URET and DESIREE shared data through a Host Automation Gateway (HAG) so that JEDI/URET operated as if connected to ERAM and DESIREE was able to display conflict probe and trial planning information on the D-side position or a fixed display above the R-side radar display.

The TGF is a dynamic, real-time air traffic simulation capability designed to generate realistic aircraft targets for HITL simulations. The TGF models aircraft performance characteristics and maneuvers aircraft based upon scripted flight plan data and simulation pilot commands. TGF also consists of multiple simulation pilot workstations operated by trained personnel who communicate with controllers and enter flight plan changes based upon controller instructions.

2.3.3 Airspace

We selected different sectors from two ARTCCs to test the HPR and Generic Sector concepts in a series of experiments. In the first two experiments, we used three sectors from Cleveland ARTCC (ZOB) to test the HPR concept. We combined the Jamestown sector (ZOB-79) with the adjacent North sector and the adjacent South sector to serve as the simulation airspace. In the third experiment, we used two sectors from Kansas City ARTCC (ZKC) to test the Generic Sector concept. We selected ZKC-21 and ZKC-7 to form a combined sector for the simulation airspace.

2.3.3.1 High Performance Routes

The Jamestown combined sectors are located on the East side of Cleveland ARTCC and support heavy traffic flow between Chicago and New York City. These sectors handle arrival traffic into the New York metropolitan area; including John F. Kennedy, La Guardia, and Newark airports (see Figure 1). The research team chose these sectors to investigate the transition from HA HPRs to lower altitude airspace. This airspace simulates realistic sectors and traffic situations to evaluate HPRs and the human performance issues associated with them.

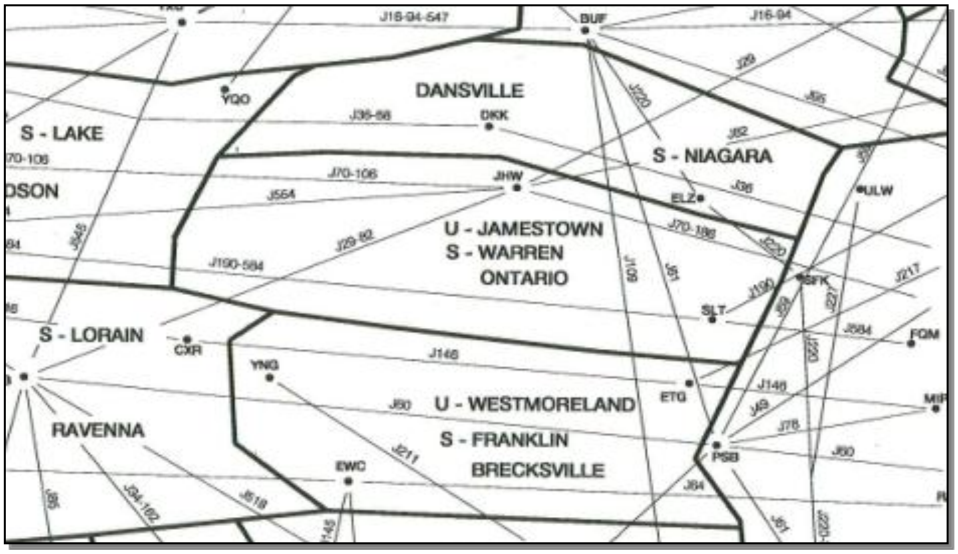


Figure 1. Cleveland ARTCC Jamestown (ZOB-79) and adjacent sectors.

2.3.3.2 Generic Sector Operations

The ZKC-21 and ZKC-7 combined sectors are located on the Northwest side of Kansas City ARTCC. MITRE identified these sectors as relatively simple sectors with routine traffic flow that represent good candidates for Generic Sector operations (see Figure 2). We used the sectors to determine whether controllers could safely and efficiently manage traffic in HA Generic Sectors with support tools and minimal sector training.

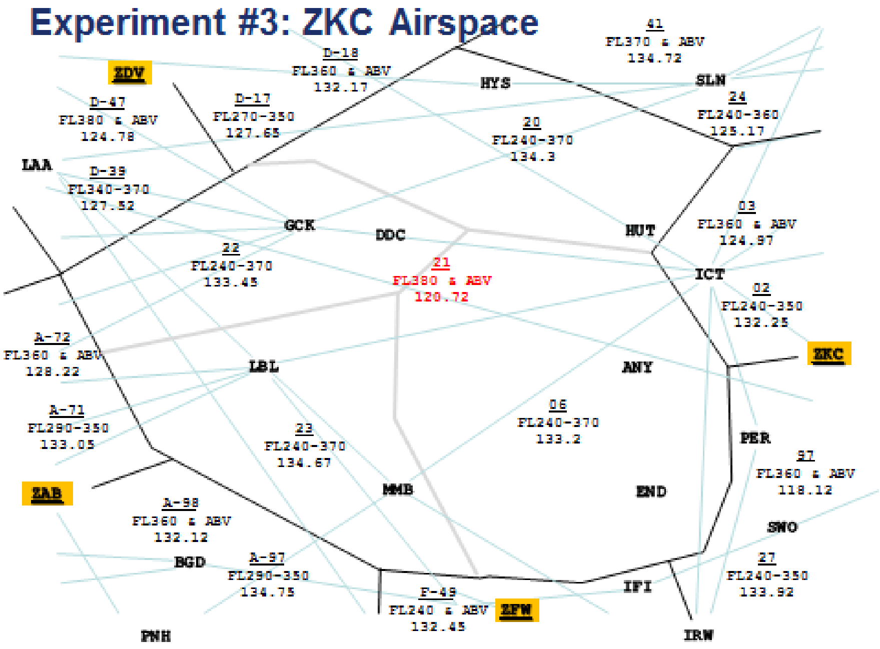


Figure 2. Kansas City ARTCC ZKC-21 and ZKC-7 generic sectors.

2.3.4 Traffic Scenarios

For the HPR concept, traffic scenarios were built based on actual traffic samples from Cleveland ARTCC. The samples were extensively modified in volume and routes to meet the study requirements. All experimental scenarios were 45 minutes in duration. Traffic levels differed between the first and second half of the scenario. The Monitor Alert Parameter (MAP) value for the sector was set at 22 aircraft. In the first half of each scenario, traffic quickly increased to approximately 85% of the sector MAP value. In the second half of each scenario, traffic continued to build to a maximum of 115% of the sector MAP value.

For the Generic Sector concept, traffic scenarios were built based on actual traffic samples from Kansas City ARTCC in November, 2011. We did not change the scenario traffic in any way from the original samples in order to simulate actual traffic operations as accurately as possible. All experimental scenarios were 30 minutes in duration.

2.3.4.1 High Performance Routes

To test two aspects of the HPRs concept, we designed two experiments and generated eight ZOB-79 experimental traffic scenarios. In Experiment 1, four of the scenarios were used to investigate managing arrival traffic into a metroplex environment. In Experiment 2, the remaining four scenarios were used to evaluate alternative lane usage strategies.

Experiment 1 was conducted using R-side/D-side teams. Experiment 1 included two conditions and had the goal of investigating the transition between HPRs and lower altitude airspace. The baseline condition required controllers to use jet routes (i.e., jetways) as they exist today. The second condition simulated a midterm future concept in which HPRs replace many of the inefficient jet routes. Each controller performed the ZOB-79 experimental scenarios as an R-side and D-side, accounting for a total of four scenarios. In all scenarios, controllers had to follow Letters of Agreement (LOAs). Figures 3 and 4 show Experiment 1's configuration of ZOB-79 airspace, along with descriptions of the LOAs that had to be followed for the baseline (jet routes) condition and HPRs condition, respectively. As shown in Figure 4, aircraft that were equipped to support Optimum Profile Descent (OPD) were able to leave the sector at altitude in the HPRs condition and were issued a "Descend Via" command, which cleared the aircraft for an OPD arrival—however, an OPD arrival procedure was never initiated in ZOB-79. In the jet routes condition, OPD equipped aircraft had to comply with the LOAs, as there was no way to separate the OPD equipped aircraft from the non-OPD equipped aircraft in order to clear OPD-equipped aircraft for an OPD arrival.

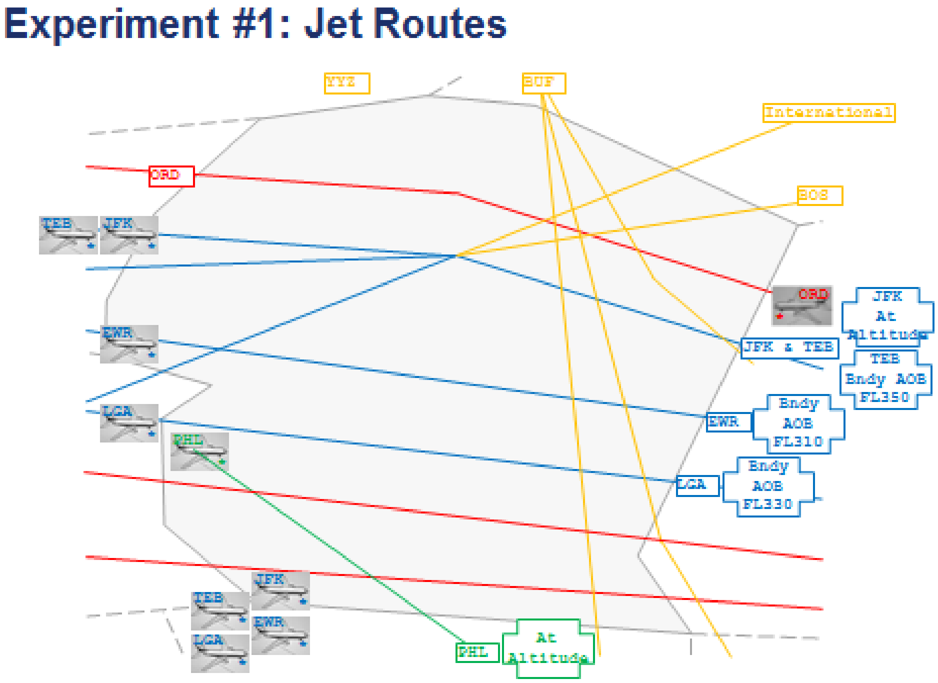


Figure 3. Jamestown (ZOB-79) during Experiment 1: Jet Routes condition.

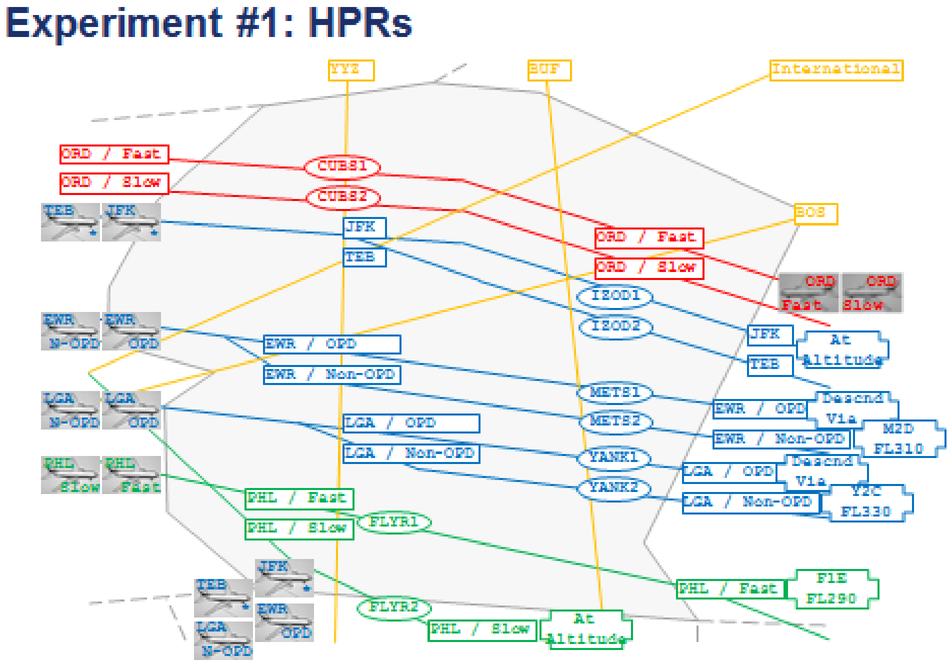


Figure 4. Jamestown (ZOB-79) during Experiment 1: HPRs condition.

Experiment 2 was conducted using a single R-side controller, included four conditions, and had the goal of investigating alternative lane usage options for HPRs. The conditions included a baseline (jet routes) and lane usage conditions based on aircraft speed, destination, and equipment. The four ZOB-79 lane usage scenarios used overflight traffic to evaluate these alternative lane management conditions. Because these were overflight scenarios, aircraft could remain at altitude through the sector in all conditions. In all scenarios, there was a weather pattern located at the southeast corner of ZOB-79. The weather did not move from its initial position and did not introduce additional variables (e.g., wind) throughout the scenario. Due to weather, all Philadelphia aircraft were re-routed south of ZOB-79 and all LaGuardia aircraft were re-routed north of the weather. Having weather in these scenarios not only made this sector more manageable for a single R-side controller, by reducing it in terms of usable size, but also allowed us to test a single sorting procedure in each scenario. Figures 5, 6, 7, and 8 show Experiment 2's configuration of ZOB-79 airspace, along with descriptions of the procedures that had to be followed, for the baseline (jet routes), HPRs by destination, HPRs by speed, and HPRs by OPD equipment conditions, respectively.

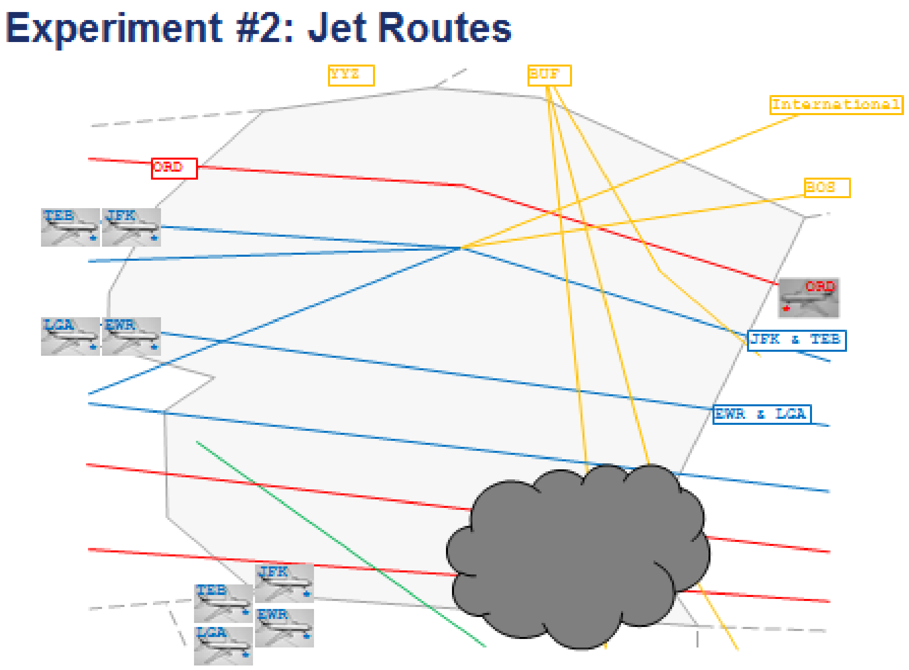


Figure 5. Jamestown (ZOB-79) during Experiment 2: Jet Routes condition.

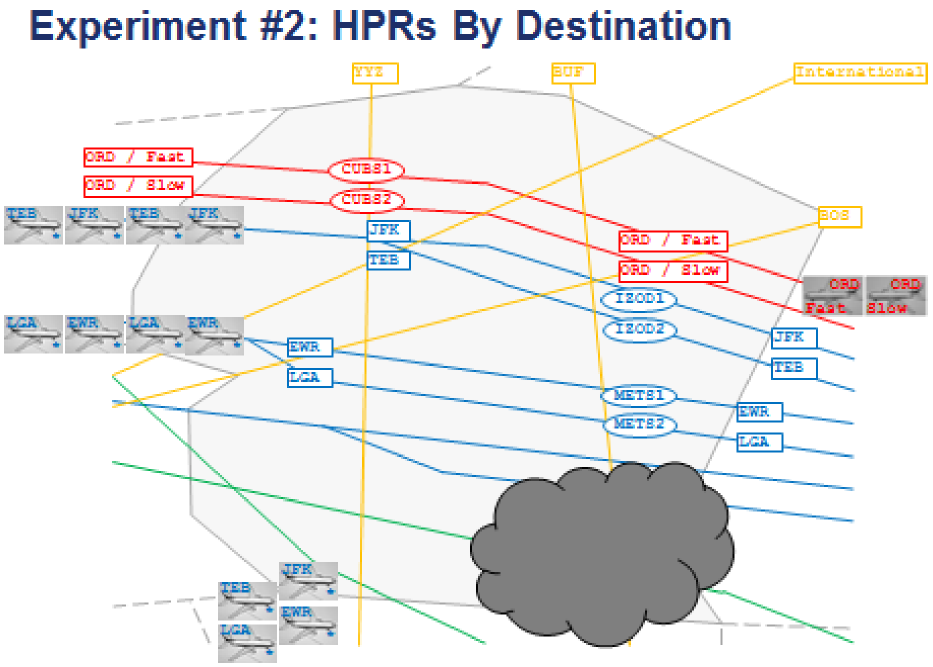


Figure 6. Jamestown (ZOB-79) during Experiment 2: HPRs by Destination condition.

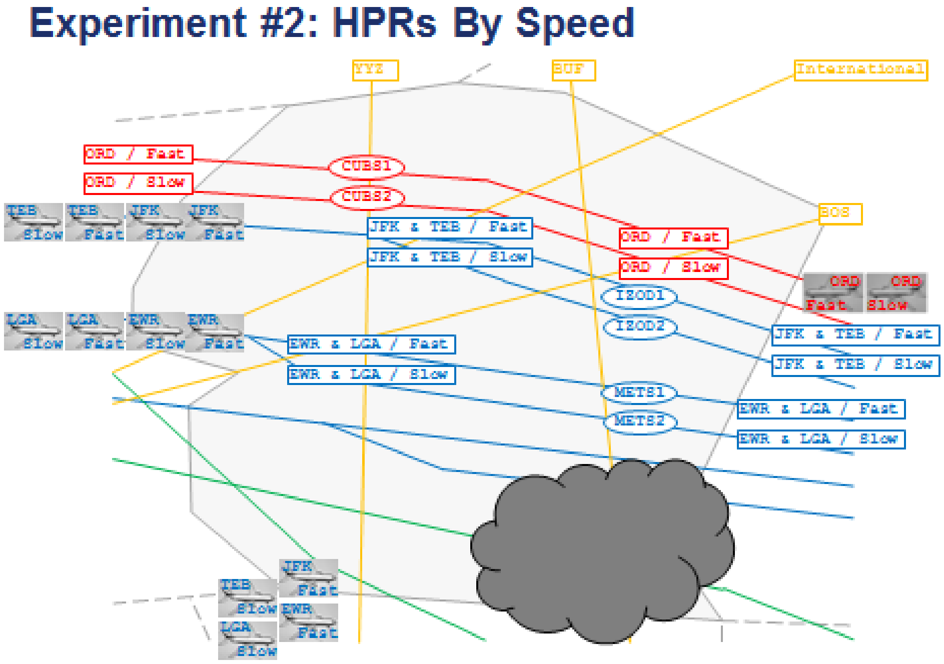


Figure 7. Jamestown (ZOB-79) during Experiment 2: HPRs by Speed condition.

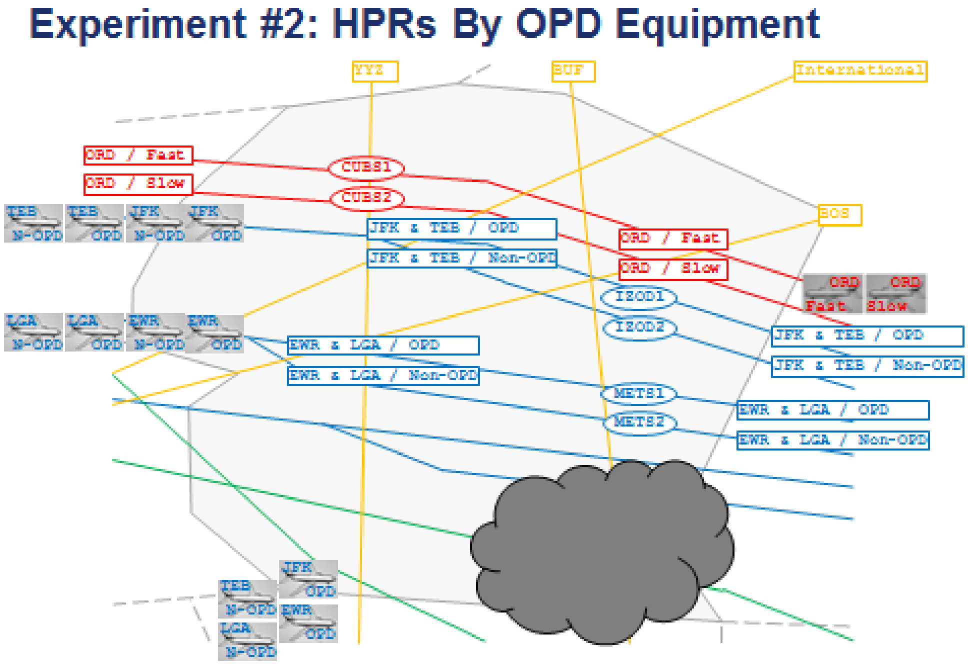


Figure 8. Jamestown (ZOB-79) during Experiment 2: HPRs by OPD Equipment condition.

Each group of participants was trained with ten practice scenarios. During these scenarios, the controllers learned the sectors, the HPR procedures, and the Data Communications (Data Comm) system. Each group of controllers performed six of the practice scenarios in Experiment 1, four as R-side/D-side teams and two as R-side only controllers. Controllers performed the remaining four scenarios in Experiment 2 as R-side only controllers. All ZOB-79 practice scenarios were 45 minutes in duration and simulated slightly lower traffic levels compared with the experimental scenarios.

2.3.4.2 Generic Sector Operations

We designed Experiment 3 to test the Generic Sector concept in the two Kansas City ARTCC sectors that we selected. We used four experimental traffic scenarios based on live traffic feeds from November 2011. The traffic was representative of that anticipated in the midterm timeframe.

The controllers had no training scenarios because the purpose of the simulation was to investigate ease of learning without extensive training. Prior to the first scenario, the controllers received sector briefings and Controller Information Tool (CIT) training (discussed in Section 2.4.3). Each group of controllers performed four test scenarios as four R-side only positions.

2.4 Equipment

2.4.1 Controller Workstations

We configured the controller workstations for both R-side/D-side team operations and R-side only operations depending upon the experimental conditions. The R-side controller workstation consisted of a high-resolution (2,048 x 2,048) 29" radar display, keyboard, trackball, and Keypad Selection Device (KSD). The D-side controller workstation consisted of a high-resolution (2,048 x 2,048) 29" display, keyboard, and mouse. When controllers worked in R-side/D-side teams, the JEDI/URET prototype was deployed on the D-side controller display. When controllers worked in the R-side only configuration, the JEDI/URET prototype was deployed on a fixed display above the controller's radar display. The controllers used a Voice Switching and Control System (VSCS) panel to communicate with the simulation pilots. In addition, controllers used a Workload Assessment Keypad (WAK) to record their workload ratings during the simulation.

2.4.2 Communications System

2.4.2.1 Voice Communications

Controllers used the RDHFL communications system that emulates the user interface of the VSCS currently used in the field. The communications system consists of a Push-to-Talk (PTT) capability with individual relay switchboxes, headsets with microphones, and PTT handsets or foot pedals. The communications system records the time, position, and switch status for every PTT transmission during a simulation.

2.4.2.2 Data Communications

During training scenarios, approximately 50% of aircraft were Data Comm equipped. For all test scenarios, Data Comm equipage rates were set at 30% to approximate the level anticipated in the NextGen midterm environment. For these sessions, the aircraft were equipped with Data Comm Segment I Services including altitude clearances, speed changes, heading changes, and transfer-of-communication services (see Figure 9).

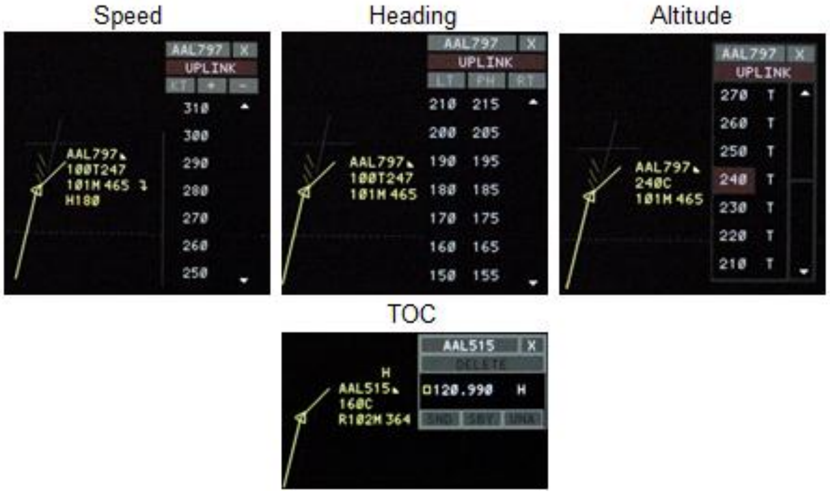
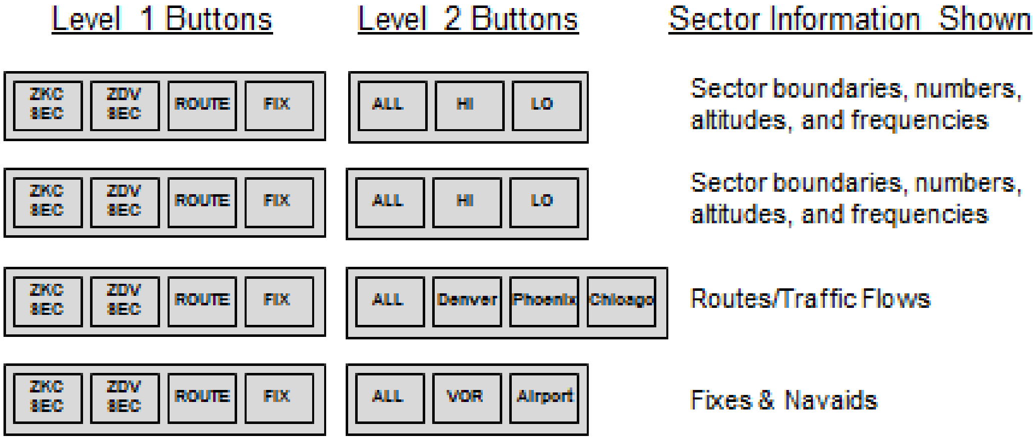


Figure 9. Data communications services.

2.4.3 Controller Information Tool

The ERAM CIT is an R-side display tool developed by NASA to support the Generic Sector concept (Mogford, 2010). It shows necessary sector-specific information for all sectors in the area, including a map with sector boundaries, sector names, sector traffic routes, altitudes, radio frequencies, and fix names and locations. The tool is operated by a series of software toggle buttons that show or hide the individual sector information. In the Generic Sector simulation runs, participants had access to the full CIT as it is depicted in Figure 10. In the HPR simulation runs, participants used a limited CIT that provided only HPR routes and fix names organized by destination.



Level 1 Buttons	Level 2 Buttons	Sector Information Shown
ZKC SEC, ZDV SEC, ROUTE, FIX	ALL, HI, LO	Sector boundaries, numbers, altitudes, and frequencies
ZKC SEC, ZDV SEC, ROUTE, FIX	ALL, HI, LO	Sector boundaries, numbers, altitudes, and frequencies
ZKC SEC, ZDV SEC, ROUTE, FIX	ALL, Denver, Phoenix, Chicago	Routes/Traffic Flows
ZKC SEC, ZDV SEC, ROUTE, FIX	ALL, VOR, Airport	Fixes & Navaids

Figure 10. Controller information tool.

2.4.4 Workload Assessment Keypad

Controllers used the RDHFL WAK devices to provide workload ratings using the Air Traffic Workload Input Technique (ATWIT). ATWIT is an unobtrusive and reliable technique for collecting controller workload ratings as controllers work traffic in a simulation (Stein, 1985; Stein, 1991). The WAK consists of a touch panel display with 10 buttons labeled from 1 to 10. The WAK is connected to a computer that controls the device and records workload ratings. The system is programmable allowing researchers to select the timing parameters for the study. The system prompts controllers for workload ratings at a selected time interval by emitting a tone and illuminating the keypad buttons. Controllers provide their workload ratings by pressing one of the 10 buttons, where 1 indicates *very low* workload and 10 indicates *very high* workload. If controllers do not respond before the timeout period, the system records a code indicating there was no response. In this simulation, we selected 2 minutes as the rating time interval and 20 seconds as the timeout period.

2.4.5 Audio-Visual Recording System

We used the RDHFL audio-video recording system to record controller voice communications and actions during the simulation. We positioned an overhead video camera above each team to record controllers' upper body and arm actions. The audio-video recording serves as a record of the simulation that the researchers can review if needed.

2.4.6 Simulation Pilot Workstations

The present study required between six and twelve simulation pilot workstations linked together in a network with the controller workstations. Each simulation pilot workstation consisted of a computer monitor, keyboard, and mouse. A section of the computer monitor depicted a situation display of the airspace and aircraft in the simulation similar to the controller display. The remaining display area contained a list of aircraft assigned to the simulation pilot, flight information, and a user interface to enter flight plan changes into the system. Each simulation pilot was responsible for several aircraft during the simulation. The simulation pilots used the RDHFL communications system to talk to controllers.

2.5 Materials

2.5.1 Informed Consent Statement

Each participant read and signed the Informed Consent Statement before beginning the study. The Informed Consent Statement described the purpose of the study and the rights and responsibilities of the participants, and assured participants that their data would be confidential and anonymous (see Appendix A).

2.5.2 Biographical Questionnaire

Each participant completed the Biographical Questionnaire before beginning the experiment. The purpose of the Biographical Questionnaire was to collect general descriptive information about the participants including gender, age, and level of ATC experience (see Appendix B).

2.5.3 Post-Scenario Questionnaire

The participants completed the Post-Scenario Questionnaire (PSQ) after each test scenario. The purpose of the PSQ was to collect data regarding the controller's experience in the traffic scenario just completed. The controllers provided ratings about their performance, workload, and situation awareness. Controllers also provided ratings about the experimental conditions tested in the scenario, such as the generic procedures and support tools (if any). The PSQ included ratings and open-ended questions about the support tools and their effects on safety, capacity, and efficiency. The controllers were able to comment about anything they experienced during the scenario that they considered relevant to the study (see Appendix C).

2.5.4 Exit Questionnaire

The participants completed the Exit Questionnaire after completing all traffic scenarios. The purpose of the Exit Questionnaire was to collect data regarding the controller's experience in the entire study. The controllers provided ratings about the realism of the simulation including the airspace, traffic scenarios, and ATC equipment. Controllers also provided ratings that compared the experimental conditions tested in each experiment. The Exit Questionnaire included ratings and open-ended questions. The controllers were able to comment about anything they experienced that they considered relevant to the study (see Appendix D).

2.5.5 Observer Rating Form

After each test scenario, the SMEs used the Observer Rating Form to provide performance ratings for each of the R-side/D-side controller teams or for individual controllers when they operated in the R-side only configuration. The Observer Rating Form was developed by ERPs and SMEs in the RDHFL to evaluate new ATC concepts and procedures by observing controller performance in HITL simulations (Sollenberger, Stein, & Gromelski, 1997; Vardaman & Stein, 1998). The Observer Rating Form consists of several rating scales designed to assess different aspects of ATC performance, such as resolving aircraft conflicts, sequencing aircraft, prioritizing tasks, communicating effectively, and maintaining situation awareness (see Appendix E). SMEs filled out a PSQ after each scenario and filled out the Exit Questionnaire at the completion of the study.

2.6 Experimental Design

2.6.1 Independent Variables

We used two different experiments to investigate the HPR concepts and a third experiment to demonstrate the Generic Sector concept. In Experiment 1, we used ZOB-79 to investigate the transition from HA HPRs to lower altitude airspace. We compared the baseline traffic scenarios with jet routes to the HPR scenarios. In Experiment 2, we used ZOB-79 to investigate different lane usage strategies for HPRs. We compared the baseline traffic scenarios with jet routes to HPRs using lanes for different aircraft destinations, different aircraft speeds, and for aircraft that were either equipped or not equipped to support OPD operations. Aircraft that were equipped for OPD operations were identified by their equipment suffixes (/Q indicated OPD-equipped). In addition, for all scenarios throughout Experiments 1 and 2, pilots were instructed to report OPD equipment on board when they checked in with controllers upon entering the sector. In both experiments, scenarios were designed to examine the experimental conditions in both medium- and high-traffic levels.

In Experiment 3, we used ZKC-21 and ZKC-7 to investigate whether controllers could manage traffic safely and efficiently in an unfamiliar, generic sector with support tools and minimal sector training.

2.6.2 Simulation Measures

2.6.2.1 System Effectiveness Measures

The RDHFL simulation software has an extensive data collection system that records aircraft track and status information during the simulation. We analyzed the aircraft track and status data to produce objective system effectiveness measures in the critical areas of safety, capacity, efficiency, and communications (Buckley, DeBaryshe, Hitchner, & Kohn, 1983; Stein & Buckley, 1992). The primary measure of safety was loss of aircraft separation. The measures of capacity were the number of aircraft accepted, number of handoffs initiated, and number of aircraft under control. The measures for efficiency were the aircraft time and distance flown through each sector as well as the number of control commands (e.g., altitude, heading, and speed commands) issued. We used the data recorded by the communications system to analyze the frequency and duration of controller and pilot communications.

2.6.2.2 Human Factors Measures

The PSQ was the main source of subjective data, and it measured controller performance, workload, and situation awareness in each of the experimental conditions. We used the WAK and ATWIT to provide an additional measure of workload using a real-time technique as controllers performed the traffic scenarios.

2.6.2.3 Observer Ratings

The SMEs used the Observer Rating Form to provide subjective ratings of controller performance in each of the experimental conditions. The SMEs were experienced observers who were used to training controllers and evaluating ATC performance. SMEs often detect controller actions that affect safety, capacity, and efficiency that cannot be measured by objective techniques.

2.6.2.4 Support Tools Usage

The RDHFL simulation software also records keyboard data entry and trackball input. We analyzed the keyboard and trackball data to determine how often the controllers were using the CIT and other controller functions. The PSQ and Exit Questionnaire included questions about CIT effectiveness and acceptability.

2.7 Procedure

2.7.1 Daily Schedule

Table 2 shows the daily schedule of activities for the participants in the current study (also see Table 3 for a description of each scenario listed in the schedule). Each group of participants consisted of four controllers who were released from their facility for two weeks to participate in the HA Airspace Study and the Conflict Resolution Advisory Study. The controllers traveled to the FAA WJHTC on Monday and participated in the HA Airspace Study during the first week. The controllers stayed over the weekend on off-duty travel and participated in the Conflict Resolution Advisory Study during the second week.

Table 2. Daily Schedule of Activities

Tuesday		Wednesday		Thursday		Friday	
Time	Activity	Time	Activity	Time	Activity	Time	Activity
8:00-9:00	Project Briefing	8:00-8:45	T1B: Baseline	8:00-8:45	T2: Jet Routes	8:00-8:45	E2: Destination
9:00-9:15	Break	8:45-9:15	Break	8:45-9:15	Break	8:45-9:15	Break
9:15-10:15	Sector Briefing	9:15-10:00	E1: Baseline	9:15-10:00	T2: Speed	9:15-10:00	E2: Equipment
10:15-10:30	Break	10:00-10:30	Break	10:00-10:30	Break	10:00-10:30	Break
10:30-11:15	T1A: Baseline	10:30-11:15	E1: Baseline	10:30-11:15	T2: Destination	10:30-11:15	E3: Generic-1
11:15-12:45	Lunch	11:15-12:45	Lunch	11:15-12:45	Lunch	11:15-12:45	Lunch
12:45-1:30	T1A: Baseline	12:45-1:30	T1B: HPR	12:45-1:30	T2: Equipment	12:45-1:30	E3: Generic-2
1:30-2:00	Break	1:30-2:00	Break	1:30-2:00	Break	1:30-2:00	Break
2:00-2:45	T1A: HPR	2:00-2:45	E1: HPR	2:00-2:45	E2: Jet Routes	2:00-2:45	E3: Generic-3
2:45-3:15	Break	2:45-3:15	Break	2:45-3:15	Break	2:45-3:15	Break
3:15-4:00	T1A: HPR	3:15-4:00	E1: HPR	3:15-4:00	E2: Speed	3:15-4:00	E3: Generic-4
4:00-4:30	Discussion	4:00-4:30	Discussion	4:00-4:30	Discussion	4:00-4:30	Discussion

Note. T refers to training scenario and E refers to experimental scenario. The number after the T or E indicates the experiment number for the scenario.

On Tuesday, Wednesday, Thursday, and Friday the controllers participated in the experiment and performed training and experimental scenarios. The daily schedule was the same across participants with the exception that the order of conditions within each of the three experiments was counterbalanced across participants. At the end of each day, we held a group meeting to answer the participants' questions and discuss their experiences in the simulation. On the first day of the study, we briefed the participants about the project goals and sectors they were operating in the simulation. The participants completed the Informed Consent Statement and the Biographical Questionnaire. On the last day of the study, we conducted an exit briefing, and the participants completed the Exit Questionnaire.

2.7.2 Training and Experimental Sessions

Table 3 shows a summary of the training and experimental sessions. Tuesday was a training session for Experiment 1. The participants performed four practice scenarios to become familiar with the simulation equipment and procedures. The controllers began training in the ZOB-79 sector and practiced the Baseline and HPR experimental conditions. The controllers conducted traffic as R-side/D-side teams and switched positions between scenarios to gain training in both positions. On Wednesday, the participants performed one more practice scenario in the morning and one in the afternoon as R-side only controllers before beginning the experimental scenarios. The controllers performed the experimental scenarios as R-side/D-side teams and switched positions between simulation runs. We counterbalanced the presentation order of the Baseline and HPR conditions between the two teams of controllers.

Table 3. Summary of Training and Experimental Sessions

Session	Purpose	Sector	Experimental Conditions	Configuration
T1A	HPR Arrival Training	ZOB-79	Baseline, HPR	R-side/D-side Teams
T1B	HPR Arrival Training	ZOB-79	Baseline, HPR	R-side Only
E1	HPR Arrival Experiment	ZOB-79	Baseline, HPR	R-side/D-side Teams
T2	HPR Lane Usage Training	ZOB-79	Jet Routes, Speed, Destination, Equipment	R-side Only
E2	HPR Lane Usage Experiment	ZOB-79	Jet Routes, Speed, Destination, Equipment	R-side Only
E3	Generic Sector Experiment	ZKC-21, ZKC-7	Generic-1, Generic-2, Generic-3, Generic-4	R-side Only

Note. T refers to training scenario and E refers to experimental scenario. The number after the T or E indicates the experiment number for the scenario; HPR = High Performance Route; ZOB = Cleveland ARTCC; ZKC = Kansas City ARTCC.

On Thursday and Friday, we conducted Experiment 2, which used ZOB-79 to investigate alternative lane usages for HPRs. The participants practiced the jet routes, HPRs by speed, HPRs by destination, and HPRs by OPD equipment experimental conditions. The controllers performed all of the practice and experimental scenarios in the R-side only position. We counterbalanced the presentation order of the four experimental conditions across each group of four controllers.

On Friday afternoon, we conducted Experiment 3 to demonstrate the Generic Sector concept. Prior to beginning the first scenario run, controllers were trained to use the CIT. The participants performed four experimental scenarios with the ZKC-21 and ZKC-7 sectors. The controllers performed all of the experimental scenarios in the R-side only position. We counterbalanced the presentation order of the scenarios across each group of four controllers.

For all experiments, participants used WAKs during all training and experimental scenarios. We used the audio-video recording system during the experimental sessions. After each experimental scenario, controllers completed the PSQ and SMEs completed the PSQ as well as the Observer Rating Form.

3. RESULTS

3.1 Experiment 1

In Experiment 1, participants controlled traffic in two-person teams. We analyzed workload ratings, PTT transmissions, and questionnaire responses separately for R-side and D-side positions. For all other measures, we did not differentiate between R-side and D-side positions. Scenarios lasted for 45 minutes, and system metrics were analyzed for the time interval between 2 minutes and 44 minutes so that data was not contaminated while participants were acclimating themselves to the scenario run or the when participants were winding down in anticipation of the scenario's end. The medium-traffic scenario interval was defined as the time between 2 minutes and 23 minutes. The high-traffic scenario interval was defined as the time between 23 minutes and 44 minutes. Unless otherwise noted, data were analyzed using a 2 (condition) x 2 (traffic level) repeated measures Analysis of Variance (ANOVA). If a statistically significant interaction was shown (as indicated by a *p*-value less than 0.05), follow-up, Bonferroni-adjusted paired *t*-tests were run to determine the nature of the interaction.

Due to the design of the airspace in the HPRs condition, Philadelphia aircraft flying on the Flyr1 HPR spent more time in the controller's sector than they did in the jet routes condition despite the Flyr1 providing a more efficient route to Philadelphia. We removed Philadelphia aircraft from the majority of our analyses because they represent a case where the jet routes condition would appear more efficient than the HPRs condition using our metrics (e.g., time and distance in the controller's sector), even if the HPRs condition is more efficient when the aircraft's entire route is taken into account. We note below which analyses factored out the Philadelphia aircraft.

3.1.1 Voice Communications

Voice communications contribute to participants' workload, so reducing the number and duration of voice communications is desirable. D-side controller-initiated transmissions were not analyzed, as very few D-side controllers initiated any PTT transmissions, making it impossible to make comparisons across conditions. The subsequent analyses focus only on R-side controller-initiated transmissions and pilot-initiated transmissions. Figures 11 and 12 present the means, by traffic level and condition, for the number and duration of controller-initiated PTT transmissions. Figures 13 and 14 present the means, by traffic level and condition, for the number and duration of pilot-initiated PTT transmissions.

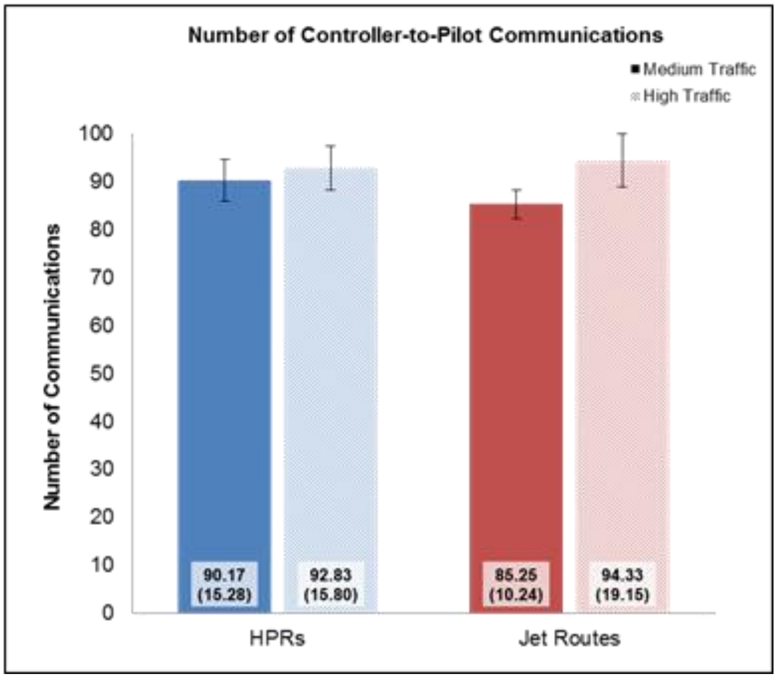


Figure 11. Number of R-side controller-initiated transmissions by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

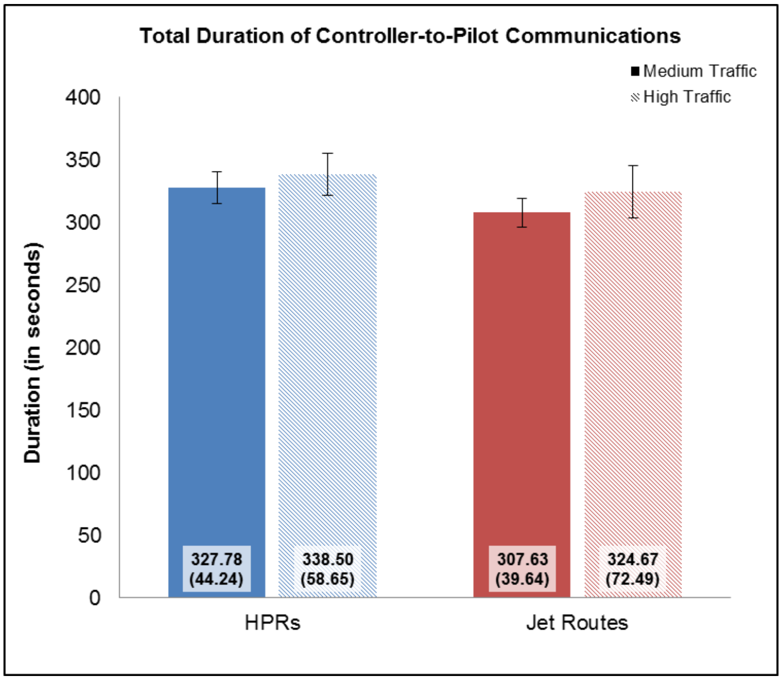


Figure 12. Total duration of R-side controller-initiated transmissions by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

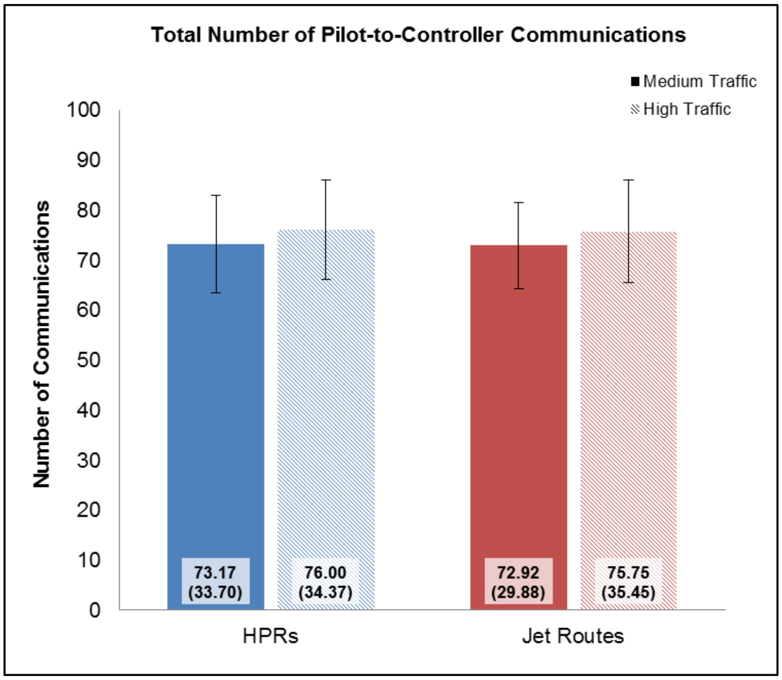


Figure 13. Number of pilot-initiated transmissions by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

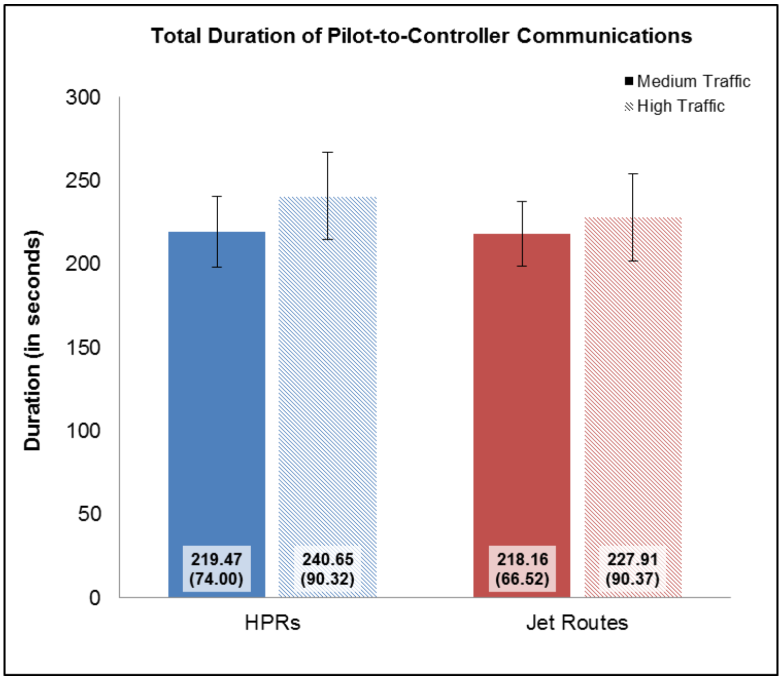


Figure 14. Duration of pilot-initiated transmissions by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

Participants made marginally more controller-initiated transmissions under high levels of traffic compared with medium levels of traffic, $F(1, 11) = 3.28, p = 0.098, \eta_\rho^2 = 0.23$. There were no effects of traffic level on the total duration of controller-initiated transmissions, $F(1, 11) = 2.20, p = 0.17, \eta_\rho^2 = 0.17$, the number of pilot-initiated transmissions, $F(1, 11) = 1.27, p = 0.28, \eta_\rho^2 = 0.10$, or the total duration of pilot-initiated transmissions, $F(1, 11) = 2.88, p = 0.12, \eta_\rho^2 = 0.21$.

The total duration of controller-initiated transmission was marginally longer in the HPR scenarios compared with the jet route scenarios, $F(1, 11) = 3.91, p = 0.07, \eta_\rho^2 = 0.26$. There were no significant differences between HPRs and jet routes for the number of controller-initiated transmissions, $F(1, 11) = 0.46, p = 0.51, \eta_\rho^2 = 0.04$, the number of pilot-initiated transmissions, $F(1, 11) = 0.01, p = 0.91, \eta_\rho^2 = 0.001$ or the duration of pilot-initiated transmissions, $F(1, 11) = 0.95, p = 0.35, \eta_\rho^2 = 0.08$.

There were no interactions between traffic level and condition for the number of controller-initiated transmissions, $F(1, 11) = 0.91, p = 0.36, \eta_\rho^2 = 0.08$, the number of pilot-initiated transmissions, $F(1, 11) < 0.001, p = 1.00, \eta_\rho^2 < 0.001$, the duration of controller-initiated transmissions, $F(1, 11) = 0.12, p = 0.74, \eta_\rho^2 = 0.01$, or the duration of pilot initiated transmissions, $F(1, 11) = 1.02, p = 0.34, \eta_\rho^2 = 0.09$.

In summary, for the number of both controller- and pilot-initiated transmissions as well as for the total duration of pilot-initiated transmissions, there were no notable differences between HPRs and jet routes. There was a nonsignificant trend in the duration data showing that participants spent slightly more time using voice communications in the HPRs condition compared with the jet routes condition under high levels of traffic. A number of factors could have led to this trend including unfamiliarity with the HPR routes and fixes, the way we implemented HPR procedures in the scenarios, or the number of syllables necessary to voice the HPR route names and fixes compared with jet routes.

3.1.2 Losses of Separation

Losses of separation were defined as incidents where aircraft were separated by less than 5 nmi laterally and 1,000 ft. vertically. An ATC SME reviewed all potential losses of separation and categorized the incidents as either occurring due to system or simulation pilot error or due to controller error. Across all experimental scenarios in Experiment 1, one incident was attributed to controller error. Two aircraft were on converging courses at FL370. After Conflict Alert activated, the participant attempted to turn both aircraft to maintain separation but was not successful. In addition, the participant descended one aircraft to FL310 but not in time to achieve separation. This loss of separation occurred in the HPRs condition in the 37th minute of the scenario, when the traffic level was high, and lasted for 24 seconds.

3.1.3 JEDI/URET Conflict Probe Notifications

We looked at the number and duration of conflict probe notifications as potential measures of safety; however, it is important to note that a controller doing a good job keeping aircraft separated may run aircraft close together, resulting in many notifications. We examined the number and duration of conflict probe notifications as well as the number and duration for each type of notification (red or yellow). Figures 15, 16, 17, 18, 19, and 20 display the number and duration of conflict probe notifications overall, with red notifications, and with yellow notifications, respectively.

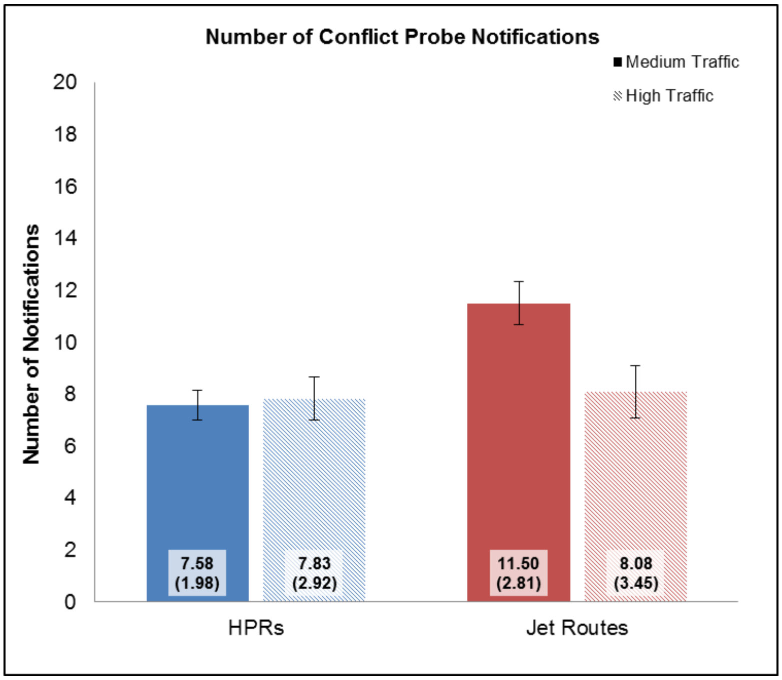


Figure 15. Number of conflict probe notifications by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

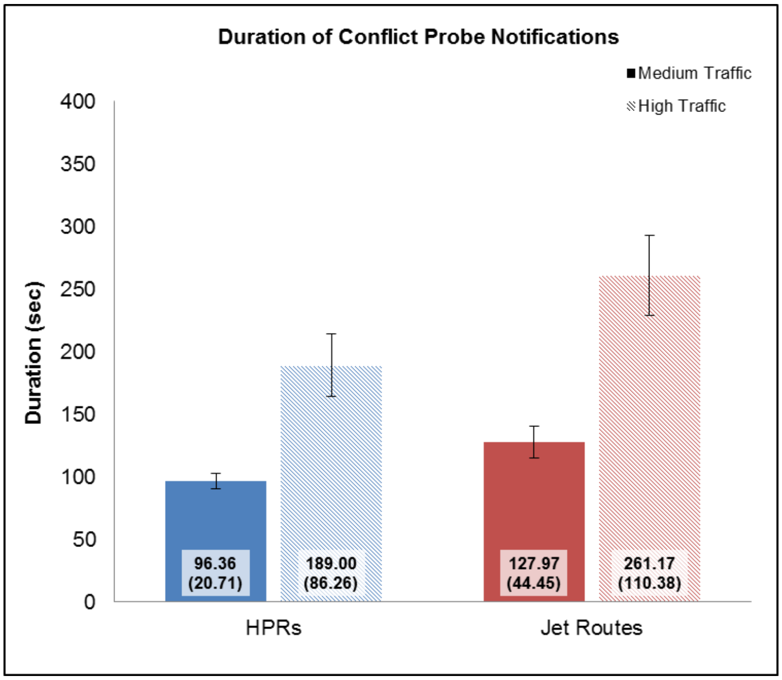


Figure 16. Duration of conflict probe notifications by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

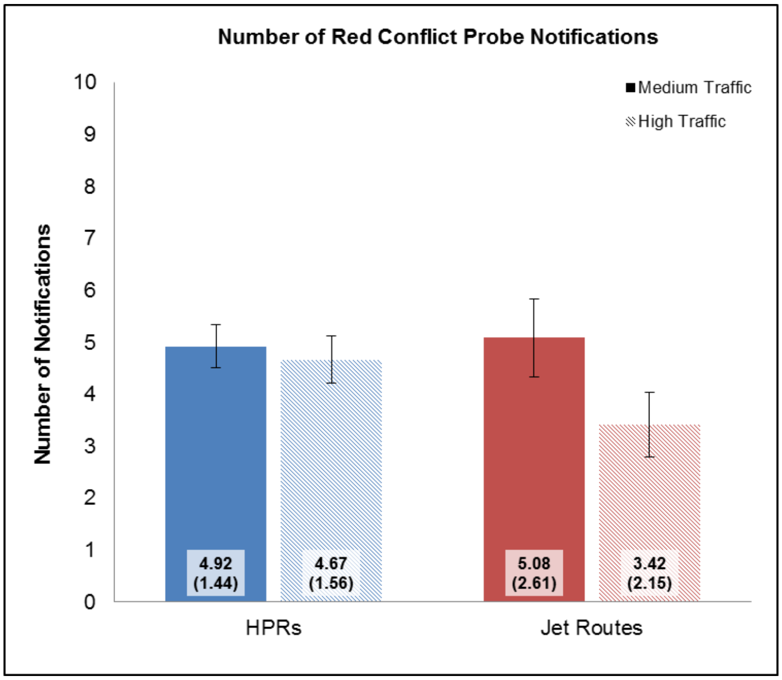


Figure 17. Number of red conflict probe notifications by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

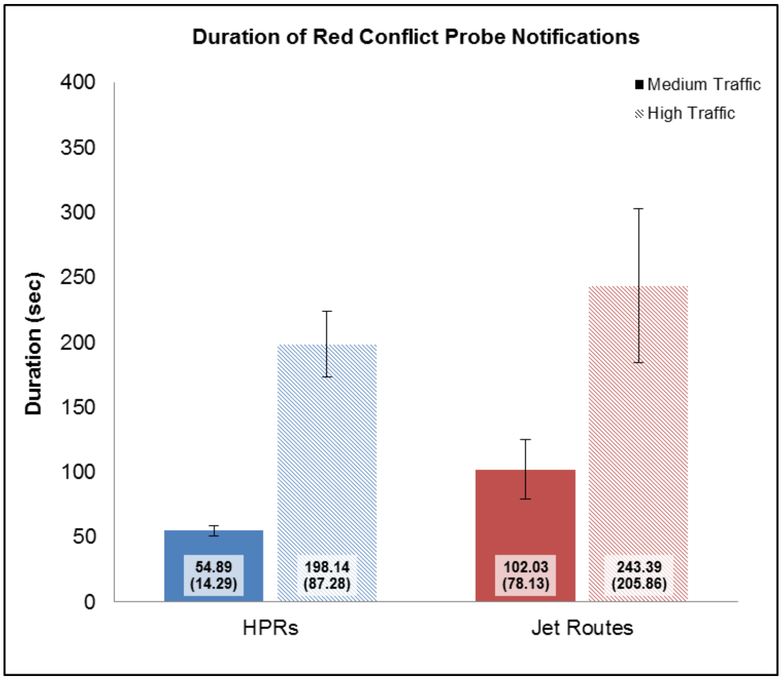


Figure 18. Duration of red conflict probe notifications by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

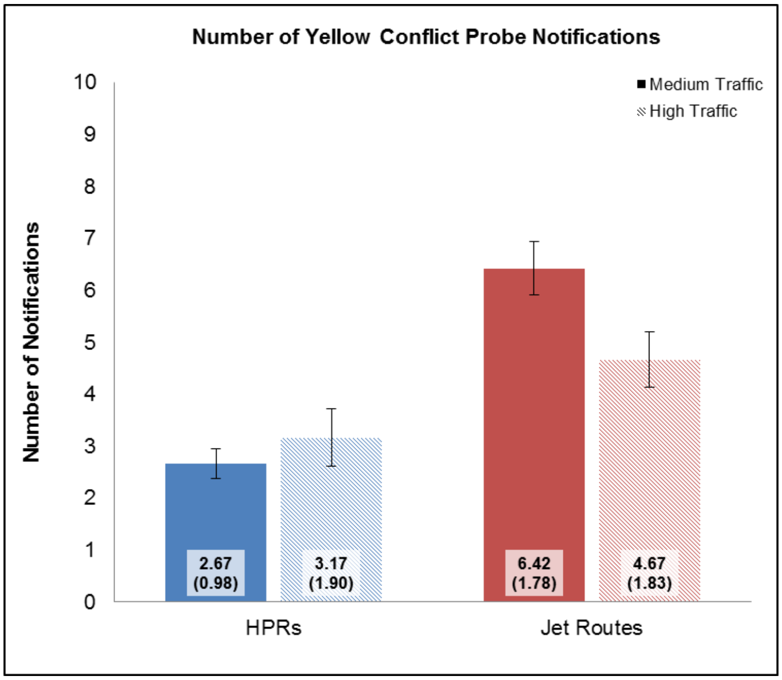


Figure 19. Number of yellow conflict probe notifications by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

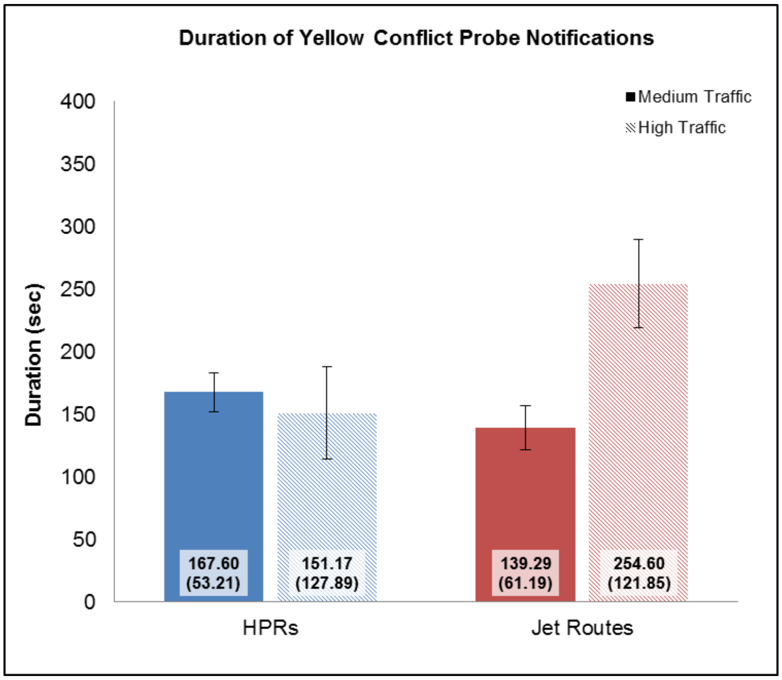


Figure 20. Duration of yellow conflict probe notifications by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

There were fewer conflict probe notifications (collapsed across color) under high-traffic levels compared with medium-traffic levels, $F(1, 11) = 10.55$, $p = 0.008$, $\eta_\rho^2 = 0.49$, but the duration of the notifications was longer under high-traffic levels compared with medium-traffic levels, $F(1, 11) = 27.40$, $p < 0.001$, $\eta_\rho^2 = 0.71$. There were fewer notifications when participants used HPRs compared with jet routes, $F(1, 11) = 5.71$, $p = 0.036$, $\eta_\rho^2 = 0.34$. In addition, conflict probe notifications were shorter when participants used HPRs compared with jet routes, $F(1, 11) = 6.17$, $p = 0.03.$, $\eta_\rho^2 = 0.36$. Finally, there was a marginal interaction between traffic level and condition for the number of notifications, $F(1, 11) = 4.52$, $p = 0.057$, $\eta_\rho^2 = 0.29$, and no interaction for the duration of notifications, $F(1, 11) = 1.62$, $p = 0.23$, $\eta_\rho^2 = 0.13$.

There were significantly more yellow notifications under medium-traffic levels compared with high-traffic levels, $F(1, 11) = 11.30$, $p = 0.006$, $\eta_\rho^2 = 0.51$, but the duration of the yellow notifications was significantly longer under high-traffic levels, $F(1, 11) = 6.60$, $p = 0.03$, $\eta_\rho^2 = 0.38$. Similarly, there were marginally more red notifications under medium-traffic levels compared with high-traffic levels, $F(1, 11) = 4.71$, $p = 0.053$, $\eta_\rho^2 = 0.30$, but the duration of the red notifications was significantly longer under high-traffic levels, $F(1, 11) = 14.07$, $p = 0.003$, $\eta_\rho^2 = 0.56$.

There were fewer yellow notifications when participants used HPRs compared with jet routes, $F(1, 11) = 19.96$, $p = 0.001$, but no difference between conditions for the duration of yellow notifications, $F(1, 11) = 1.71$, $p = 0.22$, $\eta_\rho^2 = 0.13$. There were no differences between conditions for the number of red notifications, $F(1, 11) = 1.07$, $p = 0.32$, $\eta_\rho^2 = 0.09$, or duration of red notifications, $F(1, 11) = 1.77$, $p = 0.21$, $\eta_\rho^2 = 0.14$.

There were interactions between traffic level and condition for both the number, $F(1, 11) = 6.41$, $p = 0.03$, $\eta_\rho^2 = 0.37$, and duration, $F(1, 11) = 10.99$, $p = 0.007$, $\eta_\rho^2 = 0.50$, of yellow notifications. For the number of yellow notifications, the difference between conditions was larger under medium-traffic levels. For the duration of yellow notifications, under medium-traffic levels durations were shorter and under high-traffic levels durations were longer for jet routes compared with HPRs. Finally, there were no interactions between traffic level and condition for the number, $F(1, 11) = 1.39$, $p = 0.26$, $\eta_\rho^2 = 0.11$, or duration, $F(1, 11) < 0.001$, $p = 0.97$, $\eta_\rho^2 < 0.001$, of red notifications.

In summary, HPRs reduced the number of conflict probe notifications. When these notifications were broken down by type (yellow or red), it was clear that the primary difference between HPRs and jet routes was the number of yellow conflict probe notifications. While there was an overall difference in the duration of notifications between conditions, this difference did not remain when the analysis focused on each type of notification.

It is possible that these differences observed across conditions reflect a safety benefit for HPRs compared with jet routes; however, it is just as likely that participants in both conditions operated traffic safely and the difference was one of strategy (e.g., running aircraft tighter and more efficiently spaced in the jet routes condition led to more notifications) rather than safety.

3.1.4 Number of Aircraft Under Control

The number of aircraft a participant can manage is a potential indicator of the effectiveness of the condition being tested. It is important to note that this measure is likely to be influenced by ceiling effects (making it difficult to see differences between conditions), as it was impossible for participants to manage more aircraft than the number that we built into a scenario. We defined the number of aircraft under control as the number of aircraft that were within the geographic bounds of the participant's sector. For this analysis, we removed all Philadelphia aircraft from the aircraft counts. Figure 21 displays the means, by traffic level and condition, for the number of aircraft under control.

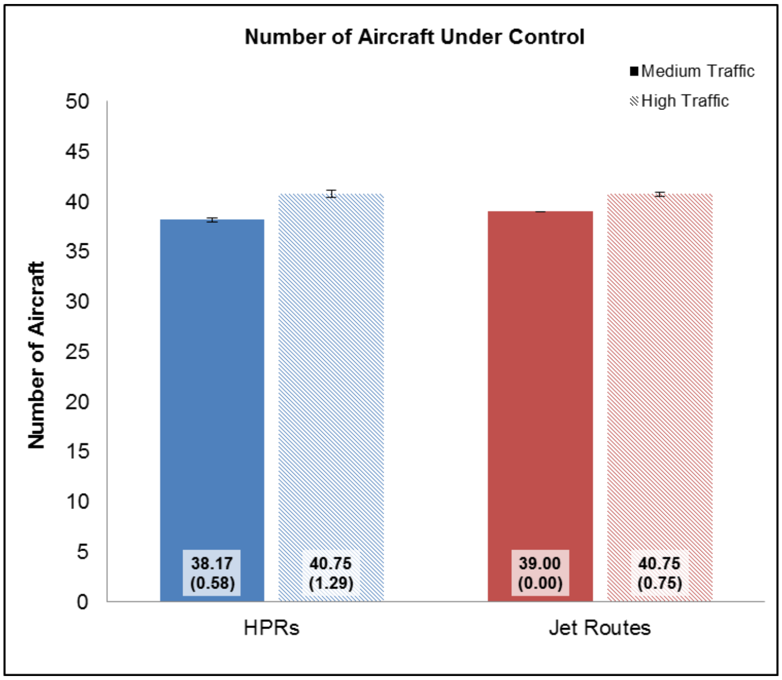


Figure 21. Number of aircraft under control by traffic level and condition.
Means (*SDs*) are presented at bottom of figure.

More aircraft were under control under high-traffic levels compared with medium-traffic levels, $F(1, 11) = 58.09, p < 0.001, \eta_\varrho^2 = 0.95$, as expected given that we intentionally increased traffic levels halfway through the scenarios. More aircraft were under control in the jet routes condition compared to the HPRs condition, $F(1, 11) = 7.86, p = 0.02, \eta_\varrho^2 = 0.02$; however, there was also a marginal interaction between traffic level and condition, $F(1, 11) = 4.66, p = 0.054, \eta_\varrho^2 = 0.61$. The trend in the data shows that when using HPRs, participants managed fewer aircraft under medium levels of traffic, but the same number of aircraft under high levels of traffic, compared with the number of aircraft they managed using jet routes.

3.1.5 Number of Handoffs Accepted and Handoffs Initiated

As an indicator of system capacity, we measured the number of handoffs participants accepted and initiated. For these analyses, we removed all Philadelphia aircraft. Figures 22 and 23 display the means, by traffic level and condition, for the number of handoffs accepted and initiated.

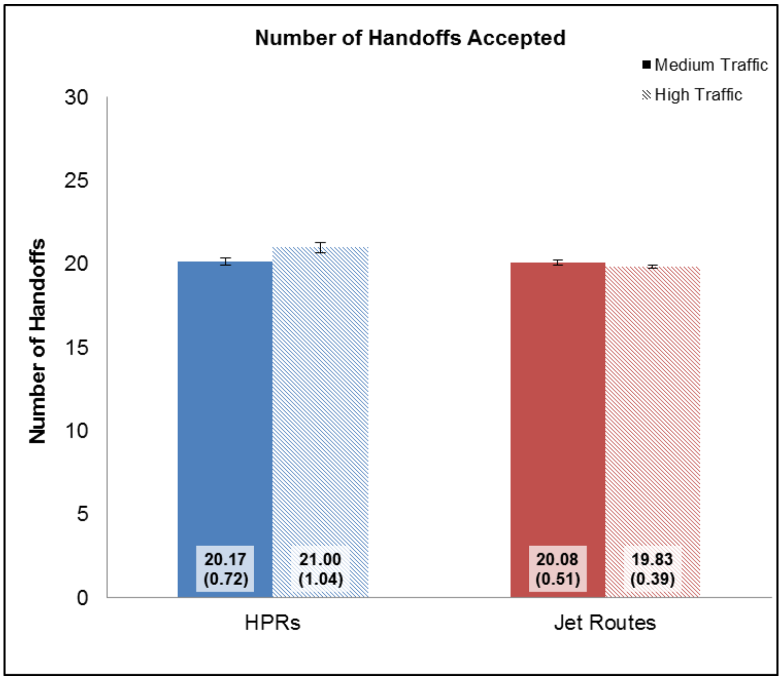


Figure 22. Number of handoffs accepted by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

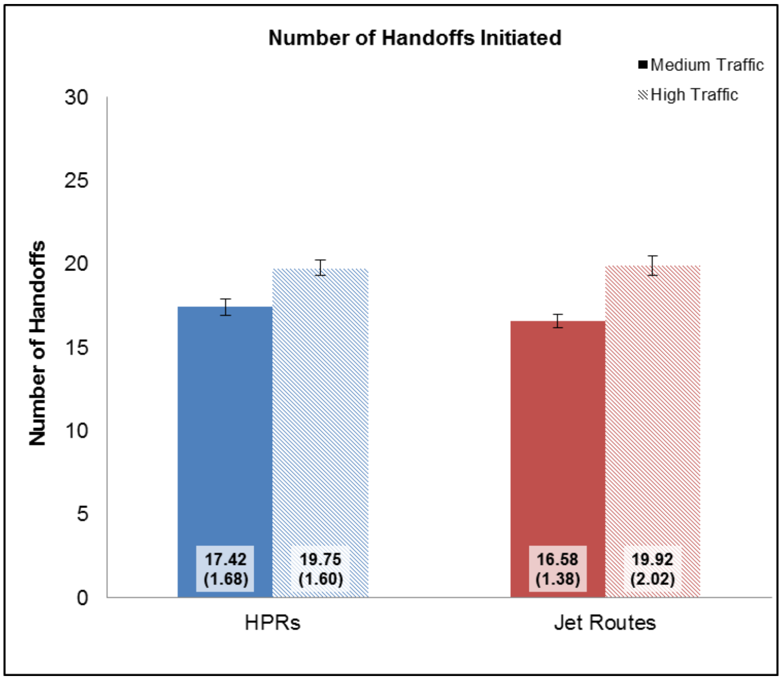


Figure 23. Number of handoffs initiated by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

There were no effects of traffic level on the number of handoffs accepted, $F(1, 11) = 1.96$, $p = 0.19$, $\eta_p^2 = 0.15$; however, fewer handoffs were initiated when traffic was at a medium level, $F(1, 11) = 42.96$, $p < 0.001$, $\eta_p^2 = 0.80$, compared with when traffic was at a high level. Participants accepted more handoffs in the HPRs condition compared with the jet routes condition, $F(1, 11) = 12.69$, $p = 0.004$, $\eta_p^2 = 0.54$, but initiated similar numbers of handoffs across both conditions, $F(1, 11) = 0.83$, $p = 0.38$, $\eta_p^2 = 0.07$. Finally, there was an interaction between traffic level and condition for the number of handoffs accepted, $F(1, 11) = 9.16$, $p = 0.01$, $\eta_p^2 = 0.45$, but no interaction for the number of handoffs initiated, $F(1, 11) = 0.77$, $p = 0.40$, $\eta_p^2 = 0.07$. The trend in the data shows that under medium-traffic levels, participants accepted similar numbers of handoffs in both conditions, $t(11) = 0.43$, $p = 0.67$, $d = 0.12$. Under high-traffic levels, participants accepted more handoffs when using HPRs compared with jet routes, $t(11) = 3.92$, $p = 0.002$, $d = 1.13$. On average, participants accepted one more handoff in the HPRs condition than they did in the jet routes condition when under high levels of traffic.

3.1.6 Aircraft Time and Distance in Sector

Aircraft time and distance in the participant's sector are measures of system efficiency and potential fuel consumption. If a participant has the time and resources, he or she can reroute aircraft with direct routings and decrease the time and distance the aircraft must travel through the sector. In addition, if one route structure is more efficient than the other (e.g., HPRs versus jet routes), we would expect to see a reduction in the time and distance each aircraft flies through the participant's sector. For these analyses, we removed all Philadelphia aircraft. Figures 24 and 25 display the means, by traffic level and condition, for aircraft time and distance in the participant's sector.

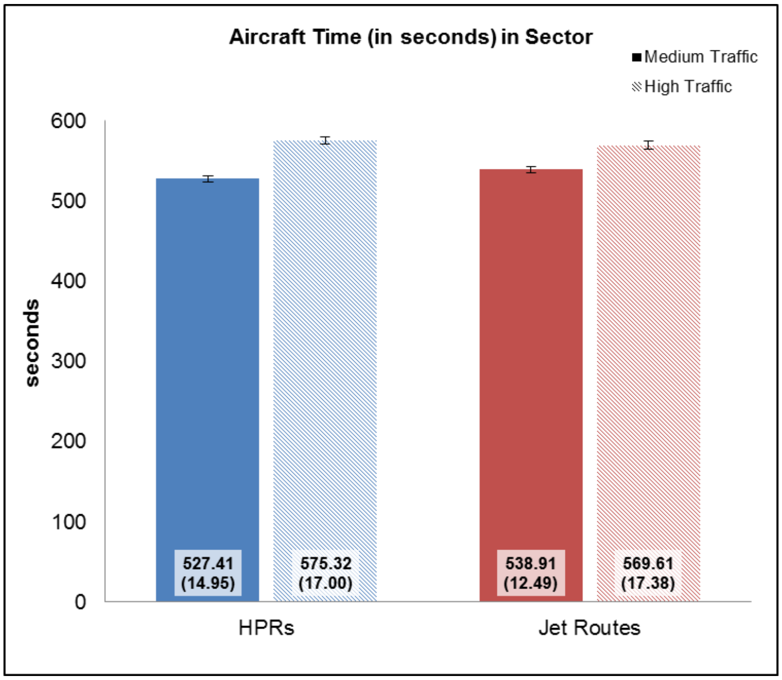


Figure 24. Aircraft time (in sec) in sector by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

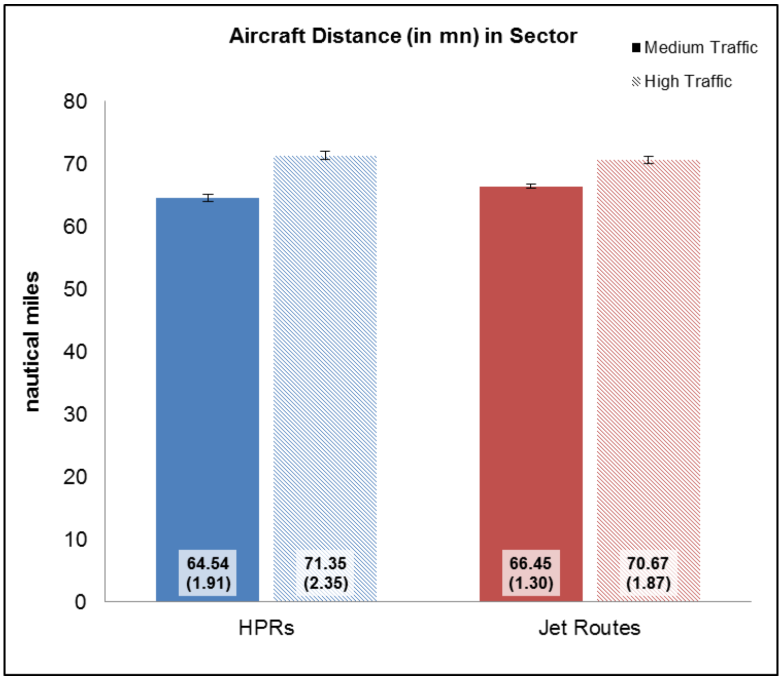


Figure 25. Aircraft distance (in nmi) in sector by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

Under high-traffic levels, aircraft flew longer amounts of time, $F(1, 11) = 41.19$, $p < 0.001$, $\eta_\rho^2 = 0.79$, and distance, $F(1, 11) = 54.51$, $p < 0.001$, $\eta_\rho^2 = 0.83$, through the sector compared with medium-traffic levels. Overall, there were no differences between HPRs and jet routes for aircraft time, $F(1, 11) = 0.55$, $p = 0.47$, $\eta_\rho^2 = 0.05$, or distance, $F(1, 11) = 1.46$, $p = 0.25$, $\eta_\rho^2 = 0.12$; however, there were interactions between traffic level and condition for both aircraft time, $F(1, 11) = 17.38$, $p = 0.002$, $\eta_\rho^2 = 0.61$, and distance, $F(1, 11) = 51.03$, $p < 0.001$, $\eta_\rho^2 = 0.82$. Follow up, Bonferroni-adjusted paired *t*-tests showed that under medium levels of traffic, aircraft flew marginally (at the Bonferroni-adjusted alpha level) shorter amounts of time, $t(11) = 2.54$, $p = 0.028$, $d = 0.73$, and significantly shorter distances, $t(11) = 3.51$, $p = 0.005$, $d = 1.01$, in the HPRs condition compared with the jet routes condition. Under high levels of traffic, aircraft flew comparable amounts of time, $t(11) = 1.33$, $p = 0.21$, $d = 0.38$, and distance, $t(11) = 1.25$, $p = 0.24$, $d = 0.36$, across conditions.

These data show that under medium levels of traffic, HPRs were more efficient than jet routes in terms of distance flown in the controller's sector (and marginally more efficient in terms of time flown through the sector). These data suggest that HPRs have the potential to increase efficiency in the NAS and at minimum they certainly will not negatively impact efficiency, even under high levels of traffic. It is possible that the benefits of HPRs under high levels of traffic were not realized here due to the high workload participants encountered in the HPRs scenarios.

3.1.7 Controller Interactions

Participants noted in their questionnaire responses that they spent too much time getting Flight Plan Readouts in the HPRs condition. We analyzed controller interactions to determine if participants used these commands more frequently in the HPRs condition. Figure 26 displays the means, by traffic level and condition, for the number of Flight Plan Readout commands.

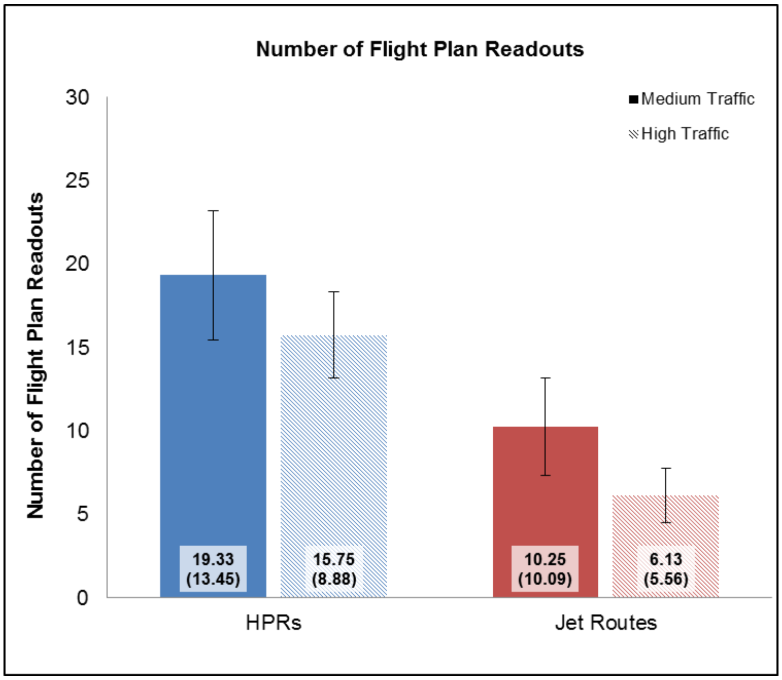


Figure 26. Number of flight plan readouts by traffic level and condition. Means (*SDs*) are presented at the bottom of the figure.

Participants checked more flight plans under medium levels of traffic compared with high levels of traffic, $F(1, 11) = 5.91$, $p = 0.03$, $\eta_\varrho^2 = 0.35$. In addition, they checked more flight plans in the HPRs condition compared with the jet routes condition, $F(1, 11) = 26.54$, $p < 0.001$, $\eta_\varrho^2 = 0.71$. There was no interaction between traffic level and condition, $F(1, 11) = 0.04$, $p = 0.85$, $\eta_\varrho^2 = 0.004$.

The main finding is that participants used Flight Plan Readouts more often in the HPRs condition than in the jet routes condition. This is most likely because information about whether or not an aircraft was OPD equipped, which was critical for determining whether an aircraft was filed on the appropriate HPR route, was accessible only through the Flight Plan Readout. OPD equipage did not matter to participants in the jet routes condition, so Flight Plan Readouts were not used as often.

3.1.8 Altitude, Speed, Heading, and Route Amendment Commands

As a measure of controller efficiency, we measured the number of altitude, speed, heading, and route amendment commands issued during each scenario. For these analyses, we removed all Philadelphia aircraft. Figures 27, 28, 29, and 30 display the means, by traffic level and condition, for the number of altitude, speed, heading, and route amendment commands.

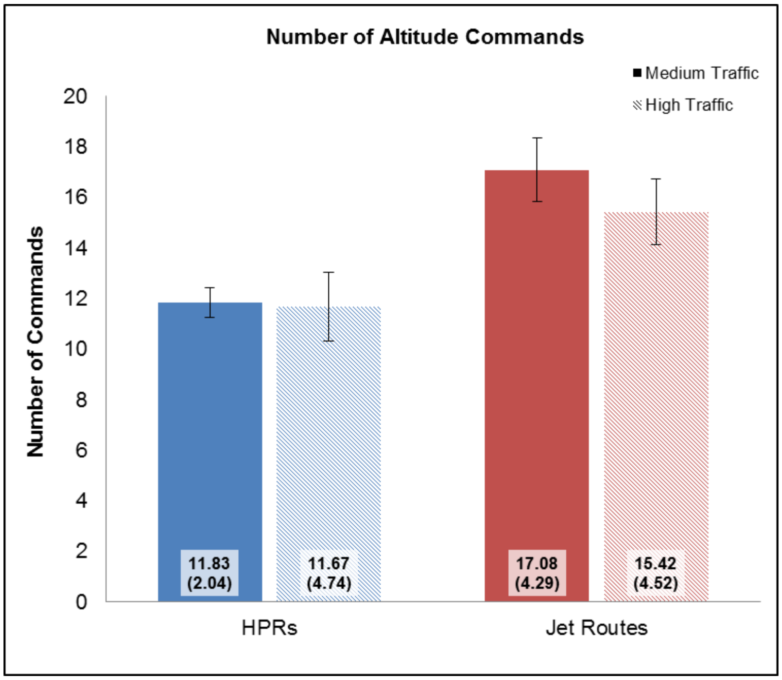


Figure 27. Number of altitude commands by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

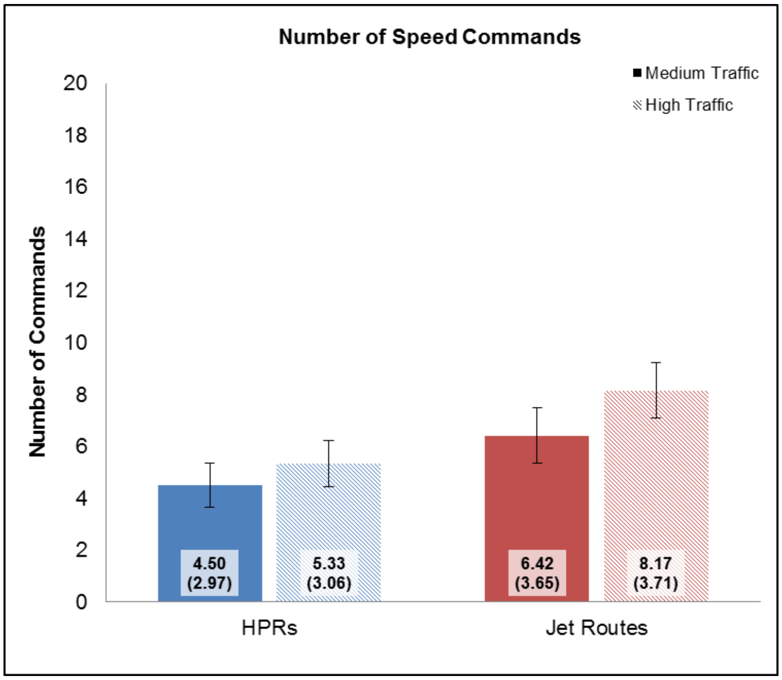


Figure 28. Number of speed commands by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

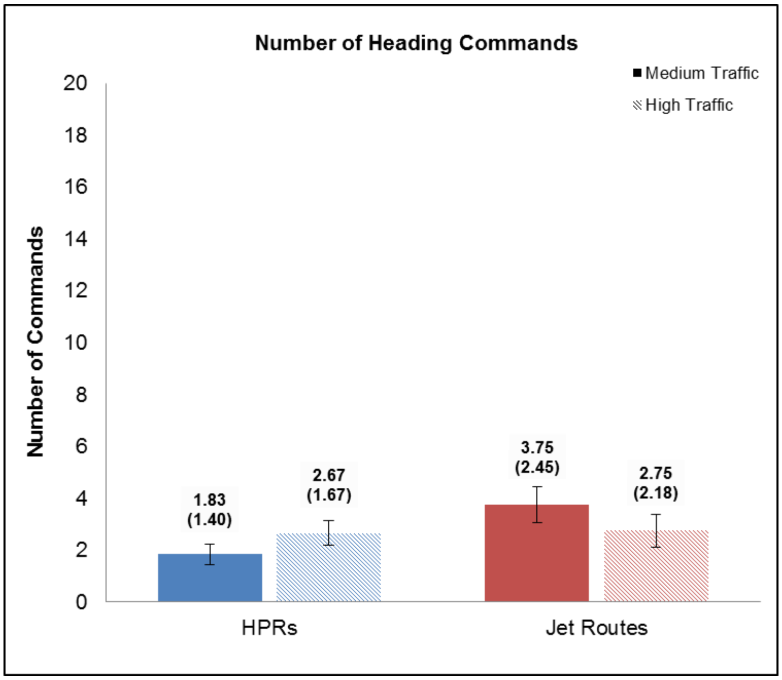


Figure 29. Number of heading commands by traffic level and condition. Means (*SDs*) are presented above the error bars.

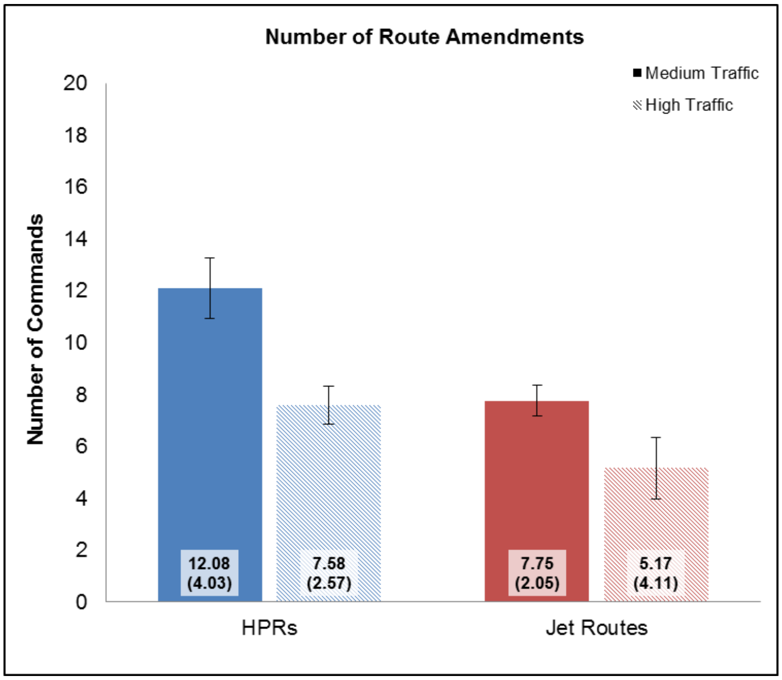


Figure 30. Number of route amendments by traffic level and condition. Means (*SDs*) are presented at bottom of figure.

There were no effects of traffic level on the number of altitude commands, $F(1, 11) = 0.56$, $p = 0.47$, $\eta_\rho^2 = 0.05$, or heading commands, $F(1, 11) = 0.02$, $p = 0.88$, $\eta_\rho^2 = 0.002$. Under high-traffic levels, participants issued marginally more speed commands, $F(1, 11) = 3.63$, $p = 0.08$, $\eta_\rho^2 = 0.25$, and issued significantly more route commands, $F(1, 11) = 29.98$, $p < 0.001$, $\eta_\rho^2 = 0.73$, compared with medium-traffic levels.

Independent of traffic level, in the jet routes condition, participants issued significantly more altitude commands, $F(1, 11) = 11.62$, $p = 0.006$, $\eta_\rho^2 = 0.51$, and speed commands, $F(1, 11) = 16.16$, $p = 0.002$, $\eta_\rho^2 = 0.60$, and marginally more heading commands, $F(1, 11) = 4.48$, $p = 0.06$, $\eta_\rho^2 = 0.29$, compared with the HPRs condition. In the HPRs condition, participants issued significantly more route amendments, $F(1, 11) = 26.35$, $p < 0.001$, $\eta_\rho^2 = 0.71$, compared to the jet routes condition.

There were no interactions between traffic level and condition for the number of altitude commands, $F(1, 11) = 1.12$, $p = 0.31$, $\eta_\rho^2 = 0.09$, speed commands, $F(1, 11) = 0.58$, $p = 0.46$, $\eta_\rho^2 = 0.05$, or route amendments, $F(1, 11) = 1.47$, $p = 0.251$, $\eta_\rho^2 = 0.12$. There was a statistically significant interaction between traffic level and condition for the number of heading commands, $F(1, 11) = 6.37$, $p = 0.03$, $\eta_\rho^2 = 0.37$. Under medium-traffic levels, participants issued more heading commands in the jet routes condition compared with the HPRs condition, $t(11) = 3.03$, $p = 0.01$, $d = 0.87$; however, there were no differences between conditions under high-traffic levels, $t(11) = 0.15$, $p = 0.88$, $d = 0.04$.

These data show that using HPRs minimizes the need for altitude, speed, and heading commands. However, participants issued more route amendments in the HPRs condition. The use of route amendments could potentially be minimized as well if aircraft are filed for the appropriate HPR route. While most aircraft were filed for the correct HPR route in Experiment 1, there were some that were not.

3.1.9 CIT Interactions

In Experiment 1, the CIT was limited in functionality. Participants were able to use six different route buttons, one for each of the major traffic streams (ORD, EWR, PHL, LGA, TEB, and JFK). When the participant pressed a route button (e.g., PHL), the pair of HPRs associated with that traffic stream would become illuminated and the fixes along the HPRs would be labeled (e.g., the PHL button would highlight the Flyr1 and Flyr2 as well as the fixes along each route). Seven of the 12 participants used the CIT during the scenarios. The other participants most likely relied on macros and the Draw function to remember information about and use the HPR lanes and fixes.

Figures 31 and 32 show the number and total duration of CIT interactions for all Experiment 1 HPR scenarios by traffic level. The total duration of CIT interactions was calculated for each participant by summing the duration of each CIT interaction. All participants are included in the calculations of the means and standard deviations (where any participants who did not use the tool had a “0” for the number and duration of CIT interactions).

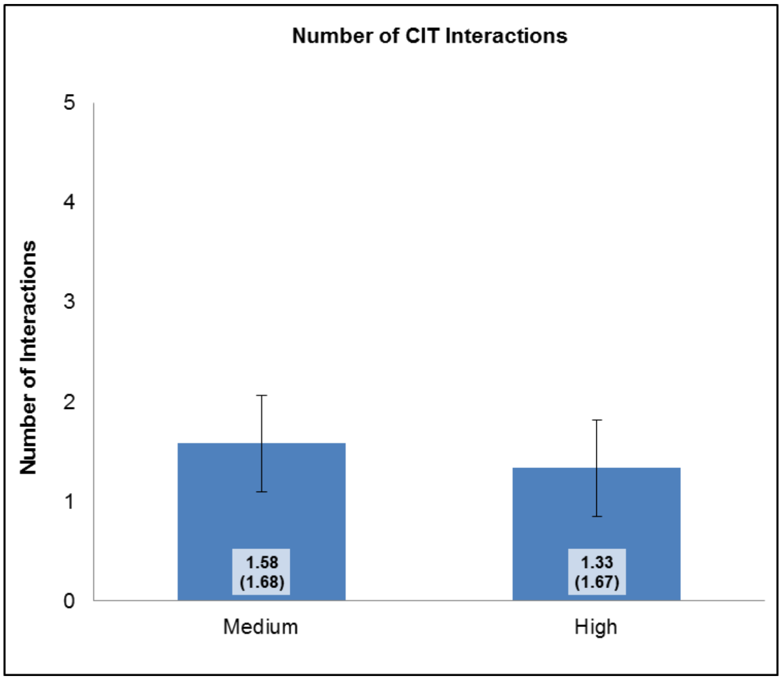


Figure 31. Number of CIT interactions by traffic level. Means (*SD*s) are presented at bottom of figure.

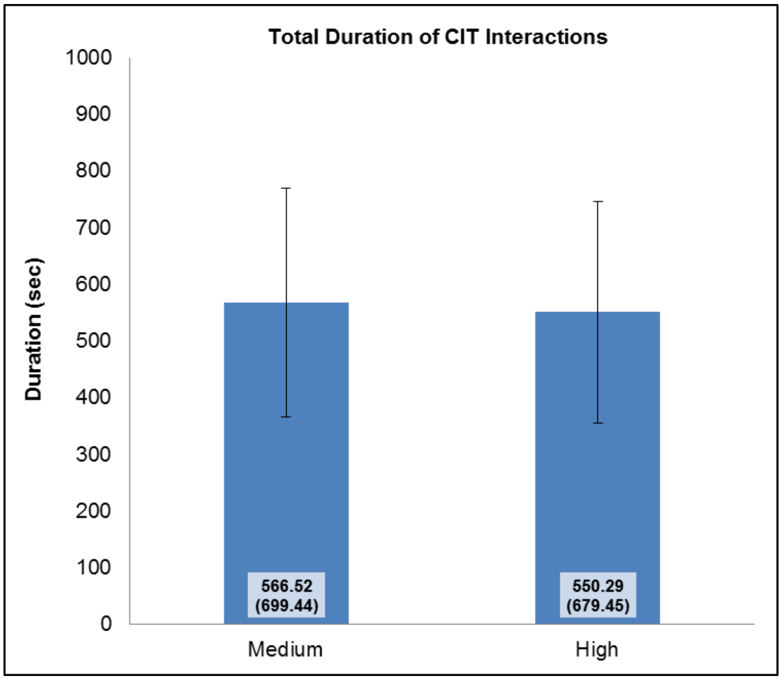


Figure 32. Total duration of CIT interactions (in sec) by traffic level. Means (*SD*s) are presented at bottom of figure.

3.1.10 Air Traffic Workload Input Technique

Participants were prompted to give workload ratings every 2 minutes across each 45-minute scenario, totaling 22 ratings per scenario. R-side participants failed to respond to 22% of the prompts in HPR scenarios and 11% of the prompts in jet routes scenarios. D-side participants failed to respond to 11% of the prompts in HPR scenarios and 5% of the prompts in jet routes scenarios. Figures 33 and 34 show means of R-side and D-side controllers' workload ratings, by traffic level and condition.

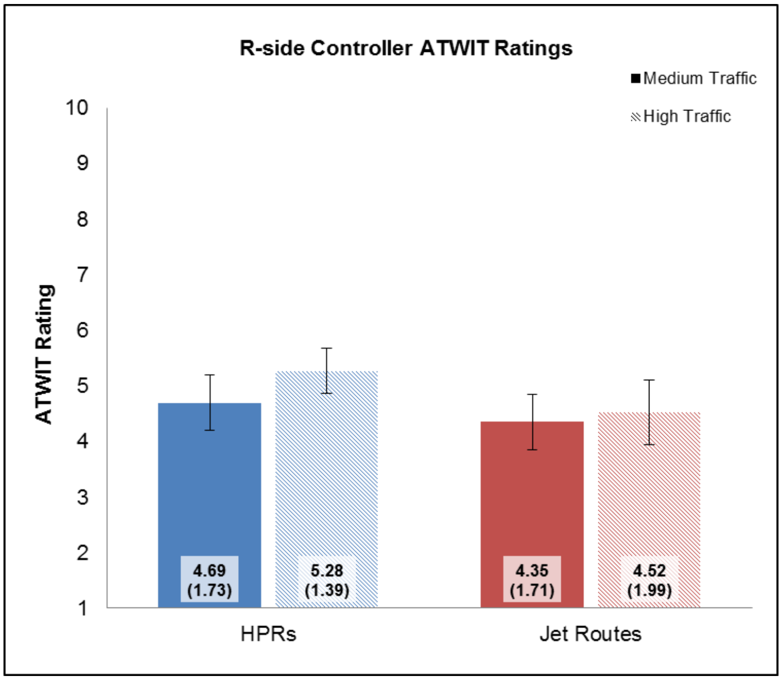


Figure 33. R-side controller workload ratings by traffic level, and condition. Means (*SD*s) are presented at bottom of figure.

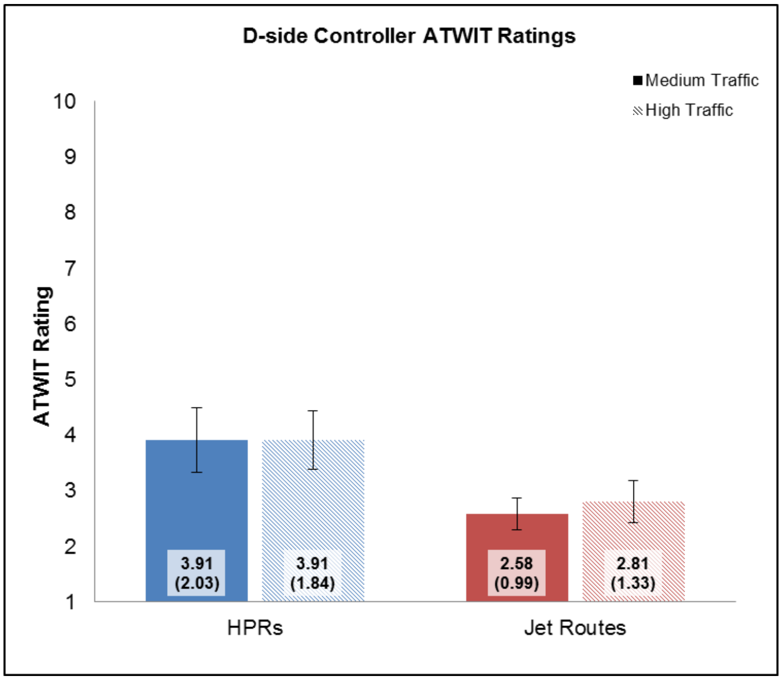


Figure 34. D-side controller workload ratings by traffic level, and condition. Means (*SD*s) are presented at bottom of figure.

As expected, when participants were on the R-side, workload ratings were significantly higher when traffic was high compared with when it was at a medium level, $F(1, 11) = 6.09$, $p = 0.03$, $\eta_{\varrho}^{2} = 0.36$. There was no effect of traffic level when participants were on the D-side, $F(1, 11) = 0.52$, $p = 0.49$, $\eta_{\varrho}^{2} = 0.05$. On both the R-side and D-side, participants' workload ratings were significantly higher when they controlled traffic using HPRs compared with jet routes, R-side: $F(1, 11) = 8.18$, $p = 0.02$, $\eta_{\varrho}^{2} = 0.43$; D-side: $F(1, 11) = 7.80$, $p = 0.02$, $\eta_{\varrho}^{2} = 0.42$. There were no interactions between traffic level and condition when participants were on the R-side, $F(1, 11) = 1.45$, $p = 0.25$, $\eta_{\varrho}^{2} = 0.12$, or the D-side, $F(1, 11) = 0.39$, $p = 0.54$, $\eta_{\varrho}^{2} = 0.03$.

The main finding is that workload ratings were higher when participants controlled traffic using HPRs compared to jet routes. There are a number of plausible explanations for higher workload. Here we describe three of the most plausible. One, participants were not completely comfortable with the HPR structure and needed additional training. Two, the HPRs concept combined with the large sector and large number of unfamiliar LOAs led to higher workload. Three, the HPRs concept independent of other factors led to higher workload. We suspect that the difference observed here is due to a combination of the first and second explanations. Participants most likely needed more training, and that combined with the large number of complexities introduced in Experiment 1 led to higher workload with HPRs.

3.1.11 Post-Scenario Questionnaire Responses

Table F1 (see Appendix F) displays the means and standard deviations for each condition as well as the *p*-values resulting from statistical analyses comparing HPRs to jet routes for each of the PSQ questions. Statistically significant effects are noted with an asterisk (marginal effects are noted with a ✝).

Participants gave lower performance ratings, higher workload ratings, lower situation awareness ratings, and higher difficulty ratings in the HPRs condition compared with the jet routes condition. R-side participants estimated that they could have handled more aircraft in the jet routes condition compared with the HPRs condition (refer to Table F1 to see which specific questions showed these differences). As discussed previously, it is possible that the amount of complexity introduced in Experiment 1 combined with the need for additional training on the HPRs procedures led participants to experience lower levels of performance and situation awareness and higher levels of workload and difficulty in the HPRs condition.

3.1.12 Over-the-Shoulder Observer Responses

Tables F2 and F3 (see Appendix F) display the means and standard deviations for each condition as well as the *p*-values resulting from statistical analyses comparing HPRs to jet routes for each of the PSQ questions and OTS evaluation items. Statistically significant effects are noted with an asterisk (marginal effects are noted with a †).

OTS observers gave higher workload ratings, lower performance ratings (for only one question: performance for identifying aircraft conflicts), lower estimates for the number of additional aircraft a participant could have controlled, and higher difficulty ratings in the HPRs condition compared with the jet routes condition. However, they also gave more positive ratings to HPRs in regards to their effectiveness for separating aircraft safely, moving aircraft efficiently, and number of aircraft participants could control. The inconsistencies suggest that OTS observers recognized certain benefits of HPRs but observed that participants had difficulties using the procedures in the scenarios (possibly due to the reasons we outlined above).

3.1.13 Exit Questionnaire Responses

Participants and OTS observers filled out the Exit Questionnaires at the completion of the simulation. Participants' responses to the Exit Questionnaire questions relevant to Experiment 1 are reported here. Figures 35, 36, 37, and 38 show the frequency of each type of response (HPRs, jet routes, No Difference) for each Experiment 1 related question For each question, a chi-square test of goodness-of-fit test was conducted to determine if the frequencies of each response type were different from each other. If there was a significant difference between response types, follow-up, Bonferroni-adjusted chi-square tests for goodness-of-fit were conducted to determine which responses were significantly different from each other.

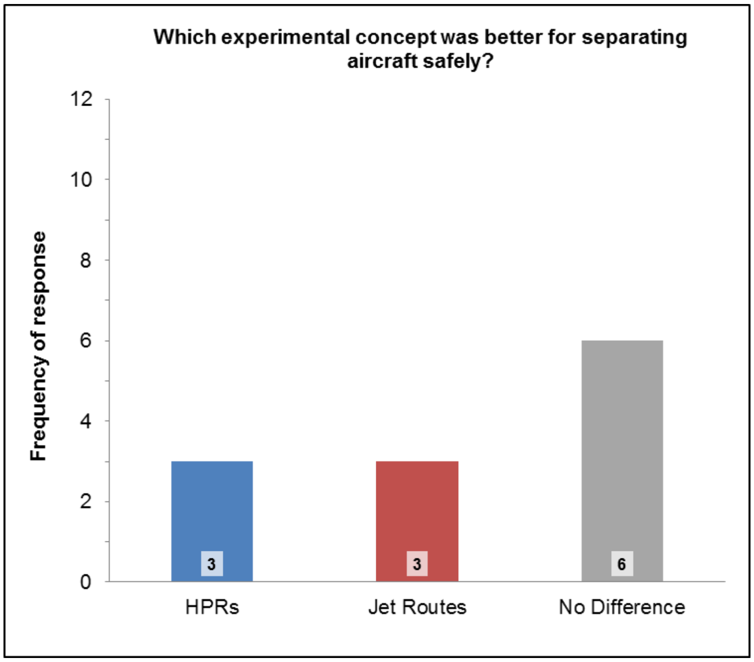


Figure 35. Frequency of participant responses to the question, "Which experimental concept was better for separating aircraft safely?" Frequencies are presented at bottom of figure.

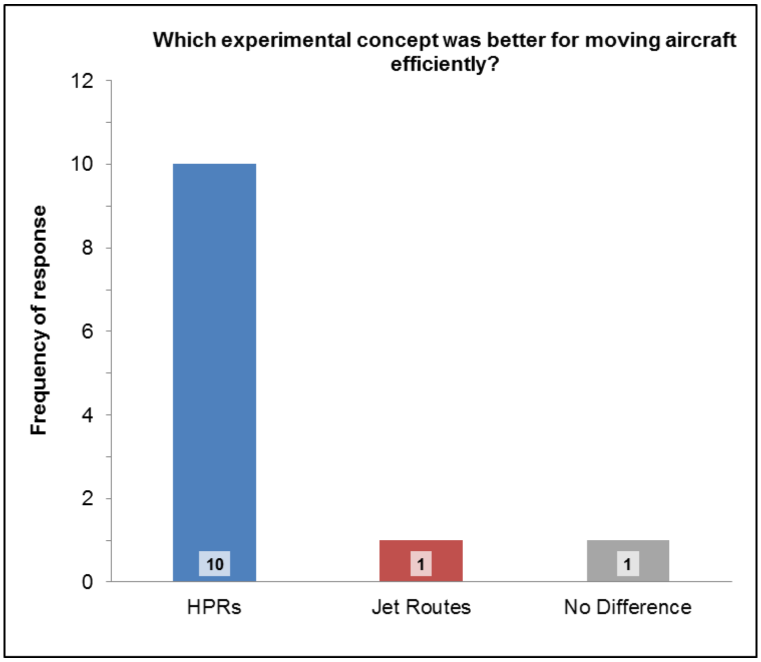


Figure 36. Frequency of participant responses to the question, "Which experimental concept was better for moving aircraft efficiently?" Frequencies are presented at bottom of figure.

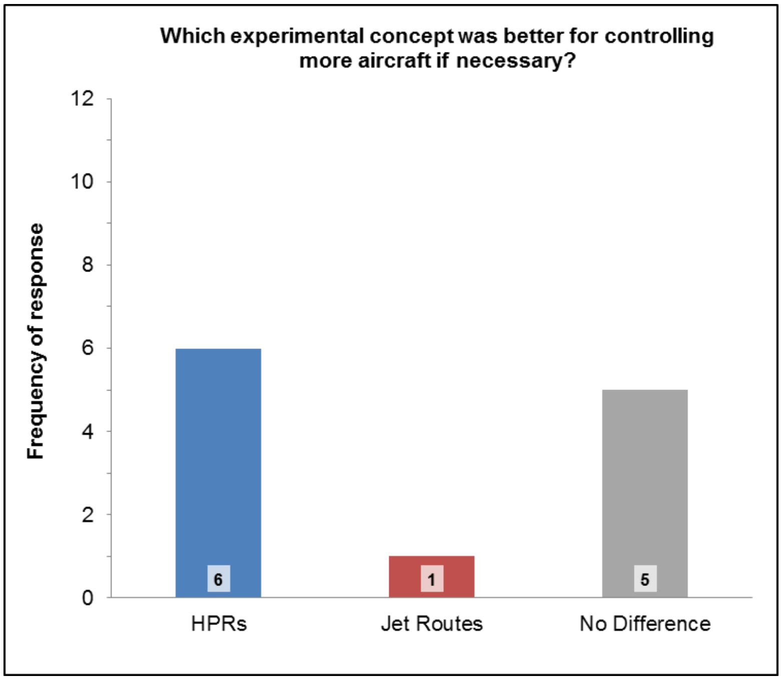


Figure 37. Frequency of participant responses to the question, "Which experimental concept was better for controlling more aircraft?" Frequencies are presented at bottom of figure.

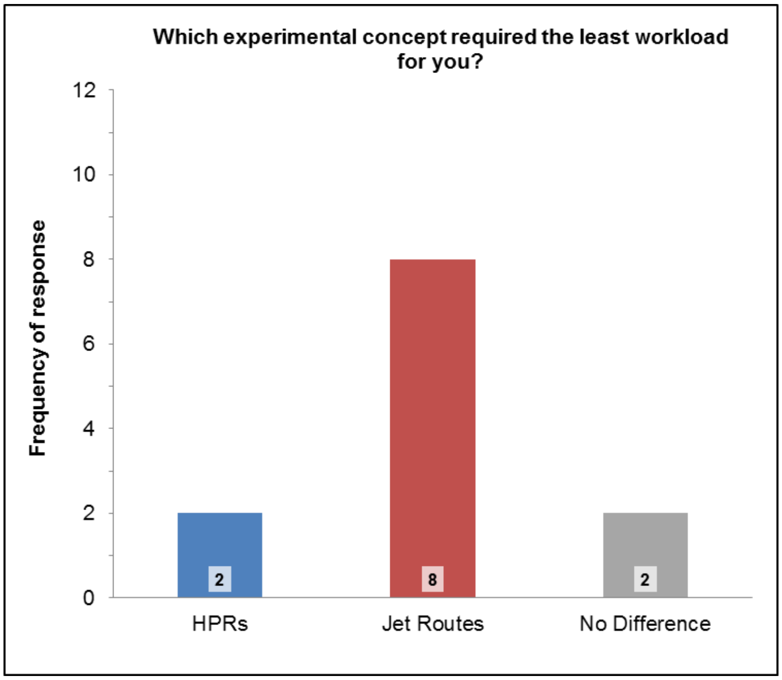


Figure 38. Frequency of participant responses to the question, "Which experimental concept required the least workload for you?" Frequencies are presented at bottom of figure.

There were no differences between responses for the question asking about which concept was better for separating aircraft safely, $\chi^2(2, N = 12) = 1.50, p = 0.47$, or for the question asking about which concept was better for controlling more aircraft, $\chi^2(2, N = 12) = 3.50, p = 0.17$ (though the trend in the data for this question does suggest that of participants who preferred one concept to the other, most preferred HPRs). There were differences between responses for the question asking about which concept was better for moving aircraft efficiently, $\chi^2(2, N = 12) = 13.50, p = 0.001$. Follow-up tests indicated that more participants preferred HPRs compared with jet routes, $\chi^2(2, N = 12) = 7.36, p = 0.007$, and No Difference, $\chi^2(2, N = 12) = 7.36, p = 0.007$. There was also a marginal difference between responses for the question asking about which concept required the least workload, $\chi^2(2, N = 12) = 6.00, p = 0.050$. The trend in the data suggests that more participants preferred the jet routes response compared with HPRs and No Difference.

3.2 Experiment 2

In Experiment 2, participants controlled traffic independently. As in Experiment 1, system metrics were analyzed for the time interval between 2 minutes and 44 minutes. The medium-traffic scenario interval was defined as the time between 2 minutes and 23 minutes. The high-traffic scenario interval was defined as the time between 23 minutes and 44 minutes. Unless otherwise noted, data were analyzed using a 4 (condition) x 2 (traffic level) repeated measures ANOVA. If a statistically significant interaction was shown, follow-up Bonferroni-adjusted paired *t*-tests were run to interpret the interaction. If a statistically significant main effect of condition was shown without a significant interaction, then the data were collapsed across traffic levels and follow-up, Bonferroni-adjusted paired *t*-tests were run to determine which pairs of conditions differed from each other.

3.2.1 Voice Communications

Figures 39 and 40 display the means, by traffic level and condition, for the number and duration of controller-initiated PTT transmissions. Figures 41 and 42 display the means, by traffic level and condition, for the number and duration of pilot-initiated PTT transmissions.

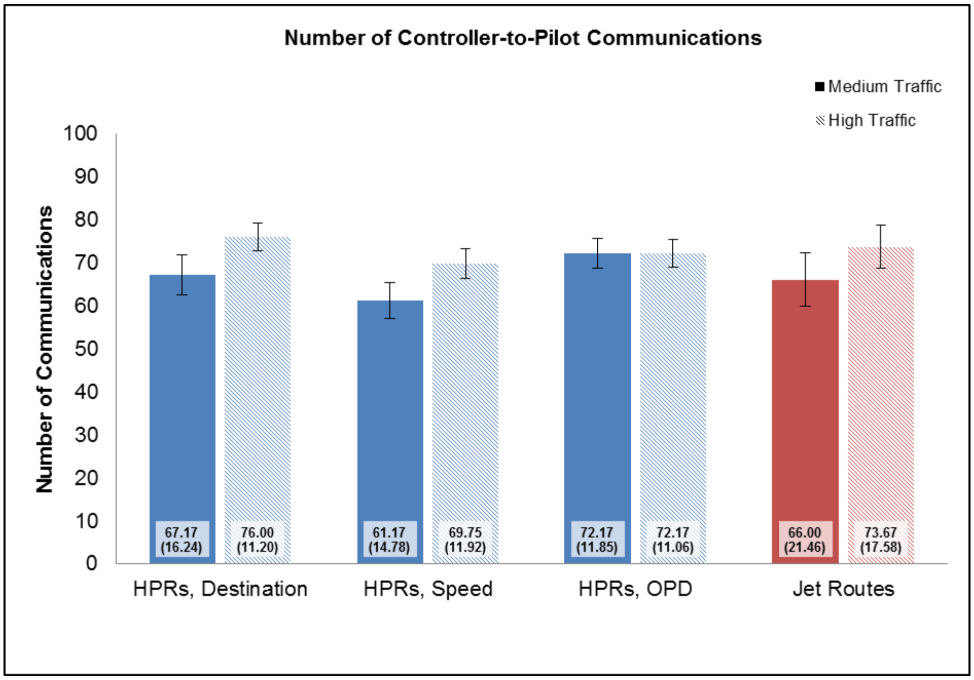


Figure 39. Number of controller-initiated transmissions by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

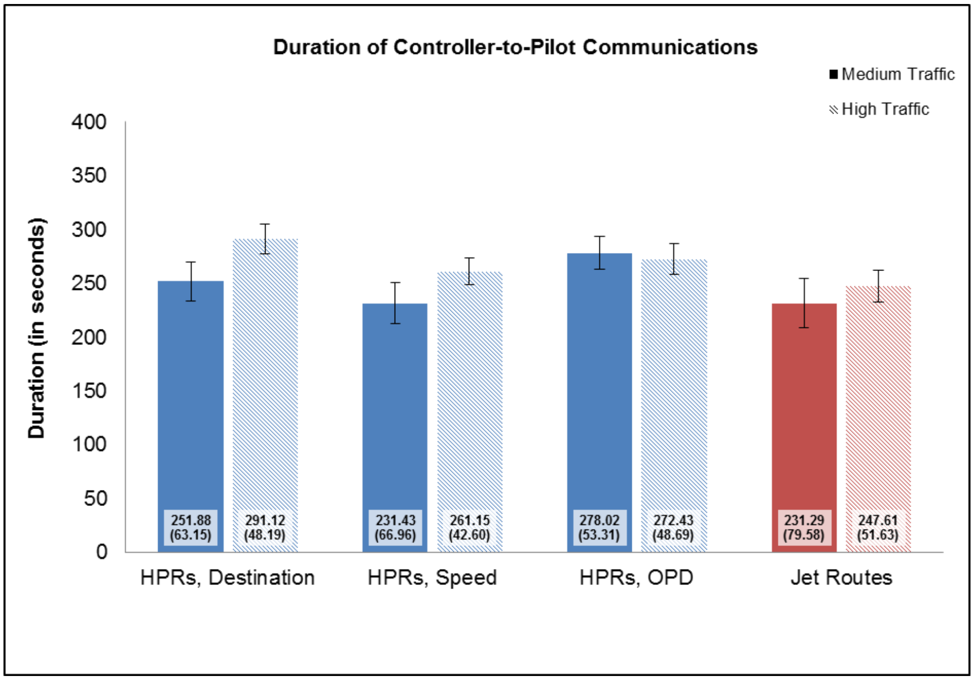


Figure 40. Total duration of controller-initiated transmissions by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

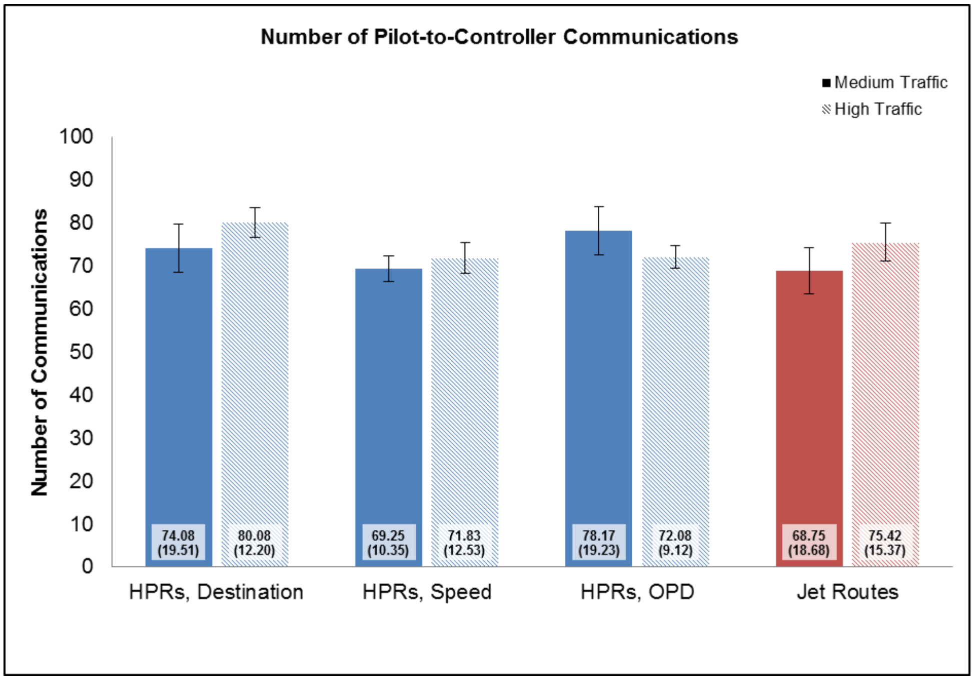


Figure 41. Number of pilot-initiated transmissions by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

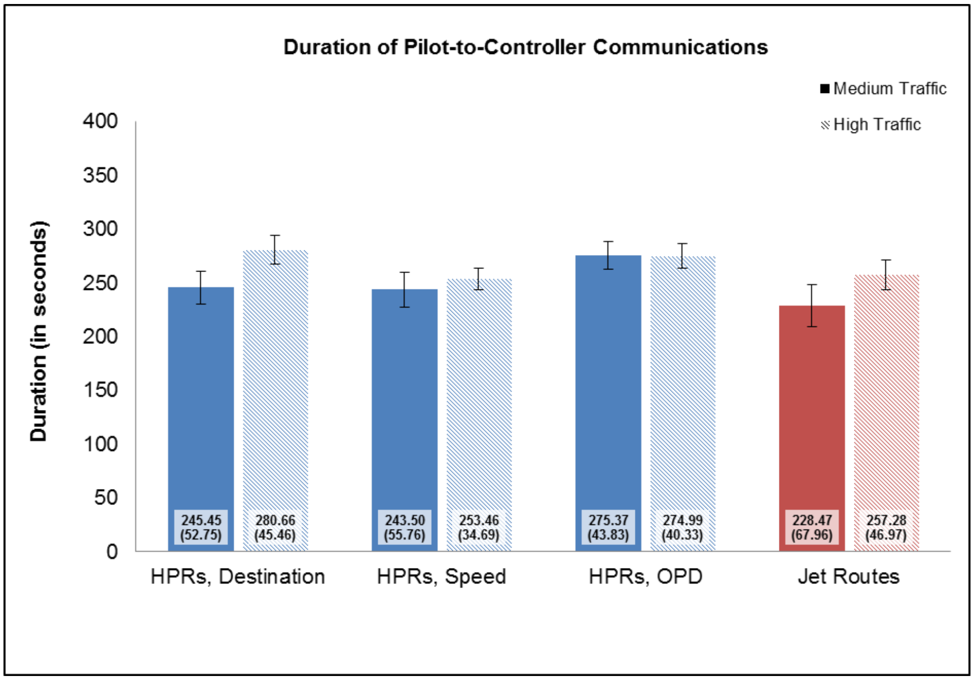


Figure 42. Total duration of pilot-initiated transmissions by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

Participants made marginally more PTT transmissions under high levels of traffic compared with medium levels of traffic, $F(1, 11) = 4.04$, $p = 0.07$, $\eta_{\varrho}^{2} = 0.27$. In addition, the total duration of controller-initiated transmissions was marginally longer under high levels of traffic, $F(1, 11) =$

4.03, $p = 0.07$, $\eta_\varrho^2 = 0.27$. There were no effects of traffic level on the number of pilot-initiated transmissions, $F(1, 11) = 0.54$, $p = 0.48$, $\eta_\varrho^2 = 0.05$, or the total duration of pilot-initiated transmissions, $F(1, 11) = 2.81$, $p = 0.12$, $\eta_\varrho^2 = 0.20$.

There were no significant differences between conditions for the number of controller-initiated transmissions, $F(1.70, 18.73) = 1.35$, $p = 0.27$, $\eta_\varrho^2 = 0.11$, or pilot-initiated transmissions, $F(3, 33) = 1.36$, $p = 0.27$, $\eta_\varrho^2 = 0.11$.[1] There were significant main effects of condition on the total duration of controller-initiated transmissions, $F(1.72, 18.94) = 4.69$, $p = 0.03$, $\eta_\varrho^2 = 0.30$,[2] as well as pilot-initiated transmissions, $F(3, 33) = 3.47$, $p = 0.03$, $\eta_\varrho^2 = 0.24$. Follow-up, Bonferroni-adjusted paired *t*-tests for showed that the duration of transmissions was longer in the HPRs by OPD equipment condition compared with the HPRs by speed condition for both controller-initiated transmissions, $t(11) = 4.19$, $p = 0.002$, $d = 1.21$, and pilot-initiated transmissions, $t(11) = 3.36$, $p = 0.006$, $d = 0.97$. There were no other statistically significant pairwise comparisons.

There were no interactions between traffic level and condition for the number of controller-initiated transmissions, $F(3, 33) = 0.86$, $p = 0.47$, $\eta_\varrho^2 = 0.07$, the number of pilot-initiated transmissions, $F(3, 33) = 1.51$, $p = 0.23$, $\eta_\varrho^2 = 0.12$, the total duration of controller-initiated transmissions, $F(3, 33) = 1.13$, $p = 0.35$, $\eta_\varrho^2 = 0.09$, or the total duration of pilot initiated transmissions, $F(3, 33) = 0.97$, $p = 0.41$, $\eta_\varrho^2 = 0.08$.

In summary, when participants used HPRs to sort aircraft by OPD equipment, the total duration of the controller- and pilot-initiated transmissions was longer than it was when participants sorted aircraft by speed. Given that there were no other statistically significant differences between conditions, it is difficult to make any additional concrete claims about the benefits of one condition over another.

3.2.2 Losses of Separation

Losses of separation were defined as incidents where aircraft were separated by less than 5 nmi laterally and 1,000 ft. vertically. As in Experiment 1, ATC SME reviewed all potential losses of separation and categorized the incidents as either occurring due to system or simulation pilot error or due to controller error. Across all experimental scenarios in Experiment 2, one incident was attributed to controller error. A participant cleared an aircraft to descend and maintain FL310, and neglected to observe another aircraft at FL310 approximately 3 nmi in front of the descending aircraft. The participant attempted to stop the descending aircraft at FL320 but was unsuccessful. This loss of separation occurred in the jet routes condition in the 37th minute of the scenario when the traffic level was high and lasted for 1 minute and 48 seconds.

3.2.3 JEDI/URET Conflict Probe Notifications

As in Experiment 1, we examined the number and duration of conflict probe notifications as well as the number and duration for each type of notification (red or yellow). Figures 43, 44, 45, 46, 47, and 48 display the number and duration of conflict probe notifications overall, with red notifications, and with yellow notifications, respectively.

---

[1] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 18.63$, $p = 0.002$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.57$).

[2] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 11.93$, $p = 0.04$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.57$).

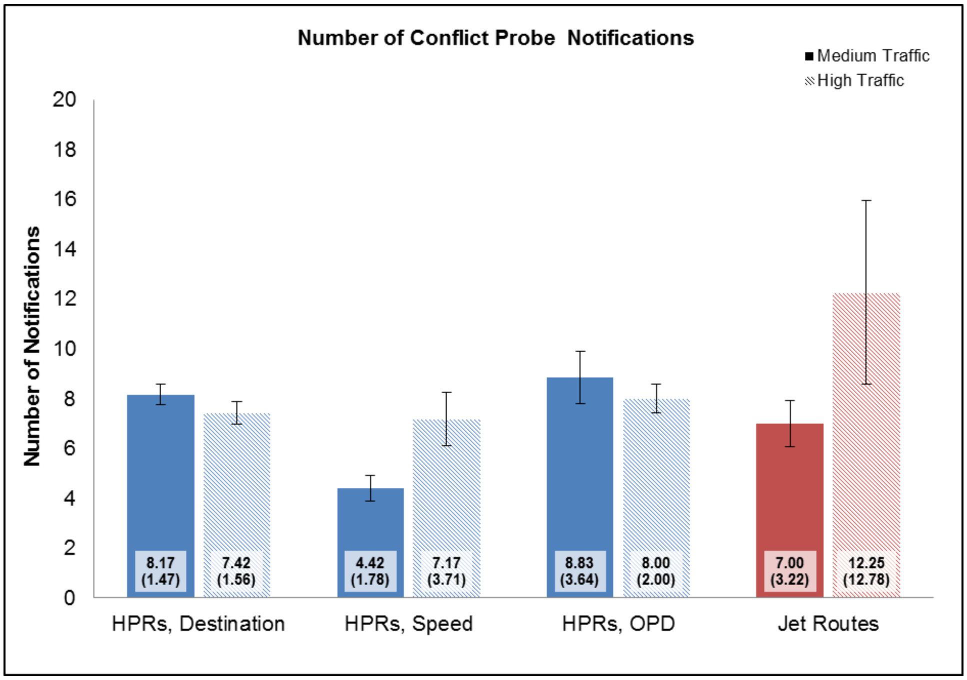


Figure 43. Number of conflict probe notifications by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

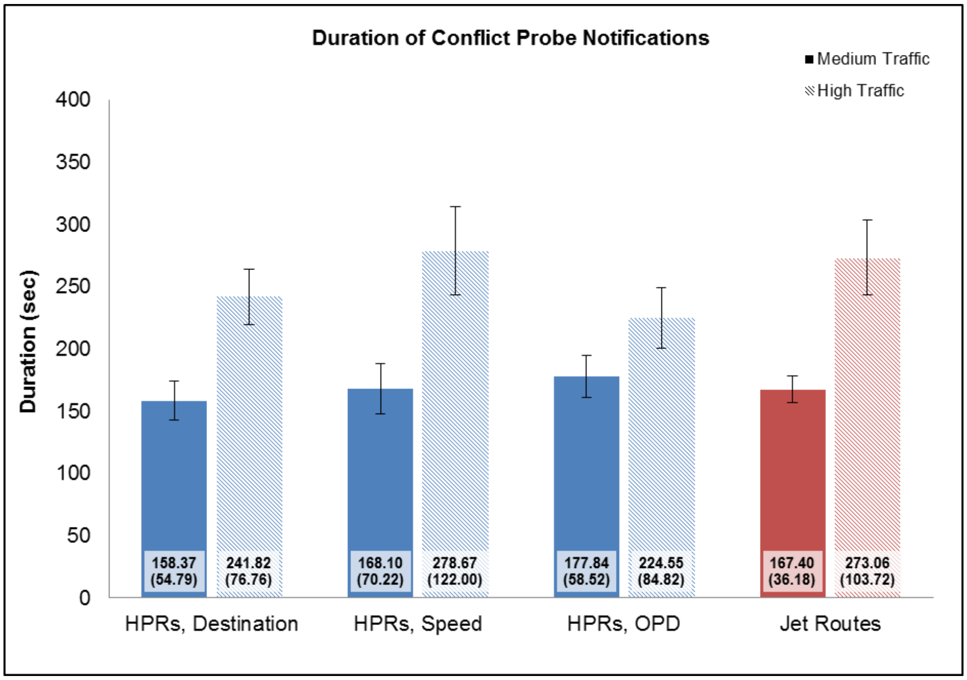


Figure 44. Duration of conflict probe notifications by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

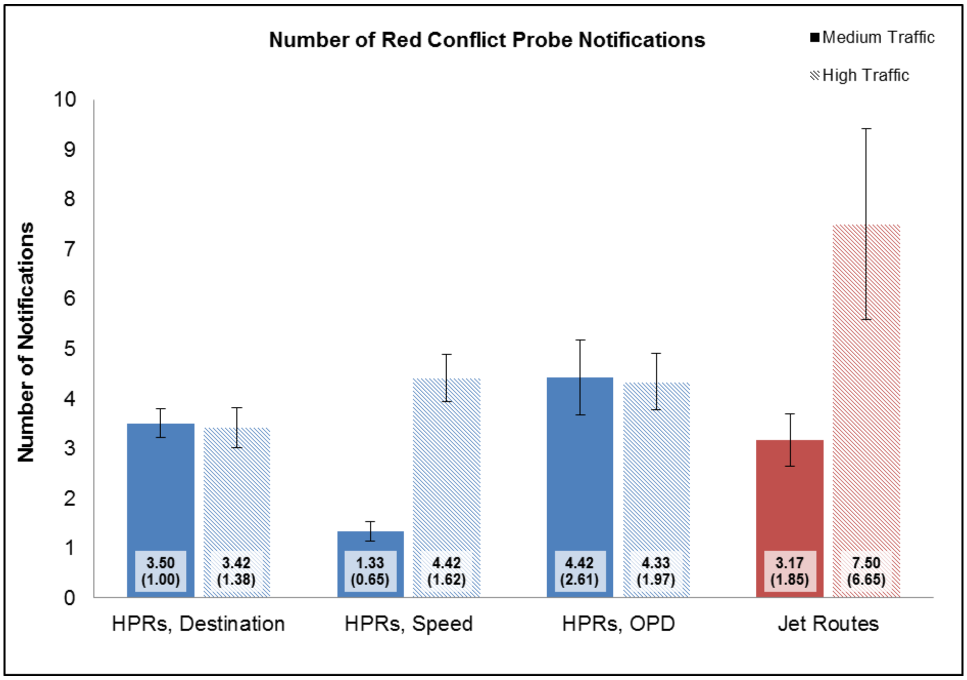


Figure 45. Number of red conflict probe notifications by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

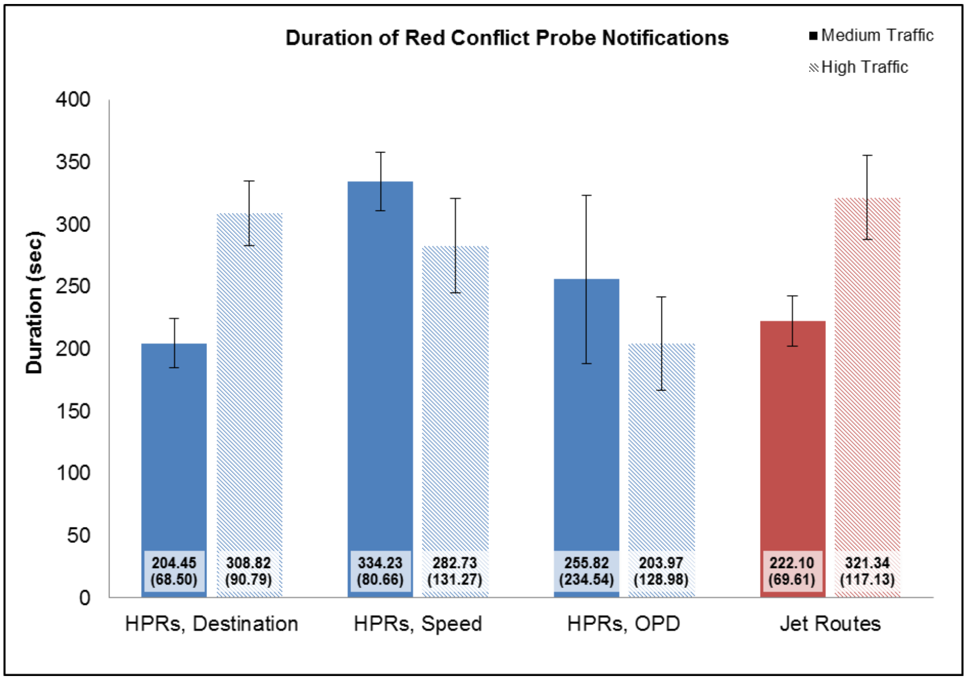


Figure 46. Duration of red conflict probe notifications by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

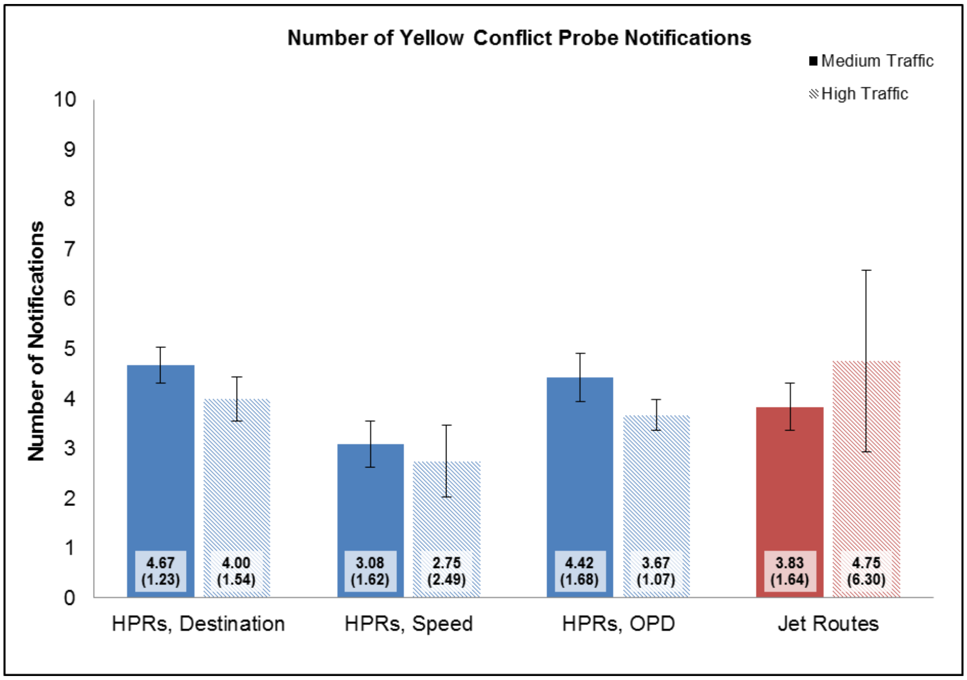


Figure 47. Number of yellow conflict probe notifications by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

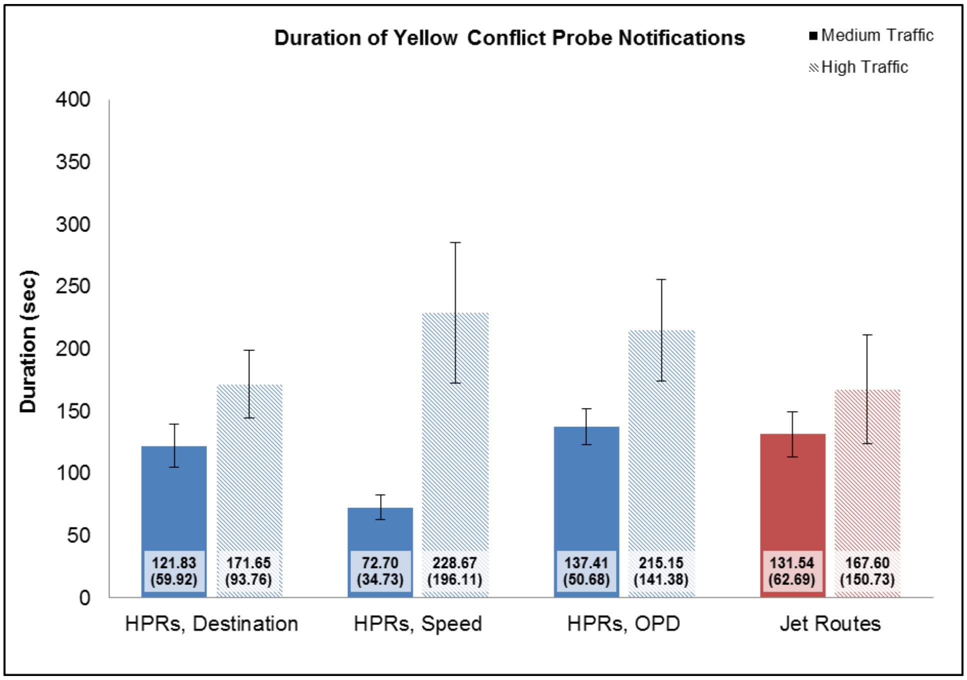


Figure 48. Duration of yellow conflict probe notifications by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

There were marginally more conflict probe notifications (collapsed across color) under high-traffic levels compared with medium-traffic levels, $F(1, 11) = 3.58$, $p = 0.085$, $\eta_\rho^2 = 0.25$. In addition, the duration of the notifications was longer under high-traffic levels compared with medium-traffic levels, $F(1, 11) = 50.43$, $p < 0.001$, $\eta_\rho^2 = 0.82$. There were no differences between conditions for the number of notifications,[3] $F(1.20, 13.25) = 1.99$, $p = 0.18$, $\eta_\rho^2 = 0.15$, or duration of notifications, $F(3, 33) = 0.52$, $p = 0.67$, $\eta_\rho^2 = 0.05$. Finally, there were no interactions between traffic level and condition for the number of notifications,[4] $F(1.49, 16.33) = 2.42$, $p = 0.13$, $\eta_\rho^2 = 0.18$, or duration of notifications, $F(3, 33) = 0.81$, $p = 0.50$, $\eta_\rho^2 = 0.07$.

There was no difference between traffic levels for the number of yellow notifications, $F(1, 11) = 0.18$, $p = 0.68$, $\eta_\rho^2 = 0.02$, but the duration of the yellow notifications was longer under high-traffic levels, $F(1, 11) = 20.88$, $p = 0.001$, $\eta_\rho^2 = 0.66$. There were more red notifications under high-traffic levels compared with medium-traffic levels, $F(1, 11) = 16.75$, $p = 0.002$, $\eta_\rho^2 = 0.60$; however, the duration of the notifications was similar across traffic levels, $F(1, 11) = 0.89$, $p = 0.37$, $\eta_\rho^2 = 0.08$.

There were no differences between conditions for the number of yellow notifications,[5] $F(1.468, 16.15) = 1.20$, $p = 0.32$, $\eta_\rho^2 = 0.10$, or the duration of yellow notifications, $F(3, 33) = 0.28$, $p = 0.84$, $\eta_\rho^2 = 0.02$. However, there was a difference between conditions for the number of red notifications,[6] $F(1.403, 15.43) = 3.08$, $p = 0.04$, $\eta_\rho^2 = 0.22$, but not the duration of red notifications, $F(3, 33) = 1.63$, $p = 0.20$, $\eta_\rho^2 = 0.13$.

There were no interactions between traffic level and condition for the number of yellow notifications,[7] $F(1.618, 17.799) = 0.54$, $p = 0.66$, $\eta_\rho^2 = 0.05$, or the duration of yellow notifications, $F(3, 33) = 1.46$, $p = 0.24$, $\eta_\rho^2 = 0.12$. There were interactions between traffic level and condition for the number of red notifications,[8] $F(1.464, 16.108) = 5.59$, $p = 0.003$, $\eta_\rho^2 = 0.34$, and the duration of red notifications, $F(3, 33) = 3.00$, $p = 0.04$, $\eta_\rho^2 = 0.21$. Follow-up, Bonferroni-adjusted paired t-tests showed that under medium-traffic levels, there were fewer red notifications in the HPRs by speed condition compared with the HPRs by destination condition, $t(11) = 5.92$, $p < 0.001$, $d = 1.71$, and the HPRs by OPD equipment condition, $t(11) = 3.89$, $p = 0.003$, $d = 1.12$. There were no statistically significant pairwise comparisons for the duration of red notifications.

There were few differences between conditions for the number and duration of conflict probe notifications. The main finding was that the HPRs by speed condition yielded the fewest red notifications compared with the other HPR lane usage conditions (and the nonsignificant trend is in the same direction when comparing HPRs by speed to the jet routes condition). This finding suggests that the HPRs by speed condition led to safer separation; however, it is important to point

---

[3] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 30.44$, $p < 0.001$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.40$).

[4] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 20.22$, $p = 0.001$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.50$).

[5] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 16.35$, $p = 0.006$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.49$).

[6] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 28.75$, $p < 0.001$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.47$).

[7] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 17.09$, $p = 0.005$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.54$).

[8] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 19.40$, $p = 0.002$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.49$).

out the alternative possibility that this difference between conditions could be a reflection of separation strategies rather than safety.

3.2.4 Number of Aircraft Under Control

Figure 49 displays the means, by traffic level and condition, for the number of aircraft under control.

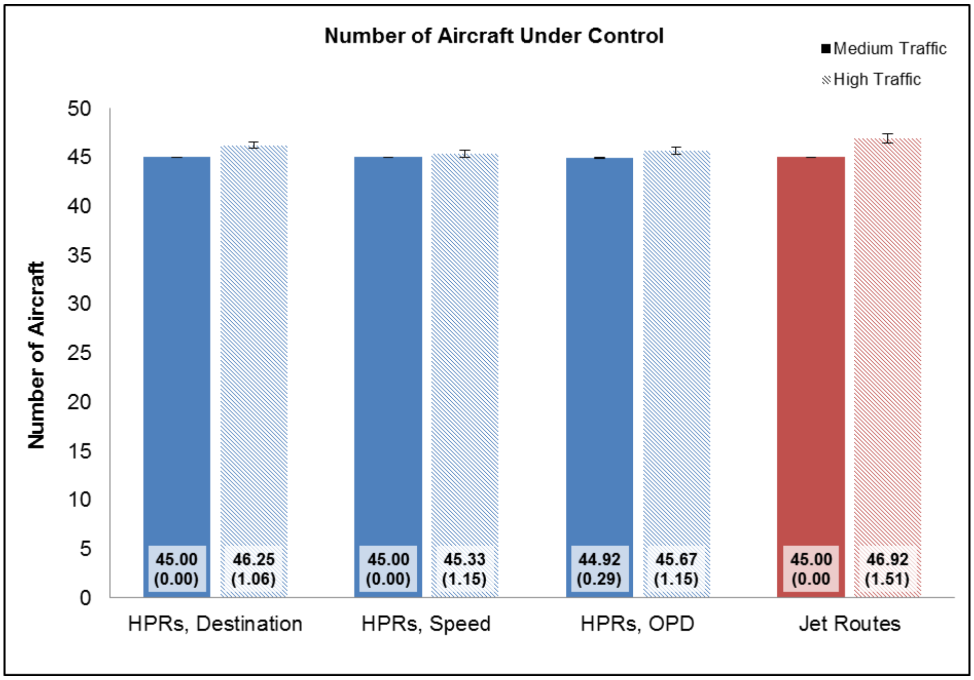


Figure 49. Number of aircraft under control by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

There were more aircraft under control under high-traffic levels compared with medium-traffic levels, $F(1, 11) = 19.50$, $p = 0.001$, $\eta_{\varrho}^{2} = 0.64$, as expected given that we intentionally increased traffic halfway through the scenarios. There was a main effect of condition, $F(3, 33) = 5.89$, $p = 0.002$, $\eta_{\varrho}^{2} = 0.35$. Finally, there was an interaction between traffic level and condition, $F(3, 33) = 4.90$, $p = 0.006$, $\eta_{\varrho}^{2} = 0.31$. When traffic was at a medium level, participants controlled equivalent numbers of aircraft, but when traffic was at a high level, participants controlled significantly more aircraft in the jet routes condition compared with the HPRs by speed condition, $t(11) = 3.98$, $p = 0.002$, $d = 1.15$. There were no other statistically significant differences between conditions. Given the close relationship between number of aircraft under control and the number of aircraft we built into the scenario, it is not surprising that we saw few differences between conditions here.

3.2.5 Number of Handoffs Accepted and Handoffs Initiated

Figures 50 and 51 display the means, by traffic level and condition, for the number of handoffs accepted and initiated.

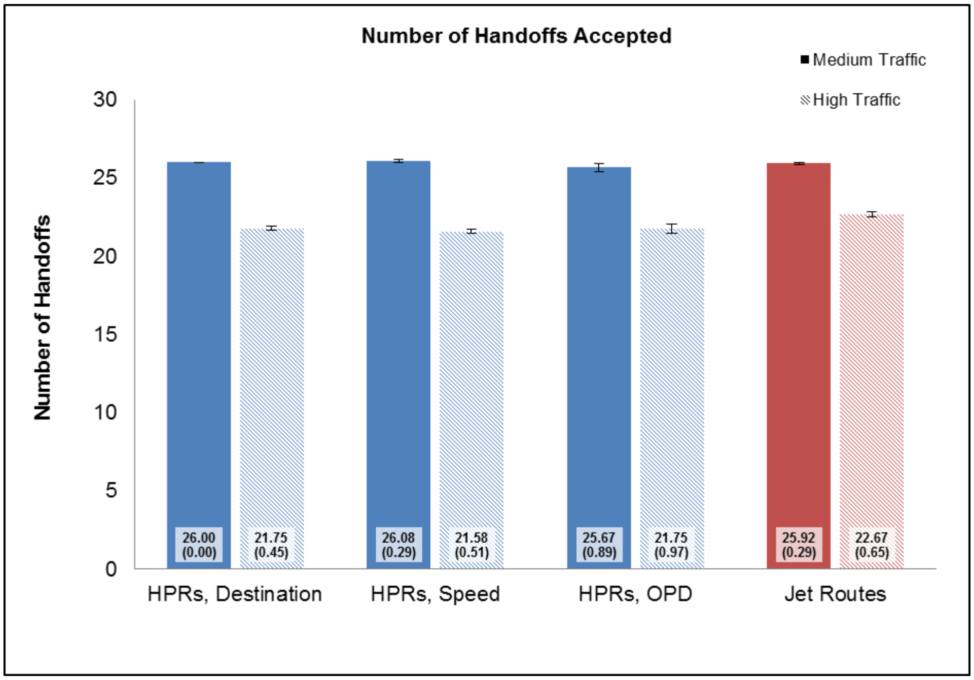


Figure 50. Number of handoffs accepted by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

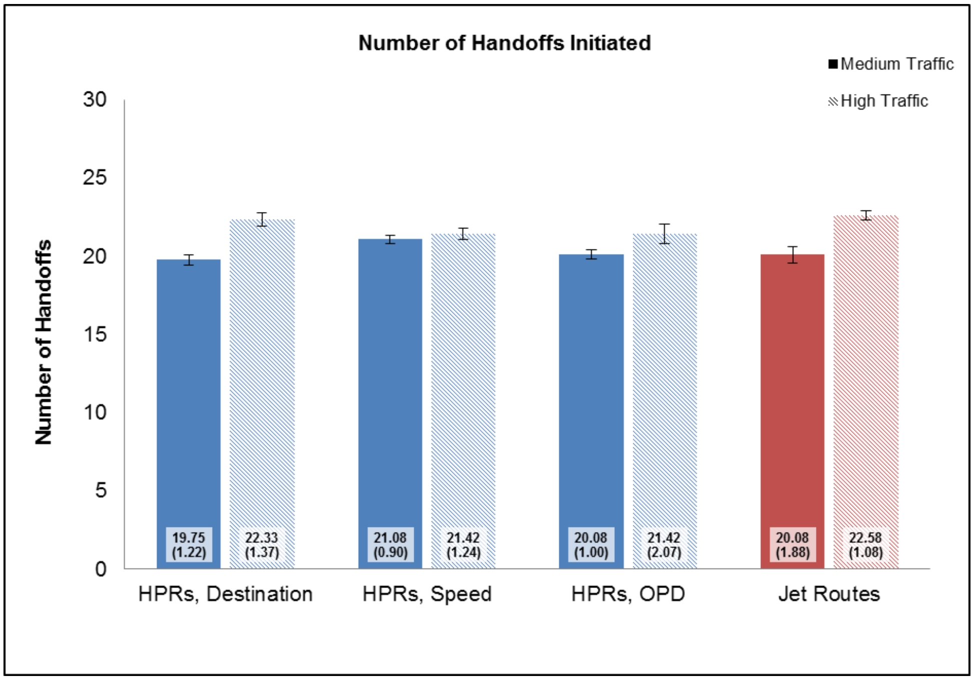


Figure 51. Number of handoffs initiated by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

Under medium-traffic levels, participants accepted more handoffs, $F(1, 11) = 1242.39$, $p < 0.001$, $\eta_{\rho}^{2} = 0.90$, and initiated fewer handoffs, $F(1, 11) = 15.82$, $p = 0.002$, $\eta_{\rho}^{2} = 0.59$, compared with high-traffic levels.

There was a main effect of condition on the number of handoffs accepted, $F(3, 33) = 11.25$, $p < 0.001$, $\eta_{\rho}^{2} = 0.51$, but no interaction between traffic level and condition, $F(1.33, 14.57) = 2.93$, $p = 0.10$, $\eta_{\rho}^{2} = 0.47$.[9] Follow-up Bonferroni-adjusted paired *t*-tests showed that more handoffs were accepted in the jet routes condition compared with the HPRs by destination condition, $t(11) = 4.02$, $p = 0.002$, $d = 1.16$, the HPRs by speed condition, $t(11) = 4.01$, $p = 0.002$, $d = 1.16$, and the HPRs by OPD equipment condition, $t(11) = 5.63$, $p < 0.001$, $d = 1.63$.

There was no main effect of condition on the number of handoffs initiated, $F(3, 33) = 1.16$, $p = 0.33$, $\eta_{\rho}^{2} = 0.10$; however, there was an interaction between traffic level and condition, $F(3, 33) = 3.52$, $p = 0.03$, $\eta_{\rho}^{2} = 0.24$. Follow-up, Bonferroni-adjusted paired *t*-tests showed no statistically significant differences between conditions.

Participants accepted more handoffs when using jet routes compared to all of the HPR lane usage conditions, though the numerical differences between the conditions' means is relatively small in relation to the total number of aircraft participants controlled (in all of these comparisons, the average difference between conditions was less than 1 aircraft).

3.2.6 Aircraft Time and Distance in Sector

Figures 52 and 53 display the means, by traffic level and condition, for aircraft time and distance in the participant's sector.

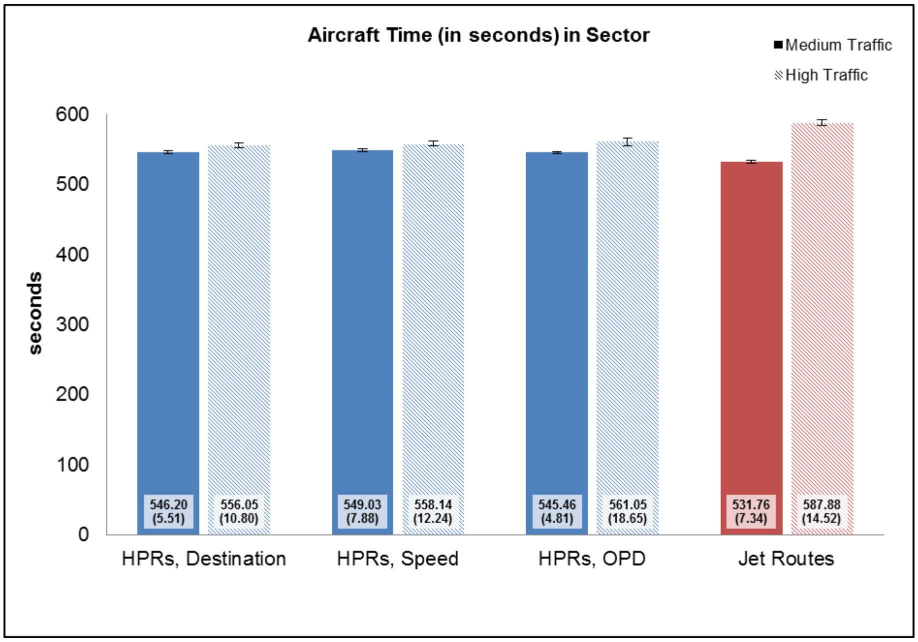


Figure 52. Aircraft time (in sec) in sector by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

---

[9] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 23.21$, $p < 0.01$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.44$).

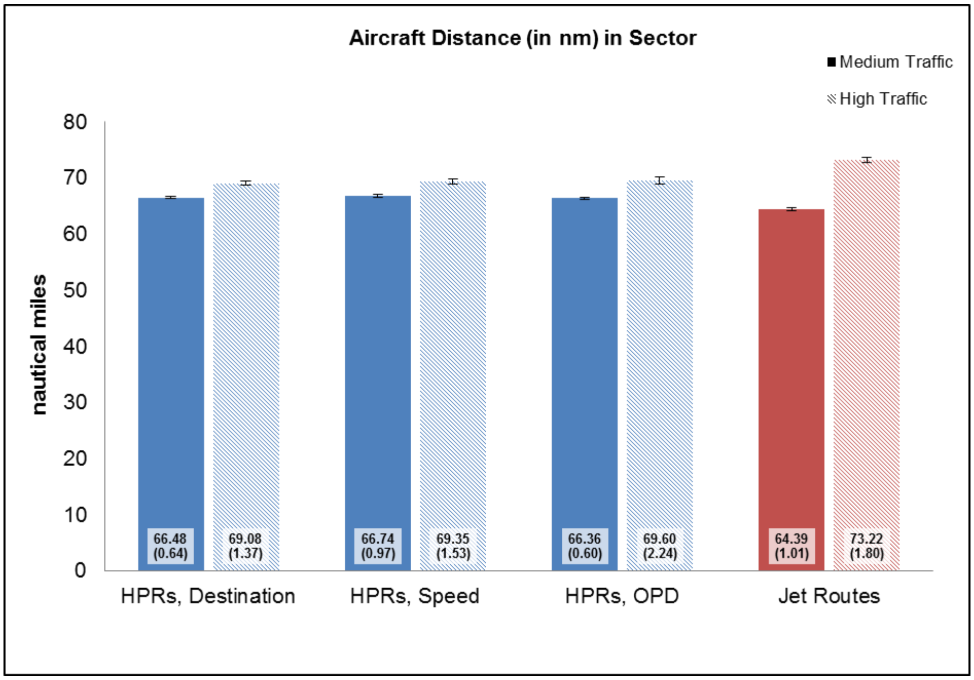


Figure 53. Aircraft distance (in nmi) in sector by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

Under high-traffic levels, aircraft flew longer amounts of time, $F(1, 11) = 103.50$, $p < 0.001$, $\eta_{\varrho}^{2} = 0.90$, and distance, $F(1, 11) = 258.65$, $p < 0.001$, $\eta_{\varrho}^{2} = 0.96$, compared with medium-traffic levels. There were also main effects of condition on aircraft time, $F(3, 33) = 4.29$, $p = 0.01$, $\eta_{\varrho}^{2} = 0.28$, and distance, $F(3, 33) = 4.16$, $p = 0.01$, $\eta_{\varrho}^{2} = 0.27$. Finally, there were interactions between traffic level and condition on both aircraft time, $F(3, 33) = 32.00$, $p < 0.001$, $\eta_{\varrho}^{2} = 0.74$, and distance, $F(3, 33) = 36.99$, $p < 0.001$, $\eta_{\varrho}^{2} = 0.77$.

Follow-up, Bonferroni-adjusted paired *t*-tests showed that under medium-traffic levels, aircraft flew through the sector in shorter amounts of time and distance in the jet routes condition compared with the HPRs by destination condition, time: $t(11) = 7.36$, $p < 0.001$, $d = 2.13$, distance: $t(11) = 8.56$, $p < 0.001$, $d = 2.47$; HPRs by speed condition, time: $t(11) = 5.98$, $p < 0.001$, $d = 1.73$, distance: $t(11) = 6.29$, $p < 0.001$, $d = 1.82$; and HPRs by OPD equipment condition, time: $t(11) = 6.17$, $p < 0.001$, $d = 1.78$, distance: $t(11) = 7.35$, $p < 0.001$, $d = 2.12$. Under high-traffic levels, aircraft flew through the sector in longer amounts of time and distance in the jet routes condition compared with the HPRs by destination condition: time, $t(11) = 9.39$, $p < 0.001$, $d = 2.71$, and distance, $t(11) = 9.44$, $p < 0.001$, $d = 2.73$; HPRs by speed condition: time, $t(11) = 5.98$, $p < 0.001$, $d = 1.73$, and distance, $t(11) = 6.19$, $p < 0.001$, $d = 1.79$; and HPRs by OPD equipment condition: time, $t(11) = 5.06$, $p < 0.001$, $d = 1.46$, and distance, $t(11) = 5.54$, $p < 0.001$, $d = 1.60$.

In summary, all HPR lane usage conditions were more efficient than jet routes under high-traffic levels, but jet routes were more efficient under medium-traffic levels. These results demonstrate the potential for HPRs by a single sorting procedure to be beneficial when traffic volumes are high, most likely because it allows an aircraft to get around other aircraft that may be in its way without reducing speed, changing altitude, or flying a heading. However, when traffic volumes are lower, moving an aircraft to another lane adds additional time and distance to the route without providing a benefit (as it is less likely that other aircraft will be in that aircraft's way).

These results are somewhat inconsistent with what we found in Experiment 1, though there were a number of differences between the two studies that could explain the inconsistencies. Participants were controlling traffic in a smaller sector (due to weather), were using fewer LOAs, and were using HPRs for only one type of sorting procedure. In addition, participants were more familiar with HPRs procedures by the time they reached Experiment 2, so they may have used the procedures more optimally here.

3.2.7 Controller Interactions

As in Experiment 1, we analyzed controller interactions to determine if participants used Flight Plan Readouts more frequently in some conditions compared with others. Figure 54 displays the means, by traffic level and condition, for the number of Flight Plan Readout commands.

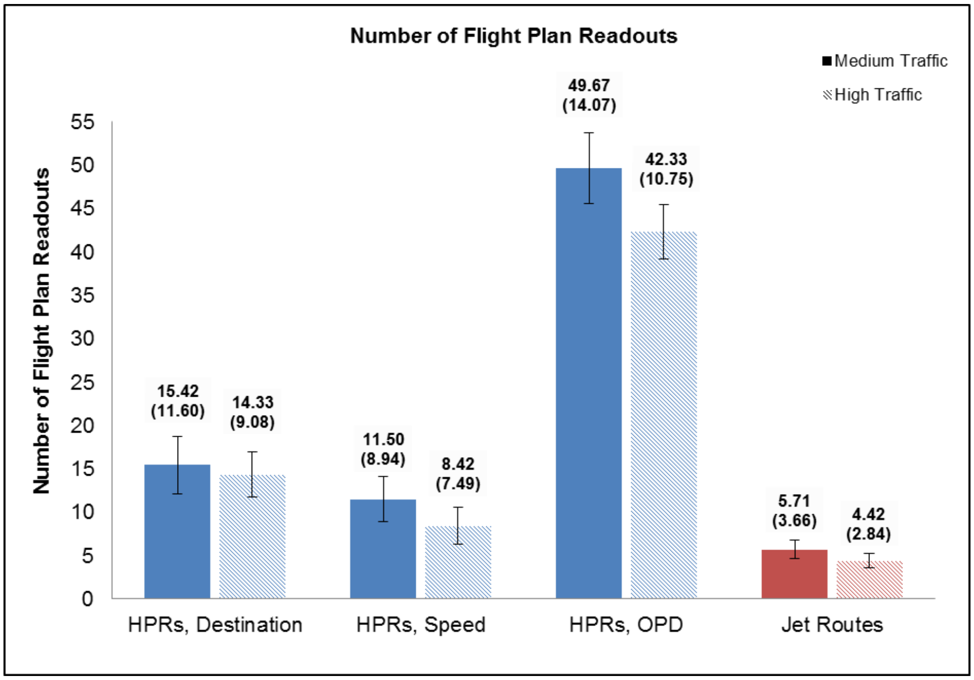


Figure 54. Number of flight plan readouts by traffic level and condition. Means (*SD*s) are presented above the error bars.

Participants checked marginally more flight plans under medium levels of traffic compared with high levels of traffic, $F(1, 11) = 4.45, p = 0.06, \eta_p^2 = 0.29$. In addition, there was a main effect of condition on number of Flight Plan Readouts,[10] $F(1.50, 16.50) = 96.34, p < 0.001, \eta_p^2 = 0.90$. There was no interaction between traffic level and condition,[11] $F(1.37, 15.12) = 1.11, p = 0.33, \eta_p^2 = 0.09$. Follow-up, Bonferroni adjusted paired *t*-tests showed that participants checked flight plans the most in the HPRs by OPD equipment condition versus HPRs by destination, $t(11) = 8.17, p < 0.001, d = 2.36$; versus HPRs by speed, $t(11) = 11.01, p < 0.001, d = 3.18$; versus jet routes, $t(11) = 14.78, p < 0.001, d = 4.27$, followed by the HPRs by destination condition versus HPRs by speed, $t(11) = 5.41, p < 0.001, d = 1.56$; versus jet routes, $t(11) = 4.18, p = 0.002, d = 1.21$. The HPRs by

---

[10] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 24.56, p < 0.001$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.50$).

[11] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 20.81, p = 0.001$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.46$).

speed and jet routes conditions were not significantly different from each other at the Bonferroni-adjusted significance level, though the trend in the data shows that participants checked flight plans more often in the HPRs by speed condition compared with the jet routes condition.

In summary, participants checked flight plans more often in the HPRs by OPD equipment condition than any other condition. They needed to check flight plans to find out if an aircraft was OPD equipped or not. In addition, participants checked flight plans slightly more often in the other HPRs conditions compared to the jet routes condition.

3.2.8 Altitude, Speed, Heading, and Route Amendment Commands

As a measure of controller efficiency, we analyzed the number of altitude, speed, heading, and route amendment commands issued during each scenario. Figures 55, 56, 57, and 58 display the means, by traffic level and condition, for the number of issued altitude, speed, heading, and route amendment commands.

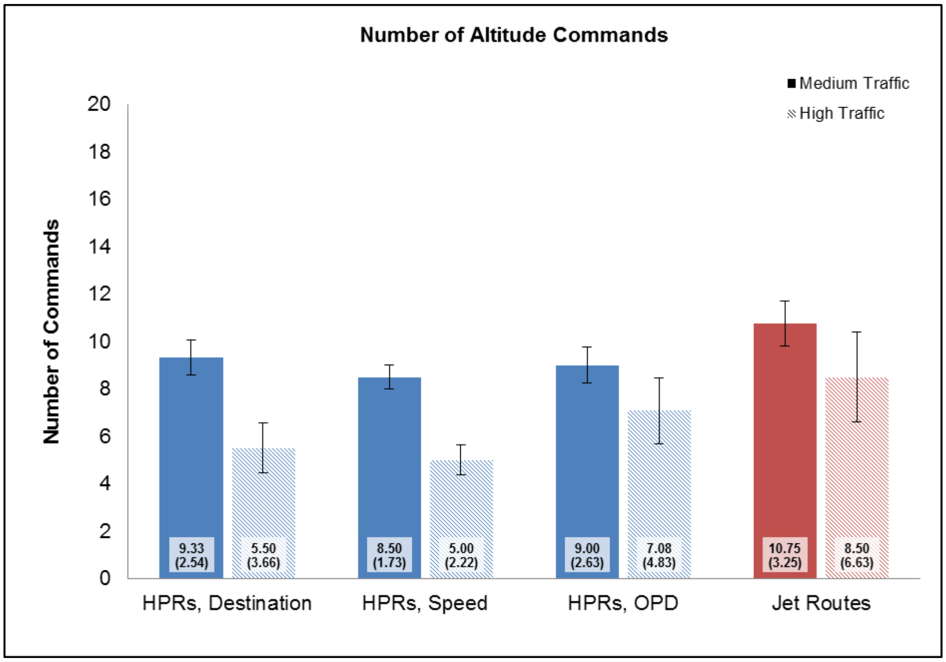


Figure 55. Number of altitude commands by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

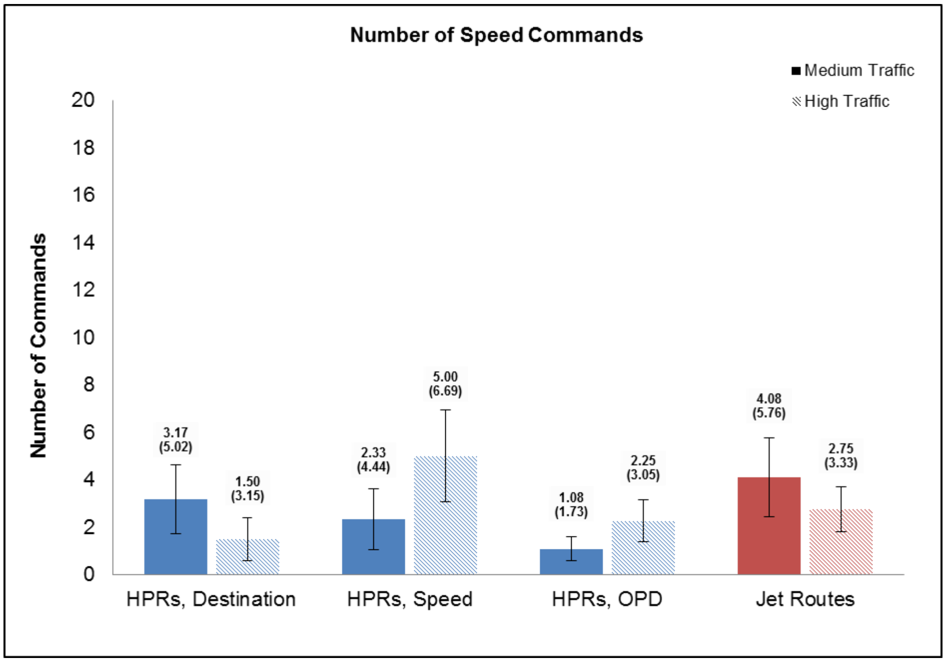


Figure 56. Number of speed commands by traffic level and condition. Means (*SD*s) are presented above the error bars.

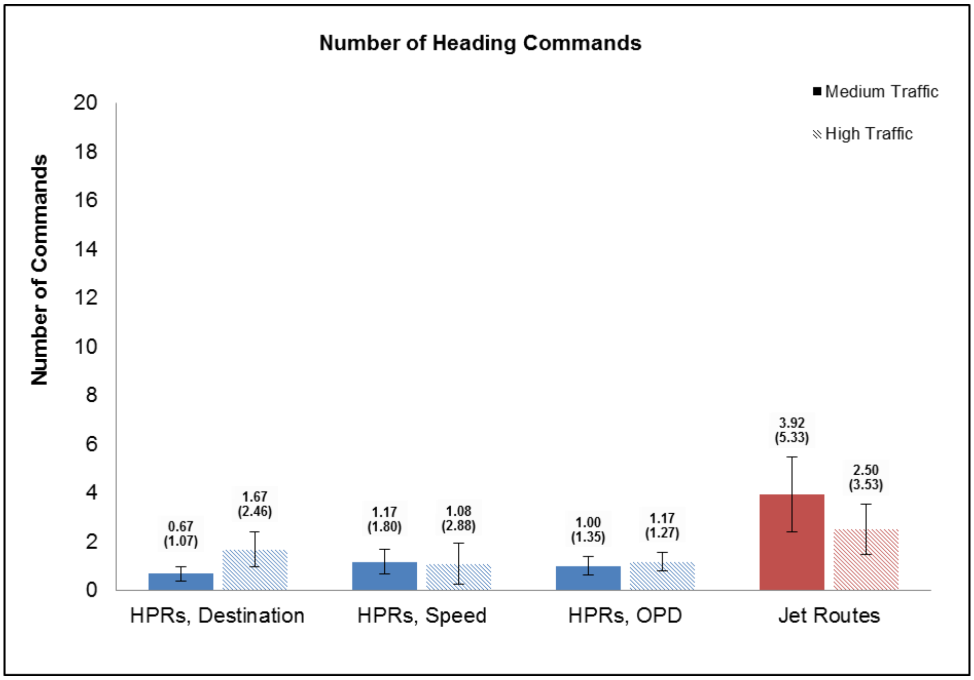


Figure 57. Number of heading commands by traffic level and condition. Means (*SD*s) are presented above the error bars.

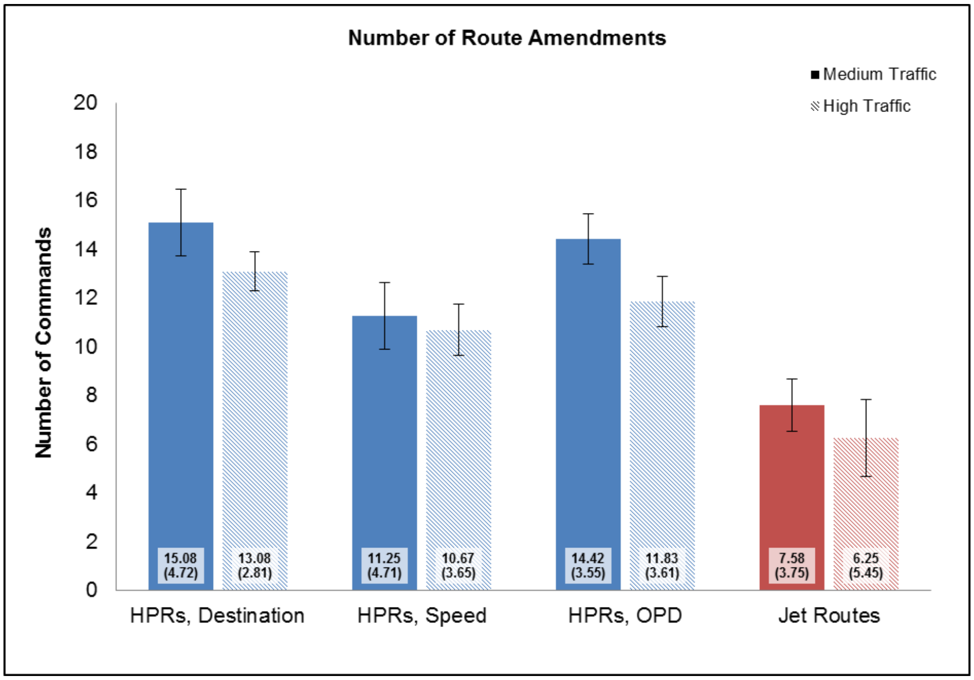


Figure 58. Number of route amendments by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

Under medium-traffic levels, participants issued more altitude commands, $F(1, 11) = 10.89$, $p = 0.007$, $\eta_p^2 = 0.50$, and marginally more route amendments, $F(1, 11) = 4.13$, $p = 0.07$, $\eta_p^2 = 0.27$, compared with high-traffic levels. There was no effect of traffic level on number of speed commands issued, $F(1, 11) = 0.21$, $p = 0.66$, $\eta_p^2 = 0.02$, or the number of heading commands issued, $F(1, 11) = 0.056$, $p = 0.82$, $\eta_p^2 = 0.005$.

There was a marginal main effect of condition on the number of altitude commands issued, $F(1.51, 16.57) = 2.99$, $p = 0.089$, $\eta_p^2 = 0.21$,[12] and no main effects of condition on the number of speed commands issued, $F(1.43, 15.74) = 1.07$, $p = 0.35$, $\eta_p^2 = 0.09$,[13] or the number of heading commands issued, $F(1.56, 17.12) = 2.08$, $p = 0.16$, $\eta_p^2 = 0.16$.[14] There was a statistically significant main effect of condition on the number of route amendments issued, $F(3, 33) = 28.04$, $p < 0.001$, $\eta_p^2 = 0.72$. Follow-up Bonferroni-adjusted paired *t*-tests showed that fewer route amendments were issued in the jet routes condition compared with each HPR condition, versus HPRs by destination: $t(11) = 6.68$, $p < 0.001$, $d = 1.93$; versus HPRs by speed: $t(11) = 4.63$, $p = 0.001$, $d = 1.34$; versus HPRs by OPD equipment: $t(11) = 8.01$, $p < 0.001$, $d = 2.31$. Among the HPR conditions, participants issued significantly fewer route amendments in the HPRs by speed condition compared with the HPRs by destination condition, $t(11) = 4.19$, $p = 0.002$, $d = 1.21$. No other differences were statistically significant.

---

[12] Mauchly's test showed that the assumption of sphericity had been violated, χ (5) = 17.495, $p = 0.004$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.50$).

[13] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 20.88$, $p = 0.001$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.48$).

[14] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 16.854$, $p = 0.005$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.52$).

There were no statistically significant interactions between traffic level and condition for the number of altitude commands issued, $F(3, 33) = 0.81$, $p = 0.50$, $\eta_{\varrho}^2 = 0.07$, or the number of route amendments issued, $F(3, 33) = 0.62$, $p = 0.60$, $\eta_{\varrho}^2 = 0.05$. There were marginal interactions between traffic level and condition for the number of speed commands issued, $F(1.73, 18.98) = 3.57$, $p = 0.054$, $\eta_{\varrho}^2 = 0.25$, as well as for the number of heading commands issued, $F(3, 33) = 2.84$, $p = 0.053$, $\eta_{\varrho}^2 = 0.21$.

Consistent with Experiment 1, fewer route amendments were issued in the jet routes condition, which is expected given that the HPRs conditions required participants to sort aircraft onto the different HPR lanes. If aircraft were filed for the correct HPRs, fewer route amendments would need to be issued. Unlike in Experiment 1 where there were clear differences between conditions for the number of altitude, speed, and heading commands issued, we show few differences between conditions here. In Experiment 1, participants had to descend aircraft according to the LOAs across all traffic streams when using jet routes, but only had to descend non-OPD equipped aircraft when using HPRs. Experiment 2 was an overflight scenario where no aircraft had to cross the sector boundary at a particular altitude. This difference in experimental procedure explains the lack of a difference in altitude commands observed here. For the number of speed and heading commands issued, the data is consistent with the pattern observed in Experiment 1; though, we may not have observed significant differences between conditions here due to floor effects in the data.

3.2.9 CIT Interactions

As in Experiment 1, the CIT was limited in functionality during Experiment 2. Participants were shown six different route buttons, one for each of the major traffic streams (ORD, EWR, PHL, LGA, TEB, and JFK). Only 3 out of 12 participants used the CIT during the scenarios. Figures 59 and 60 show the number and total duration of CIT interactions for all Experiment 2 HPR conditions, separated by traffic level. As in Experiment 1, the total duration of CIT interactions was calculated for each participant by summing the duration of each CIT interaction. All participants are included in the calculations of the means and standard deviations (where any participants who did not use the tool had a 0 for number and duration of CIT interactions).

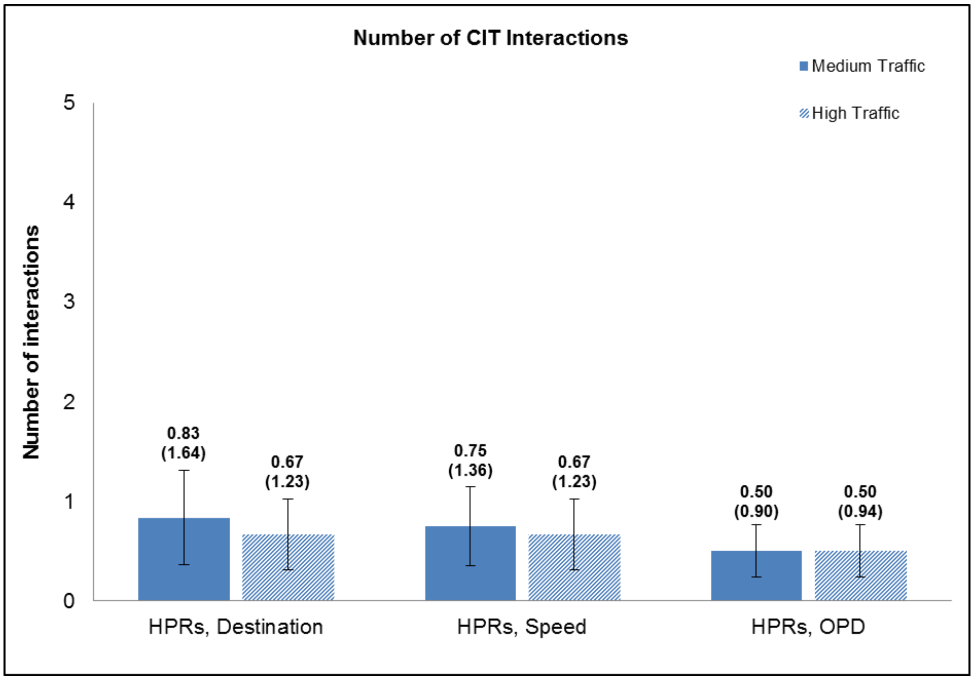


Figure 59. Number of CIT interactions by traffic level and condition. Means (*SD*s) are presented above the error bars.

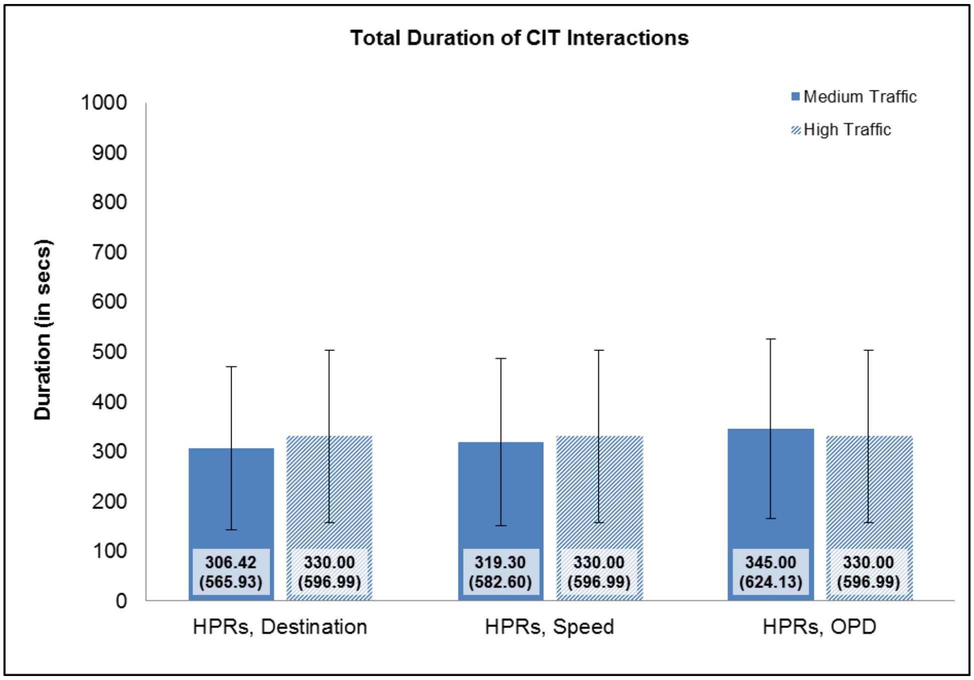


Figure 60. Total duration of CIT interactions by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

3.2.10 Air Traffic Workload Input Technique

Participants were prompted to give workload ratings every 2 minutes across each 45-minute scenario, totaling 22 ratings per scenario. Participants failed to respond to 11% of the prompts in HPR by destination scenarios, 15% of the prompts in HPR by speed scenarios, 9% of the prompts in HPR by OPD equipment scenarios, and 13% of the prompts in jet routes scenarios. Figure 61 shows means of workload ratings by position, traffic level, and condition.

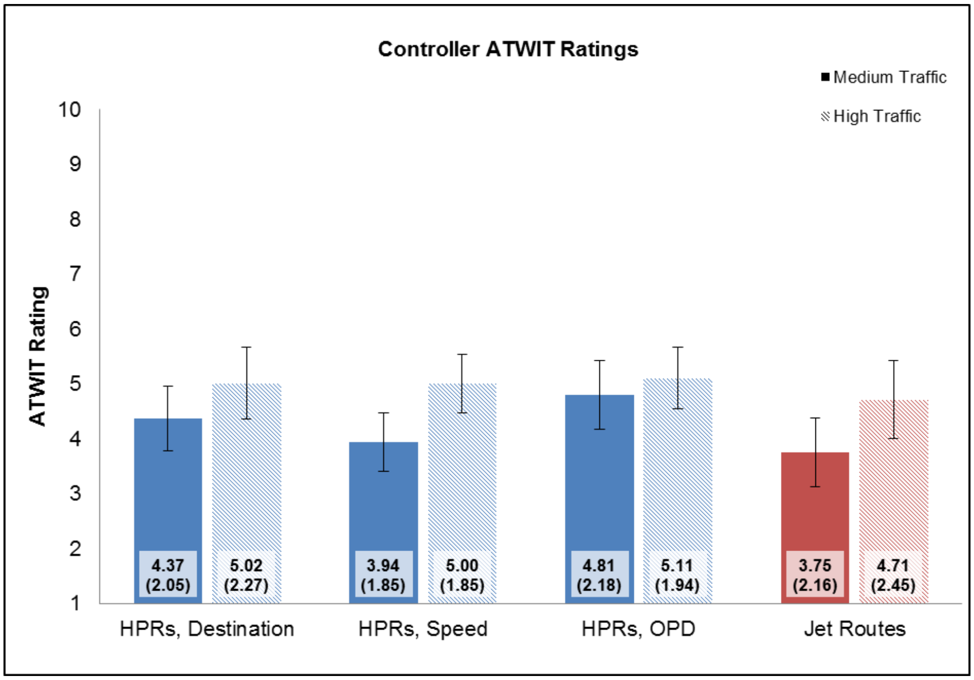


Figure 61. Workload ratings by traffic level and condition. Means (*SD*s) are presented at bottom of figure.

As expected, workload ratings were higher when traffic was high compared with when it was at a medium level, $F(1, 11) = 26.56$, $p < 0.001$, $\eta_p^2 = 0.71$. There was no effect of condition on workload ratings, $F(3, 33) = 2.04$, $p = 0.13$, $\eta_p^2 = 0.16$, but there was a marginal interaction between traffic level and condition, $F(3, 33) = 2.34$, $p = 0.09$, $\eta_p^2 = 0.18$.

The main finding was that there were no statistically significant differences in workload ratings between the four conditions. The reason why there were no statistically significant differences between the conditions was due to higher variability in average workload ratings in all conditions compared with the variability observed in Experiment 1. It is important to note that, qualitatively, participants still gave lower workload ratings in the jet routes condition compared with the HPRs by destination and HPRs by OPD equipment conditions—consistent with Experiment 1. Participants gave similar workload ratings in the HPRs by speed condition compared with the jet routes condition. To summarize, while the trends in the data reported here match with those reported in Experiment 1 (with the exception of the HPRs by speed condition), the ratings are not as consistent across participants (as shown by higher variability across conditions), suggesting that at least for some participants, the difference in workload between jet routes and each of the HPRs conditions was minimal.

On their own, these data are not adequate for suggesting that there were no differences in subjective workload across conditions given that the trends in the data were consistent with those observed in Experiment 1. However, the workload ratings gathered from questions on the PSQ (see next Section 3.2.11) also showed no statistically significant differences across conditions. Together, these two sets of results suggest that participants do not perceive a difference in workload across the four conditions. There are three potential reasons why this would be the case in Experiment 2 even though Experiment 1 showed significantly higher workload ratings in both the ATWIT and PSQ ratings. The first is that participants may have been more comfortable with the HPRs concept when running the Experiment 2 scenarios because they had already spent a day and a half using it. The second is that we reduced the size of the airspace in Experiment 2. The third is that we simplified the HPRs concept in Experiment 2 by only asking participants to sort aircraft onto HPRs in one way (e.g., by speed). In Experiment 1, participants sorted aircraft onto HPRs in three different ways, so they had the additional challenge of keeping track of which lanes were used for which sorting procedures.

3.2.11 Post-Scenario Questionnaire Responses

Table F4 (see Appendix F) displays the means and standard deviations for each condition as well as the *p*-values resulting from statistical analyses comparing the conditions for each of the PSQ questions. Statistically significant effects are noted with an asterisk (marginal effects are noted with a †) and statistically significant pairwise comparisons are indicated with superscripts ("1" indicates that the condition is significantly different from HPRs by destination, "2" indicates that the condition is significantly different from HPRs by speed, "3" indicates that the condition is significantly different from HPRs by OPD equipment, and "4" indicates that the condition is significantly different from jet routes). For all questions, no statistically significant differences between conditions were observed.

3.2.12 Over-the-Shoulder Observer Responses

Tables F5 and F6 (see Appendix F) display the means and standard deviations for each condition as well as the *p*-values resulting from statistical analyses comparing the conditions for each of the OTS observer responses to the PSQ and OTS Observer Rating Form. Statistically significant effects are noted with an asterisk (marginal effects are noted with a †). Superscripts indicate statistically significant pairwise comparisons (1 indicates that the condition is significantly different from HPRs by destination, 2 indicates that the condition is significantly different from HPRs by speed, 3 indicates that the condition is significantly different from HPRs by OPD equipment, and 4 indicates that the condition is significantly different from jet routes).

Consistent with participants' subjective ratings, OTS observers gave similar performance, workload, situation awareness, and difficulty ratings across conditions. In addition, OTS observers did not rate participants' performance differently for any of the OTS Observer Form items. The only differences observed between conditions were for questions on the PSQ asking about the effectiveness of different HPR lane usage procedures and jet route procedures. OTS observers rated both HPRs by destination and HPRs by speed as more effective than jet routes for separating aircraft safely and moving aircraft efficiently. They rated HPRs by speed as more effective than jet routes for controlling more aircraft. These differences were not observed in the participants' ratings for the same questions.

3.2.13 Exit Questionnaire Responses

Participants' responses to the Exit Questionnaire questions relevant to Experiment 2 are reported here. Figures 62, 63, 64, 65, and 66 show the frequency of each type of response (HPRs by destination, HPRs by speed, HPRs by OPD equipment, jet routes, and no difference) for each Experiment 2 question. Some participants did not follow directions and chose more than one response for a question. These participants' responses are removed from the graphs and analyses reported below, as it is impossible to determine which concept was preferred. For each question, a chi-square test of goodness-of-fit was conducted to determine if the frequencies of each response type were different from each other. If there was a significant difference between response types, follow-up, Bonferroni-adjusted chi-square tests for goodness-of-fit were conducted to determine which responses were significantly different from each other.

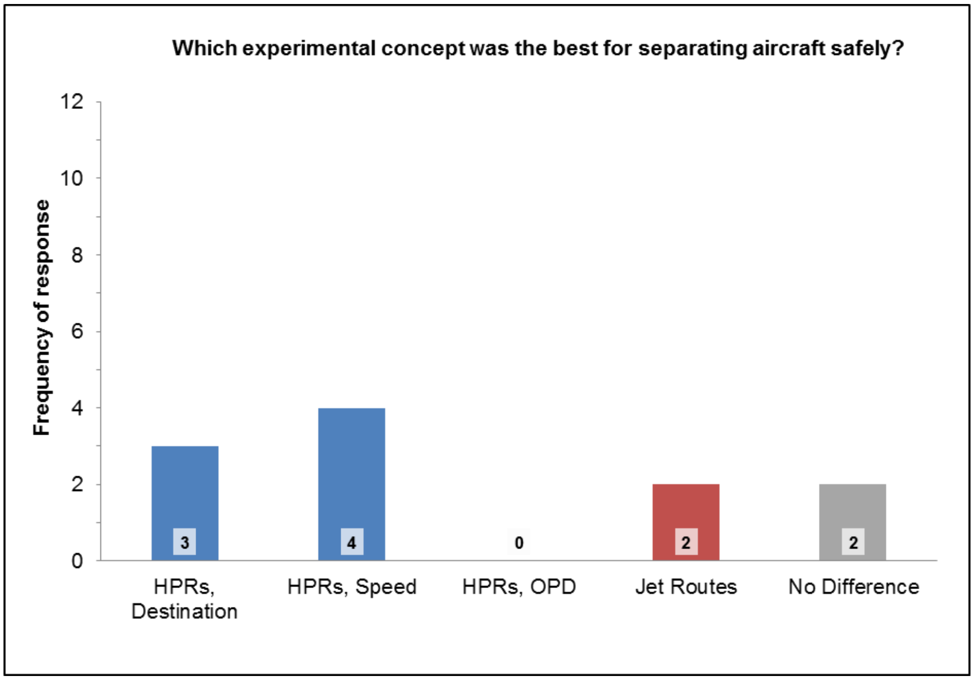


Figure 62. Frequency of participant responses to the question, "Which experimental concept was the best for separating aircraft safely?" Frequencies are presented at bottom of figure.

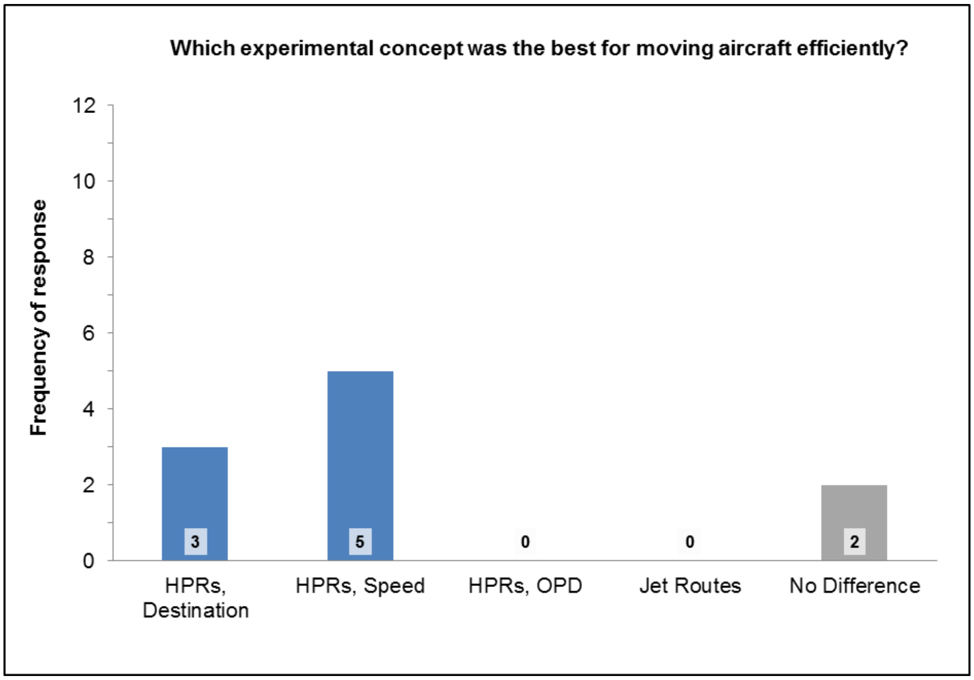


Figure 63. Frequency of participant responses to the question, "Which experimental concept was the best for moving aircraft efficiently?" Frequencies are presented at bottom of figure.

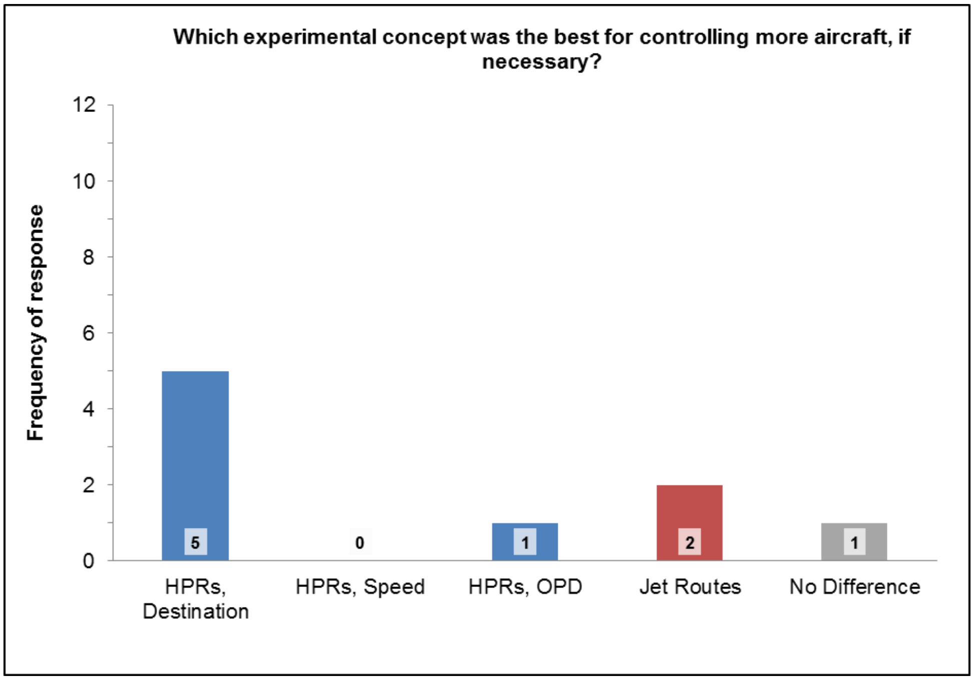


Figure 64. Frequency of participant responses to the question, "Which experimental concept was the best for controlling more aircraft?" Frequencies are presented at bottom of figure.

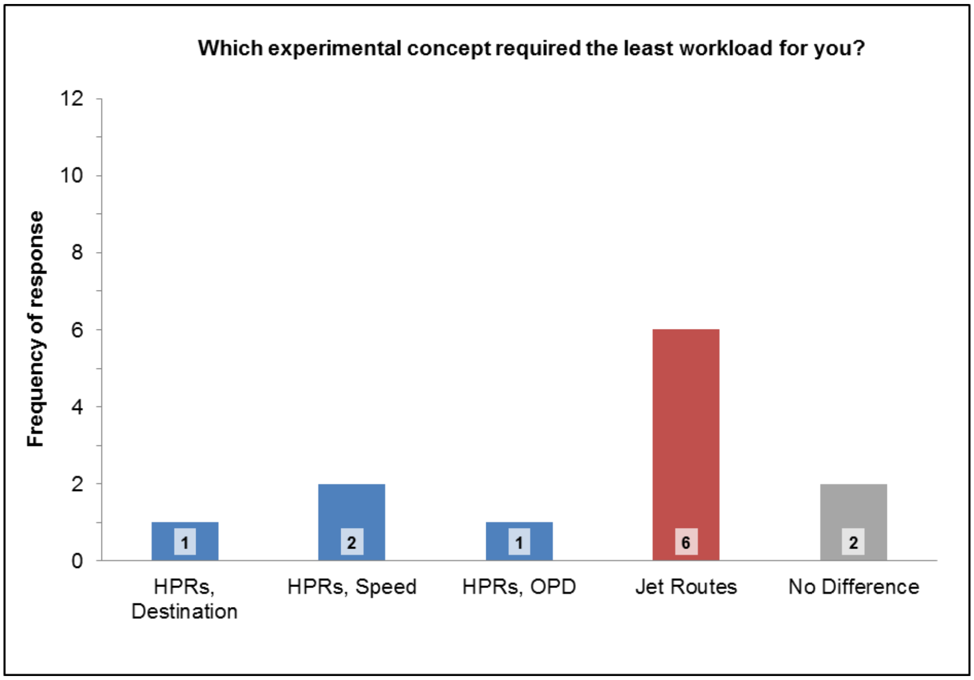


Figure 65. Frequency of participant responses to the question, "Which experimental concept required the least workload?" Frequencies are presented at bottom of figure.

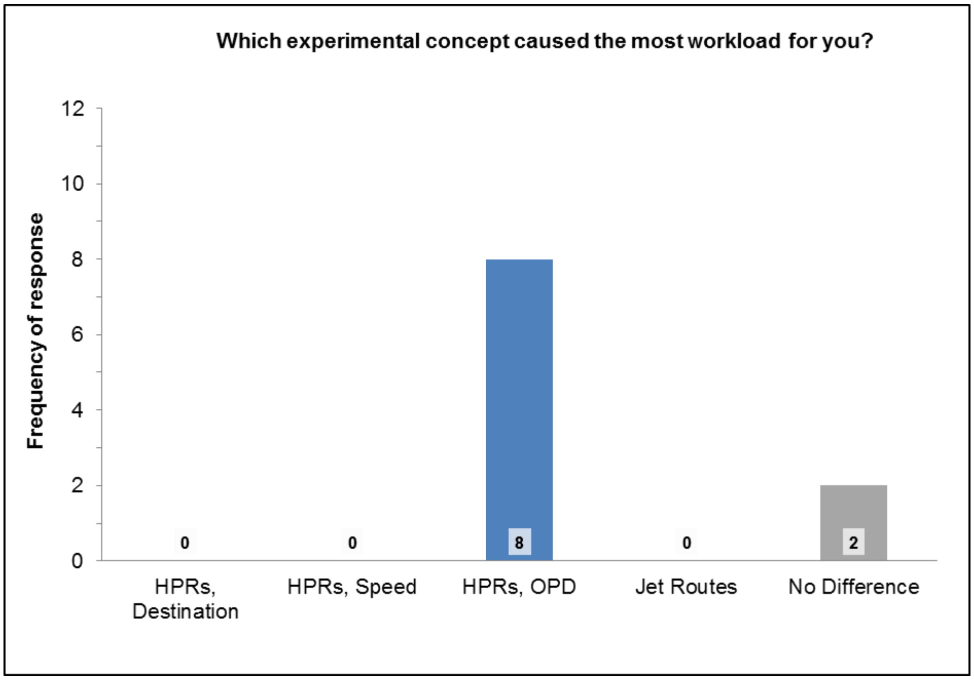


Figure 66. Frequency of participant responses to the question, "Which experimental concept caused the most workload?" Frequencies are presented at bottom of figure.

There were no differences between responses for the question asking about which concept was best for separating aircraft safely, $\chi^2(4, N = 11) = 4.00, p = 0.41$, for the question asking about which concept was best for moving aircraft efficiently, $\chi^2(4, N = 10) = 9.00, p = 0.06$ (though the trend in the data suggests that participants preferred the HPRs by speed response followed by HPRs by destination response), for the question asking about which concept was best for controlling more aircraft, $\chi^2(4, N = 10) = 8.22, p = 0.08$ (though the trend in the data shows a slight preference for HPRs by destination), and for the question asking about which concept required the least workload, $\chi^2(4, N = 12) = 7.17, p = 0.13$ (though the trend in the data suggests that participants preferred the jet routes response). There were differences between responses for the question asking about which concept caused the most workload, $\chi^2(4, N = 12) = 21.91, p < 0.001$. More participants chose the HPRs by OPD equipment response compared with any other response (for all comparisons between HPRs by OPD equipment and each other response, all χ^2's > 15.30 and all p's < 0.001).

3.3 Experiment 3

In Experiment 3, participants controlled traffic independently. Experiment 3 scenarios lasted for 30 minutes, and system metrics were analyzed for the time interval between 2 minutes and 29 minutes. Unless otherwise noted, data were analyzed in two different ways. First, we ran a one-way, repeated measures ANOVA to look at changes over time. If a statistically significant main effect of scenario run was shown, then follow-up, Bonferroni-adjusted paired *t*-tests were run to determine which pairs of runs differed from each other. Second, we ran a paired *t*-test on Scenario Run 1 versus Scenario Run 4 look at differences between participants' first attempt to control traffic and their final attempt to control traffic in the generic sector.

3.3.1 Voice Communications

Figures 67 and 68 present the means, by scenario run, for the number and duration of controller-initiated PTT transmissions. Figures 69 and 70 present the means, by scenario run, for the number and duration of pilot-initiated PTT transmissions.

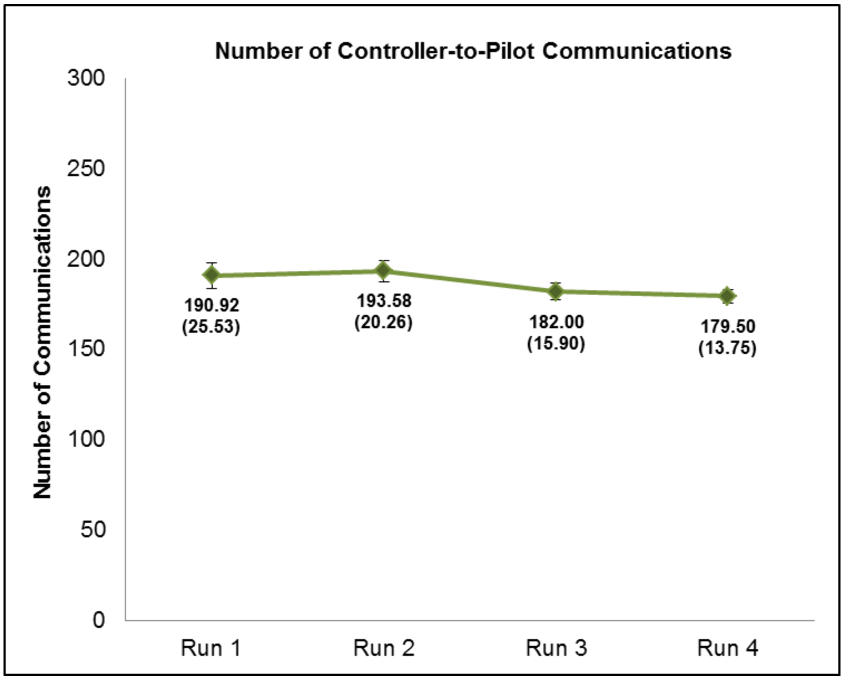


Figure 67. Number of controller-initiated transmissions by scenario run. Means (*SD*s) are presented below the error bars.

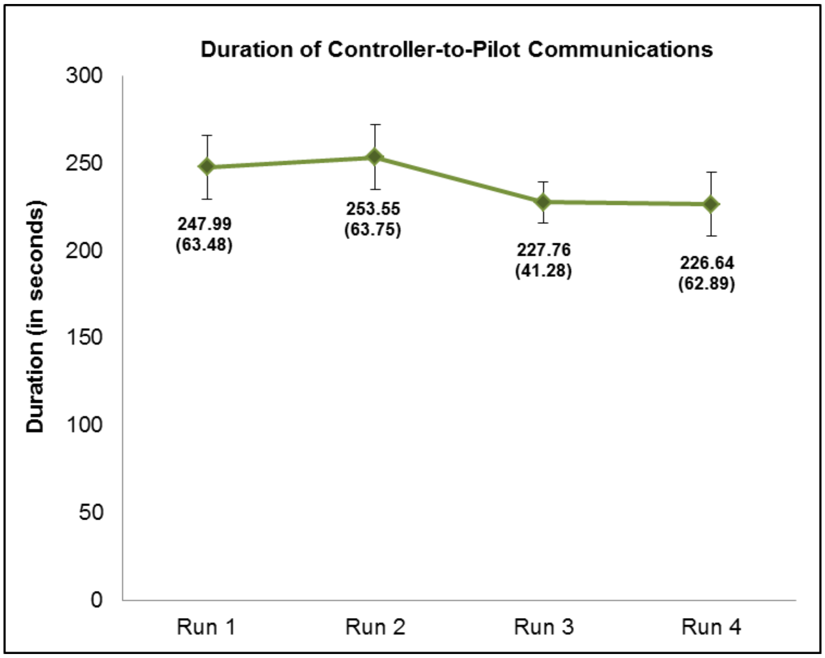


Figure 68. Total duration of controller-initiated transmissions by scenario run. Means (*SD*s) are presented below the error bars.

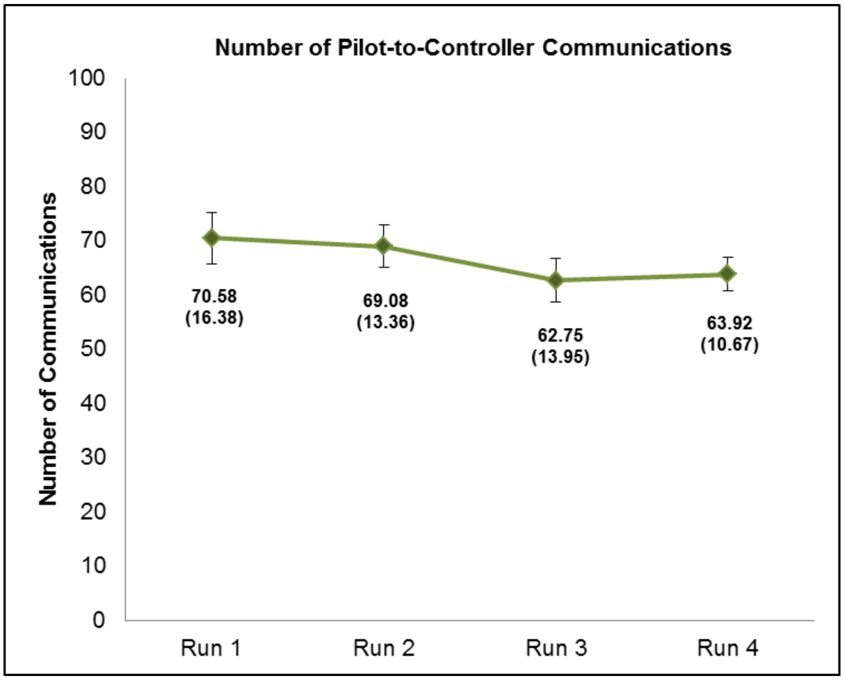


Figure 69. Number of pilot-initiated transmissions by scenario run. Means (*SD*s) are presented below the error bars.

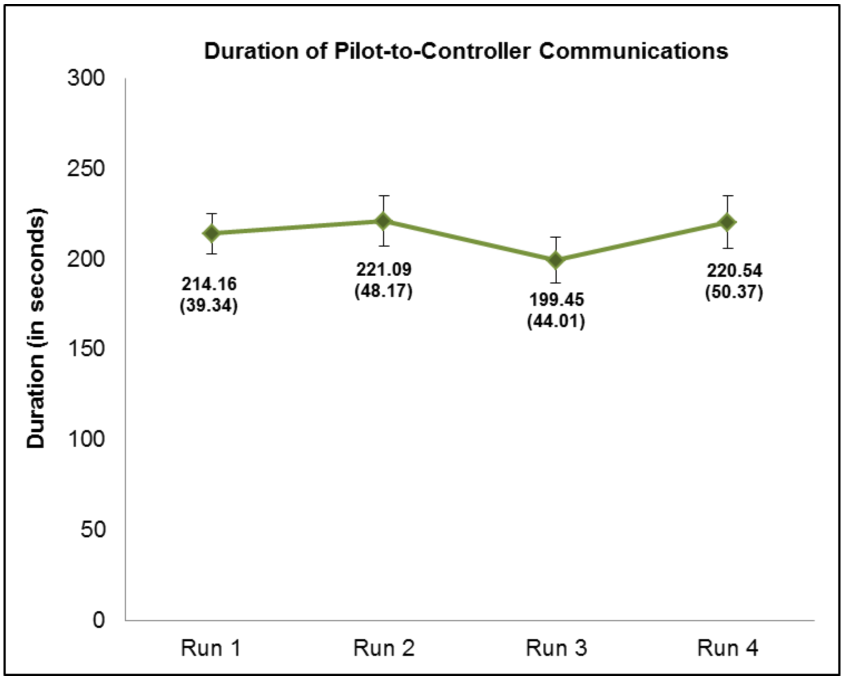


Figure 70. Total duration of pilot-initiated transmissions by scenario run. Means (*SD*s) are presented below the error bars.

There was a main effect of scenario run on the number of controller-initiated transmissions, $F(3, 33) = 3.31$, $p = 0.03$, $\eta_{\rho}^{2} = 0.23$; however, follow-up, Bonferroni-adjusted paired *t*-tests showed no statistically significant differences between scenario runs. There were no differences between scenario runs for the number of pilot-initiated transmissions, $F(3, 33) = 1.07$, $p = 0.38$, $\eta_{\rho}^{2} = 0.09$. In addition, there were no significant differences between scenario runs for the total duration of controller-initiated transmissions, $F(3, 33) = 1.07$, $p = 0.38$, $\eta_{\rho}^{2} = 0.09$, or pilot-initiated transmissions, $F(3, 33) = 0.63$, $p = 0.60$, $\eta_{\rho}^{2} = 0.05$.

There were no differences between the first run and the final run for the number of controller-initiated transmissions, $t(11) = 1.61$, $p = 0.14$, $d = 0.47$, or pilot-initiated transmissions, $t(11) = 1.23$, $p = 0.24$, $d = 0.36$. There were also no differences between the first run and the final run for the total duration of controller-initiated, $t(11) = 1.19$, $p = 0.26$, $d = 0.34$, or pilot-initiated transmissions, $t(11) = 0.33$, $p = 0.74$, $d = 0.10$.

In summary, there were no statistically significant changes in the number or total duration of controller- or pilot-initiated transmissions as participants gained more experience in the generic sector.

3.3.2 Losses of Separation

Losses of separation were defined as incidents where aircraft were separated by less than 5 nmi laterally and 1,000 ft. vertically. As in the previous experiments, an ATC SME reviewed all potential losses of separation and categorized the incidents as either occurring due to system or simulation pilot error or due to controller error. Across all experimental scenarios in Experiment 3, there were no losses of separation.

3.3.3 JEDI/URET Conflict Probe Notifications

We examined the number and duration of conflict probe notifications as well as the number and duration for each type of notification (red or yellow). Figures 71, 72, 73, 74, 75, and 76 display the number and duration of conflict probe notifications, overall, with red notifications and with yellow notifications, respectively.

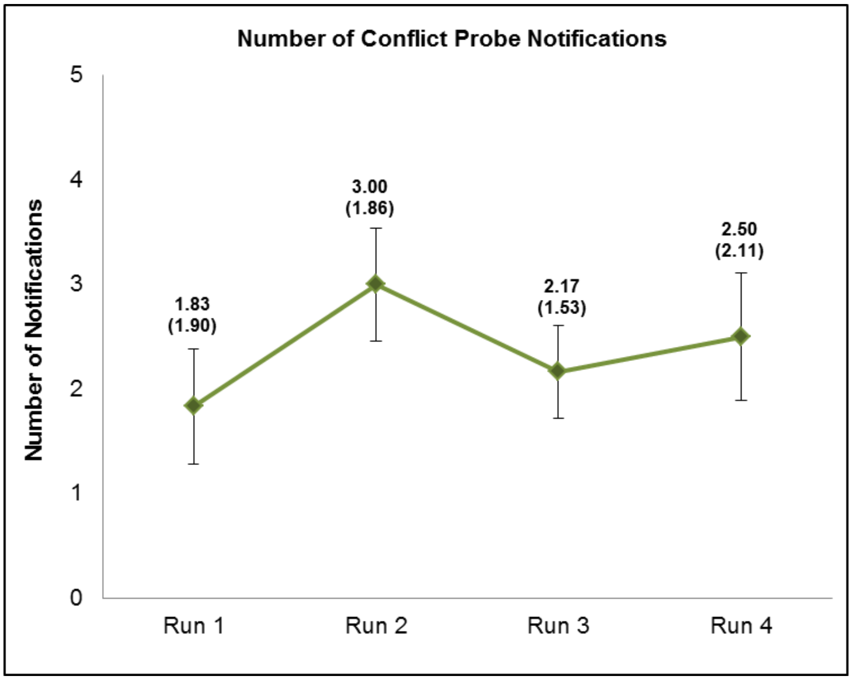


Figure 71. Number of conflict probe notifications by scenario run. Means (*SD*s) are presented above the error bars.

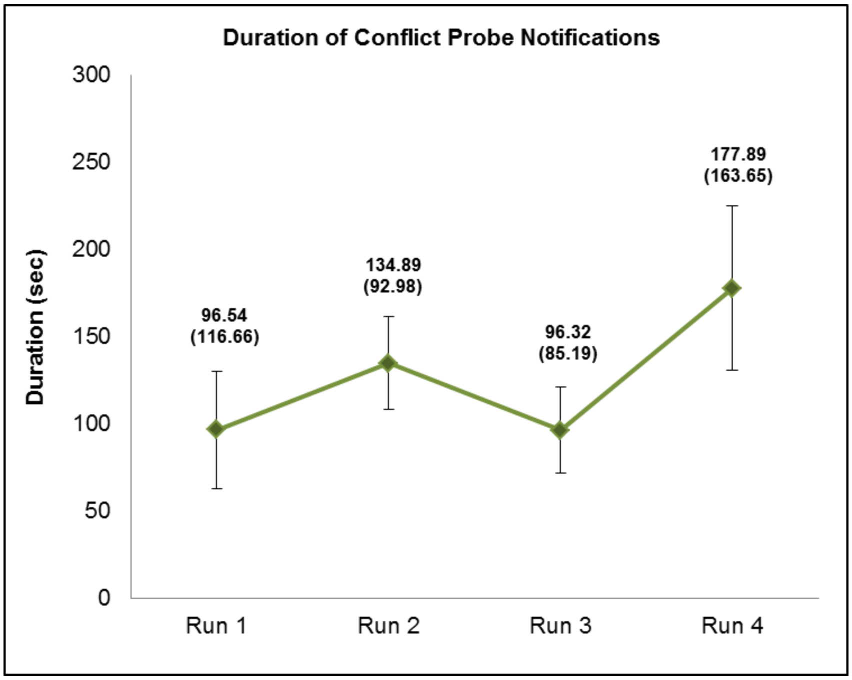


Figure 72. Duration of conflict probe notifications by scenario run. Means (*SD*s) are presented above the error bars.

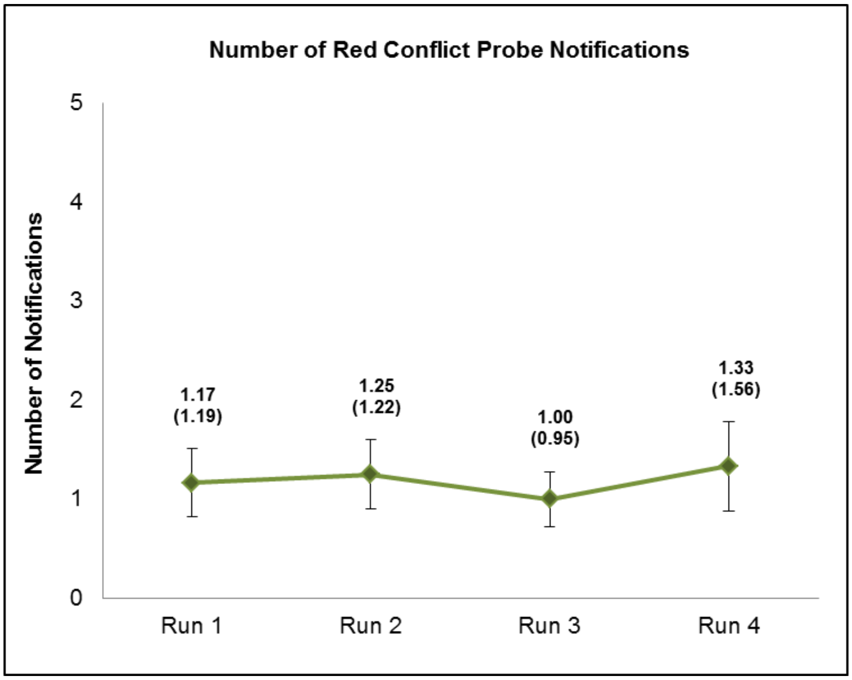


Figure 73. Number of red conflict probe notifications by scenario run. Means (*SD*s) are presented above the error bars.

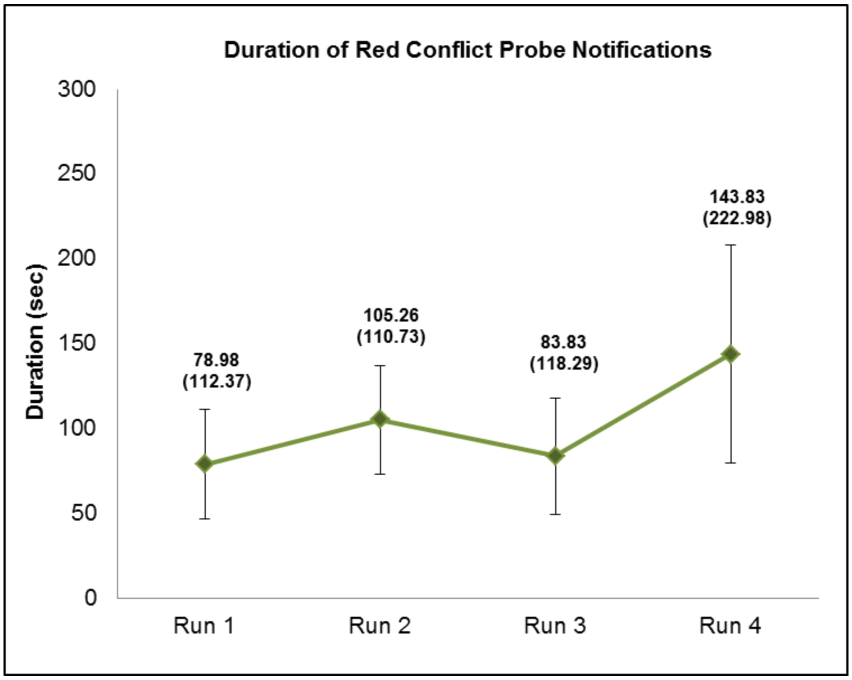


Figure 74. Duration of red conflict probe notifications by scenario run. Means (*SD*s) are presented above the error bars.

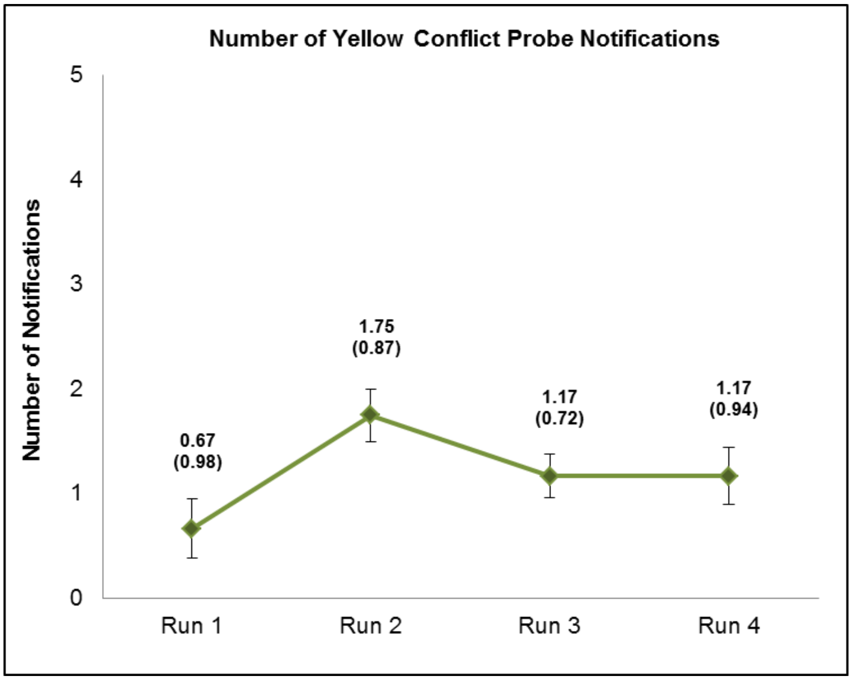


Figure 75. Number of yellow conflict probe notifications by scenario run. Means (*SD*s) are presented above the error bars.

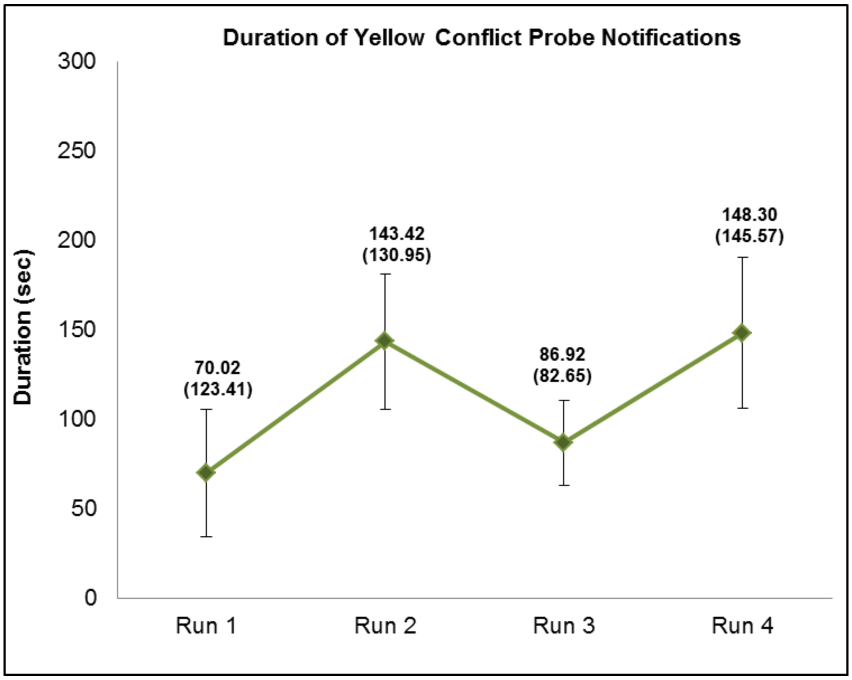


Figure 76. Duration of yellow conflict probe notifications by scenario run. Means (*SD*s) are presented above the error bars.

There were no differences between scenario runs for the number, $F(3, 33) = 0.73$, $p = 0.54$, $\eta_p^2 = 0.06$, or duration, $F(3, 33) = 1.19$, $p = 0.33$, $\eta_p^2 = 0.10$, of conflict probe notifications. In addition, there were no differences between the first run and the final run for the number, $t(11) = 0.69$, $p = 0.51$, $d = 0.20$, or duration, $t(11) = 1.23$, $p = 0.24$, $d = 0.36$, of conflict probe notifications.

There was a marginal effect of scenario run on the number of yellow notifications, $F(3, 33) = 2.76$, $p = 0.058$, $\eta_p^2 = 0.20$, most likely due to the lower number of notifications during Run 1. There was no effect of scenario run on the duration of yellow notifications, $F(3, 33) = 1.24$, $p = 0.31$, $\eta_p^2 = 0.10$. In addition, there were no differences between the first run and the final run for the number, $t(11) = 1.25$, $p = 0.24$, $d = 0.36$, or duration, $t(11) = 1.19$, $p = 0.26$, $d = 0.34$, of yellow notifications. There were no differences between scenario runs for the number, $F(3, 33) = 0.13$, $p = 0.94$, $\eta_p^2 = 0.01$, or duration, $F(3, 33) = 0.41$, $p = 0.74$, $\eta_p^2 = 0.04$, of red notifications. Finally, there were no differences between the first run and the final run for the number, $t(11) = 0.23$, $p = 0.82$, $d = 0.07$, or duration, $t(11) = 0.79$, $p = 0.44$, $d = 0.23$, of red notifications.

There were no differences between scenario runs for the number or duration of conflict probe notifications, suggesting that gaining experience with the generic sector does not change the number or duration of conflict probe notifications. More importantly, the number of notifications was relatively low across all scenarios, suggesting that participants were controlling safe sectors.

3.3.4 Number of Aircraft Under Control

Figure 77 displays the means (and standard deviations), by scenario run, for the number of aircraft under control.

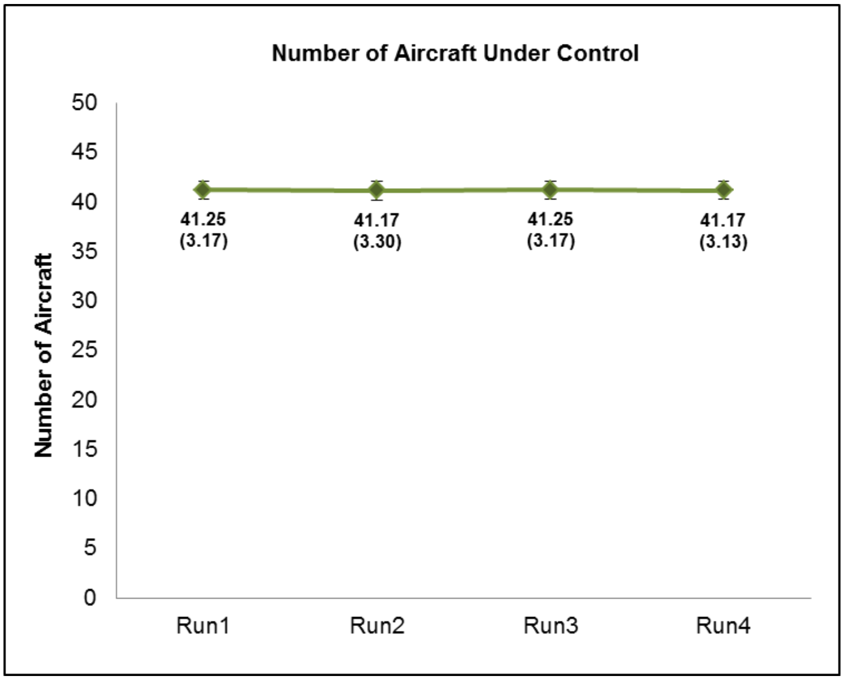


Figure 77. Number of aircraft under control by scenario run. Means (*SD*s) are presented below the error bars.

There were no statistically significant differences across scenario runs for the number of aircraft under control, $F(3, 33) = 0.002$, $p = 0.99$, $\eta_p^2 < 0.001$. In addition, there was no statistically significant difference between the first run and the final run, $t(11) = 0.05$, $p = 0.96$, $d = 0.01$. Gaining experience with the generic sector did not affect how many aircraft participants handled.

3.3.5 Number of Handoffs Accepted and Handoffs Initiated

Figures 78 and 79 display the means (and standard deviations), by scenario run, for the number of handoffs accepted and initiated.

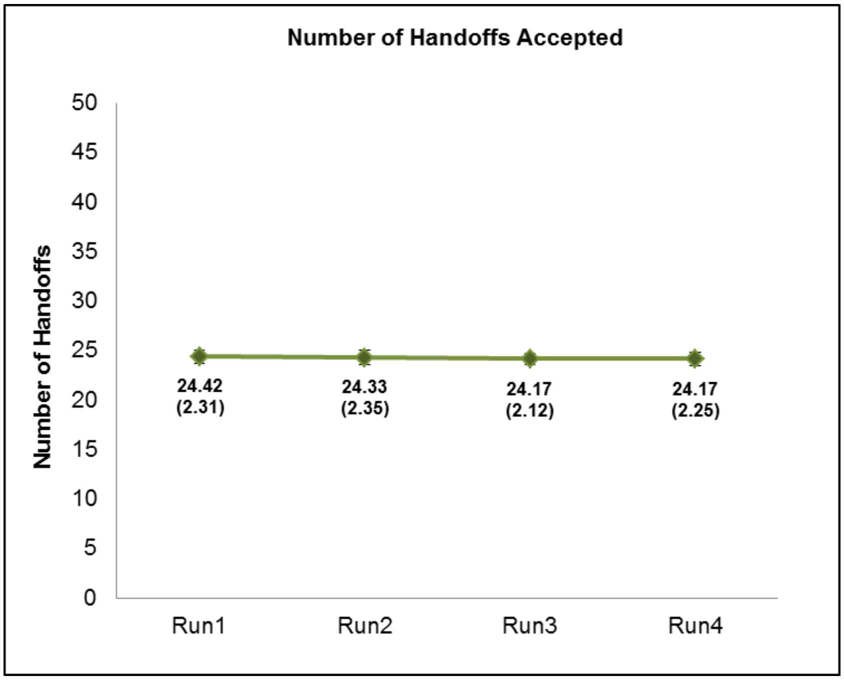


Figure 78. Number of handoffs accepted by scenario run. Means (*SD*s) are presented below the error bars.

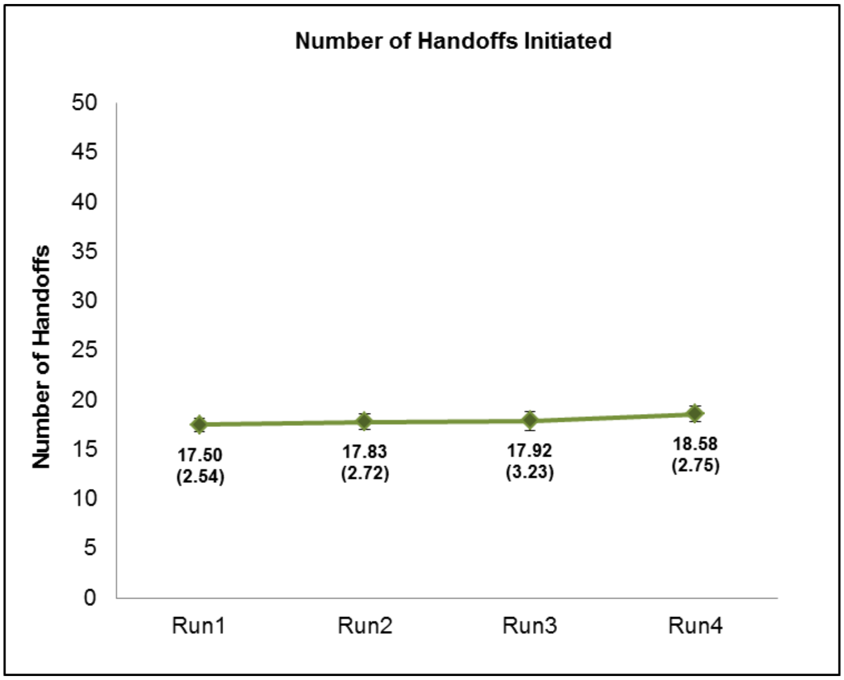


Figure 79. Number of handoffs initiated by scenario run. Means (*SD*s) are presented below the error bars.

There were no effects of scenario run on the number of handoffs accepted, $F(3, 33) = 0.03$, $p = 0.99$, $\eta_{\varrho}^{2} = 0.003$ or the number of handoffs initiated, $F(3, 33) = 0.26$, $p = 0.85$, $\eta_{\varrho}^{2} = 0.02$. In addition, there were no statistically significant differences between the first run and the final run for the number of handoffs accepted, $t(11) = 0.28$, $p = 0.79$, $d = 0.08$, or initiated, $t(11) = 0.89$, $p = 0.39$, $d = 0.26$. Gaining experience with the generic sector did not affect the number of handoffs participants accepted or initiated.

3.3.6 Aircraft Time and Distance in Sector

Figures 80 and 81 display the means, by scenario run, for aircraft time and distance in the participant's sector.

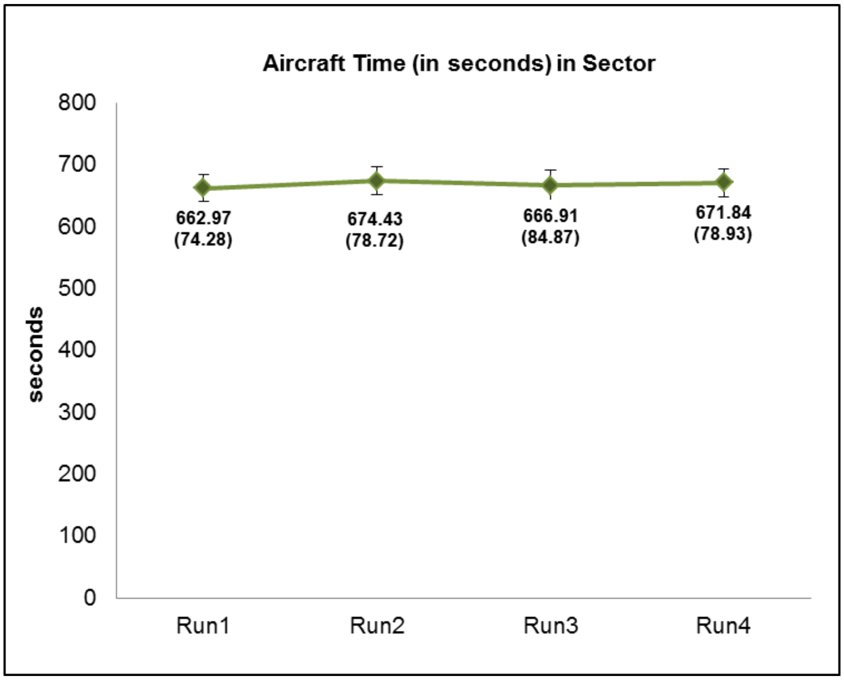


Figure 80. Time in sector by scenario run. Means (*SD*s) are presented below the error bars.

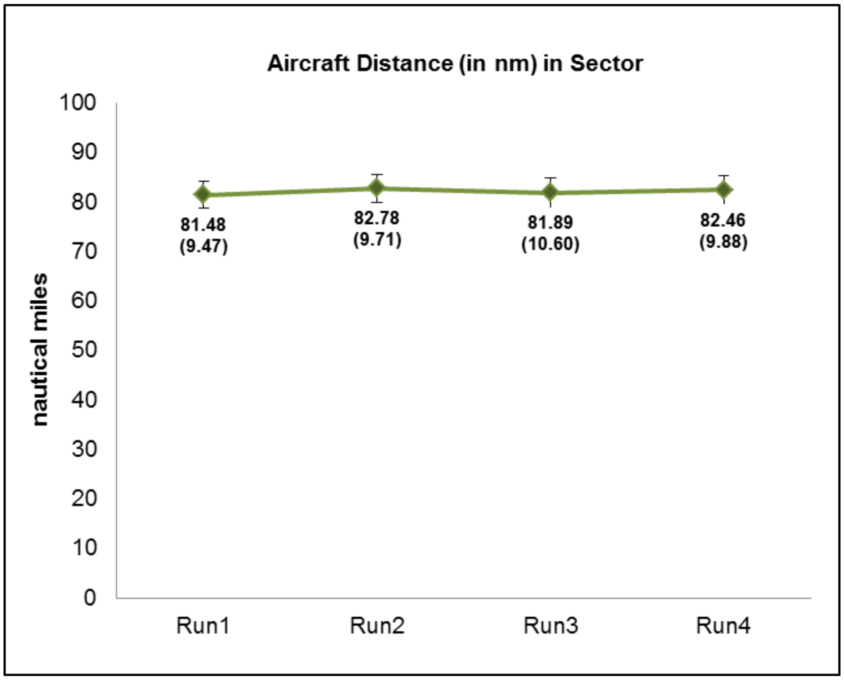


Figure 81. Distance in sector by scenario run. Means (*SD*s) are presented below the error bars.

There were no statistically significant differences between scenario runs for either aircraft time, $F(1.073, 11.80) = 0.04$, $p = 0.87$, $\eta_G^2 = 0.003$,[15] or aircraft distance in the participant's sector, $F(1.05, 11.52) = 0.03$, $p = 0.87$, $\eta_G^2 = 0.003$.[16] In addition, there were no statistically significant differences between the first run and the final run for either aircraft time, $t(11) = 0.20$, $p = 0.84$, $d = 0.06$, or aircraft distance, $t(11) = 0.18$, $p = 0.86$, $d = 0.05$.

Gaining experience with the generic sector did not affect participants' ability to reduce aircraft flight time or distance in the sector. What we cannot conclude is whether we observed no differences across scenarios because participants were able to provide efficient routing for aircraft across all four scenarios or because participants were never able to learn efficient routing, and consequently, did not save time or distance across all four scenarios.

3.3.7 Altitude, Speed, Heading, and Route Amendment Commands

As a measure of controller efficiency, we measured the number of altitude, speed, heading, and route amendment commands that participants issued during each scenario. Figures 82, 83, 84, and 85 display the means, by scenario run, for the number of altitude, speed, heading, and route amendment commands issued.

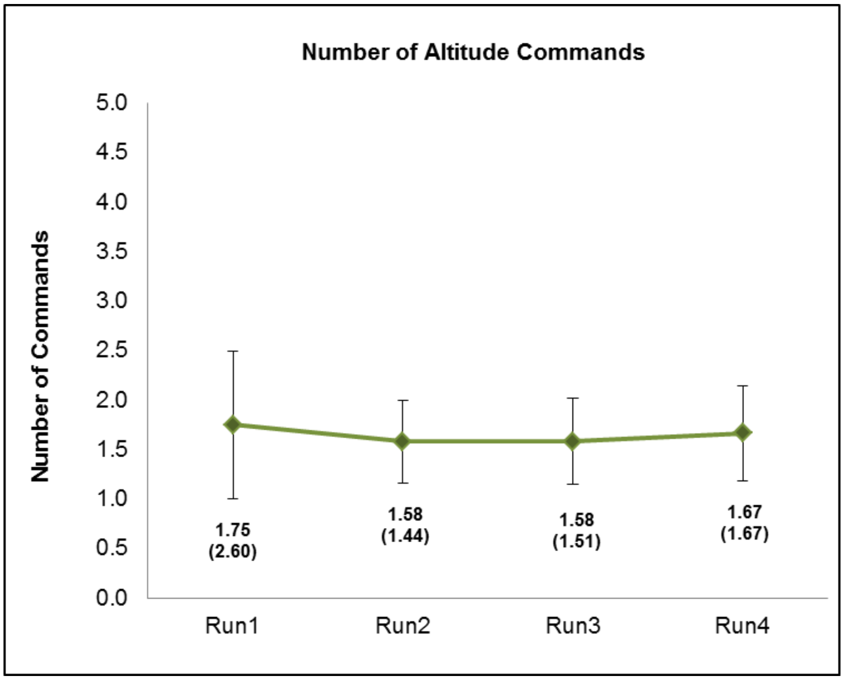


Figure 82. Number of altitude commands by scenario run. Means (*SD*s) are presented below the error bars.

---

[15] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 54.27$, $p < 0.001$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.36$).

[16] Mauchly's test showed that the assumption of sphericity had been violated, $\chi^2(5) = 59.91$, $p < 0.001$, so degrees of freedom were adjusted using the Greenhouse-Geisser estimate of sphericity ($\varepsilon = 0.35$).

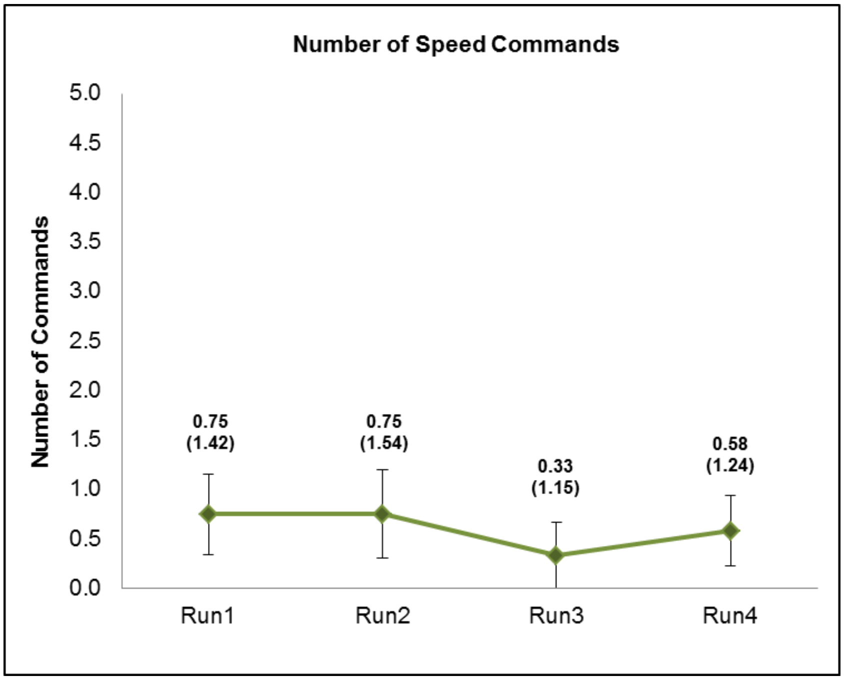


Figure 83. Number of speed commands by scenario run. Means (*SD*s) are presented above the error bars.

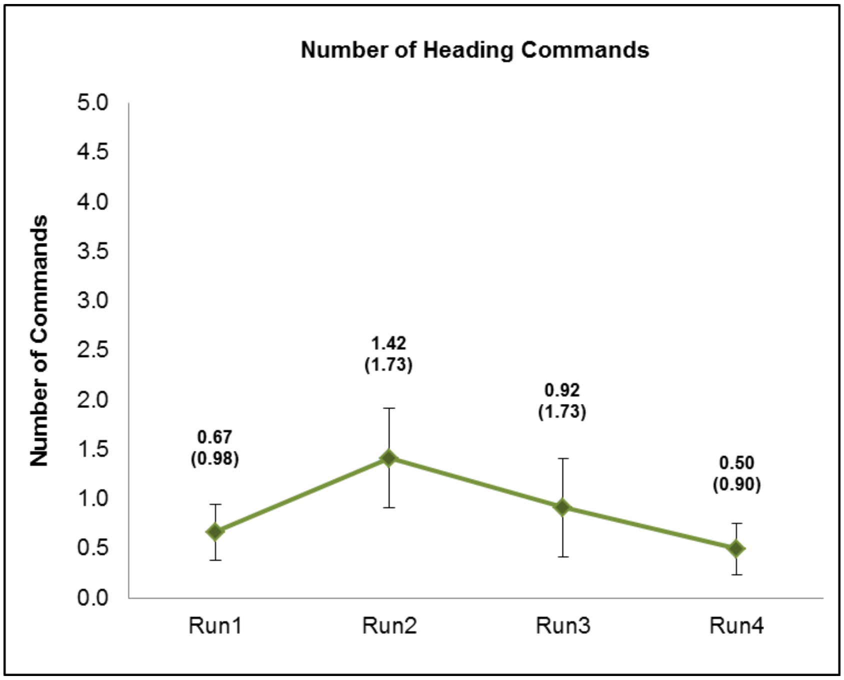


Figure 84. Number of heading commands by scenario run. Means (*SD*s) are presented above the error bars.

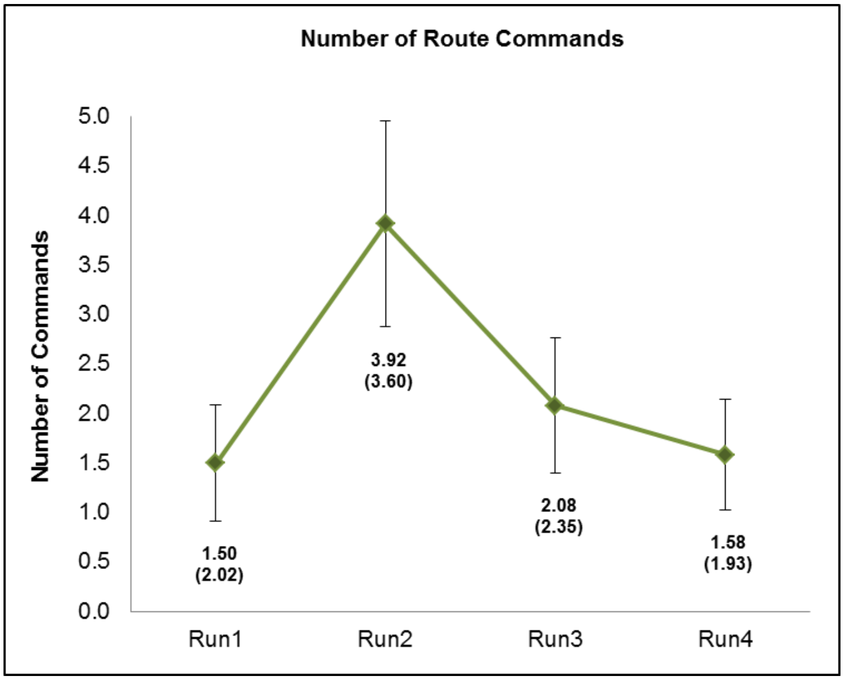


Figure 85. Number of route amendments commands by scenario run.
Means (*SD*s) are presented below the error bars.

There were no main effects of scenario run on the number of altitude commands issued, $F(3, 33) = 0.03$, $p = 0.99$, $\eta_p^2 = 0.003$, speed commands issued, $F(1.70, 18.70) = 0.61$, $p = 0.53$, $\eta_p^2 = 0.05$, and heading commands issued, $F(3, 33) = 0.95$, $p = 0.43$, $\eta_p^2 = 0.08$. There was a marginal main effect of scenario run on number of route amendments issued, $F(3, 33) = 2.56$, $p = 0.07$, $\eta_p^2 = 0.19$, likely due to the second scenario run where participants seemed to be issuing more route amendments. In addition, there were no statistically significant differences between the first run and the final run for number of altitude commands issued, $t(11) = 0.10$, $p = 0.92$, $d = 0.03$, speed commands issued, $t(11) = 0.56$, $p = 0.59$, $d = 0.16$, heading commands issued, $t(11) = 0.39$, $p = 0.70$, $d = 0.11$, and route amendments issued, $t(11) = 0.14$, $p = 0.89$, $d = 0.04$. Given that there were no statistically significant effects observed, we can conclude that gaining experience with the generic sector did not affect participants' use of control commands.

3.3.8 CIT Interactions

Figures 86 and 87 present the means, by scenario run, for the number and total duration of CIT interactions. We calculated the total duration of CIT interactions for each participant by summing the duration of each CIT interaction.

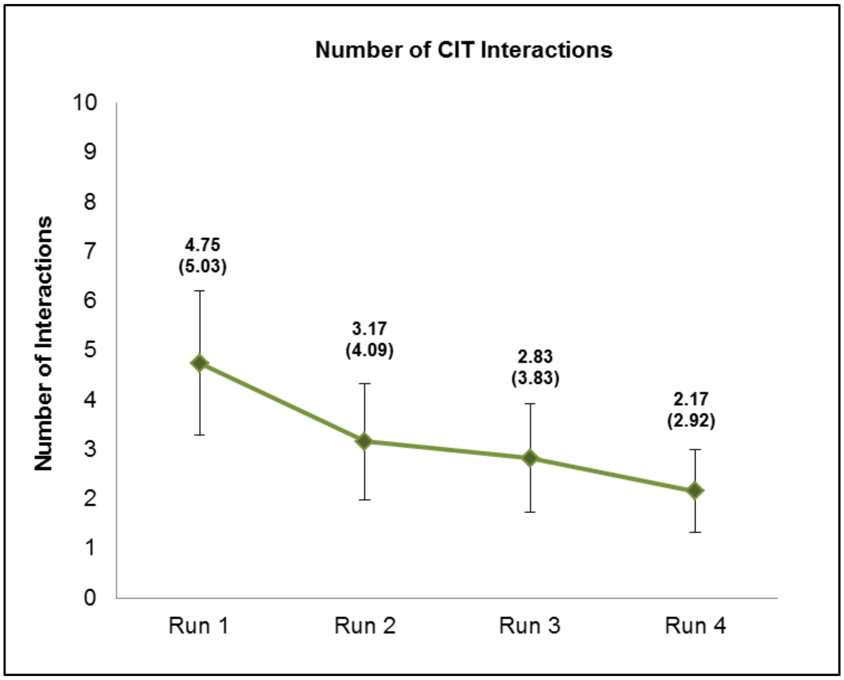


Figure 86. Number of CIT interactions by scenario run. Means (*SD*s) are presented above the error bars.

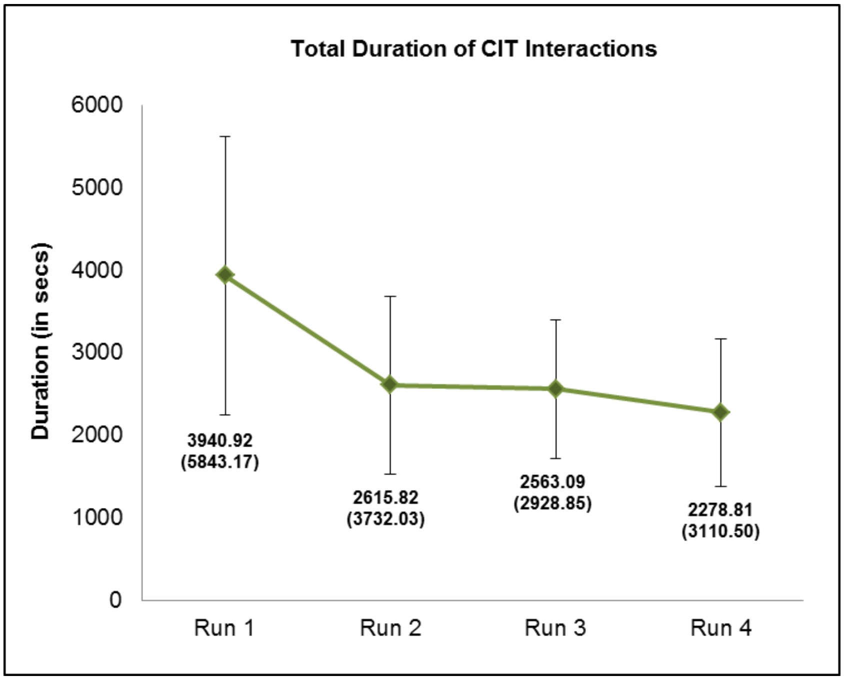


Figure 87. Total duration of CIT interactions (in seconds) by scenario run. Means (*SD*s) are presented below the error bars.

Ten out of twelve participants used the CIT during at least one scenario run. There were no main effects of scenario run on the number of CIT interactions, $F(3, 33) = 1.77$, $p = 0.17$, $\eta_\varrho^2 = 0.14$[17] or on the total duration of CIT interactions, $F(3, 33) = 1.66$, $p = 0.20$, $\eta_\varrho^2 = 0.13$.[18] There was a statistically significant difference between the first run and the final run for the number of CIT interactions, $t(11) = 2.43$, $p = 0.03$, $d = 0.70$,[19] but no statistically significant difference between the first run and final run for the total duration of CIT interactions, $t(11) = 1.53$, $p = 0.16$, $d = 0.44$.[20]

Between the first and the final scenario runs, the number of interactions with CIT functions decreased. One possible explanation for this difference is that more experience with the generic sector led to decreased reliance on the tool. An alternative explanation of the effect is that participants tried out the various functions of the CIT during the first scenario in an attempt to learn to use the tool. By the final scenario, they had learned to use the tool efficiently and had learned which functions were most helpful and which were least helpful. Experience with the generic sector and experience with the CIT were confounded in this study, so both explanations are equally plausible.

Despite a difference between the first and final run for the number of interactions with CIT functions, there were no statistically significant differences between scenario runs for the duration the CIT functions were used (though the trend in the data shows a decrease in duration as participants gained more experience). One explanation for the lack of a difference is that there was high variability in the duration of CIT interactions across participants, suggesting that participants used the CIT differently. This coincides with what was observed during the simulation, which was that some participants kept certain pieces of information on their scope for the entire scenario run while others turned CIT functions on and off depending on whether they were needed during the scenario.

Figures 88 and 89 illustrate the percentage of the total number and total duration of interactions for each CIT function for the first scenario run, Figures 90 and 91 present those data for the second scenario run, Figures 92 and 93 present those data for the third scenario run, and Figures 94 and 95 present those data for the fourth scenario run.

---

[17] Removing the two participants who never used the CIT during Experiment 3 does not change the results, $F(3, 27) = 1.80$, $p = 0.17$, $\eta_\varrho^2 = 0.17$.

[18] Removing the two participants who never used the CIT during Experiment 3 does not change the results, $F(3, 27) = 1.68$, $p = 0.20$, $\eta_\varrho^2 = 0.16$.

[19] Removing the two participants who never used the CIT during Experiment 3 does not change the results, $t(9) = 2.55$, $p = 0.03$, $d = 0.81$.

[20] Removing the two participants who never used the CIT during Experiment 3 does not change the results, $t(9) = 1.53$, $p = 0.16$, $d = 0.48$.

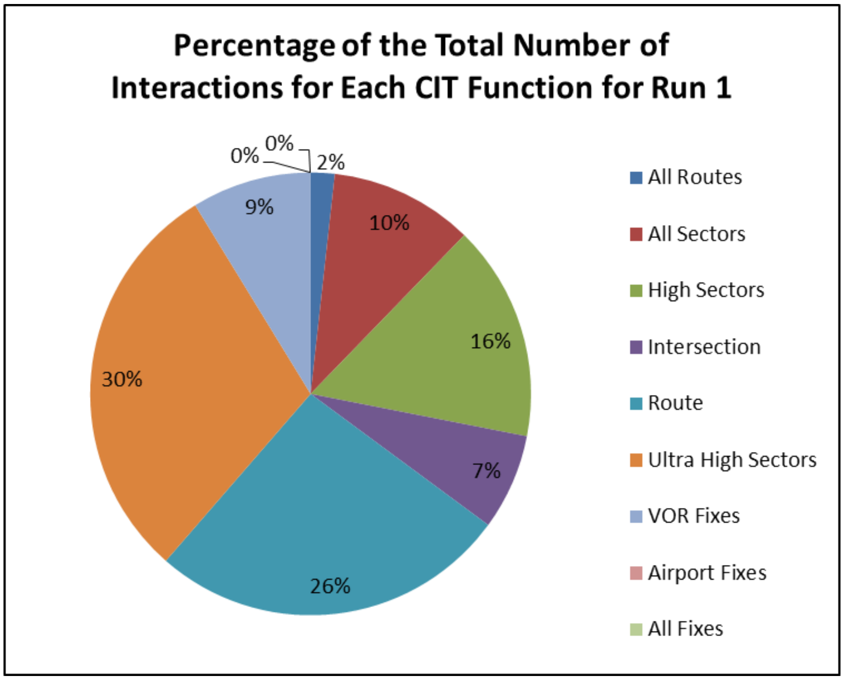


Figure 88. Percentage of the total number of interactions for each CIT function for Run 1.

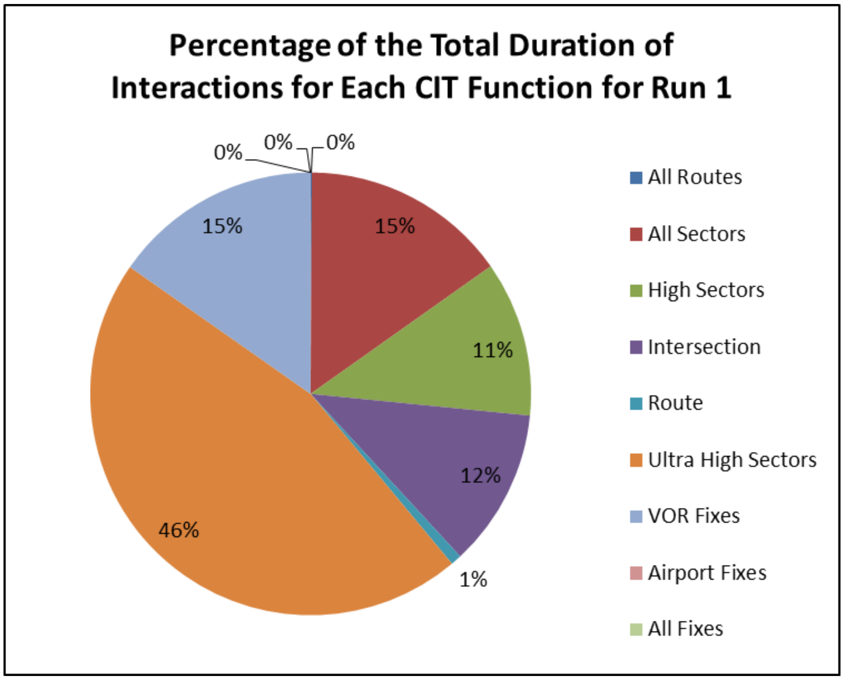


Figure 89. Percentage of the total duration of interactions for each CIT function for Run 1.

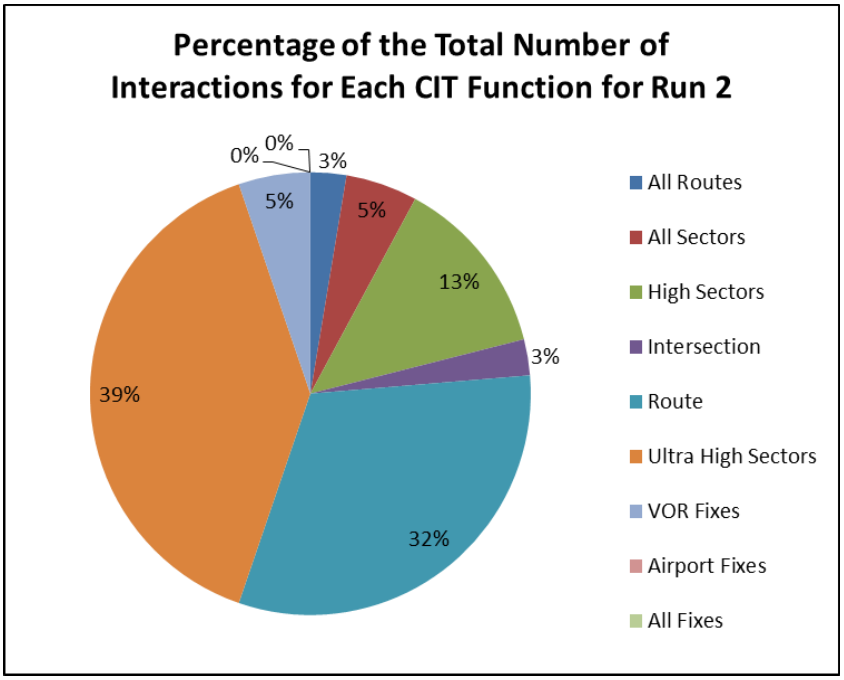


Figure 90. Percentage of the total number of interactions for each CIT function for Run 2.

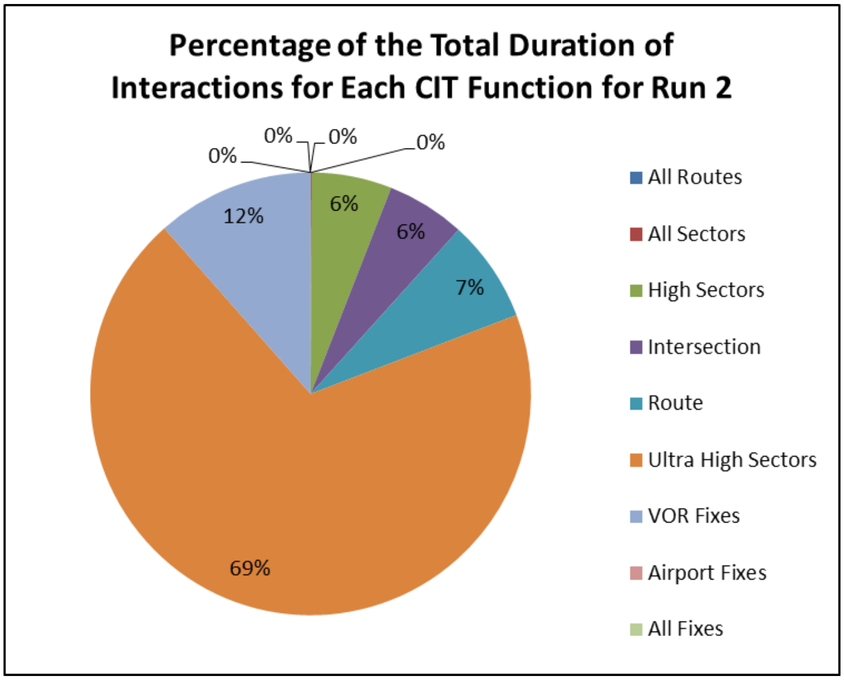


Figure 91. Percentage of the total duration of interactions for each CIT function for Run 2.

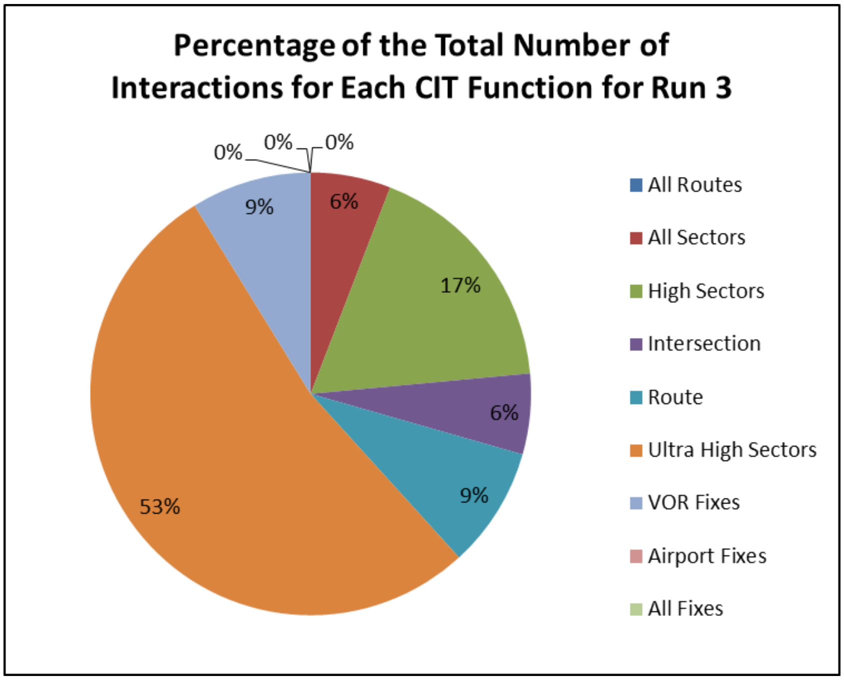


Figure 92. Percentage of the total number of interactions for each CIT function for Run 3.

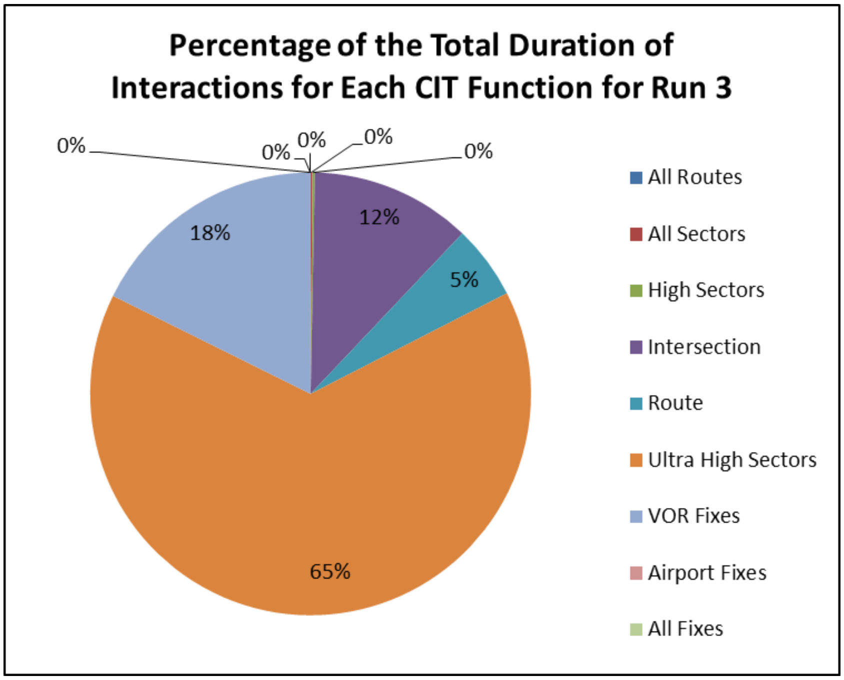


Figure 93. Percentage of the total duration of interactions for each CIT function for Run 3.

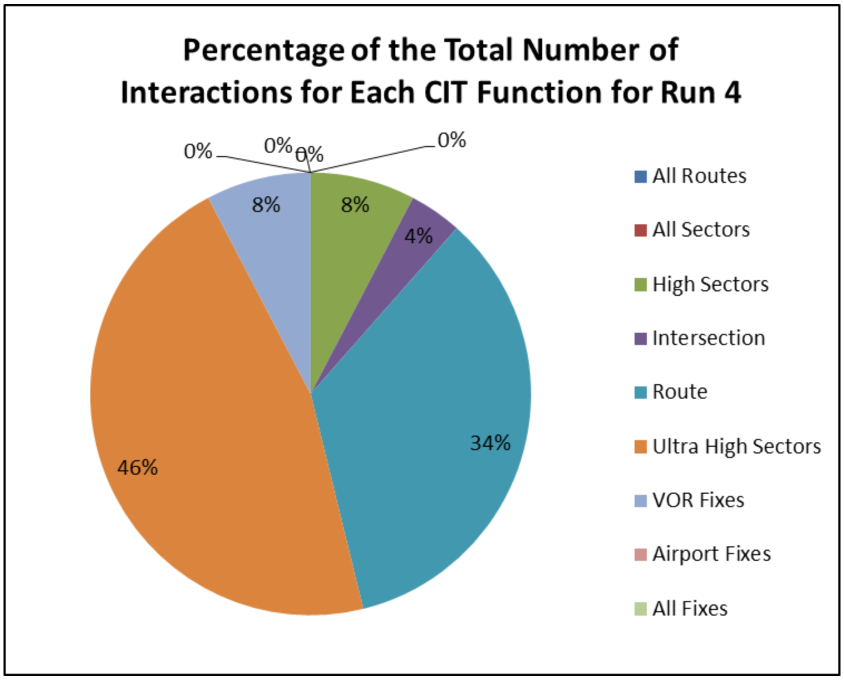


Figure 94. Percentage of the total number of interactions for each CIT function for Run 4.

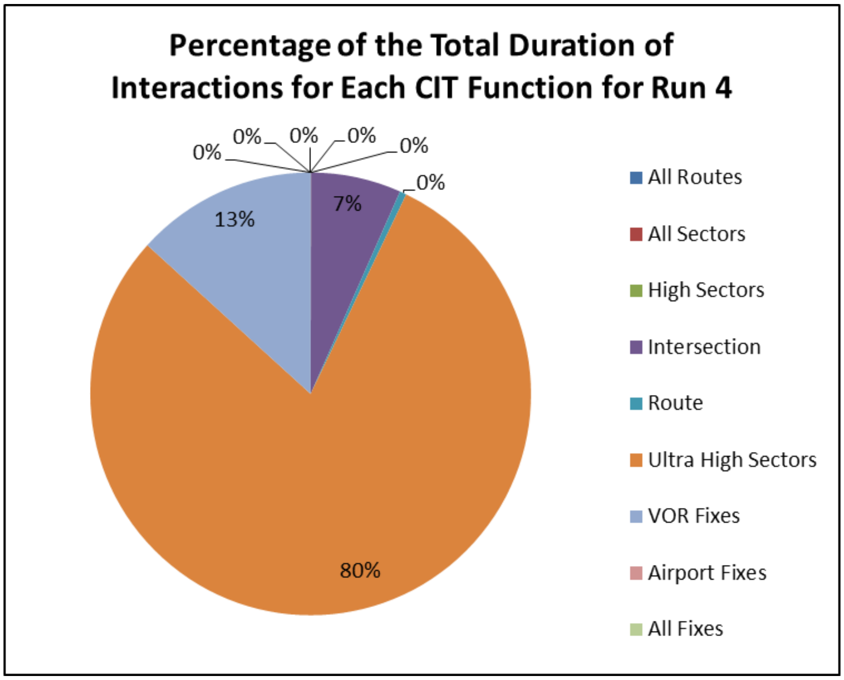


Figure 95. Percentage of the total duration of interactions for each CIT function for Run 4.

The trend in the data shows that for all four scenario runs, participants used the "Ultra High Sectors" information button more often and for longer periods of time than other CIT functions. The individual route display function ("Route") was also used often, though participants only kept this information on their scope for short amounts of time. Participants' use of the other types of functions (e.g., VOR Fixes, High Sectors) seemed to diminish across scenario runs.

Participants noted in their PSQs, Exit Questionnaires, and during the debriefing that the CIT's hierarchical structure made it cumbersome and difficult to access the information through the CIT. We examined the number of mouse clicks it took participants to turn a CIT function on and off to gain a better understanding of the number of actions participants had to use to get the information they wanted.

Participants could access multiple functions at the same time without going back to the initial CIT Menu, so we looked at the number of clicks to turn a function on or off, separated by whether or not the participant had already opened the CIT Menu before beginning navigation to a function.

On average, participants used 4.55 clicks ($SD = 1.44$) to turn on the function they wanted to use if they began at the CIT Menu and 2.01 clicks ($SD = 0.33$) if the CIT Menu was already open. Participants used an average of 4.71 clicks ($SD = 4.46$) to turn off the function they wanted to use if they began at the CIT Menu and 1.64 clicks ($SD = 0.24$) if the CIT Menu was already open.

Because it took a substantial number of mouse clicks to get to the function the participant wanted to use, some participants set up their scope with the functions they wanted before the scenario began and never removed them from the screen. However, these participants noted that without the ability to change the brightness levels of the CIT, this method is not ideal.

In summary, for the participants who used the CIT, the information presented for the ultra-high sectors seemed to be the most helpful for them. The information presented to participants from this function included sector boundaries, numbers, altitudes, and frequencies. The individual route display function also seemed helpful for short amounts of time. It is possible that participants would have used the CIT more often and for more information if the brightness settings had been controllable and if the information was easier to access.

3.3.9 Air Traffic Workload Input Technique

Participants were prompted to give workload ratings every 2 minutes across each 30-minute scenario, totaling 15 ratings per scenario. Participants failed to respond to 12% of the prompts in the first scenario run, 9% of the prompts in the second scenario run, 8% of the prompts in the third scenario run, and 6% of the prompts in the fourth scenario run. Figure 96 shows means of workload ratings by scenario run.

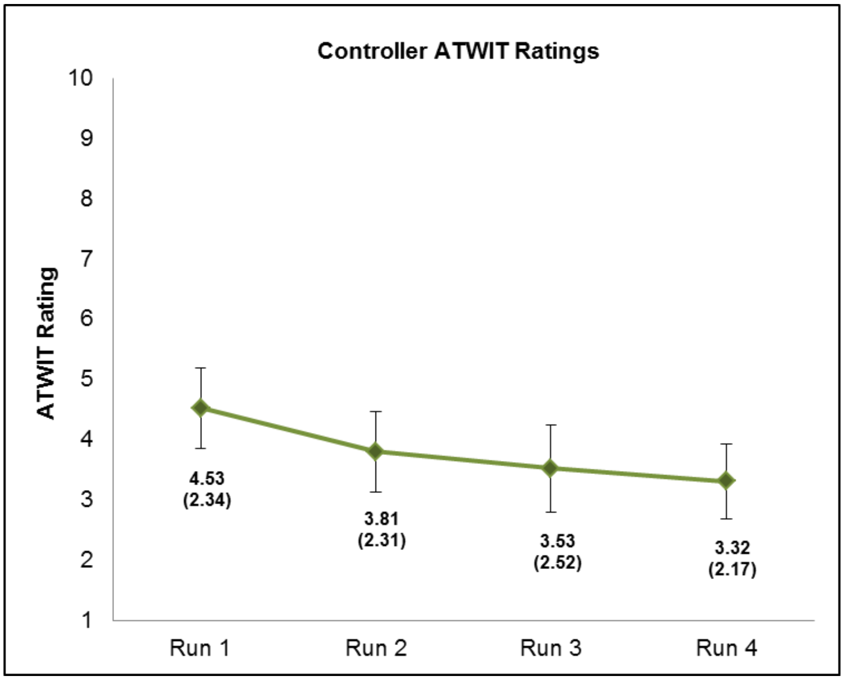


Figure 96. Workload ratings by scenario run. Means (*SD*s) are presented below the error bars.

There was an effect of scenario run on workload ratings, $F(3, 33) = 7.22$, $p = 0.001$, $\eta_{G}^{2} = 0.40$. Follow-up, Bonferroni-adjusted paired *t*-tests showed that workload ratings during the first scenario run were significantly higher than ratings during the third run, $t(11) = 3.78$, $p = 0.003$, $d = 1.09$, and fourth run, $t(11) = 3.40$, $p = 0.006$, $d = 0.98$. Across all four runs, participants' average workload ratings were at low to medium levels, suggesting that participants were not overwhelmed by controlling traffic in a new, unfamiliar sector. In addition, by the third half-hour run, participants' workload ratings were significantly reduced compared with the first run.

3.3.10 Post-Scenario Questionnaire Responses

Table F7 (see Appendix F) displays the means and standard deviations for each scenario run as well as the *p*-values resulting from statistical analyses comparing the scenario runs for each of the PSQ questions. Statistically significant effects are noted with an asterisk (marginal effects are noted with a †) and statistically significant pairwise comparisons are indicated with superscripts (1 indicates that the condition is significantly different from Run 1, 2 indicates that the condition is significantly different from Run 2, 3 indicates that the condition is significantly different from Run 3, and 4 indicates that the condition is significantly different from Run 4).

Consistent with the objective system performance data, there were very few items showing statistically significant differences in ratings across the scenario runs. Participants rated their performance as high across all four scenarios, their workload as relatively low, and their situation awareness as high. The only differences in ratings were for questions asking about what participants had learned as they gained more experience with the sector. Participants reported that they learned more about the sector characteristics after the third and fourth scenario runs compared with the first scenario run (and second scenario run when compared with the fourth scenario run). Participants reported learning more about the traffic patterns after the fourth scenario run compared with the first and second runs. Finally, they reported higher levels of confidence after the fourth scenario run compared with the first. By the end of the third scenario run, eleven of the twelve participants reported that they felt they had controlled a safe sector.

In summary, these results suggest that participants felt they were controlling traffic safely and efficiently, and with minimal workload, from the start of their first exposure to the generic sector. Their performance, workload, situation awareness, and difficulty ratings did not fluctuate across scenario runs because participants were at ceiling for their performance and situation awareness ratings and almost at floor for their workload and difficulty ratings.

3.3.11 Over-the-Shoulder Observer Responses

Tables F8 and F9 (see Appendix F) display the means and standard deviations of OTS observer responses to the PSQ and OTS Observer Rating Form, by scenario run, as well as the *p*-values resulting from statistical analyses comparing the scenario runs for each of the items. Statistically significant effects are noted with an asterisk (marginal effects are noted with a †).

The OTS observers' PSQ and OTS Observer Form ratings reflected improvements in participants' efficiency as they gained more experience with the sector. Specifically, ratings were higher after the final scenario run compared with the first for performance with avoiding traffic flow delays and moving aircraft efficiently. Observers' ratings also reflected that participants' situation awareness improved as they gained experience with the sector. Like the participants, observers noted increases in knowledge of sector characteristics. By the end of the second run, OTS observers had noted that all twelve participants were operating a safe sector.

3.3.12 Exit Questionnaire Responses

Participants' responses to the Exit Questionnaire questions relevant to Experiment 3 are reported here. All of the questions about Experiment 3 were about the effectiveness of the CIT. Table 4 shows the means (and standard deviations) of participants' ratings for each question (ratings were on a scale of 1-9 for the first three questions and 1-10 for the fourth question).

Table 4. Means (*SDs*) of Questionnaire Responses

Question:	Rating Scale	Response
1. What effect, if any, did the Controller Information Tool have on your ability to quickly learn the sector map?	Neg. Effect ①②③④ ⑤ ⑥⑦⑧⑨ Pos. Effect | None	6.00 (1.13)
2. What effect, if any, did the Controller Information Tool have on your ability to quickly learn the traffic routes?	Neg. Effect ①②③④ ⑤ ⑥⑦⑧⑨ Pos. Effect | None	5.67 (1.15)
3. What effect, if any, did the Controller Information Tool have on your ability to effectively control traffic in the sector?	Neg. Effect ①②③④ ⑤ ⑥⑦⑧⑨ Pos. Effect | None	5.25 (0.87)
4. How difficult was it to use the CIT interface?	Not Difficult ①②③④⑤⑥⑦⑧⑨⑩ Extremely Difficult	4.17 (2.21)

In general, participants' responses show that the CIT had little to no effect on their performance and learning. This is most likely because many participants used it infrequently during the scenarios. Participants' responses to open-ended questions indicated that many participants liked the idea of the CIT and thought it was a great tool in concept; however, they needed control over brightness and needed the information to be more easily accessible in order to use it effectively while controlling traffic. For example, one participant noted, "I liked all the information that was made available." Another participant made a similar comment but also noted that he disliked the lack of brightness control and that he had to use too many entries to get to the information. In addition, one of the OTS observers (a front-line manager) commented, "I did not like that [the CIT] was not already in the field. This is a great tool. I like that it can be easily toggled on and off as needed." Seven of the twelve participants indicated specific features that they liked about the tool, which included routes, fix names, NAVAIDS, and frequencies. These comments suggest that the CIT has the potential to improve learning and provide controllers with the information they need in an effective way. However, based on controllers' feedback, further development is needed to improve the user interface as well as the way that controllers access the information they need.

4. CONCLUSIONS

In three experiments, we investigated two HA Airspace concepts for the midterm, the HPRs concept and the Generic Sector concept, using a high-fidelity HITL simulation. In Experiment 1, we compared the jet routes condition to the HPRs condition to examine how each affects the transition between HA Airspace and lower altitude airspace. In Experiment 2, we compared the jet routes condition to three different HPR lane usage conditions (destination, speed, OPD equipment) to observe the effects of different lane usage strategies. In Experiment 3, we looked at whether controllers could safely and efficiently manage aircraft in an unfamiliar sector with the help of the CIT.

4.1 Experiment 1

The primary benefits of HPRs are that they should be more efficient and should increase airspace capacity compared to traditional jet routes. The results of the study give some support for these benefits. In Experiment 1, HPRs slightly decreased aircraft flight distances and time in the sector relative to the jet routes condition under medium-traffic levels. However, there were no differences in flight distances or time in the second half of the scenarios when traffic levels increased. The HPRs were more direct compared to jet routes, and controllers were able to use them to decrease aircraft flight distances and time, which generally reduces aircraft fuel consumption. Although it is surprising that the HPRs did not show the same benefits under high-traffic conditions, they did not hurt efficiency.

Controllers issued fewer altitude, speed, and heading control instructions while using the HPRs compared with jet routes. This suggests that controllers used the HPRs to maintain aircraft separation with fewer control instructions, allowing aircraft to fly through the sector more efficiently with less maneuvering. The researchers designed the HPRs to support future concepts for transition into lower altitude airspace, such as OPD routes. The HPRs allowed aircraft to maintain altitude through the sector and descend via OPD routes closer to their destination airports. The jet routes condition required controllers to issue more altitude commands to "step down" aircraft much farther from the terminal area. The results also indicated that controllers used fewer speed commands and marginally fewer heading commands. In typical operations, controllers use speed commands to prevent faster aircraft from overtaking slower aircraft on the same route, which causes some inefficiency for the faster aircraft. In the simulation, the HPRs allowed aircraft to maintain their speeds through the sector and enabled controllers to use different lanes to separate slower aircraft from faster ones.

In general, there were no differences between the HPRs and jet routes conditions in terms of the airspace capacity measures. Although the HPRs increased the number of routes in the sector, there were no more aircraft accepted into the sector or handed off to the next sector compared to jet routes. This result is expected given the limitations of the simulation. The researchers designed the traffic scenarios to schedule a fixed number of aircraft into the sector during the simulation runs. Under these conditions, the controllers were not able to accept more aircraft even if the sector was below maximum capacity. In actual ATC operations, traffic flow and sector capacity can be increased with more routes as long as controller workload remains acceptable and safety is maintained.

It is important to evaluate NextGen concepts in simulations to ensure that controllers can use the future concepts and maintain aircraft safety. There was only one incident leading to loss of aircraft separation; it occurred in the HPRs condition under high-traffic levels but did not suggest that the HPR concept compromised safety. There was no difference between HPRs and the jet routes condition in terms of the number of JEDI/URET red alerts. In fact, there were fewer yellow alerts when controllers used HPRs as compared to jet routes. However, the number of conflict probe alerts generated by the system is not necessarily an indicator of safety. The JEDI/URET tool is designed as an early warning for potential loss of aircraft separation. Controllers may receive alerts and resolve conflicts long before loss of aircraft separation occurs.

While the objective simulation measures of efficiency, capacity, and safety showed some benefits for the HPR concept, the controllers' subjective measures and evaluations of the HPRs were less positive. The controllers' workload ratings were higher in the HPRs as compared to the jet routes condition. The controllers' questionnaire ratings of their situation awareness and performance were lower in the HPRs condition. Many controllers expressed concern about the difficulty in using the HPRs. They reported that recognizing the appropriate aircraft for different HPR lane usages increased their workload. They responded that the different procedures for HPR lane usage added to the complexity of their task. This participant feedback is important to help the researchers identify the human performance issues using the proposed HPR concept.

It is possible that many of the reported issues are due, at least in part, to insufficient training with the new HPR procedures. In the study, training time was limited to only a few hours of practice runs before actual test runs began. Although the researchers prepared an extensive training plan with graphical aids to support controllers during simulation, most participants reported that more training time was needed. In Experiment 1, the simulation sector was very large and consisted of different procedures for HPR lane usage within the sector. This simulation environment may have been too complex given the limited training time. In Experiment 2, the researchers limited the sector complexity by reducing the sector size and using only a single procedure for HPR lane usage (destination, speed, or OPD equipment). In addition, researchers simplified the sector by simulating overflight operations, instead of the more complex arrival operations from Experiment 1.

4.2 Experiment 2

The results of Experiment 2 indicated that the HPRs decreased aircraft flight distances and time in the sector compared to the jet routes condition under high-traffic levels. However, aircraft flight distances and times were similar across conditions under medium-traffic levels. This result suggests that HPRs may be more beneficial in high-traffic conditions, where more aircraft have the opportunity to take advantage of the direct routes. This result is in contrast with Experiment 1, where the HPRs did not show any distance or time benefits in high-traffic conditions. However, the HPR procedures were more complex and the sector was much larger in Experiment 1. Also, the arrival procedures made operations more difficult in Experiment 1 compared to the overflight procedures in Experiment 2.

Unlike in Experiment 1, there were no differences in the number of altitude, speed, and heading commands for the HPRs and jet routes conditions. As expected, in Experiment 2, controllers issued fewer overall control instructions because of the less complex overflight operations. There was no need to descend aircraft and less "compression" of the traffic flow due to arrival procedures. In fact, there was little need for any control instructions to keep aircraft separated. In these overflight operations, the researchers expected that using HPR lanes for aircraft flying at different speeds would show benefits, but controllers issued very few speed instructions. It may be that there were not enough overtake situations in the traffic scenarios to cause conflicts on the jet routes and present an advantage for HPR speed lanes.

Similar to Experiment 1, there were no differences between HPRs and jet routes in terms of the airspace capacity measures. Again, this is likely due to the limitations of the simulation. In Experiment 2, there was only one incident that led to a loss of aircraft separation and the incident occurred in the jet routes condition. There were few differences in the number of JEDI/URET alerts generated by the system. However, the HPR speed lane usage showed fewer red alerts compared to destination and OPD equipment HPR lane usage.

Overall, controller workload ratings were no different between HPRs and jet routes. This result differs from Experiment 1, where workload ratings were higher using HPRs. The result suggests that the workload differences are minimal with the less complex operations. However, an individual analysis of the workload ratings indicated that there were a few controllers who still felt that workload was higher using HPRs. The questionnaire results indicated that there were no longer any differences in situation awareness and performance ratings for HPRs and jet routes.

The results of Experiment 2 indicate that there were some benefits of HPRs mostly due to decreased aircraft flight distances and time in sector. However, the most remarkable results were that controllers were more comfortable with the HPR procedures in the second experiment. The controllers' subjective ratings of workload, situation awareness, and performance using HPRs were no different from jet routes. This suggests that the controllers were becoming more fluent using HPRs because of their experience in Experiment 1. In addition, the less complex overflight operations, the reduced sector size, and the simplified procedures for HPR lane usage helped controllers to use the HPR concept with less difficulty.

4.3 Experiment 3

The primary goal of the third experiment was to demonstrate that controllers could safely operate a relatively simple sector without training to evaluate the Generic Sector concept. The results of the experiment give some support for this concept. The controllers performed four 30-minute traffic scenarios using the same ZKC sector. The researchers expected that controller performance would rapidly improve with each traffic scenario indicating that the controllers were able to quickly learn the operations in the generic sector. The results indicated that there were no significant differences between any of the scenarios in terms of the efficiency, capacity, and safety simulation measures. This suggests that either the controllers did not receive enough experience with the sector to improve performance or that controllers did not need much experience to control traffic in the sector effectively. Although the simulation measures appeared to indicate that controllers were able to handle traffic operations safely, we were unable to compare the results of our simulation with actual field data for the same measures. This comparison could determine whether controllers who were unfamiliar with the generic sector could operate traffic as effectively as experienced controllers.

The results of the controllers' subjective ratings and questionnaire responses indicated support for the generic sector concept. The controllers' workload ratings decreased slightly over the four scenarios and the difference between the first and third scenarios was statistically significant. Although there were no differences in controllers' ratings of situation awareness and performance, the participants reported that they were able to learn the sector quickly and to operate traffic safely. The OTS observers confirmed that the controllers were able to learn the sector and safely control traffic.

The CIT was designed to support controllers and provide readily accessible information about the generic sector. Feedback indicated that the CIT has the potential to be useful for learning sector information. On the other hand, controllers encountered usability issues that decreased the CIT's effectiveness and, consequently, the amount that controllers used the tool. Multiple participants commented that they enjoyed having "information readily available on the scope. One front-line manager (who was an OTS observer in the study) said, "I did not like that it was not already in the field. This is a great tool. I like that it can be easily toggled on and off as needed." Despite the positive feedback regarding the usefulness of the tool, controllers did not use the tool as much as expected. The most frequently used CIT function was the sector information button showing sector boundaries, numbers, altitudes, and radio frequencies. Although researchers designed the CIT to easily show and hide information on the radar display with a toggle button, few controllers used this feature. Instead, controllers either set up their radar display with the sector information and never toggled it off or did not use the CIT at all. Many controllers reported that the information was too bright and lacked a dimmer switch, which would have made the CIT more useful. Some controllers commented that the ERAM drawing tools can be used to show sector information just as easily and allowed for adjustable brightness settings. In addition, the CIT interface required multiple button clicks to level down to the functions that controllers wanted to display. The controllers suggested that an interface with faster access to function buttons would improve the CIT. The positive feedback from controllers and front-line managers highlights the potential usefulness of the CIT and suggests that there is a need for such a tool—as it would be particularly helpful for learning sector information. However, the small amount of use and the feedback regarding the tool's design reflects that the CIT requires further research and development to correct the usability issues and improve its effectiveness.

THIS PAGE IS BLANK INTENTIONALLY.

References

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

Federal Aviation Administration. (2012). *Operational concept for mid-term high altitude airspace* [Version 2.0, Coordination copy]. Washington, DC.

Joint Planning and Development Office. (2007). *Concept of operations for the Next Generation Air Transportation System* [Version 2.0]. Retrieved from http://www.jpdo.gov/library/NextGen_v2.0.pdf

Mogford, R. (2010). Generic airspace concepts and research. In *Proceedings of the 10th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference,* Orlando, FL.

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 Technical Center.

Stein, E. S. (1985). *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. (1991). *Air traffic controller memory: A field survey* (DOT/FAA/CT-TN90/60). Atlantic City International Airport, NJ: Federal Aviation Administration Technical Center.

Stein, E. S., & Buckley, E. P. (1992). *Human factors at the FAA Technical Center: Bibliography 1958-1994* (DOT/FAA/CT-TN94/50). Atlantic City International Airport, NJ: Federal Aviation Administration 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

AOS	Area of Specialization
ARTCC	Air Route Traffic Control Center
ATC	Air Traffic Control
ATM	Air Traffic Management
ATWIT	Air Traffic Workload Input Technique
CIT	Controller Information Tool
CPC	Certified Professional Controller
Data Comm	Data Communications
DESIREE	Distributed Environment for Simulation, Rapid Engineering, and Experimentation
D-side	Data-side
ERAM	En Route Automation Modernization
ERP	Engineering Research Psychologist
FAA	Federal Aviation Administration
HA	High Altitude
HITL	Human-In-The-Loop
HPR	High Performance Route
JEDI	Java En Route Development Initiative
LOA	Letter of Agreement
MAP	Monitor Alert Parameter
MITRE	Massachusetts Institute of Technology Research and Engineering Corporation
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
NextGen	Next Generation Air Transportation System
OPD	Optimum Profile Descent
OTS	Over-The-Shoulder
PSQ	Post-Scenario Questionnaire
PTT	Push-To-Talk
RDHFL	Research Development and Human Factors Laboratory
RNAV	Area Navigation
R-side	Radar-side
SME	Subject Matter Expert
TGF	Target Generation Facility
URET	User Request Evaluation Tool
VSCS	Voice Switching Control System
WAK	Workload Assessment Keypad
WJHTC	William J. Hughes Technical Center
ZKC	Kansas City ARTCC
ZOB	Cleveland ARTCC

Appendix A: Informed Consent Statement

Informed Consent Statement

I, ______________________________, understand that this study, entitled "The High Altitude Airspace Study" is sponsored by the Federal Aviation Administration (FAA) and is being directed by Dr. Randy Sollenberger.

Nature and Purpose:

I have been recruited to volunteer as a participant in this project. The purpose of the study is to investigate two Next Generation Air Transportation System (NextGen) concepts for High Altitude (HA) Airspace. The two concepts we will investigate are High Performance Routes (HPR) and Generic Sectors in a high-fidelity, human-in-the-loop simulation. The researchers will use the results of the study to evaluate the operational viability of the concepts and to identify human performance issues.

Experimental Procedures:

A group of four controllers will be released from their operational facility for two weeks to participate in the HA Airspace Study and the Conflict Resolution Advisory Study. In the first week, the controllers will travel to the FAA William J. Hughes Technical Center (WJHTC) on Monday and participate in the HA Airspace Study. The controllers will stay over the weekend on off-duty travel. In the second week, the controllers will participate in the Conflict Resolution Advisory Study and depart on Friday.

On Tuesday, Wednesday, Thursday, and Friday of the first week the controllers will participate in the study and perform air traffic scenarios in our laboratory's ATC simulator. The participants will work from 8:00 AM to 4:30 PM each day with a rest break after each traffic scenario and a midday lunch break. At the end of each day, we will have a group meeting to answer the participants' questions and discuss their experiences in the simulation. On the first day of the study, we will brief the participants about the project goals and sectors they will be operating in the simulation. On the last day of the High Altitude study, we will conduct an exit briefing to gather feedback from participants about the entire study.

The study will consist of three experiments to investigate HPRs and Generic Sectors in HA Airspace. The study will use different sectors and air traffic scenarios for each experiment. The participants will perform several practice scenarios in each sector before starting the experimental scenarios. In the first experiment, the controllers will operate a ZOB sector to compare a baseline condition using jet routes to an HPR condition. The second experiment will use another ZOB sector to compare jet routes to three alternative lane usages for HPRs. The last experiment will use two ZKC sectors to demonstrate the Generic Sector concept. In the first experiment, the controllers will operate in R-side/D-side teams. In the last two experiments, the controllers will each operate an R-side only position.

After each test scenario, the controllers will complete a questionnaire to evaluate their performance, workload, and situation awareness. In addition, subject matter experts will make over-the-shoulder observations during the simulation to evaluate the effects of the experimental conditions on controller performance. Finally, the simulation software will record aircraft track and status data to produce measures of safety, capacity, efficiency, and communications. We will use the laboratory's audio-visual recording system during the study.

Discomfort and Risks:

I understand that I will not be exposed to any foreseeable risks. The work that I will perform in the study is safe and includes operating traffic scenarios, completing questionnaires, and providing feedback to the researchers about my simulation experience.

Confidentiality:

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

Benefits:

I understand that the only benefit to me is that I will be able to provide the researchers with valuable feedback and insight about my experiences in the simulation. My data will help the FAA to safely implement the NextGen concept examined in the study.

Participant Responsibilities:

I am aware that to participate in this study I must be a certified professional controller who is qualified at an air traffic control facility and holds a current medical certificate. I will control simulated traffic and answer any questions asked during the study to the best of my abilities. I will not discuss the content of the experiment with anyone until the study is completed.

Participant's Assurances:

I understand that my participation in this study is completely voluntary, and I have the freedom to withdraw at any time without penalty. I also understand that the researchers in this study may terminate my participation if they feel this to be in my best interest.

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

Dr. Sollenberger has adequately answered all the questions I have asked about this study. I understand that Dr. Sollenberger or another member of the research team will be available to answer any other questions that I may have as the study proceeds.

If I have questions about this study or need to report any adverse effects from the research procedures, I will contact Dr. Sollenberger at (609) 485-7169.

Compensation and Injury:

I agree to immediately report any injury or suspected adverse effect to Dr. Randy Sollenberger at (609) 485-7169. 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 statement. 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 statement.

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.

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

2. What is your **age**?	_____ years _____ months
3. How long have you worked as an **ATCS (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

Post-Scenario Questionnaire

Instructions:

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

Performance

1. Rate your **overall level of ATC performance** during this scenario.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

2. Rate your **performance** for **identifying aircraft conflicts** during this scenario.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

3. Rate your **performance** for **separating aircraft** during this scenario.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

4. Rate your **performance** for **moving aircraft efficiently and avoiding traffic flow delays**.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

5. Rate your **performance** for **moving aircraft efficiently with minimal fuel consumption**.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

Workload

6. Rate your **overall workload** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

7. Rate your **workload** due to **scanning for aircraft conflicts** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

8. Rate your **workload** due to **separating aircraft effectively** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

9. Rate your **workload** due to **ensuring smooth traffic flow** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

10. Rate your **workload** due to **communicating to pilots** during this scenario.	Extremely Low	①②③④⑤⑥⑦⑧⑨⑩	Extremely High

Situation Awareness

11. Rate your **overall level of situation awareness** during this scenario.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

12. Rate your **situation awareness** for **detecting aircraft conflicts** during this scenario.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

13. Rate your **situation awareness** for **detecting aircraft that are causing traffic flow delays** during this scenario.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

14. Rate your **situation awareness** for **identifying opportunities for efficient aircraft routing** during this scenario.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

15. Rate your **situation awareness** for **detecting pilot errors** during this scenario.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

Airspace Capacity

16. What was the maximum number of aircraft on your frequency at one time?	_______ specify number of aircraft

17. How many more aircraft could you have safely handled on your frequency at one time?	_______ specify number of aircraft

Simulation Pilots and Scenario Difficulty

18. Rate the **performance of the simulation pilots** in terms of their responding to control instructions and providing readbacks.	Extremely Poor	①②③④⑤⑥⑦⑧⑨⑩	Extremely Good

19. Rate the **difficulty** of this scenario.	Not Difficult	①②③④⑤⑥⑦⑧⑨⑩	Extremely Difficult

Experiment 1
HPR Transition to Low Altitude Airspace

20. What effect, if any, did the **Jet Routes or HPR** procedures have on your ability **to separate aircraft safely**?	Negative Effect	①②③④ ⑤ ⑥⑦⑧⑨ | None	Positive Effect

21. What effect, if any, did the **Jet Routes or HPR** procedures have on your ability **to move aircraft efficiently**?	Negative Effect	①②③④ ⑤ ⑥⑦⑧⑨ | None	Positive Effect

22. What effect, if any, did the **Jet Routes or HPR** procedures have on the number of aircraft you could **control**?	Negative Effect	①②③④ ⑤ ⑥⑦⑧⑨ | None	Positive Effect

23. How **difficult** was it to use the **Jet Routes or HPR** procedures?	Not Difficult	①②③④⑤⑥⑦⑧⑨⑩	Extremely Difficult

For HPRs

24. How **difficult** was it **to find the information you needed** to confirm that aircraft were filed on the appropriate HPR lanes?	Not Difficult	①②③④⑤⑥⑦⑧⑨⑩	Extremely Difficult

25. How (or where) did you find the information to confirm that aircraft were filed correctly onto the appropriate speed, destination, and equipment HPR lanes?

____

____

____

26. Briefly describe any procedural or human performance issues you experienced during this scenario.

____

____

____

____

____

Experiment 2
HPR Alternative Lane Usages

20. What effect, if any, did the **Jet Routes or HPR Lane Usage** procedures have on your ability **to separate aircraft safely**?	Negative Effect	①②③④ ⑤ ⑥⑦⑧⑨ None	Positive Effect

21. What effect, if any, did the **Jet Routes or HPR Lane Usage** procedures have on your ability **to move aircraft efficiently**?	Negative Effect	①②③④ ⑤ ⑥⑦⑧⑨ None	Positive Effect

22. What effect, if any, did the **Jet Routes or HPR Lane Usage** procedures have on your ability **to control more aircraft, if necessary**?	Negative Effect	①②③④ ⑤ ⑥⑦⑧⑨ None	Positive Effect

23. How **difficult** was it to use the **Jet Routes or HPR Lane Usage** procedures?	Not At All Difficult	①②③④⑤⑥⑦⑧⑨⑩	Extremely Difficult

For HPRs

24. How **difficult** was it **to find the information you needed** to sort aircraft onto the HPRs?	Not Difficult	①②③④⑤⑥⑦⑧⑨⑩	Extremely Difficult

25. How would you describe your strategy to sort aircraft onto slow/fast HPR lanes (e.g., cutoff speed, speed differences, when overtaking, only when necessary)?

____

____

____

26. Briefly describe any procedural or human performance issues you experienced during this scenario.

____

____

____

____

____

Experiment 3
Generic Sectors

20. How **well** were you able **to learn the sector characteristics** during this scenario?	Not At All Well	①②③④⑤⑥⑦⑧⑨⑩	Extremely Well

21. How **well** were you able **to learn the traffic patterns** during this scenario?	Not At All Well	①②③④⑤⑥⑦⑧⑨⑩	Extremely Well

22. How **confident** are you that you would be able **to control actual traffic in this sector**?	Not At All Confident	①②③④⑤⑥⑦⑧⑨⑩	Extremely Confident

23. Rate your **change in performance** from the previous scenario; if first scenario please indicate.	First Scenario	Worse	-3 -2 -1 0 +1 +2 +3 Same	Better

24. Did you feel like you were **controlling a safe sector**?	① Yes	② No

25. Briefly describe any sector learning or knowledge issues you experienced during this scenario.

____

____

____

____

____

____

____

____

____

____

All Experiments
General Comments

A. Please provide any additional comments or clarifications about your experience in this scenario.

Appendix D: Exit Questionnaire

Exit Questionnaire

Instructions:

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

Simulation Realism and Research Apparatus Ratings

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

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

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

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

5. Please provide any comments or suggestions for improvement regarding our simulation capabilities.

____

____

____

____

____

____

Experiment 1
HPR Transition to Low Altitude Airspace

6. Which experimental concept was **better for separating aircraft safely**?	① Jet Routes ② HPRs ③ No Difference

7. Which experimental concept was **better for moving aircraft efficiently**?	① Jet Routes ② HPRs ③ No Difference

8. Which experimental concept was **better for controlling more aircraft, if necessary**?	① Jet Routes ② HPRs ③ No Difference

9. Which experimental concept required the **least workload** for you?	① Jet Routes ② HPRs ③ No Difference

10. Which experimental concept was better overall, if any, and why?

____

____

____

____

____

____

Experiment 2
HPR Alternative Lane Usages

11. Which experimental concept was **the best for separating aircraft safely**?	① Jet Routes ② HPRs, Speed ③ HPRs, Destination ④ HPRs, Equipment ⑤ No Difference

12. Which experimental concept was **the best for moving aircraft efficiently**?	① Jet Routes ② HPRs, Speed ③ HPRs, Destination ④ HPRs, Equipment ⑤ No Difference

13. Which experimental concept was **the best for controlling more aircraft, if necessary**?	① Jet Routes ② HPRs, Speed ③ HPRs, Destination ④ HPRs, Equipment ⑤ No Difference

14. Which experimental concept required the **least workload** for you?	① Jet Routes ② HPRs, Speed ③ HPRs, Destination ④ HPRs, Equipment ⑤ No Difference

15. Which experimental concept caused the **most workload** for you?	① Jet Routes ② HPRs, Speed ③ HPRs, Destination ④ HPRs, Equipment ⑤ No Difference

16. Which experimental concept was the best overall, if any, and why?

___

___

___

Experiment 3
Generic Sectors

Controller Information Tool

16. What effect, if any, did the **Controller Information Tool** have on your ability **to quickly learn the sector map**?	Negative Effect	①②③④ ⑤ ⑥⑦⑧⑨ | None	Positive Effect
17. What effect, if any, did the **Controller Information Tool** have on your ability **to quickly learn the traffic routes**?	Negative Effect	①②③④ ⑤ ⑥⑦⑧⑨ | None	Positive Effect
18. What effect, if any, did the **Controller Information Tool** have on your ability **to effectively control traffic in the sector**?	Negative Effect	①②③④ ⑤ ⑥⑦⑧⑨ | None	Positive Effect
19. How difficult was it to use the CIT interface?	Not Difficult	①②③④⑤⑥⑦⑧⑨⑩	Extremely Difficult

20. Were there features of the CIT you preferred to use over others?

21. Was there anything missing from the CIT that you wish was there?

22. Briefly describe what you liked and did not like about the CIT.

All Experiments
General Comments

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

Appendix E: Over-the-Shoulder Observer Rating Form

Over-the-Shoulder Observer Rating Form

Instructions

This form is designed to be used by Supervisory Air Traffic Control Specialists (SATCSs) to evaluate the effectiveness of controllers working in simulation environments. SATCSs will observe and rate the performance of controllers in several different performance dimensions using the scale below as a general purpose guide. Use the entire scale range as much as possible. Take extensive notes on what you see. Do not depend on your memory. Write down your observations. Space is provided after each scale for comments. You may make preliminary ratings during the course of the scenario. However, wait until the scenario is finished before making your final ratings and remain flexible until the end when you have had an opportunity to see all the available behavior. At all times please 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, then you may leave a specific rating blank. Also, please write down any comments that may help improve this evaluation form. Do not write your name on the form itself. You will not be identified by name. An observer code known only to yourself and the researchers conducting this study will be assigned to you. The observations you make do not need to be restricted to the performance areas covered in this form and may include other areas that you think are important.

Assumptions

ATC is a complex activity that contains both observable and unobservable behavior. There are so many complex behaviors involved that no observational rating form can cover everything. A sample of the behaviors is the best that can be achieved, and a good form focuses on those behaviors that controllers themselves have identified as the most relevant in terms of their overall 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.

Rating Scale Descriptors

Remove this Page and keep it available while doing ratings

Scale	Quality	Supplementary
1	**Least Effective**	Unconfident, Indecisive, Inefficient, Disorganized, Behind the power curve, Rough, Leaves some tasks incomplete, Makes mistakes
2	**Poor**	May issue conflicting instructions, Doesn't plan completely
3	**Fair**	Distracted between tasks
4	**Low Satisfactory**	Postpones routine actions
5	**High Satisfactory**	Knows the job fairly well
6	**Good**	Works steadily, Solves most problems
7	**Very Good**	Knows the job thoroughly, Plans well
8	**Most Effective**	Confident, Decisive, Efficient, Organized, Ahead of the power curve, Smooth, Completes all necessary tasks, Makes no mistakes

I - MAINTAINING SAFE AND EFFICIENT TRAFFIC FLOW
II - MAINTAINING ATTENTION AND SITUATION AWARENESS
III – PRIORITIZING
IV – PROVIDING CONTROL INFORMATION
V – TECHNICAL KNOWLEDGE
VI – COMMUNICATING

I - Maintaining Safe and Efficient Traffic Flow

1. Maintaining Separation and Resolving Potential Conflicts.................. 1 2 3 4 5 6 7 8
    - 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
2. Sequencing Aircraft Efficiently.. 1 2 3 4 5 6 7 8
    - using efficient and orderly spacing techniques for arrival, departure, and en route aircraft
    - maintaining safe arrival and departure intervals that minimize delays
3. Using Control Instructions Effectively/Efficiently.............................. 1 2 3 4 5 6 7 8
    - 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
4. Overall Safe and Efficient Traffic Flow Scale Rating........................... 1 2 3 4 5 6 7 8

II - Maintaining Attention and Situation Awareness

5. Maintaining Awareness of Aircraft Positions.. 1 2 3 4 5 6 7 8
    - avoiding fixation on one area of the radar scope when other areas need attention
    - using scanning patterns that monitor all aircraft on the radar scope
6. Giving and Taking Handoffs in a Timely Manner................................ 1 2 3 4 5 6 7 8
    - ensuring that handoffs are initiated in a timely manner
    - ensuring that handoffs are accepted in a timely manner
    - ensuring that handoffs are made according to procedures
7. Ensuring Positive Control.. 1 2 3 4 5 6 7 8
    - tailoring control actions to situation
    - using effective procedures for handling heavy, emergency, and unusual traffic situations
8. Detecting Pilot Deviations from Control Instructions........................ 1 2 3 4 5 6 7 8
    - ensuring that pilots follow assigned clearances correctly
    - correcting pilot deviations in a timely manner
9. Correcting Own Errors in a Timely Manner... 1 2 3 4 5 6 7 8
    - acting quickly to correct errors
    - changing an issued clearance when necessary to expedite traffic flow
10. Overall Attention and Situation Awareness Scale Rating.................... 1 2 3 4 5 6 7 8

III – PRIORITIZING

11.	Taking Actions in an Appropriate Order of Importance • resolving situations that need immediate attention before handling low priority tasks • issuing control instructions in a prioritized, structured, and timely manner	**1 2 3 4 5 6 7 8**
12.	Preplanning Control Actions • scanning adjacent sectors to plan for future and conflicting traffic	**1 2 3 4 5 6 7 8**
13.	Handling Control Tasks for Several Aircraft • shifting control tasks between several aircraft when necessary • communicating in timely fashion while sharing time with other actions	**1 2 3 4 5 6 7 8**
14.	Overall Prioritizing Scale Rating	**1 2 3 4 5 6 7 8**

IV – PROVIDING CONTROL INFORMATION

15.	Providing Essential Air Traffic Control Information • providing mandatory services and advisories to pilots in a timely manner • exchanging essential information	**1 2 3 4 5 6 7 8**
16.	Providing Additional Air Traffic Control Information • providing additional services when workload permits • exchanging additional information	**1 2 3 4 5 6 7 8**
17.	Providing Coordination • providing effective and timely coordination • using proper point-out procedures	**1 2 3 4 5 6 7 8**
18.	Overall Providing Control Information Scale Rating	**1 2 3 4 5 6 7 8**

V – TECHNICAL KNOWLEDGE

19. Showing Knowledge of LOAs and SOPs **1 2 3 4 5 6 7 8**
    - controlling traffic as depicted in current LOAs and SOPs
    - performing handoff procedures correctly

20. Showing Knowledge of Aircraft Capabilities and Limitations **1 2 3 4 5 6 7 8**
    - using appropriate speed, vectoring, and/or altitude assignments to separate aircraft with varied flight capabilities
    - issuing clearances that are within aircraft performance parameters

21. Showing Effective Use of Equipment **1 2 3 4 5 6 7 8**
    - updating data blocks
    - using equipment capabilities

22. Overall Technical Knowledge Scale Rating **1 2 3 4 5 6 7 8**

VI – COMMUNICATING

23. Using Proper Phraseology **1 2 3 4 5 6 7 8**
    - using words and phrases specified in the 7110.65
    - using phraseology that is appropriate for the situation
    - using minimum necessary verbiage

24. Communicating Clearly and Efficiently **1 2 3 4 5 6 7 8**
    - speaking at the proper volume and rate for pilots to understand
    - speaking fluently while scanning or performing other tasks
    - ensuring clearance delivery is complete, correct and timely
    - speaking with confident, authoritative tone of voice

25. Listening to Pilot Readbacks and Requests **1 2 3 4 5 6 7 8**
    - correcting pilot readback errors
    - acknowledging pilot or other controller requests promptly
    - processing requests correctly in a timely manner

26. Overall Communicating Scale Rating **1 2 3 4 5 6 7 8**

HPR Specific Rating Considerations

1. How effective was the participant's HPR usage?	Not Effective	①②③④⑤⑥⑦⑧⑨⑩	Extremely Effective

2. Please characterize the participant's HPR usage strategy (e.g., use of fast/slow lanes, fixes used, what information they used)

3. How well did participants follow the LOAs for the sector?	Not Well	①②③④⑤⑥⑦⑧⑨⑩	Extremely Well

4. Explain any problems in the use of LOAs

Appendix F: Post-Scenario Questionnaire and OTS Observer Rating Form Responses

Table F1. Means (*SD*) of Participants' PSQ Responses by Position and Condition for Experiment 1

Question:	**R-side**			**D-side**		
	HPRs	**Jet Routes**	***p*-value**	**HPRs**	**Jet Routes**	***p*-value**
1. Overall ATC **performance**	6.67 (2.10)	7.92 (1.31)	0.09†	7.00 (2.26)	8.25 (1.36)	0.02*
2. **Performance** for identifying aircraft conflicts	5.92 (2.47)	8.50 (1.17)	0.01*	7.42 (2.47)	8.67 (1.07)	0.08†
3. **Performance** for separating aircraft	6.83 (2.08)	8.25 (1.14)	0.07†	7.75 (2.53)	8.50 (1.24)	0.33
4. **Performance** for moving aircraft efficiently and avoiding traffic flow delays	7.08 (1.31)	7.83 (1.70)	0.04*	6.83 (2.86)	7.92 (2.07)	0.053†
5. **Performance** for moving aircraft efficiently with minimal fuel consumption	6.75 (1.76)	7.25 (2.26)	0.29	6.83 (3.10)	7.08 (2.64)	0.60
6. Overall **workload**	7.75 (1.36)	6.42 (1.62)	0.02*	6.42 (2.27)	4.42 (2.61)	0.004*
7. **Workload** due to scanning for aircraft conflicts	6.92 (1.98)	6.25 (1.96)	0.10	6.83 (2.33)	4.92 (2.61)	0.004*
8. **Workload** due to separating aircraft effectively	6.67 (1.23)	6.17 (1.40)	0.27	5.92 (2.02)	3.92 (2.64)	0.04*
9. **Workload** due to ensuring smooth traffic flow	7.25 (1.06)	6.25 (1.54)	0.09†	6.25 (2.99)	3.58 (2.31)	0.01*
10. **Workload** due to communicating to pilots	7.08 (2.15)	5.83 (2.12)	0.08†	4.83 (3.51)	4.00 (2.89)	0.48
11. Overall **situation awareness**	6.75 (1.76)	7.92 (1.16)	0.07†	6.42 (2.61)	8.17 (1.53)	0.02*
12. **Situation awareness** for detecting aircraft conflicts	6.25 (1.82)	8.17 (0.94)	0.01*	7.08 (2.64)	8.25 (1.48)	0.09†
13. **Situation awareness** for detecting aircraft causing traffic flow delays	6.73 (2.05)	7.91 (1.38)	0.10	7.27 (2.10)	8.00 (1.67)	0.31

(table continues)

Table F1. Means (*SD*) of Participants' PSQ Responses by Position and Condition for Experiment 1 (continued)

Question:	R-side			D-side		
	HPRs	**Jet Routes**	***p*-value**	**HPRs**	**Jet Routes**	***p-value***
14. **Situation awareness** for identifying opportunities for efficient aircraft routing	6.58 (1.68)	7.58 (1.73)	0.02*	6.58 (2.47)	7.67 (2.02)	0.07†
15. **Situation awareness** for detecting pilot errors	7.64 (1.21)	7.91 (2.26)	0.68	7.27 (2.37)	8.00 (2.05)	0.38
16. Maximum number of aircraft on your frequency	22.73 (5.53)	21.55 (5.94)	0.34	24.09 (5.09)	22.09 (4.66)	0.06†
17. How many more aircraft could you have safely handled on your frequency	4.42 (3.99)	7.25 (5.83)	0.03*	5.50 (5.82)	6.92 (4.03)	0.42
18. Performance of simulation pilots	5.50 (1.98)	6.50 (1.83)	0.32	5.75 (2.14)	6.33 (1.61)	0.36
19. Overall **difficulty**	7.67 (0.89)	6.00 (1.81)	0.002*	7.67 (1.83)	5.17 (1.95)	0.001*
20**. Effect of Jet routes or HPRs** on separating aircraft safely	4.82 (2.32)	5.18 (1.40)	0.67	4.45 (2.21)	5.27 (0.79)	0.24
21. **Effect of Jet routes or HPRs** on moving aircraft efficiently	5.45 (1.97)	4.82 (1.33)	0.43	5.09 (2.55)	5.00 (1.00)	0.91
22. **Effect of Jet routes or HPRs** on number of aircraft you could control	4.45 (2.30)	5.27 (1.62)	0.42	4.73 (2.15)	5.00 (0.89)	0.64
23. Difficulty of using **Jet Route or HPR procedures**	5.73 (2.28)	2.36 (1.75)	0.002*	5.82 (2.82)	3.36 (1.57)	0.02*
24. How difficult to find information to confirm aircraft were filed on appropriate HPR lane	5.27 (2.53)	---	---	5.82 (2.99)	---	---

Note. $^{*}p < .05$. $^{\dagger}p < .10$.

Table F2. Means (*SD*) of OTS Observers' PSQ Responses by Condition for Experiment 1

PSQ Question:	HPRs	Jet Routes	*p*-value
1. Overall ATC **performance**	7.75 (1.83)	8.46 (0.86)	0.14
2. **Performance** for identifying aircraft conflicts	7.61 (1.80)	8.60 (1.02)	0.047*
3. **Performance** for separating aircraft	8.07 (2.00)	8.93 (0.76)	0.08†
4. **Performance** for moving aircraft efficiently and avoiding traffic flow delays	8.07 (1.45)	8.38 (1.04)	0.45
5. **Performance** for moving aircraft efficiently with minimal fuel consumption	8.00 (1.49)	8.18 (1.15)	0.68
6. Overall **workload**	7.82 (1.31)	7.13 (1.04)	0.049*
7. **Workload** due to scanning for aircraft conflicts	7.71 (1.22)	7.03 (1.06)	0.051†
8. **Workload** due to separating aircraft effectively	7.31 (1.30)	6.57 (1.19)	0.07†
9. **Workload** due to ensuring smooth traffic flow	7.81 (1.27)	7.00 (1.39)	0.01*
10. **Workload** due to communicating to pilots	7.63 (1.60)	6.56 (1.33)	0.03*
11. Overall **situation awareness**	7.61 (1.98)	8.22 (1.01)	0.20
12. **Situation awareness** for detecting aircraft conflicts	7.58 (2.09)	8.51 (0.85)	0.08†
13. **Situation awareness** for detecting aircraft causing traffic flow delays	8.11 (1.23)	8.40 (0.93)	0.45
14. **Situation awareness** for identifying opportunities for efficient aircraft routing	8.13 (1.42)	8.33 (0.94)	0.59
15. **Situation awareness** for detecting pilot errors	8.33 (1.62)	8.89 (0.58)	0.21
16. Maximum number of aircraft on your frequency	28.83 (3.71)	28.07 (3.68)	0.16
17. How many more aircraft could you have safely handled on your frequency	4.67 (4.23)	6.64 (3.77)	0.02*
18. Performance of simulation pilots	5.96 (2.24)	7.10 (0.93)	0.09†
19. Overall **difficulty**	7.79 (1.12)	6.78 (1.07)	0.001*

(table continues)

Table F2. Means (*SD*) of OTS Observers' PSQ Responses by Condition for Experiment 1 (continued)

PSQ Question:	**HPRs**	**Jet Routes**	***p*-value**
20**. Effect of Jet Routes or HPRs** on separating aircraft safely	6.03 (1.09)	4.94 (0.42)	0.008*
21. **Effect of Jet Routes or HPRs** on moving aircraft efficiently	6.50 (0.77)	4.61 (0.53)	< 0.001*
22. **Effect of Jet Routes or HPRs** on number of aircraft you could control	6.40 (1.05)	4.97 (0.55)	0.001*
23. Difficulty of using **Jet Route or HPR procedures**	4.71 (1.85)	4.51 (1.09)	0.68
24. How difficult to find information to confirm aircraft were filed on appropriate HPR lane	4.96 (1.64)	---	---

Note. $^{*}p < .05$. $^{\dagger}p < .10$.

Table F3. Means (*SD*) of OTS Observers' OTS Observer Rating Form Responses by Condition for Experiment 1

OTS Observer Rating Form Question:	HPRs	Jet Routes	*p*-value
1. Maintaining separation and resolving potential conflicts	6.33 (1.43)	6.97 (0.57)	0.13
2. Sequencing aircraft efficiently	6.47 (1.14)	6.89 (0.67)	0.32
3. Using control instructions effectively/efficiently	6.65 (1.04)	7.01 (0.59)	0.25
4. Overall safe and efficient traffic flow	6.47 (1.34)	7.01 (0.51)	0.21
5. Maintaining awareness of aircraft positions	6.53 (0.88)	6.86 (0.52)	0.28
6. Giving and taking handoffs in a timely manner	6.39 (0.98)	6.72 (0.91)	0.34
7. Ensuring positive control	6.67 (1.38)	7.13 (0.55)	0.18
8. Detecting pilot deviations from control instructions	6.83 (1.14)	7.24 (0.44)	0.19
9. Correcting own errors in a timely manner	6.76 (1.18)	7.24 (0.44)	0.14
10. Overall attention and situation awareness scale rating	6.53 (1.17)	7.13 (0.40)	$0.07^{\dagger}$
11. Taking actions in an appropriate order of importance	6.63 (0.91)	6.97 (0.57)	0.15
12. Preplanning control actions	6.28 (1.15)	6.67 (0.92)	$0.08^{\dagger}$
13. Handling control tasks for several aircraft	6.74 (1.01)	7.13 (0.53)	0.15
14. Overall prioritizing scale rating	6.61 (0.90)	7.03 (0.45)	$0.09^{\dagger}$
15. Providing essential air traffic control information	6.53 (1.04)	6.94 (0.74)	0.02*
16. Providing additional air traffic control information	6.18 (0.96)	6.53 (0.82)	$0.08^{\dagger}$
17. Providing coordination	6.47 (0.77)	6.97 (0.56)	0.01*
18. Overall providing control information	6.51 (0.76)	7.01 (0.49)	0.01*
19. Showing knowledge of LOAs and SOPs	6.61 (0.81)	6.82 (0.67)	0.46

(table continues)

Table F3. Means (*SD*) of OTS Observers' OTS Observer Rating Form Responses by Condition for Experiment 1 (continued)

OTS Observer Rating Form Question:	HPRs	Jet Routes	*p*-value
20. Showing knowledge of aircraft capabilities and limitations	6.03 (1.09)	4.94 (0.42)	0.01*
21. Showing effective use of equipment	6.97 (0.82)	7.11 (0.67)	0.52
22. Overall technical knowledge scale rating	6.97 (0.68)	7.10 (0.62)	0.49
23. Using proper phraseology	6.74 (0.93)	6.97 (0.92)	0.19
24. Communicating clearly and efficiently	7.10 (0.69)	7.22 (0.66)	0.20
25. Listening to pilot readbacks and requests	6.86 (1.26)	7.21 (0.61)	0.12
26. Overall communicating scale rating	6.99 (0.97)	7.15 (0.83)	0.23
27. How effective was the participant's HPR usage	8.08 (1.37)	---	---
28. How well did participants follow the LOAs for the sector	8.18 (1.25)	8.68 (0.89)	0.29

Note. $^{*}p < .05$. $^{\dagger}p < .10$.

Table F4. Means (*SD*) of Participants' PSQ Responses by Condition for Experiment 2

Question:	HPRs, Destination	HPRs, Speed	HPRs, OPD Equipment	Jet Routes	*p*-value
1. Overall ATC **performance**	7.83 (2.04)	7.58 (2.19)	7.25 (1.71)	7.00 (2.37)	0.76
2. **Performance** for identifying aircraft conflicts	7.75 (2.09)	7.33 (2.27)	7.33 (1.87)	7.75 (2.60)	0.91
3. **Performance** for separating aircraft	8.00 (2.13)	7.58 (2.27)	7.75 (1.86)	8.08 (2.50)	0.93
4. **Performance** for moving aircraft efficiently and avoiding traffic flow delays	7.27 (1.56)	7.64 (1.91)	6.82 (2.27)	7.55 (1.37)	0.67
5. **Performance** for moving aircraft efficiently with minimal fuel consumption	7.42 (1.83)	7.50 (1.98)	6.42 (1.93)	6.58 (2.15)	0.32
6. Overall **workload**	7.75 (1.36)	7.50 (2.07)	7.83 (1.27)	6.92 (2.68)	0.48
7. **Workload** due to scanning for aircraft conflicts	7.58 (1.51)	7.17 (2.48)	7.92 (1.24)	6.83 (1.85)	0.47
8. **Workload** due to separating aircraft effectively	6.83 (2.25)	6.67 (2.02)	8.17 (1.03)	7.22 (1.13)	0.19
9. **Workload** due to ensuring smooth traffic flow	6.42 (2.35)	6.58 (1.73)	6.83 (2.69)	6.17 (3.10)	0.89
10. **Workload** due to communicating to pilots	6.25 (2.18)	6.33 (2.27)	6.17 (2.29)	6.33 (2.46)	0.99
11. Overall **situation awareness**	7.58 (1.62)	7.25 (2.49)	7.42 (1.08)	7.00 (2.56)	0.89
12. **Situation awareness** for detecting aircraft conflicts	7.25 (1.76)	7.00 (2.41)	7.75 (1.48)	7.50 (2.35)	0.79
13. **Situation awareness** for detecting aircraft causing traffic flow delays	7.27 (1.62)	7.64 (1.69)	6.73 (1.56)	6.36 (2.54)	0.40
14. **Situation awareness** for identifying opportunities for efficient aircraft routing	6.75 (2.09)	7.25 (1.76)	6.92 (1.98)	6.17 (2.04)	0.49
15. **Situation awareness** for detecting pilot errors	7.67 (1.72)	7.17 (2.33)	7.33 (2.39)	6.92 (2.35)	0.81
16. Maximum number of aircraft on your frequency	25.91 (5.28)	26.55 (6.77)	28.00 (5.33)	26.82 (5.22)	0.31
17. How many more aircraft could you have safely handled on your frequency	7.25 (6.05)	7.17 (6.28)	3.92 (4.48)	8.50 (7.32)	0.14

(table continues)

Table F4. Means (*SD*) of Participants' PSQ Responses by Condition for Experiment 2 (continued)

Question:	**HPRs, Destination**	**HPRs, Speed**	**HPRs, OPD Equipment**	**Jet Routes**	***p*-value**
18. Performance of simulation pilots	7.08 (1.68)	6.42 (2.19)	6.08 (2.07)	6.50 (2.68)	0.68
19. Overall **difficulty**	6.92 (1.83)	6.67 (1.87)	7.33 (1.50)	6.17 (2.79)	0.36
20**. Effect of Jet Routes or HPR Lane Usage** procedures on separating aircraft safely	5.36 (1.63)	5.73 (1.19)	5.36 (1.21)	4.64 (1.21)	0.15
21. **Effect of Jet Routes or HPR Lane Usage** procedures on moving aircraft efficiently	5.82 (2.32)	6.27 (1.42)	5.36 (1.57)	4.91 (1.43)	0.22
22. **Effect of Jet Routes or HPR Lane Usage** procedures on controlling more aircraft, if necessary	5.82 (1.89)	6.27 (1.68)	4.91 (1.58)	4.82 (1.33)	0.08[†]
23. Difficulty of using **Jet Routes or HPR Lane Usage procedures**	3.90 (2.73)	3.90 (2.60)	3.80 (2.44)	2.80 (2.15)	0.49
24. How difficult to find information to sort aircraft onto HPRs	3.10 (2.69)	3.90 (2.81)	3.40 (1.84)	---	0.58

Table F5. Means (*SD*) of OTS Observers' PSQ Responses by Condition for Experiment 2

PSQ Question:	HPRs, Destination	HPRs, Speed	HPRs, OPD Equipment	Jet Routes	*p*-value
1. Overall ATC performance	7.92 (2.75)	8.58 (1.56)	8.42 (2.23)	8.42 (2.23)	0.88
2. Performance for identifying aircraft conflicts	8.58 (2.27)	8.33 (1.92)	8.75 (2.26)	8.92 (1.98)	0.93
3. Performance for separating aircraft	8.92 (1.73)	8.67 (2.02)	8.83 (1.99)	9.00 (1.60)	0.98
4. Performance for moving aircraft efficiently and avoiding traffic flow delays	8.25 (2.60)	8.33 (1.67)	8.58 (1.78)	8.08 (1.26)	0.95
5. Performance for moving aircraft efficiently with minimal fuel consumption	8.67 (1.78)	8.25 (1.76)	8.58 (1.31)	7.50 (3.40)	0.56
6. Overall workload	7.83 (1.80)	7.25 (1.82)	7.58 (1.78)	7.92 (1.93)	0.65
7. Workload due to scanning for aircraft conflicts	7.83 (1.80)	7.42 (2.15)	7.67 (2.27)	7.33 (2.02)	0.87
8. Workload due to separating aircraft effectively	7.08 (1.98)	7.25 (1.91)	7.33 (2.23)	7.25 (1.91)	0.99
9. Workload due to ensuring smooth traffic flow	8.17 (1.47)	7.33 (2.10)	7.67 (1.78)	7.17 (2.04)	0.35
10. Workload due to communicating to pilots	6.92 (2.15)	6.42 (2.19)	6.50 (2.43)	6.58 (2.64)	0.91
11. Overall situation awareness	7.58 (2.50)	8.33 (2.02)	8.33 (2.27)	8.42 (1.78)	0.67
12. Situation awareness for detecting aircraft conflicts	8.08 (2.19)	8.75 (1.91)	8.67 (2.31)	8.83 (1.40)	0.78

(table continues)

Table F5. Means (*SD*) of OTS Observers' PSQ Responses by Condition for Experiment 2 (continued)

PSQ Question:	HPRs, Destination	HPRs, Speed	HPRs, OPD Equipment	Jet Routes	*p*-value
13. Situation awareness for detecting aircraft causing traffic flow delays	8.58 (1.83)	8.33 (2.02)	8.75 (1.36)	9.00 (1.41)	0.78
14. Situation awareness for identifying opportunities for efficient aircraft routing	8.58 (1.88)	8.92 (0.79)	8.67 (1.72)	8.92 (1.51)	0.92
15. Situation awareness for detecting pilot errors	8.42 (2.19)	9.00 (1.48)	8.75 (1.29)	9.17 (1.19)	0.61
16. Maximum number of aircraft on your frequency	29.42 (6.02)	29.00 (5.39)	29.83 (5.39)	29.75 (5.43)	0.51
17. How many more aircraft could you have safely handled on your frequency	5.18 (4.79)	4.64 (2.77)	3.91 (3.94)	4.36 (6.14)	0.77
18. Performance of simulation pilots	7.75 (1.96)	8.17 (1.11)	8.08 (1.98)	7.58 (2.50)	0.78
19. Overall difficulty	7.33 (1.61)	6.83 (2.21)	7.42 (1.83)	7.42 (1.83)	0.76
20. Effect of Jet Routes or HPR Lane Usage procedures on separating aircraft safely	6.33 (1.30) ^{1}	7.25 (1.48) ^{2}	6.25 (1.82)	4.67 (1.23) ^{12}	0.002*
21. Effect of Jet Routes or HPR Lane Usage procedures on moving aircraft efficiently	6.58 (1.44) ^{1}	7.50 (1.24) ^{2}	6.75 (1.66)	5.00 (1.41) ^{12}	0.001*
22. Effect of Jet Routes or HPR Lane Usage procedures on controlling more aircraft, if necessary	6.17 (1.99)	7.17 (1.34) ^{1}	6.42 (1.98)	4.92 (1.56) ^{1}	0.01*
23. Difficulty of using Jet Routes or HPR Lane Usage procedures	3.67 (2.19)	3.67 (1.78)	5.08 (2.23)	3.67 (2.53)	0.24
24. How difficult to find information to sort aircraft onto HPRs	3.20	3.00	5.30	---	0.06†

Note. *$p < .05$. $^{\dagger}p < .10$. The superscript "1" indicates that the condition is significantly different from HPRs by destination, "2" is significantly different from HPRs by speed, "3" is significantly different from HPRs by OPD equipment, and "4" is significantly different from jet routes.

Table F6. Means (*SD*) of OTS Observers' OTS Observer Rating Form Responses by Condition for Experiment 2

OTS Observer Form Question:	HPRs, Destination	HPRs, Speed	HPRs, OPD Equipment	Jet Routes	*p*-value
1. Maintaining separation and resolving potential conflicts	6.92 (1.31)	7.25 (1.14)	7.33 (1.44)	6.83 (1.70)	0.69
2. Sequencing aircraft efficiently	6.75 (1.91)	7.17 (0.94)	7.08 (0.67)	7.00 (0.67)	0.88
3. Using control instructions effectively/efficiently	6.58 (1.88)	7.17 (0.83)	7.42 (0.90)	6.92 (1.73)	0.41
4. Overall safe and efficient traffic flow	6.75 (1.54)	7.17 (0.94)	6.92 (1.16)	7.00 (1.60)	0.86
5. Maintaining awareness of aircraft positions	6.83 (1.47)	7.00 (1.13)	7.00 (0.74)	7.17 (0.94)	0.82
6. Giving and taking handoffs in a timely manner	6.25 (2.09)	7.25 (0.97)	6.42 (2.07)	6.58 (1.93)	0.22
7. Ensuring positive control	7.17 (1.11)	7.42 (1.00)	7.17 (1.47)	6.92 (1.73)	0.73
8. Detecting pilot deviations from control instructions	7.17 (1.40)	7.33 (0.98)	7.42 (0.79)	7.33 (1.15)	0.95
9. Correcting own errors in a timely manner	7.25 (0.75)	7.33 (0.98)	7.33 (0.78)	6.92 (1.62)	0.72
10. Overall attention and situation awareness scale rating	6.73 (1.49)	7.18 (0.87)	6.91 (1.22)	7.18 (1.60)	0.71
11. Taking actions in an appropriate order of importance	7.08 (1.56)	7.17 (1.03)	7.42 (0.67)	7.50 (0.67)	0.67
12. Preplanning control actions	6.58 (1.62)	6.83 (1.47)	6.75 (1.66)	6.75 (1.60)	0.91
13. Handling control tasks for several aircraft	7.00 (1.54)	7.42 (0.51)	7.33 (0.78)	7.33 (1.07)	0.64
14. Overall prioritizing scale rating	6.91 (1.58)	7.18 (0.98)	7.09 (0.94)	7.18 (1.08)	0.97
15. Providing essential air traffic control information	6.58 (1.88)	7.25 (0.97)	6.83 (1.47)	7.00 (1.71)	0.67
16. Providing additional air traffic control information	6.17 (1.95)	6.67 (1.07)	6.17 (1.59)	6.42 (1.78)	0.75

(table continues)

Table F6. Means (*SD*) of OTS Observers' OTS Observer Rating Form Responses by Condition for Experiment 2 (continued)

OTS Observer Form Question:	HPRs, Destination	HPRs, Speed	HPRs, OPD Equipment	Jet Routes	*p*-value
17. Providing coordination	6.64 (1.91)	7.64 (0.50)	7.45 (0.69)	6.91 (1.76)	0.14
18. Overall providing control information	6.58 (1.78)	7.17 (0.83)	6.67 (1.37)	6.75 (1.66)	0.73
19. Showing knowledge of LOAs and SOPs	7.42 (1.16)	7.42 (0.79)	7.67 (0.65)	7.67 (0.65)	0.79
20. Showing knowledge of aircraft capabilities and limitations	7.27 (1.10)	7.45 (0.69)	7.64 (0.50)	7.82 (0.40)	0.25
21. Showing effective use of equipment	7.42 (0.79)	7.50 (0.67)	7.67 (0.65)	7.58 (0.51)	0.64
22. Overall technical knowledge scale rating	7.42 (0.79)	7.25 (0.75	7.58 (0.67)	7.67 (0.49)	0.25
23. Using proper phraseology	7.33 (0.98)	7.42 (0.67)	7.50 (0.67)	7.08 (1.51)	0.51
24. Communicating clearly and efficiently	7.58 (0.67)	7.58 (0.51)	7.58 (0.67)	7.33 (1.37)	0.53
25. Listening to pilot readbacks and requests	7.25 (1.36)	7.50 (0.67)	7.58 (0.67)	7.08 (1.62)	0.55
26. Overall communicating scale rating	7.33 (0.98)	7.42 (0.67)	7.58 (0.67)	7.25 (1.42)	0.63
27. How effective was the participant's HPR usage	9.10 (1.37)	8.20 (1.62)	8.90 (0.74)	---	0.33
28. How well did participants follow the LOAs for the sector	9.09 (1.51)	8.91 (0.83)	8.73 (1.10)	9.45 (1.21)	0.49

Note. $^{*}p < .05$. $^{\dagger}p < .10$. The superscript "1" indicates that the condition is significantly different from HPRs by destination, "2" is significantly different from HPRs by speed, "3" is significantly different from HPRs by OPD equipment, and "4" is significantly different from jet routes.

Table F7. Means (*SD*) of Participants' PSQ Responses by Scenario Run for Experiment 3

Question:	Run 1	Run 2	Run 3	Run 4	*p*-value
1. Overall ATC **performance**	8.08 (1.56)	7.75 (1.14)	8.67 (1.30)	8.92 (1.08)	0.04*
2. **Performance** for identifying aircraft conflicts	8.50 (1.57)	8.08 (0.90)	9.00 (0.95)	8.75 (1.42)	0.14
3. **Performance** for separating aircraft	8.33 (1.50)	8.25 (0.87)	8.92 (1.08)	8.75 (1.22)	0.31
4. **Performance** for moving aircraft efficiently and avoiding traffic flow delays	7.42 (2.75)	7.00 (2.49)	8.17 (2.44)	7.83 (2.69)	0.24
5. **Performance** for moving aircraft efficiently with minimal fuel consumption	7.00 (3.02)	7.08 (2.47)	7.83 (2.62)	7.17 (2.98)	0.51
6. Overall **workload**	4.92 (3.03)	5.67 (2.42)	4.83 (2.86)	5.58 (2.54)	0.41
7. **Workload** due to scanning for aircraft conflicts	6.17 (2.48)	6.25 (2.63)	5.42 (3.09)	5.67 (2.67)	0.57
8. **Workload** due to separating aircraft effectively	5.25 (2.80)	6.08 (2.50)	4.67 (2.64)	4.33 (2.27)	0.09†
9. **Workload** due to ensuring smooth traffic flow	4.58 (2.91)	4.75 (2.42) [1]	3.25 (2.22) [1]	3.33 (1.67)	0.03*
10. **Workload** due to communicating to pilots	4.92 (2.68)	3.83 (2.41)	4.00 (2.09)	3.75 (1.60)	0.18
11. Overall **situation awareness**	7.92 (2.68)	7.25 (1.76)	7.83 (2.69)	8.50 (1.31)	0.32
12. **Situation awareness** for detecting aircraft conflicts	7.92 (2.61)	7.75 (1.36)	8.58 (1.24)	8.58 (1.16)	0.27
13. **Situation awareness** for detecting aircraft causing traffic flow delays	7.45 (2.62)	7.00 (2.19)	7.36 (2.11)	7.82 (2.14)	0.55
14. **Situation awareness** for identifying opportunities for efficient aircraft routing	6.17 (2.86)	6.42 (2.54)	6.75 (2.86)	7.00 (2.92)	0.55
15. **Situation awareness** for detecting pilot errors	7.33 (2.81)	7.25 (2.49)	7.17 (2.95)	7.67 (2.61)	0.68

(table continues)

Table F7. Means (*SD*) of Participants' PSQ Responses by Scenario Run for Experiment 3 (continued)

Question:	Run 1	Run 2	Run 3	Run 4	*p*-value
16. Maximum number of aircraft on your frequency	20.09 (6.49)	20.91 (4.30)	22.36 (5.59)	23.09 (4.59)	0.82
17. How many more aircraft could you have safely handled on your frequency	9.25 (5.15)	8.58 (5.58)	10.00 (6.18)	10.33 (5.09)	0.28
18. Performance of simulation pilots	7.75 (1.14)	7.50 (1.00)	7.83 (1.27)	7.83 (1.03)	0.81
19. Overall **difficulty**	4.67 (2.96)	5.25 (2.42)	4.08 (2.68)	4.17 (2.25)	0.14
20. How well were you able to **learn the sector characteristics**	5.25 (2.34) ^{13}	6.17 (2.29) ^{2}	7.00 (2.26) ^{3}	7.75 (1.54) ^{12}	< 0.001*
21. How well were you able to **learn traffic patterns**	5.58 (2.19) ^{1}	5.92 (2.35) ^{2}	7.00 (2.52)	7.83 (1.34) ^{12}	< 0.001*
22. How **confident** are you that you would be able to control actual traffic	6.25 (2.73) ^{1}	7.00 (3.05)	7.75 (2.26)	8.33 (1.50) ^{1}	0.003*
23. Rate **change in performance** from previous scenario	---	0.83 (1.34)	1.42 (1.00)	1.67 (1.15)	0.23
24. Did you feel like you were controlling a safe sector? (# of "Yes" responses out of 12)	10	10	11	11	0.39

Note. $^{*}p < .05$. $^{\dagger}p < .10$. The superscript "1" indicates that the condition is significantly different from Run 1, "2" is significantly different from Run 2, "3" is significantly different from Run 3, and "4" is significantly different from Run 4.

Table F8. Means (*SD*) of OTS Observers' PSQ Responses by Scenario Run for Experiment 3

PSQ Question:	Run 1	Run 2	Run 3	Run 4	*p*-value
1. Overall ATC **performance**	8.42 (1.00) [1]	8.92 (1.24)	9.08 (1.16)	9.67 (0.49) [1]	0.03*
2. **Performance** for identifying aircraft conflicts	8.92 (1.31)	9.50 (0.67)	9.58 (0.67)	9.83 (0.39)	0.08†
3. **Performance** for separating aircraft	9.00 (1.54)	9.42 (0.79)	9.75 (0.45)	9.83 (0.39)	0.11
4. **Performance** for moving aircraft efficiently and avoiding traffic flow delays	8.17 (1.59) [1]	9.25 (0.87)	9.42 (0.67)	9.75 (0.45) [1]	< 0.001*
5. **Performance** for moving aircraft efficiently with minimal fuel consumption	8.25 (1.66)	9.17 (1.11)	9.17 (0.94)	9.58 (0.51)	0.01*
6. Overall **workload**	7.00 (1.86)	6.17 (1.90)	5.58 (2.31)	6.33 (2.42)	0.14
7. **Workload** due to scanning for aircraft conflicts	6.92 (1.93)	6.25 (1.91)	6.00 (2.49)	6.75 (2.67)	0.42
8. **Workload** due to separating aircraft effectively	6.17 (1.99)	5.83 (2.25)	5.67 (2.31)	6.33 (2.31)	0.73
9. **Workload** due to ensuring smooth traffic flow	6.67 (2.31)	6.08 (2.19)	5.50 (2.39)	6.08 (2.50)	0.50
10. **Workload** due to communicating to pilots	5.42 (2.43)	5.67 (2.15)	5.17 (2.41)	5.25 (2.77)	0.86
11. Overall **situation awareness**	8.08 (1.31) [1]	8.58 (1.31)	9.17 (0.72)	9.50 (0.52) [1]	0.001*
12. **Situation awareness** for detecting aircraft conflicts	8.58 (1.16) [1]	8.83 (1.27)	9.50 (0.67)	9.75 (0.45) [1]	0.03*

(table continues)

Table F8. Means (*SD*) of OTS Observers' PSQ Responses by Scenario Run for Experiment 3 (continued)

PSQ Question:	Run 1	Run 2	Run 3	Run 4	*p*-value
13. **Situation awareness** for detecting aircraft causing traffic flow delays	7.25 (1.82) ^{12}	8.58 (1.51) ^{1}	8.83 (1.64)	9.42 (0.79) ^{2}	< 0.001*
14. **Situation awareness** for identifying opportunities for efficient aircraft routing	6.83 (2.17) ^{12}	8.25 (2.14)	8.92 (1.44) ^{1}	9.08 (1.38) ^{2}	< 0.001*
15. **Situation awareness** for detecting pilot errors	8.25 (1.06) ^{12}	8.83 (1.19)	9.33 (0.65) ^{1}	9.50 (0.52) ^{2}	0.003*
16. Maximum number of aircraft on your frequency	23.00 (6.45)	23.83 (5.64)	24.25 (5.40)	24.92 (5.55)	0.22
17. How many more aircraft could you have safely handled on your frequency	5.82 (3.92)	10.18 (8.06)	11.00 (8.11)	9.00 (3.90)	0.14
18. Performance of simulation pilots	8.00 (0.85)	8.25 (2.05)	8.25 (1.76)	8.25 (1.54)	0.58
19. Overall **difficulty**	6.58 (2.43)	5.67 (1.87)	5.17 (2.33)	5.42 (2.15)	0.06†
20. How well were you able to **learn the sector characteristics**	6.58 (2.71)	8.00 (1.60)	8.08 (1.78)	8.58 (1.44)	0.04*
21. How well were you able to **learn traffic patterns**	6.42 (2.50)	7.58 (1.44)	8.25 (1.66)	8.58 (1.68)	0.01*
22. How **confident** are you that you would be able to control actual traffic	7.17 (3.43)	8.00 (2.26)	8.42 (2.19)	9.00 (1.65)	0.08†
23. Rate **change in performance** from previous scenario	0.83 (0.83)	1.50 (0.90)	1.17 (0.83)	0.33 (1.23)	0.15
24. Did you feel like you were controlling a safe sector? (# of "Yes" responses out of 12)	9	12	12	12	0.03*

Note. $^{*}p < .05$. $^{\dagger}p < .10$. The superscript "1" indicates that the condition is significantly different from Run 1, "2" is significantly different from Run 2, "3" is significantly different from Run 3, and "4" is significantly different from Run 4.

Table F9. Means (*SD*) of OTS Observers OTS Observer Rating Form Responses by Scenario Run for Experiment 3

OTS Observer Form Question:	Run 1	Run 2	Run 3	Run 4	*p*-value
1. Maintaining separation and resolving potential conflicts	7.58 (0.51)	7.33 (1.07)	7.58 (0.90)	7.83 (0.39)	0.23
2. Sequencing aircraft efficiently	7.36 (0.67)	7.36 (0.81)	7.73 (0.47)	7.73 (0.47)	0.13
3. Using control instructions effectively/efficiently	7.33 (0.65) [1]	7.50 (0.80)	7.75 (0.45)	7.83 (0.39) [1]	0.02*
4. Overall safe and efficient traffic flow	7.25 (0.87)	7.33 (0.89)	7.67 (0.65)	7.83 (0.39)	0.045*
5. Maintaining awareness of aircraft positions	7.17 (0.72)	7.33 (0.98)	7.67 (0.49)	7.75 (0.45)	0.045*
6. Giving and taking handoffs in a timely manner	7.08 (0.90)	7.25 (1.22)	7.75 (0.45)	7.67 (0.49)	0.09†
7. Ensuring positive control	7.25 (0.75)	7.50 (0.90)	7.67 (0.65)	7.83 (0.39)	0.15
8. Detecting pilot deviations from control instructions	7.50 (0.80)	7.75 (0.45)	7.58 (0.51)	7.83 (0.39)	0.24
9. Correcting own errors in a timely manner	7.55 (0.69)	7.64 (0.67)	7.64 (0.50)	7.91 (0.30)	0.29
10. Overall attention and situation awareness scale rating	7.25 (0.62) [1]	7.42 (0.67)	7.58 (0.51)	7.83 (0.39) [1]	0.01*
11. Taking actions in an appropriate order of importance	7.33 (0.65)	7.33 (0.98)	7.75 (0.45)	7.83 (0.39)	0.08
12. Preplanning control actions	6.83 (1.70)	7.00 (1.48)	7.33 (1.44)	7.33 (1.44)	0.04*
13. Handling control tasks for several aircraft	7.25 (0.75)	7.50 (0.80)	7.92 (0.29)	7.83 (0.39)	0.02*
14. Overall prioritizing scale rating	7.25 (0.62) [1]	7.33 (0.78)	7.67 (0.49)	7.75 (0.45) [1]	0.04*
15. Providing essential air traffic control information	7.00 (1.04)	7.33 (0.98)	7.67 (0.89)	7.75 (0.45)	0.13
16. Providing additional air traffic control information	7.25 (0.75)	7.17 (0.72)	7.42 (1.16)	7.58 (0.90)	0.49

(table continues)

Table F9. Means (*SD*) of OTS Observers OTS Observer Rating Form Responses by Scenario Run for Experiment 3 (continued)

OTS Observer Form Question:	Run 1	Run 2	Run 3	Run 4	*p*-value
17. Providing coordination	6.92 (1.24)	7.25 (0.87)	7.42 (1.16)	7.50 (0.90)	0.18
18. Overall providing control information	7.17 (0.72)	7.17 (0.83)	7.58 (0.79)	7.75 (0.45)	0.03*
19. Showing knowledge of LOAs and SOPs	6.08 (1.38) ^{1}	7.00 (1.04)	7.25 (1.06)	7.58 (0.67) ^{1}	< 0.001*
20. Showing knowledge of aircraft capabilities and limitations	7.58 (0.51)	7.50 (0.67)	7.83 (0.39)	7.83 (0.39)	0.06†
21. Showing effective use of equipment	7.25 (0.62)	7.08 (1.00)	7.58 (0.51)	7.67 (0.49)	0.045*
22. Overall technical knowledge scale rating	7.00 (0.74) ^{1}	7.08 (0.90)	7.50 (0.67)	7.75 (0.45) ^{1}	0.001*
23. Using proper phraseology	7.75 (0.45)	7.67 (0.65)	7.67 (0.65)	7.75 (0.65)	0.41
24. Communicating clearly and efficiently	7.50 (0.52)	7.83 (0.58)	7.92 (0.29)	7.92 (0.29)	0.01*
25. Listening to pilot readbacks and requests	7.75 (0.45)	7.75 (0.62)	7.83 (0.58)	7.92 (0.29)	0.49
26. Overall communicating scale rating	7.75 (0.45)	7.83 (0.58)	7.83 (0.58)	7.92 (0.29)	0.51
27. How well did participants follow the LOAs for the sector	7.83 (1.60)	8.67 (1.37)	8.33 (1.97)	9.00 (1.55)	0.43

Note. $^{*}p < .05$. $^{\dagger}p < .10$. The superscript "1" indicates that the condition is significantly different from Run 1, "2" is significantly different from Run 2, "3" is significantly different from Run 3, and "4" is significantly different from Run 4.