DOT/FAA/TC-13/55

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

Human Factors Evaluation of Conflict Resolution Advisories in the En Route Domain

Sehchang Hah, FAA Human Factors Branch
Ben Willems, FAA Human Factors Branch
Gary Mueller, FAA Human Factors Branch
Daniel R. Johnson, FAA Human Factors Branch
Hyun Woo, FAA Human Factors Branch
John DiRico, TASC, Inc.
Kevin Hallman, TASC, Inc.
Kenneth Schulz, TASC, Inc.
Sonia Alvidrez, TASC, Inc.
Karl A. Meyer, MITRE Corporation
Robert Bastholm, Spectrum Software Technology, Inc.
Scott Terrace, Behavioral Science Associates, Inc.
Jonathan R. Rein, Spectrum Software Technology, Inc.
Joshua Harrison, Drexel University

October 2013

Technical Report

U.S. Department of Transportation
Federal Aviation Administration

NOTICE

Technical Report Documentation Page

1. Report No. DOT/FAA/TC-13/55	**2. Government Accession No.**	**3. Recipient's Catalog No.**

4. Title and Subtitle Human Factors Evaluation of Conflict Resolution Advisories in the En Route Domain	**5. Report Date** October 2013 **6. Performing Organization Code** ANG-E25
7. Author(s) Sehchang Hah, Ben Willems, Gary Mueller, Daniel R. Johnson, and Hyun Woo, FAA Human Factors Branch John DiRico, Kevin Hallman, Kenneth Schulz, and Sonia Alvidrez, TASC, Inc. Karl A. Meyer, MITRE Corporation Scott Terrace, Behavioral Science Associates, Inc. Robert Bastholm and Jonathan R. Rein, Spectrum Software Technology, Inc. Joshua Harrison, Drexel University	**8. Performing Organization Report No.** DOT/FAA/TC-13/55
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 Human Factors Division 800 Independence Avenue, S.W. Washington, DC 20591	**13. Type of Report and Period Covered** Technical Report **14. Sponsoring Agency Code** ANG-C41

15. Supplementary Notes

16. Abstract

Objective: In this human-in-the-loop simulation experiment, we evaluated how Conflict Resolution Advisories (CRA) affected en route controllers. **Background**: Controllers currently use a conflict probe and trial planning tool, known as the User Request Evaluation Tool (URET), which is available on the Radar Associate Position. However, under Trajectory-Based Operations—that is, Separation Management Modern Procedures (SepMan)—several capabilities will become available to the Radar Position, including probed menus, conflict detection and trial planning, and support for multiple separation minima within a sector's airspace. The CRA Program is built upon the SepMan concept and will provide a proposed solution to a potential conflict as soon as a controller initiates the entry of a clearance. **Method**: Twelve current en route Certified Professional Controllers from Air Route Traffic Control Centers participated in the experiment. **Results**: CRA capabilities did not change controller workload nor time and distance flown by aircraft in the sector. Analysis of tactical and strategic conflict alerts show that controllers solved potential conflicts quickly when CRA was available. Most of the participants' subjective ratings favored the CRA, and participants expressed that CRA was a useful concept. **Conclusion**: The results show an advantage of CRA on some air traffic control tasks. In general, CRA was accepted by the participant controllers. **Application**: With a few modifications of the current CRA features and functions, the authors believe that CRA will be a useful automation tool for air traffic controllers.

17. Key Words Air Traffic Control, Automation, Conflict Alert, Conflict Detection, Conflict Probe, Conflict Resolution Advisories, Human Factors, Physiological Measures, Workload	**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)	**20. Security Classification (of this page)**	**21. No. of Pages**	**22. Price**
Unclassified	Unclassified	130	

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

THIS PAGE IS BLANK INTENTIONALLY.

Table of Contents

List of Illustrations

Tables **Page**

THIS PAGE IS BLANK INTENTIONALLY.

Acknowledgments

In this experiment, we focused on the effectiveness of Conflict Resolution Advisories (CRA) in air traffic control (ATC). To prepare and conduct a complex experiment like this one involves many people from diverse backgrounds and organizations. Thank you to everyone involved in the success of this experiment.

The participating controllers made this experiment possible. The 12 controllers from several Air Traffic Route Control Centers within the Continental United States traveled to our facilities, worked diligently on high-traffic simulation scenarios, answered all our questions, and volunteered to wear equipment to measure eye movements and brain activities. We appreciate their participation very much.

From MITRE, we received a tremendous support. Our goal was to evaluate their CRA product. It was a big challenge to integrate it into our in-house Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE) system for the simulation. We especially appreciate the contribution of the following MITRE managers, scientists, and engineers for this integration: Winfield S. Heagy (Management), Dr. Daniel B. Kirk (Systems Engineering), Dr. Scott H. Mills (Human Factors), Ray Newman (Human Factors), Nicholas E. Rozen (Systems Engineering), Ruby Sharma (Systems Engineering), and Monica Z. Weiland (Human Factors).

In the experiment, some of the aircraft were "Data Comm" equipped, which required our simulation pilots to perform additional tasks of using data communications to send requests to the controllers and respond to digital clearances received from controllers.

For this complex simulation, we had excellent support from the laboratory support. Albert Macias not only managed the Research Development and Human Factors Laboratory, but he also wrote the software that integrated eye-tracking data with the data collected during the experiment. Gary Mueller (a co-author) was the lead for the laboratory programming and developing group. Matt Bruckner, Joel Grochowski, Laura Hagmann, Vince Locasale, Chris Parratto, Nicole Smith, Yev Tabekman, Valentina Velez, and Matt Zeits programmed many of the new features in DESIREE. Several DESIREE team members also reduced data into a form that we could use in our analyses. Robert Bortu, Wallace Daczkowski, John Dilks, and Ed Little kept all our systems going during the experiment.

John DiRico, Jim McGhee, and Richard Ridgway (TASC, Inc.) prepared traffic scenarios as Subject Matter Experts of ATC. The Simulation and Analysis Team (Mike Paglione, Doug Frye, and Christina Young) at the FAA William J. Hughes Technical Center was our main customer and provided consistent support during the experiment as well as provided feedback on the data analysis and report.

Our simulated aircraft do not fly without the support of the Target Generation Facility (TGF) group. Samantha Fullerton, Dana Whicker, Rhoma Gordillo, and others ensured smooth operations of the TGF and simulation pilot workstations during the experiment.

Drexel University scientists helped us use the functional near-infrared (fNIR) equipment to measure the controllers' brain activities during the experiment. Dr. Hasan Ayaz, Joshua Harrison, and Dr. Kurtulus Izzetoglu spent many hours with us discussing the data collection and reduction and analysis of the fNIR data.

We also appreciate our editor, April Jackman (TASC, Inc.), who edited this report to make it more readable and presentable.

THIS PAGE IS BLANK INTENTIONALLY.

Executive Summary

Twelve Certified Professional Controllers participated in a human-in-the-loop experiment to assess the impact of Conflict Resolution Advisories (CRA) on capacity, safety, and efficiency of air traffic control. In the experiment, we used high-altitude sectors of Kansas City Air Route Traffic Control Center and traffic scenarios based on operational traffic recordings.

We tested differences between a baseline condition and two possible implementations of CRA: incorporated on the Radar Associate position only or on both the Radar and the Radar Associate positions. We also examined the effect of traffic levels on the controller' use of CRA by increasing traffic gradually from 30% to 150% of the Monitor Alert Parameter values in each run. Each test run lasted 50 minutes.

To assess differences between the baseline and CRA conditions, we collected subjective and objective data. The objective data included system data, such as time and distance aircraft travelled in the sector, and participant workload assessed through physiological measures, including functional near-infrared (or fNIR), electroencephalogram (EEG), and eye movements. The current report does not include results based on the EEG and eye-movement data.

All participants rated CRA as useful. However, some of them commented that it was not as useful when they were very busy. The majority of the participants (58%) suggested using CRA on both R- and D-sides. A few participants (25%) thought it should only be a D-side tool.

There were a few features and functions that did not receive positive ratings from participants. These features and functions must have affected the objective data, such as time and distance flown by aircraft, when they used the CRA. For example, when CRA presented wrong altitudes for direction of flight (WAFDOFs) as optimal resolutions, they disregarded them and selected a different altitude, which took extra time. In addition, most of the participants thought CRA had too much information and too many steps to go through to select a conflict resolution.

All participants indicated they did not receive enough training to learn the CRA features and functions thoroughly. Their subjective workload ratings did not show a significant difference between the CRA and baseline conditions. However, the decreasing trend of workload reduction by days was steeper in the CRA condition than in the baseline condition. We assume that this trend would have continued resulting in lower workload ratings in the CRA condition than in the baseline condition if our simulation experiment had lasted longer.

To be operationally acceptable, CRA may need to remove a WAFDOF, reduce the number of the steps to use CRA, and minimize the complexity of the information presented. Our results show that the number of tactical conflict alerts was less when CRA was available, suggesting that participants resolved potential conflicts more strategically with CRA. With a very limited amount of training, the introduction of CRA, although unfamiliar to the participants, did not increase their workload or reduce their performance.

THIS PAGE IS BLANK INTENTIONALLY.

1. INTRODUCTION

The Next Generation Air Transportation System (NextGen) includes several new technologies and concepts to prepare the National Airspace System (NAS) for an increase in air traffic while maintaining current levels of safety and efficiency. One of the obstacles with implementing Trajectory-Based Operations (TBO) is that controllers often provide pilots with open-ended clearances instead of closed clearances. Open-ended clearances do not take into account the effect on the full trajectory resulting in suboptimal solutions. Closed clearances on the other hand update the full end-to-end trajectory. Providing controllers with tools to create optimal solutions using TBO will increase the overall efficiency of the system. TBO will provide more accurate trajectory information that will result in better detection of aircraft conflicts. Based on this concept, MITRE Center for Advanced Aviation System Development (MITRE/CAASD) has developed a tool that provides Conflict Resolution Advisories (CRA) to controllers.

1.1 Background

The development of CRA is the culmination of several decades of MITRE/CAASD research that started in the Automated En Route Air Traffic Control (AERA) Program. In the following sections, we present Federal Aviation Administration (FAA) programs that provide baseline capabilities used in the CRA operation environment such as Separation Management (SepMan) and Data Communications (Data Comm). We will also discuss the Problem Analysis, Resolution, and Ranking tool developed under the AERA Program since it formed the foundation of CRA. Finally, we will discuss the work that MITRE/CAASD has conducted on CRA.

1.1.1 Separation Management – Modern Procedures

We assume that the capabilities developed under the Separation Management Program are available during the timeframe of the CRA implementation. The Separation Management Program is part of the NextGen TBO portfolio as Separation Management – Modern Procedures. MITRE/CAASD maintains a Concept of Operations (CONOPS) document for the FAA (e.g., Exum, Bolczak, Celio, & Poore, 2010) that reflects how the Separation Management – Modern Procedures relates to the NextGen Implementation Plan (FAA, 2011) and how the FAA may choose to implement the concept. SepMan addresses the following 15 areas of NAS enhancements.

1. Provide trajectory modeling enhancements.
2. Reduce missed and false alerts.
3. Reduce trajectory conformance bounds to reflect advanced aircraft Communication, Navigation, and Surveillance (CNS) capabilities.
4. Integrate problem detection on the radar console.
5. Introduce probed menus.
6. Introduce flight management computer based later offsets for current plans, trial plans, and probed menus.
7. Alert aircraft-to-aircraft problems for 3-nautical mile (nmi) separation areas.
8. Alert aircraft-off-flight-plan problems.
9. Alert for dynamic Special Activity Airspace.

10. Provide support for non-surveillance airspace.
11. Integrate manual trial planning on the radar console.
12. Facilitate the entry of clearances and flight plan amendments.
13. Improve conflict alert.
14. Use automation assisted controller-to-controller coordination.
15. Provide selective removal of altitude restrictions.

In a previous research experiment, Sollenberger, Willems, DiRico, Hale, and Deshmukh (2010) used the integration of problem detection and trial planning on the radar console and probe menus on both the radar console and the radar associate positions. They reported that these automation capabilities did not increase controller workload.

1.1.2 Data Communications

Although not necessary for the implementation of CRA, Data Comm will facilitate the execution of conflict resolution advisories because it will give controllers more flexibility with issuing clearances. The future Data Comm message set will include a much larger number of messages (Willems, Hah, & Schulz, 2010) than those used in the previous Controller-Pilot Data Link Communications (CPDLC) Program (Talotta, 1992). The CPDLC Build 1 message set only included Transfer of Communications, Initial Contact, and Menu Text. The CPDLC Build 1A message set expanded the Build 1 message set and added altitude, heading, speed, and direct-to-fix clearances in addition to pilot's downlink request of an altitude. In their En Route Data Comm experiment, Willems et al. emulated En Route Automation Modernization (ERAM) functions and included a large number of messages of the RTCA Special Committee 214 (SC214) message set. Controllers were able to create complex clearances using a set of newly adapted or developed interfaces even though they preferred simple clearances. Advanced automation like CRA may create such complex clearances, and controllers can use Data Comm to uplink them to the aircraft.

1.1.3 Problem Analysis Resolution and Ranking

To assist controllers with creating multiple trial plans and providing an optimal solution to the problem, MITRE's CAASD developed Problem Analysis Resolution and Ranking (PARR) during the late 1980s and early 1990s (Bowen, 2000). PARR was part of the AERA program under the Advanced Automation System development. In 1998, the PARR capability was integrated into the User Request Evaluation Tool (URET) prototype. The FAA improved the prototype and deployed it to the field initially as the Core Capabilities Limited Distribution (CCLD) version before deploying the operational version to all 20 Air Route Traffic Control Centers (ARTCCs). The operational version of URET no longer performs the PARR functionality but has provided controllers with strategic conflict detection and trial planning capabilities. Controllers can use trial planning to probe a solution to a potential conflict, but it is an iterative process.

1.1.4 Conflict Resolution Advisories

Based on the PARR capabilities, CAASD has developed a new set of capabilities that provide controllers with problem resolutions using vectoring, step-climbs, or a search-all function. With the search-all function, it can rank solutions along several dimensions of the solution space. The CRA CONOPS (Syeda, Bowen, Meyer, & Viets, 2011) explains how the CRA concept could be implemented into the NextGen Implementation Plan (FAA, 2011).

CAASD has conducted three mini-evaluations of each of the CRA capabilities and has proposed to conduct a Human-in-the-Loop (HITL) demonstration in November, 2011b (Mills, Kirk, & Rozen, 2011b). They focused on the Computer-Human Interface (CHI; see Mills & Rozen, 2010; Mills, Kirk, & Rozen, 2011a). The initial CRA Build 1 consisted of two enhancements for resolutions in the vertical and lateral dimension and one enhancement for resolutions along multiple dimensions.

1.1.5 Concerns of Conflict Resolution Advisories

The introduction of CRA to the controller workstation creates three types of concerns. First, CRA will rely heavily on the availability of other automation capabilities. Without these capabilities, only limited implementation may be possible. For example, if the conflict probe and trial planning functions do not become available on the radar console (R-side), implementation may only be possible on the D-side. Second, the introduction of CRA will affect the way controllers perform their jobs. In the current environment, controllers have support for conflict detection, but they are responsible for the detection and resolution of conflicts. Third, a combination of conflict probe and CRA will provide conflict detection and resolution support. As a result, the controller may become a monitor of the automation rather than an active participant in the process. We will address these issues in the paragraphs below.

CRA assumes that the underlying trajectory synthesis and conflict detection is accurate. The conflict detection algorithms may create false alerts or not issue alerts. We address these concerns in the experiment by examining the objective trajectory, conflict detection, and CRA solution data. We addressed these concerns in a post-scenario questionnaire and an exit questionnaire, and we received feedback from controllers.

CRA assumes that conflict detection, trial planning, and probed menus are available and reliable. If the conflict detection, trial planning, and probed menus are not available on the R-side, the CRA may still be implemented on the radar associate position (D-side). We addressed this by including an experimental condition where CRA was available only on the D-side. CRA also assumes the availability of automation assisted sector-to-sector coordination. Without the automated coordination capability, two-step resolutions that require coordination across sectors will be more difficult to implement.

Human factors issues include the complexity of the CRA menus and trust in the automation including CRA. Controllers may experience difficulty in learning the CRA menu because of the complexity of the menu. It could create tunneling of the visual system on the CRA menu interaction. This will create additional workload. We collected eye-tracking data to determine how long the interaction with the CRA menus took compared to interaction with other menus. Instantaneous assessment of subjective workload captured the difference in workload between the baseline and CRA conditions. The use of physiological measures assessed how cognitive workload may change. We addressed these issues in post-scenario and exit questionnaires. If controllers do not trust CRA, they may choose not to use the automation. On the other hand, if controllers fully trust CRA, they may not realize that the automation system has provided incorrect information. We collected data on how often controllers used the CRA solutions and addressed trust in post-scenario and exit questionnaires.

1.1.6 Benefits of Conflict Resolution Advisories

CRA may improve controller productivity with increased flight efficiency and a reduction in operational errors (Syeda et al., 2011). Syeda et al. listed seven potential benefits (see Table 1). Mills, Meyer, Syeda, Bowen, and Viets (2012) reported the benefits of the CRA. Based on their previous research, we expected our experiment to show more thorough analyses of the CRA benefits and concerns.

Table 1. Potential Benefits of Conflict Resolution Advisories

Benefit	***Benefit Mechanism***	***Assessment Technique***	**Assessed in Experiment?**
Reduced delays due to increased sector capacity	Project enhancements allow the controller to handle more aircraft at a fixed workload level.	• Gather qualitative workload indicators (both subjective and objective) during HITLs. • If a significant workload reduction is indicated, then conduct further HITLs (not part of baseline plan) to quantify increased sector capacity.	Yes
Reduced maneuvering due to improved intent entry	Improved intent due to increased entry of 2-part maneuvers (e.g., 2-leg vectors, step-climb maneuvers).	• Subject Matter Experts (SMEs) and HITLs provide estimates of the increased % of 2-part maneuver entry. • Simulation uses increased % to model improved trajectory accuracy, reduction in false/missed alert rates.	Yes
Reduced maneuvering due to more strategic control actions	Project enhancements support more strategic, efficient maneuvers.	• SMEs estimate number of maneuvers for baseline and with CRA. • Conduct simulations to compare number of maneuvers for baseline and with CRA. • Measure scenario time/fuel burn savings and extrapolate to NAS. • Measure separation achieved and time within sector boundaries.	Yes
Reduced altitude restrictions	Project enhancements allow permanent removal of selected altitude restrictions.	• SME identifies restrictions to be removed. • Estimate affected flights. • Estimate fuel burn/time savings.	No
Reduced number of altitude-capped flights	Project enhancements reduce workload for vertically transitioning aircraft, allowing a higher (more efficient) cruise altitude.	• SMEs identify criteria for removing cap. • Identify flights meeting criteria. • Model fuel burn, time savings from cap removal.	No
Reduced use of Altitude for Direction of Flight (AFDOF)	Project enhancements allow increased reduction in the frequency of AFDOF constraint application, allowing aircraft to fly more preferred altitudes.	• SMEs estimate % reduction in AFDOF constraint. • Analyze track data to estimate baseline AFDOF constraint fuel-burn cost. • Apply % reduction to baseline costs to estimate savings.	No
Increased use of established direct routes between city pairs	Project enhancements support controller in the increased complexity caused by more aircraft flying direct routing.	• Perform simulation to illustrate increased spatial distribution of conflicts with increased direct routings. • SMEs estimate city-pairs for established direct routings and % utilization. • Measure distance savings over conventional routing and estimate fuel-burn/time savings.	No

Note. HITL = Human-in-the-loop, CRA = Conflict Resolution Advisories, NAS = National Airspace System. Table adapted from Syeda, Bowen, Meyer, and Viets (2011).

1.2 Purpose

The purpose of the HITL experiment was to determine the impact of the introduction of CRA on controller workload, performance, and behavior and on system performance with steadily increasing traffic levels. We assessed system performance by analyzing trajectory accuracy, conflicts detected, and trajectory efficiency. The results of the experiment will assist in improving the CRA requirements and provide data to assess the benefits proposed in the CONOPS.

2. METHOD

2.1 Participants

Twelve current en route Certified Professional Controllers (CPCs) participated in the experiment. We recruited participants from Continental U.S. Air Route Traffic Control Centers, but we excluded CPCs from Kansas City Center due to using its airspace in the experiment. This exclusion ensured that all participants started at a level playing field while maintaining a large pool of potential participants. The participants worked either as an R-side or a D-side controller. The controller teams communicated with the pilots using voice and Data Comm, depending on aircraft equipage.

2.1.1 Simulation Participants

We used four simulation pilots for each of the two controlled sectors. The controller participants communicated clearances to the pilots and the pilots updated the Target Generation Facility (TGF) software to maneuver the aircraft in accordance to the clearances.

2.1.2 Subject Matter Experts

Four Subject Matter Experts (SMEs) trained participants on the airspace, procedures, Data Comm, SepMan, and CRA. They also served as Over-The-Shoulder raters during the simulation.

2.2 Airspace

The airspace used in the experiment consisted of two active high-altitude sectors. We used the high-altitude Kansas Center (ZKC) Sectors 20 and 22 (see Figure 1).

2.3 Traffic Scenarios

We created traffic scenarios based on samples extracted from operational traffic recorded from the Aircraft Situation Display to Industry (ASDI) feed. We filtered the traffic to include only aircraft that crossed a volume of airspace of 300-nmi by 300-nmi that included the sectors used in the experiment.

After comparing samples of various days, we decided to use the traffic of July 14, 2011, (from 13:00 to 14:00 for the test scenario and from 14:00 to 15:00 for the practice 33% to 150% scenario) and two other training scenarios from September 5 (practice 33% to 66%) and from September 6 (practice 33% to 100%). To meet the required traffic volume, we added aircraft manually. We assigned realistic call signs, routes, and aircraft types to the newly added aircraft by copying the information of the aircraft that traveled to the same route.

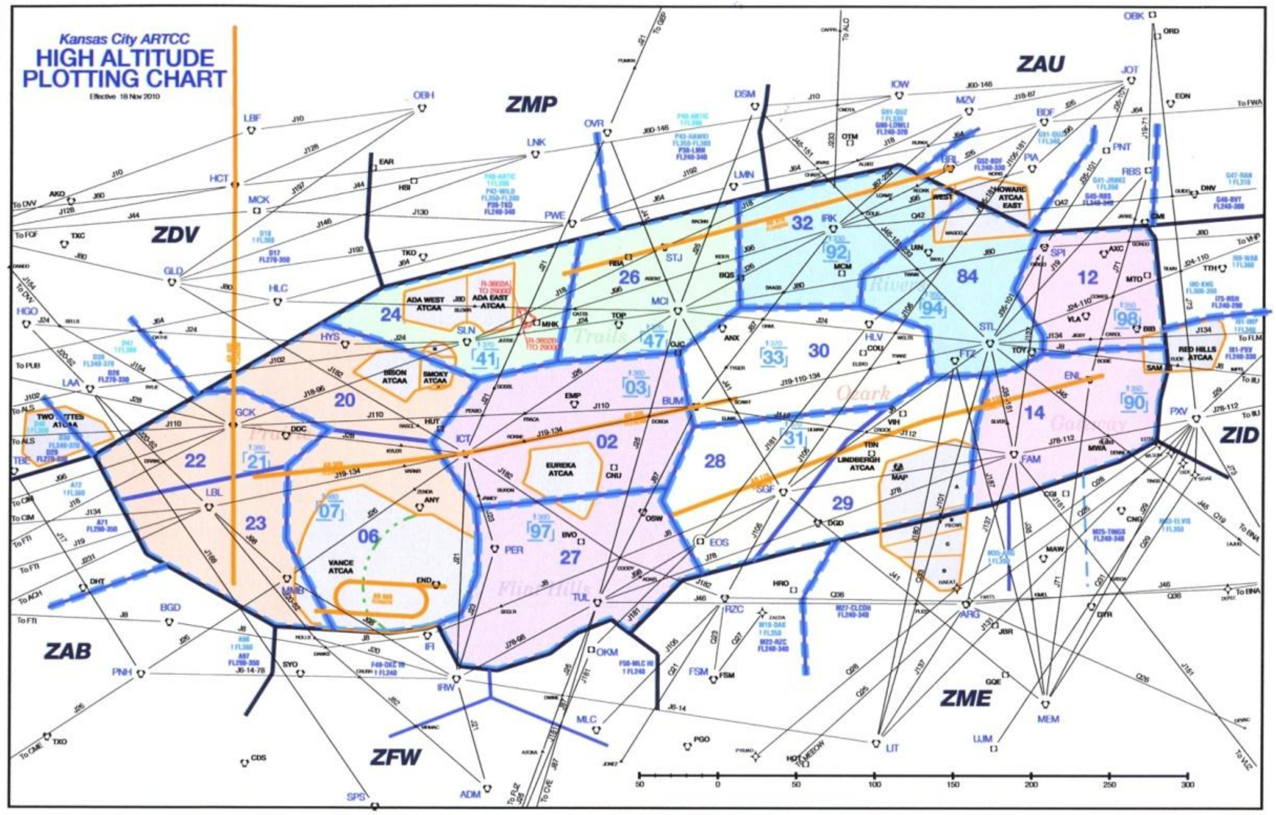


Figure 1. High-altitude sectors of ZKC, including Sectors 20 and 22 that are at the top-left corner of the center.

On the first day, we used three types of training scenarios to prevent the participants from getting overwhelmed by high-volume traffic. At first, the participants had 30-min test runs, where the traffic level increased steadily from 33% of the sector's Monitor Alert Parameter (MAP) values to 66% at 15 min and then remained at that level for the rest of the run. In the next phase of training, we used 55-min scenarios, where the traffic level increased gradually from 33% to 100%. In the last phase of the training, the traffic level increased in the same pattern as the experimental run's pattern; that is, the traffic level started at 33% and then increased gradually to 150% during the 55-min test run.

We created three separate 50-min training scenarios for the experimental days. From each of these three scenarios, we created two additional scenarios by changing call signs only. In this way, we had three sets of three training scenarios. We randomly presented them to participants.

2.4 Hardware and Software

We conducted the simulation at the Research Development and Human Factors Laboratory (RDHFL), FAA William J. Hughes Technical Center in Atlantic City International Airport, NJ. In the RDHFL experiment room, we had controller workstations and associated equipment (see Figure 2). We used simulation pilot workstations located in a separate room of the RDHFL. During the simulation, we used audio and video equipment to record the participants' communications and actions.

Figure 2. Simulation room for controllers.

2.4.1 Distributed Environment for Simulation, Rapid Engineering, and Experimentation

We used the Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE) to simulate the ERAM system. We modified the DESIREE emulation of the ERAM interface to accommodate the CRA functionality based on the capabilities developed by MITRE/ CAASD. These capabilities included an enhanced altitude menu, an enhanced heading menu, a search-all menu, and integration of reminders in appropriate ERAM and SepMan interfaces (see Figure 3 through Figure 5).

In Figure 3, the requested altitude of FL430 is a Wrong Altitude for Direction of Flight (WAFDOF) indicated by a yellow colored FP430, which currently has a temporary altitude of FL360 indicated by T360. The aircraft only has voice-communications available. CRA assumes a 0:30 minute delay between the update of the system and the execution of the maneuver. N2665 is climbing to its requested altitude of FL430, and CRA predicts two potential conflicts if the aircraft levels off at FL370. The best resolution has a temporary altitude (T) level off at FL390 and is conflict-free as indicated by its green characters. After the level-off at FL390, a conflict-free resumption of the climb to FL430 is possible in 6 minutes from now.

Figure 3. An example of the altitude CRA menu.

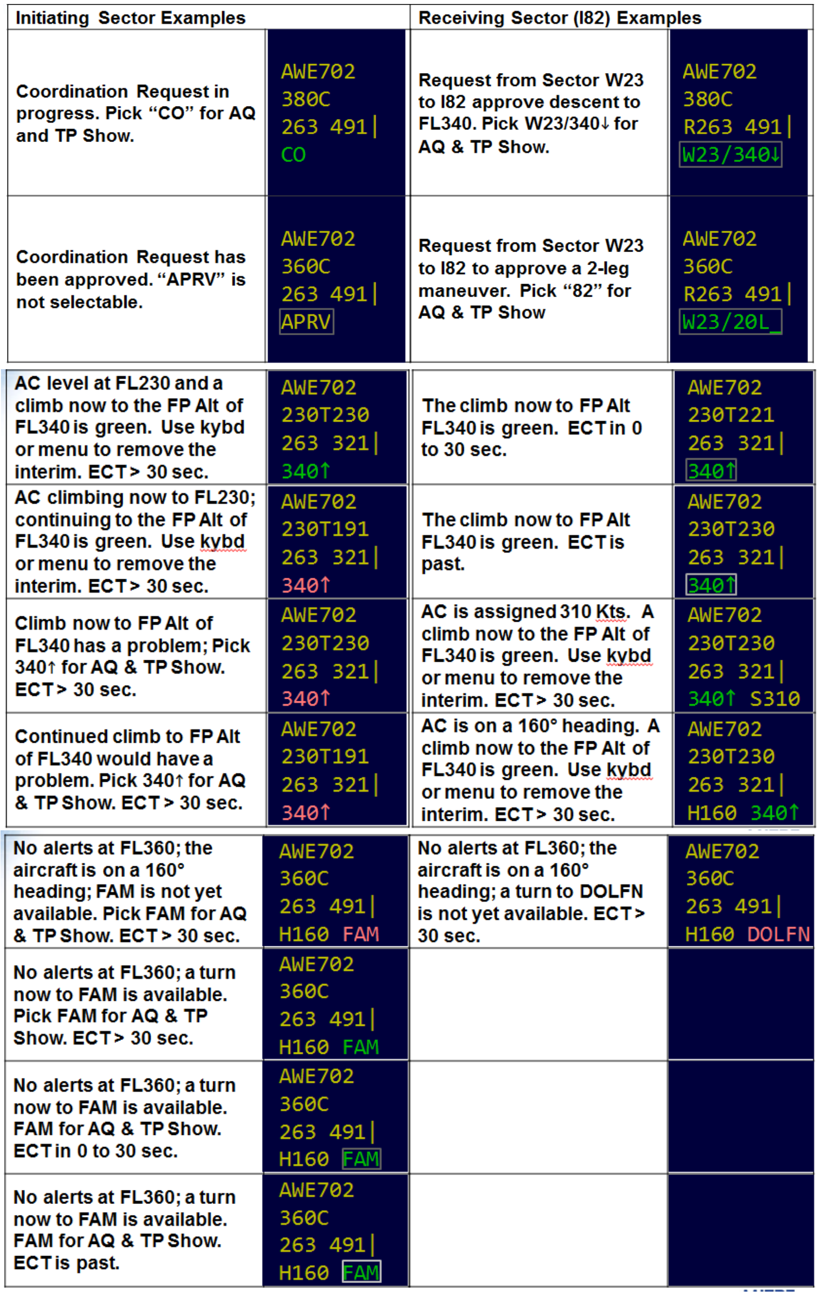
In Figure 4, the best resolution is a Right Turn (RT) maneuver; therefore the relative turn entries are shown by default. The best resolution has an initial 40-degree (40°) right turn and the route to the turn-back point (at 7 minutes from now) is conflict-free as indicated by the green characters of the minutes. The aircraft is equipped with Data Communications (DC), so the default is to send the corresponding clearance via DC. After the initial left turn, a conflict-free direct to rejoining the route at LAA is possible in 7 minutes from now—the maneuver delay is 2:45 minutes.

Figure 4. An example of the heading CRA menu.

Advisories are presented for both aircraft involved in the problem: FFT745 and FFT2487. The optimal solution as well as the aircraft to which it applies is highlighted with a white box. Resolutions to the problem that focus on the aircraft on the left have a blue background and resolutions that focus on the aircraft to the right have a black background. The rank of the resolutions is indicated on the left or the right side of the menu for the left and right aircraft, respectively. The highest ranked advisory initially turns FFT2487 five degrees (5°) to the right. The remaining portion of the advisory [2:LAA] is enclosed in brackets, denoting that the clearance (a direct to LAA) should be issued at a future time (2 minutes from now). The second best resolution focuses on FFT745 with a descent to FL330. The third ranked resolution is for FFT2487 with a descent to FL330 as well. The fourth ranked resolution is also for the FFT2487 with an increase of 20 knots indicated airspeed. The fifth ranked resolution reduces the Mach speed for FFT745 by 0.04 Mach. Finally, the sixth ranked resolution is for the FFT2487 with an initial turn of 15 degrees (15°) right followed (in 5 minutes from now) with a direct to CONAL. The advisories are initially in ranked order but can be sorted by aircraft by selecting "RANK." The default output mode (V) for voice can also be changed if either aircraft has Data Comm capability. Other CRA menus may be accessed for either aircraft and may be based on a selected trial plan.

Figure 5. An example of the search all CRA menu.

The CRA also used the fourth line of the datablock to give controllers the information about the status of clearances and enable controllers to coordinate with the neighboring sector controllers (see Figure 6). In the following text, "4a" means the first element of the fourth line and "4b" means the second element of the fourth line.



Initiating Sector Examples		Receiving Sector (I82) Examples	
Coordination Request in progress. Pick "CO" for AQ and TP Show.	AWE702 380C 263 491| CO	Request from Sector W23 to I82 approve descent to FL340. Pick W23/340↓ for AQ & TP Show.	AWE702 380C R263 491| W23/340↓
Coordination Request has been approved. "APRV" is not selectable.	AWE702 360C 263 491| APRV	Request from Sector W23 to I82 to approve a 2-leg maneuver. Pick "82" for AQ & TP Show	AWE702 360C R263 491| W23/20L_
AC level at FL230 and a climb now to the FP Alt of FL340 is green. Use kybd or menu to remove the interim. ECT > 30 sec.	AWE702 230T230 263 321| 340↑	The climb now to FP Alt FL340 is green. ECT in 0 to 30 sec.	AWE702 230T221 263 321| 340↑
AC climbing now to FL230; continuing to the FP Alt of FL340 is green. Use kybd or menu to remove the interim. ECT > 30 sec.	AWE702 230T191 263 321| 340↑	The climb now to FP Alt FL340 is green. ECT is past.	AWE702 230T230 263 321| 340↑
Climb now to FP Alt of FL340 has a problem; Pick 340↑ for AQ & TP Show. ECT > 30 sec.	AWE702 230T230 263 321| 340↑	AC is assigned 310 Kts. A climb now to the FP Alt of FL340 is green. Use kybd or menu to remove the interim. ECT > 30 sec.	AWE702 230T230 263 321| 340↑ S310
Continued climb to FP Alt of FL340 would have a problem. Pick 340↑ for AQ & TP Show. ECT > 30 sec.	AWE702 230T191 263 321| 340↑	AC is on a 160° heading. A climb now to the FP Alt of FL340 is green. Use kybd or menu to remove the interim. ECT > 30 sec.	AWE702 230T230 263 321| H160 340↑
No alerts at FL360; the aircraft is on a 160° heading; FAM is not yet available. Pick FAM for AQ & TP Show. ECT > 30 sec.	AWE702 360C 263 491| H160 FAM	No alerts at FL360; the aircraft is on a 160° heading; a turn to DOLFN is not yet available. ECT > 30 sec.	AWE702 360C 263 491| H160 DOLFN
No alerts at FL360; a turn now to FAM is available. Pick FAM for AQ & TP Show. ECT > 30 sec.	AWE702 360C 263 491| H160 FAM		
No alerts at FL360; a turn now to FAM is available. FAM for AQ & TP Show. ECT in 0 to 30 sec.	AWE702 360C 263 491| H160 FAM		
No alerts at FL360; a turn now to FAM is available. FAM for AQ & TP Show. ECT is past.	AWE702 360C 263 491| H160 FAM		

Figure 6. Various display elements on the 4th line of the datablock of the CRA display.

- **CoordClearance** – Color-coded text that indicates reception of a coordination request (either approval for request “APREQ” or clearance request); includes initiating sector ID, (a “/”) and the clearance text (truncated if necessary) and takes up the full 4th line (both 4a and 4b elements).
- **CoordIndicator** – A color-coded element (in 4a or 4b) “CO” indicates that coordination is ongoing with the aircraft (this will appear for all sectors except the receiving sector); the non-colored text Approve “APRV,” Will Comply “WLCO,” or Unable “UNBL” also is a CoordIndicator that appears following the receiving sector’s response.
- **AssignedHeading** – An element (always in 4a) that indicates a compass heading or off-angle has been assigned to the aircraft. For two-leg resolutions, the off-angle will automatically be entered, and the corresponding UnclearedFix will appear in 4b when expected.
- **AssignedSpeed** – A non-colored element (always in 4b) that indicates an absolute or relative speed has been assigned to the aircraft.
- **UnclearedFix** – A color-coded downstream fix (3 or 5 characters, or truncated FRD) that hasn’t been cleared to the aircraft; for this evaluation we will only see UnclearedFix elements for the 2nd part of a two-leg lateral amendment from CRA.
- **UnclearedAltitude** – A color-coded target altitude that hasn’t been cleared, but is expected within a parameter time (depending on color status) according to the current plan.

2.4.2 Joint En Route Decision Support System Infrastructure (JEDI)

DESIREE incorporated Conflict Detection and CRAs provided by the URET prototype maintained by MITRE/CAASD. JEDI received data through a modified Common Message Set interface and exchanged additional data with DESIREE through a separate interface. JEDI detected potential conflicts and provided data about these conflicts to DESIREE. When controllers interacted with the DESIREE interface, JEDI provided context specific Fly-Out menus. For example, if a controller clicked on the altitude field in a datablock of an aircraft with a potential conflict, JEDI provided an altitude Fly-Out menu that highlighted the JEDI-generated CRAs along with the data included in our baseline Fly-Out menus. DESIREE then placed the JEDI Fly-Out menu next to the datablock.

2.4.3 Target Generation Facility

We used the TGF to generate realistic aircraft behavior. The TGF provided track data, including position, altitude, and aircraft identification. The TGF accepted entries from the simulation pilot workstations and through its Data Comm connection with DESIREE to maneuver aircraft.

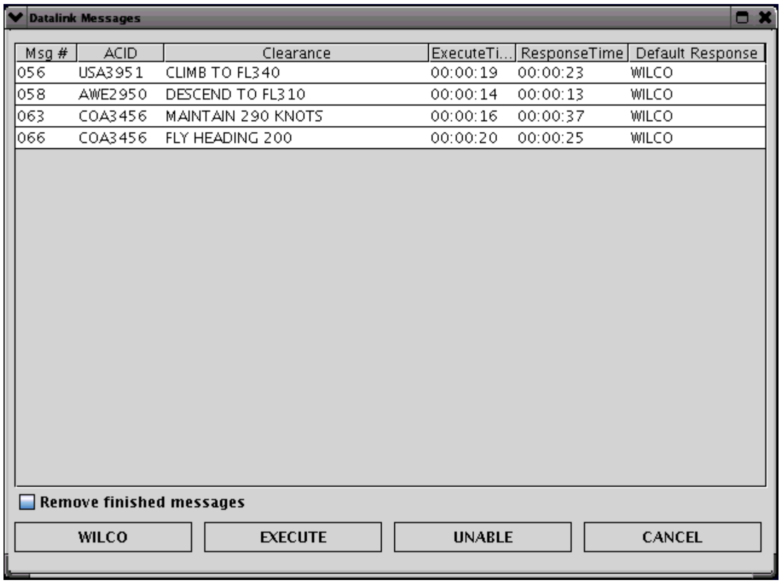
2.4.4 Voice Communications System

We used a voice communication system that mimicked the operational Voice Switching and Communications System (VSCS). Our VSCS provided the air-ground communications link between controller participants and simulation pilots using headsets and hand-held and foot-operated switches for Push-To-Talk (PTT). It also provided the ground-ground communications link between controller participants for inter-sector communications and the communication link between controllers and experimenters in the observation room behind the controller room.

2.4.5 Simulation Pilot Workstation

Three simulation pilots supported each sector. Each simulation consisted of two adjacent sectors requiring four simulation pilots for each sector. Each simulation pilot workstation consisted of a computer, keyboard, monitor, and communications equipment. The simulation pilot display showed a spatial representation of traffic and a list of aircraft assigned to them.

To accommodate the use of Data Comm, each simulation pilot workstation contained a window to display incoming Data Comm messages (see Figure 7). The TGF software uplinked messages after a variable delay (determined by the DESIREE software); however, simulation pilots could override the TGF-selected response through the simulation pilot interface.



Msg #	ACID	Clearance	ExecuteTi…	ResponseTime	Default Response
056	USA3951	CLIMB TO FL340	00:00:19	00:00:23	WILCO
058	AWE2950	DESCEND TO FL310	00:00:14	00:00:13	WILCO
063	COA3456	MAINTAIN 290 KNOTS	00:00:16	00:00:37	WILCO
066	COA3456	FLY HEADING 200	00:00:20	00:00:25	WILCO

Figure 7. Display for simulation pilots.

2.4.6 Workload Assessment Keypad

To assess instantaneously perceived workload, we used a Workload Assessment Keypad (WAK); see Figure 8. The WAK technique is an adaptation of the Air Traffic Workload Input Technique (ATWIT) that Stein (1985) developed to assess instantaneous subjective workload during simulations. ATWIT uses a 10-point scale that is anchored in operational needs. The low end of the scale reflects low workload (1-2), participants can accomplish all their tasks easily and have spare time left. At levels 3, 4, and 5 controllers will experience increasing levels of moderate workload, and they can still finish all tasks, but the chance of an error is steadily increasing and less and less spare time is available. At levels 6, 7, and 8 controllers experience high workload, and controllers have no spare time available and can barely finish all essential tasks, and will leave some unessential tasks unfinished. At levels 9 and 10 of the workload scale, participants are experiencing extremely high workload—it is likely that participants will leave essential tasks unfinished; at this point, participants will most likely focus on keeping aircraft separated.

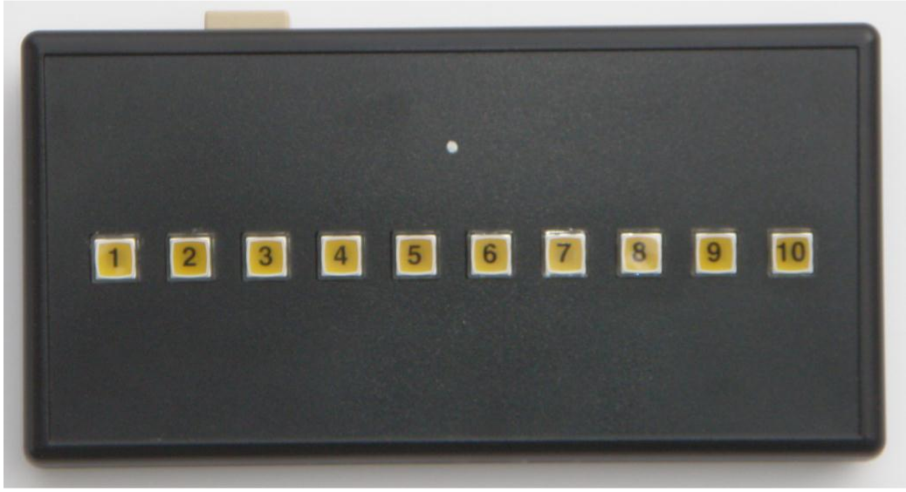


Figure 8. The workload assessment keypad.

The participants received instructions to indicate their instantaneous workload level by pressing 1 of 10 numbered buttons. We configured the WAK device to prompt participants for input every 2 minutes and 20 seconds to respond. If the device did not receive a response within 20 seconds, it recorded a code for missing data.

2.4.7 Functional Near-Infrared Spectroscopy

The portable functional Near-InfraRed (fNIR) Spectroscopy system used in this study consists of a pad that covers the forehead (see Figure 9). We put it under the eye-movement tracking system. The pad contains light sources and light detectors. The light sources contained two Light Emitting Diodes (i.e., LEDs) that shone light at two different wavelengths in the near-infrared spectrum into the prefrontal cortex of the brain. The light detectors consisted of three light-sensitive elements that detected light at the two infrared wavelengths and one detector to detect ambient light. Using the Modified Beer-Lambert law, the device calculated the concentration of oxygenated hemoglobin (oxyhemoglobin) and the concentration of de-oxygenated hemoglobin (deoxyhemoglobin). The difference in concentrations between oxyhemoglobin and deoxyhemoglobin was used as a measure of blood oxygenation. We used this blood oxygenation as a measure of cognitive workload.

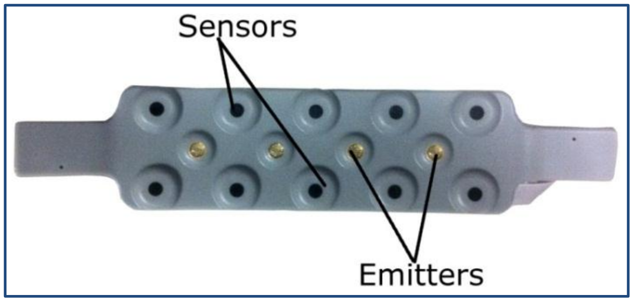


Figure 9. The fNIR pad.

Two fNIR devices, fNIR Imager 1000s, ran 16 channels at a sample rate of 2 Hz. Each fNIR system was connected to a computer running Microsoft Windows XP that collected data using Cognitive Optical Brain Imaging (COBI) Studio software.

2.4.8 Oculometer

The oculometer consisted of an eye-movement tracking system that recorded point of gaze (POG) and pupil diameter and a head tracking system to compensate for head movements (see Figure 10). The oculometer illuminated the eye in the near-infrared spectrum and determined the center of the corneal reflection and the center of the pupil. From these two points, the system created a line of sight.

Figure 10. Oculometer and fNIR device.

We used two Applied Science Laboratories H6-HS eyetrackers. Each eyetracker tracked the participant's left eye at a sample rate of 120 Hz. The eyetrackers were connected to a computer running Microsoft Windows XP that collected the data using Applied Science Laboratories EYEHEAD software.

The sampling rate was 120 Hz with an accuracy of 0.5 degrees. We used a Flock of Birds magnetic head tracker, manufactured by Ascension Technology Company to track participants' eye movements relative to their head movements. The eight scene planes were configured prior to the test. Each plane consisted of an area that contained information that would be of relevance to the participants' tasks during the trial. For example, each participant's main screen was Scene Plane 1, while his or her teammate's screen was Scene Plane 4.

2.4.9 Electroencephalogram

Two controllers (an R-side and a D-side controller) of one sector wore Electroencephalogram (EEG) equipment. We recorded EEG data with a 32 channel ActiCHamp active electrode system from Brain Products Inc. (see Figure 11). We collected data from 32 scalp positions according to the international 10-10 system. To determine the effect of muscle artifacts of eye movements on the quality of the EEG signals, we placed additional electrodes around the eyes to record eye movements and blinks on one of the two controllers. We did not use these additional electrodes on the other controller.

Figure 11. An EEG cap on the head.

Each EEG unit was a BrainVision actiCHamp equipped with 32 channels to record data at 1 MHz with a 24-bit. Four facial electrodes, connected via four auxiliary channels, which collected the data of lateral and vertical eye movements, including blinking. We connected each amplifier (two amplifiers for two controllers) to a Microsoft Windows 7, 64-bit computer. We collected data using BrainVision PyCorder software. DESIREE sent triggers to each amplifier via a parallel port from an additional Windows XP computer that synchronized DESIREE and the eye-movement system.

Each unit was equipped with 32 Ag/AgCl impedance-optimized electrodes that were set to record electrical activity at 1 MHz with a battery-powered 24-bit amplifier. We injected each electrode with SuperVisc electroconductive gel, manufactured by EASYCAP GmbH, which enabled us to collect data with no scalp preparation. We placed the electrodes in EASYCAP modular EEG recording caps. The electrodes covered the frontal, central, parietal, and occipital regions of the participants' heads. Also, we placed a ground electrode at the Fpz point (see Figure 12) and attached two electrodes (TP9 and TP10) to each participant's mastoids that served as references (see Figure 13). In addition, the R-side controller wore four facial electrodes, connected via four auxiliary channels that detected lateral and vertical eye movements, including blinking (see Figure 14).

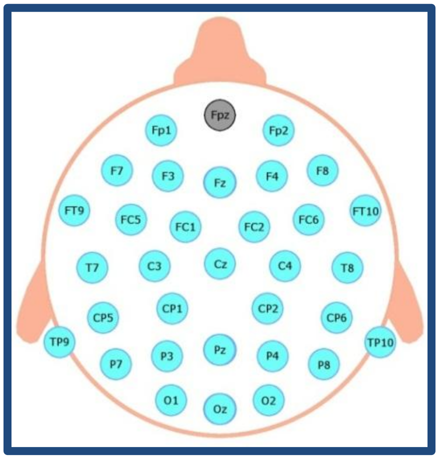


Figure 12. Electrode locations.

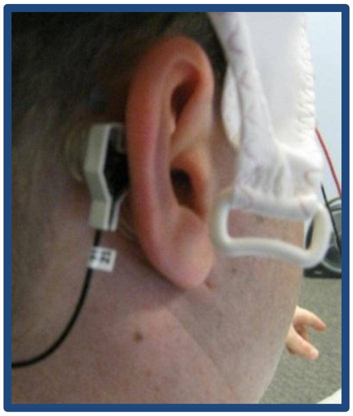


Figure 13. EEG reference electrode location.

Figure 14. Facial electrode locations of the R-side.

2.5 Procedure

The participants spent 10 days at the RDHFL. They traveled to the RDHFL on the Monday of the first week, and they left on the next Friday. In the first week, they participated in the High-Altitude Experiment, which used Cleveland ARTCC (ZOB) airspace for its High Performance Route portion of the experiment and used Kansas City ARTCC (ZKC) airspace for the generic airspace operations portion. Although the sectors in the High-Altitude experiment were super high sectors, they overlaid the sectors that we used in the CRA experiment. This gave controllers the advantage of having seen the airspace during the previous week and having some familiarity with routes and waypoints for the CRA experiment.

After the introduction to the CRA experiment, participants signed an Informed Consent Form (see Appendix A) and filled out a brief Background questionnaire (see Appendix B). On the first day of the experiment, the participants received extensive training on the airspace, systems, and procedures.

Participants worked the first training scenarios with 100% of the aircraft equipped with Data Comm without simulation pilots. We found in previous experiments that this helped participants incorporate Data Comm in their routine. During the remainder of the simulations in our experiment, 30% of the aircraft were capable of sending and receiving Data Comm messages.

Participants experienced three levels of traffic volume through three training scenarios. The first training scenario had a low traffic level (33% to 66% of the MAP value). The second scenario had a higher traffic level (33% to 100% of the MAP value). The third scenario had the highest traffic volume and was as high as the traffic volume of the experimental scenario (33% to 150% of the MAP value). We counterbalanced blocks of one training and one experimental scenario across conditions to minimize the order effect across participants.

Before each scenario, we reminded the participants of the meaning of the workload levels captured by the WAK. Participants worked traffic as they would at their own facility. During the experimental scenarios, we instrumented them with the fNIR, EEG, and the oculometer.

After each experimental scenario, participants filled out a Post-Scenario Questionnaire (see Appendix C). The participants had a 15-minute break between scenarios, a 30-minute break between blocks, and a 1-hour break for lunch.

We determined the proper size of cap that would fit securely on each participant by measuring the head circumference around the inion (see Figure 15) and the distance from the bridge of the nose to the inion as a vertical reference (see Figure 16). We instrumented the two participants that worked on Sector 22 with the EEG equipment. The R-side participant also wore facial electrodes.

Figure 15. Head circumference measurement around the inion.

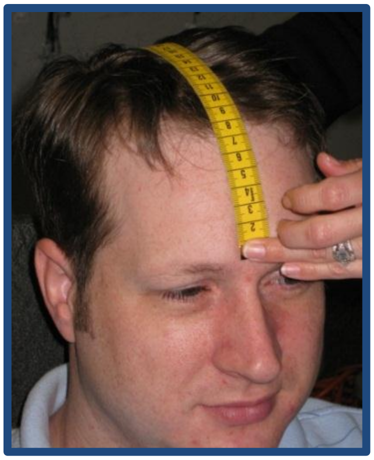


Figure 16. Vertical measurement from the nasion to the inion.

The application of the electroconductive gel took approximately 30 minutes for each participant before the first session each morning. The participants wore the caps for the duration of the day to avoid constant reapplication. For 5 minutes after the run started, we spot-checked the impedance levels of a few electrodes on each participant's head and corrected them as necessary.

During the run, triggers sent from DESIREE to the EEGs' amplifiers provided a synchronizing mark after each minute of recording. This provided points of reference for the data analyses, which requires correlation between brain activity levels and simulation events.

At the end of each run, we unplugged participants' caps from the computer, bundled all loose cables, and secured them to each participant's left shoulder. After the final run of the day, we carefully removed all head and facial electrodes. We provided towels to participants, so they could wipe the gel from their heads, as well as shower facilities and shampoo to clean their heads of any remaining gel.

For the 30-minute training sessions, participants did not wear any equipment. Each of the two participants working on Sector 20 wore fNIR and eyetracker during the test session. Prior to participants returning from their break after the training, we conducted an initial calibration for both eyetrackers using a grid of 17 calibration dots. During this process, the experimenter verified that the eyetracker was tracking accurately and recalibrated the device as necessary.

Upon returning from each break after training, we applied the fNIR and eyetracker to each participant. However, application of the equipment did not start until the EEG caps were filled with gel. This helped minimize the amount of time participants had to wear the fNIR and eye-tracking devices.

First, we applied the fNIR headband to each participant. After ensuring that it was secure and in proper position, we placed the eyetracker over the fNIR on each participant's head. After the eyetracker was secure and we found his or her pupil with the monocle, we calibrated it using the calibration dots.

After the EEGs, fNIRs, and eyetrackers were ready, an experimenter initiated DESIREE to start the experiment. As soon as it started, triggers were sent to the fNIRs and eyetracker to initiate the recording. The triggers also stopped the recording of both devices at the end of each session. Table 2 shows the schedule that we used.

Table 2. Experimental Schedule

Conflict Resolution Advisory Experiment							
Mon	**Tue, Wed, Thu, & Fri**	**Sat & Sun**	**Mon**	**Tue**	**Wed**	**Thu**	**Fri**
				Baseline	**D-side**	**Both (R- & D-sides)**	
TRAVEL	**High Altitude Experiment**		Intro (8 to 9:30) Intro to Simulator (9:30-10) Training (10 -10:45) Break (10:45 - 11) Training (11 - 11:45) Lunch (11:45 – 12:45) Training (12:45 – 13:30) Break (13:30 – 13:45) Training (13:45- 14:30) Break (14:30 – 15) Training (15 – 16) Post-scenario (16 – 16:15) Debriefing (16:15 -16:30)	Training (8 – 8:30) Break (8:30 – 9) Test (9 – 10) Break (10 – 10:15) Training (10:15 – 10:45) Break (10:45 – 11:15) Test (11:15 – 12:15) ***Lunch (12:15 – 13:15)*** Training (13:15 – 13:45) Break (13:45 – 14:15) Test (14:15 – 15:15) Break (15:15 – 15:30) Debriefing (15:30– 16:30)	Training (8 – 8:30) Break (8:30 – 9) Test (9 – 10) Break (10 – 10:15) Training (10:15 – 10:45) Break (10:45 – 11:15) Test (11:15 – 12:15) ***Lunch (12:15 – 13:15)*** Training (13:15 – 13:45) Break (13:45 – 14:15) Test (14:15 – 15:15) Break (15:15 – 15:30) Debriefing (15:30– 16:30)	Training (8 – 8:30) Break (8:30 – 9) Test (9 – 10) Break (10 – 10:15) Training (10:15 – 10:45) Break (10:45 – 11:15) Test (11:15 – 12:15) ***Lunch (12:15 – 13:15)*** Training (13:15 – 13:45) Break (13:45 – 14:15) Test (14:15 – 15:15) Break (15:15 – 15:30) Debriefing (15:30– 16:30)	**TRAVEL**

Note. In the morning of the training day, we did not use pilots.

2.6 Experimental Design and Analysis

2.6.1 Independent Variables

In this experiment, we investigated the effect of the CRA on air traffic control. Although there were several variables that might interact with the effect of CRA, we held them constant for this experiment. For example, we had included Data Comm—because we expect Data Comm to be present in en route when CRA becomes available—but we held the percentage of Data Comm equipped aircraft under control in the sector at 30%. We could still investigate how controllers used CRA differently when aircraft were Data Comm equipped by comparing aircraft with and without Data Comm within a traffic sample, but not as an independent variable which would require additional experimental conditions. Similarly, we provided controllers with SepMan capabilities, but we did not investigate how controllers used different SepMan capabilities.

The volume and complexity of traffic also affect the use of CRA. For the test scenario, we used a traffic sample that steadily increased from 33% to 150% of the MAP. We used the same traffic scenario for each experimental run, so traffic levels did not differ between scenarios, and only the call signs of aircraft were different between scenarios.

The use of traffic samples that ramped up from 33% to 150% of MAP values provided us the opportunity to investigate how the use of CRA changed with increasing traffic levels without the creation of additional scenarios that had different levels of traffic volume. The other advantage of using these ramped traffic scenarios was that we could use them to run regressions to see how a dependent variable varied with traffic level and CRA conditions. We have used this approach in the past to determine the impact of the availability of Data Comm on controller workload (Hah, Willems, & Phillips, 2006). We reported that with the 70% Data Comm equipage level, controllers could control up to 28 aircraft instead of 21 aircraft that they could handle without Data Comm.

The usefulness and effectiveness of the CRA tool may also depend on which position, R-side and/or D-side, uses it. In this experiment, we manipulated the availability of CRA in three ways: Baseline, D-side only, and R- and D-sides. The baseline was the anticipated SepMan platform without CRA. In the D-side condition, CRA was available on the D-side only. In the R-side and D-side condition, CRA was available on both sides. We describe the three experimental conditions in detail in the next section.

2.6.1.1 Baseline: No CRA

The Baseline condition had many of the anticipated SepMan capabilities, but it did not have CRA.

- Conflict Detection and Trial Planning on the R-side as well as on the D-side;
- Direct-To-Fix Fly-Out menus on the R-side;
- Probed Altitude, Speed, Heading, and Direct-To-Fix menus on the R- and D-sides (i.e., each of the values in the Fly-Out menus showed the color coding based on trial planning results);
- Electronic Coordination between sectors;
- Conflict Alert and Conflict Probe support in airspace volumes requiring 5-nmi separation as well as in volumes allowing 3-nmi separation;

- Off route indication based on aircraft/crew capabilities for Required Navigation Performance (RNP)-based routes; and
- Tools to assist controllers to detect and monitor wake turbulence separation violations. The Baseline environment also had the anticipated Data Comm capabilities for both R- and D-sides that are listed as follows:
    a. Uplink Altitude, Heading, Speed, Direct-To-Fix, Altitude Crossing Restrictions, Speed Crossing Restrictions, and Route clearances.
    b. Uplink a Transfer of Communications clearance to contact the next sector on its frequency or to monitor the next sector's frequency.
    c. Receive pilot requests.

2.6.1.2 D-side only: CRA

The D-side only condition had CRA in addition to the capabilities of the Baseline condition. The R-side controller did not have CRA but had clearance reminders resulting from the D-side use of the CRA.

2.6.1.3 Both R-side and D-side: CRA

In this condition, both R- and D-sides had CRA in addition to the Baseline condition capabilities. Both of the R- and D-sides had clearance reminders.

2.6.2 Dependent Variables

The human, monetary, and organizational resources involved in a large scale experiment are enormous. Controllers are highly regarded for their skills and are difficult to release from the facilities that need them. To maximize the return in such an investment, we collected many dependent variables. In this experiment, we collected the following data.

- **Controller behavior.** We collected all the detailed controller interactions with the simulator, pilots, and controllers in the next sector, which included keyboard input, mouse input, and voice communications. We audio and video recorded all of their actions.
- **Instantaneous subjective workload assessment.** We used the WAK to collect controller workload every 2 minutes.
- **Data of system variables.** We collected system variables including data on the number and type of controller interactions with the system, conflict alerts, conflict probe alerts, and distance and time flown in the sector. We also collected the data of aircraft trajectories and conflict probe predictions of aircraft trajectories.
- **Data of post-scenario questionnaire.** After completing each experimental run, the participants provided ratings of their previous run on major ATC tasks—that is, to identify aircraft, issue clearances, monitor air traffic situations, resolve aircraft conflicts, manage air traffic sequences, route or plan flights, assess the weather impact, and manage sector/position resources (Alexander, Alley, Ammerman, Hostetler & Jones, 1988). We also asked about the effect of the CRA on ATC. The participants had the opportunity to

provide responses to open-ended questions and to include other comments that they considered relevant (see Appendix C).

- **Post-scenario CRA assessment and exit questionnaire.** We created specific questions addressing the usability, accuracy, and impact of the CRA (see Appendices C and D). We asked participants to provide ratings and comments for each of the questions. We also collected over-the-shoulder ratings from supervisors who watched controller performance (see Appendix E).
- **Functional near-infrared spectroscopy data.** We collected fNIR Spectroscopy data to assess cognitive workload.
- **Eye-movement data.** We collected controllers' eye-movement fixations and related to them to their activities.
- **EEG data.** We collected EEG data to determine types and locations of brain activities during the experiment.

2.6.3 Research Questions

We concentrated on three research questions: Capacity, Safety, and Efficiency. We considered both controller and systems aspects. We established the following hypotheses before the data collection and reported the test results (see Appendix F for the detailed schedule).

- **Capacity.** The capacity area includes standard traffic measures of traffic through put (aircraft time and distance in the sector) as well as participants' perceived workload.
    a. A reduction in workload will lead to a reduction in capacity
        - workload will be less than baseline when CRA is available
        - baseline or D-side only when CRA is available on both positions
        - baseline when CRA is available on D-side only
    b. The controllers will issue fewer clearances per aircraft than
        - baseline when CRAs are available
        - baseline or D-side only when CRA is available on both positions
        - baseline when CRA is available on D-side only
    c. The controller workload as a function of the number of aircraft under control will increase less than
        - baseline when CRAs are available
        - baseline or D-side only when CRA is available on both positions
        - baseline when CRA is available on D-side only

- **Safety.** The safety area includes the number and duration of strategic conflict probe alerts, the number and duration of tactical conflict alerts, the number and severity of instances of loss of separation, and the measures of risk of loss of separation.
    a. The number of conflict probe alerts will be fewer than
        - baseline when CRAs are available
        - baseline or D-side only when CRA is available on both positions
        - baseline when CRA is available on D-side only
    b. The number of conflict alerts will be less than
        - baseline when CRAs are available
        - baseline or D-side only when CRA is available on both positions
        - baseline when CRA is available on D-side only
- **Efficiency.** The efficiency area includes both measures of controller and system efficiency. Measures of controller efficiency coincide to some extent with measures of workload (e.g., number and duration of controller clearances). Measures of system efficiency include time and distance flown or fuel used.
    a. The time to resolve a conflict will be less than
        - workload will be less than baseline when CRA is available
        - baseline or D-side only when CRA is available on both positions
        - baseline when CRA is available on D-side only
    b. Resolutions to potential conflicts will be more fuel efficient than
        - baseline when CRAs are available
        - baseline or D-side only when CRA is available on both positions
        - baseline when CRA is available on D-side only

3. RESULTS

We report the results of the data reduction and analysis in the following subsections. Although more reduced data sets will become available, due to time constraints, we will deliver this version of the final report with the results available to date.

We report the results of objective data first and then subjective data. In this section, we inserted some descriptions that would normally be discussed in either the Method or the Procedure section. For instance, the front parts of fNIR and coordination between sectors are introductions and methodology for those particular data sets. We preferred this format because we could present results in a more focused and tailored manner to each data set.

3.1 Workload

We analyzed the 24 WAK events from 2 min to 48 min in each 50-min test run. There were high incidences of missed responses: 14.8% averaged across all experimental runs. Missed responses may be a measure of workload; controllers may be less likely to respond to the WAK prompt when they are busier. We designed the traffic volume to increase with time in the simulation runs, so we expected to see the effects over time within runs. We aggregated data by 12-min segments within each run (i.e., four segments including six WAK prompts each) for data-smoothing purposes. Figure 17 shows the number of missed ratings per time segment by controller position and CRA presentation condition.

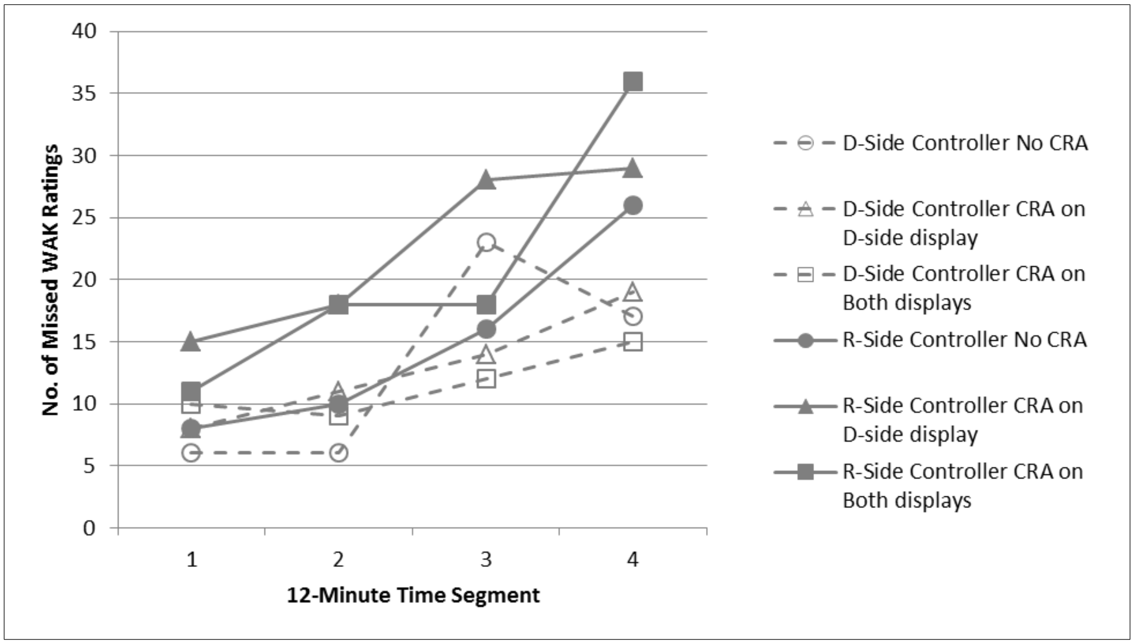


Figure 17. Missed WAK ratings.

We analyzed the missed-response data using a repeated measures analysis of variance (ANOVA). The presentation of CRA information did not significantly affect the likelihood of failing to respond to the prompt, $F(2, 16) = 1.054$, $p > 0.05$. Missed-response rates did not differ significantly between R-side controllers and D-side controllers, $F(1, 8) = 1.496$, $p > 0.05$; however, controllers missed significantly more WAK responses when working sector ZKC_20 than ZKC_22, $F(1, 8) = 6.768$, $p < 0.05$. There were more missed responses to the WAK prompt later in scenarios (Time Segments 3 and 4, when traffic was heavy), $F(3, 24) = 12.596$, $p < 0.01$.

As controllers become busier, they may take longer to respond to the WAK prompt. Controllers are instructed to treat the WAK prompt as a low priority, not to allow it to interfere with traffic separation, so we expect them to complete active tasks before responding to the prompt. Figure 18 shows latencies (elapsed time to respond to the prompt) in seconds.

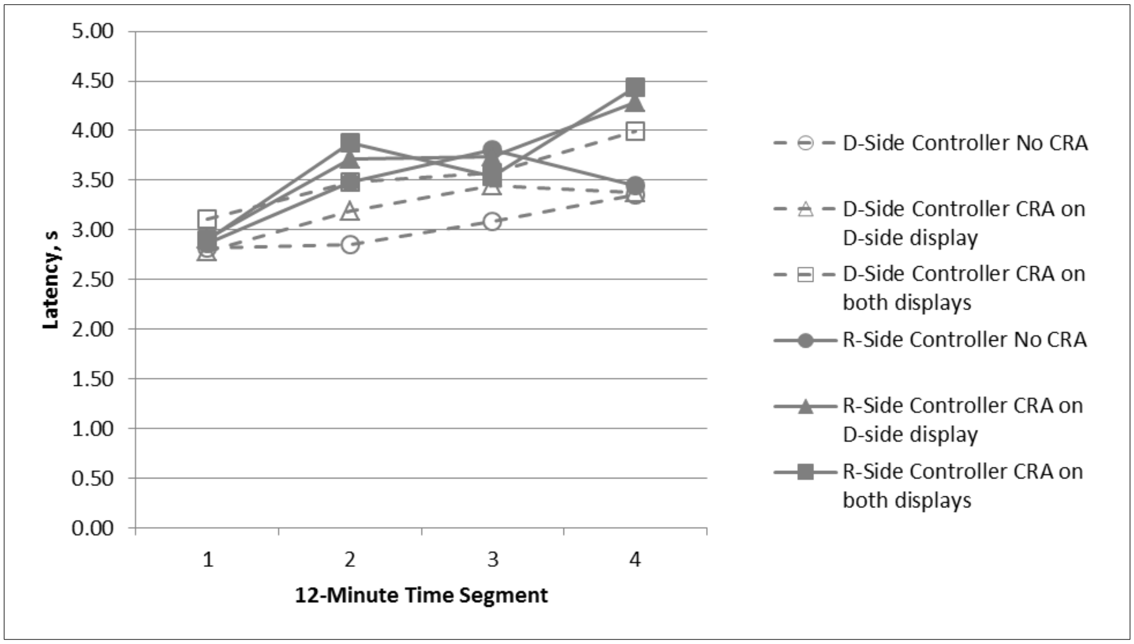


Figure 18. Response latency to WAK prompt.

We performed a repeated measures ANOVA on the latency of WAK responses. The presentation mode of CRA information significantly affected the latency of response to the prompt, $F(2, 16) = 4.395, p < 0.05$. The difference in latency between Radar- and Data-side controllers is not significant, $F(1, 8) = 2.556, p > 0.05$. Controllers took longer to respond to the WAK prompt later in the scenarios, $F(3, 24) = 5.751, p < 0.01$.

We performed a repeated measures ANOVA on the average scaled ratings across teams and conditions. The differences in ratings among the presentation modes of CRA information were not statistically significant, $F(2, 16) = 0.895, p > .05$. The difference in ratings between Radar- and Data-side controllers was not statistically significant, $F(1, 8) = 1.461, p > .05$. Workload ratings increased over the course of the simulation runs; the effect of time segment was significant, $F(3, 24) = 179.262, p < 0.01$.

Ratings averaged across both positions (see Figure 19). Differences in workload between the CRA conditions were small. The busier, later time segments clearly increased the controllers' workload substantially.

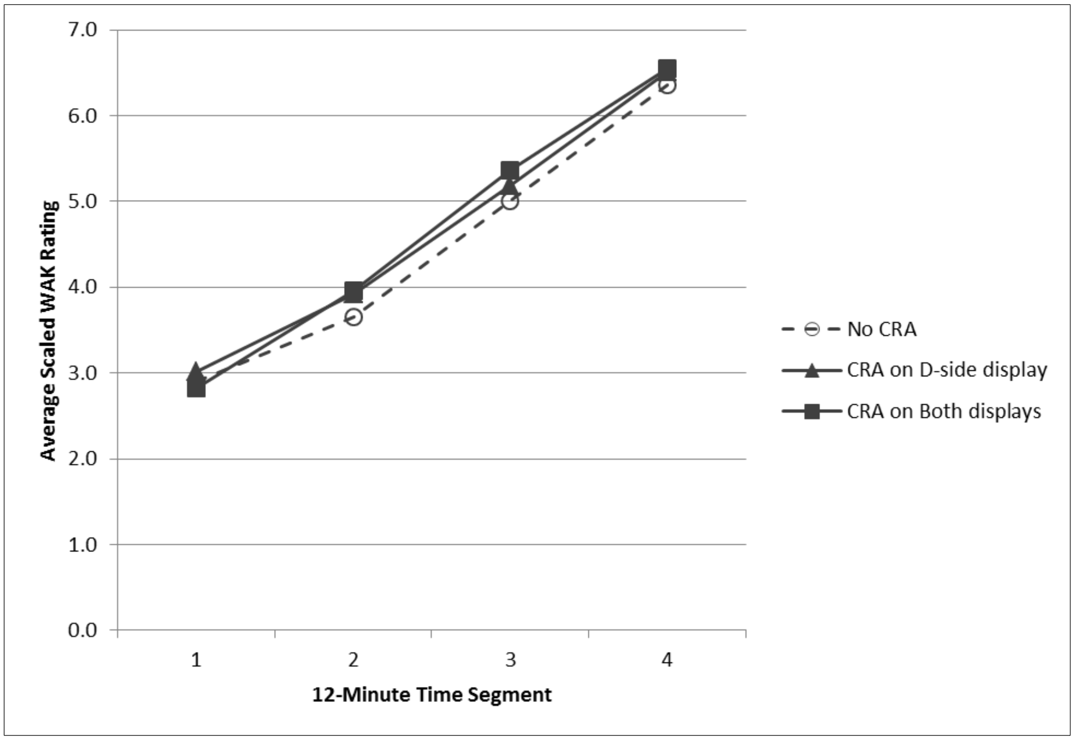


Figure 19. Average scaled WAK ratings.

We wanted to know if the initial training of participant controllers was sufficient to achieve familiarity with the airspace and proficiency with using the emulated ERAM interface and the CRA features. If controllers were still putting effort into learning new features during early test runs, we would expect them to report higher workloads initially and report declining workloads as they gained proficiency. This effect would be superimposed on any effects of experimental conditions. The counterbalancing of experimental conditions isolates them from time- or sequential-order effects. Figure 20 shows the trend of ratings across consecutive test days.

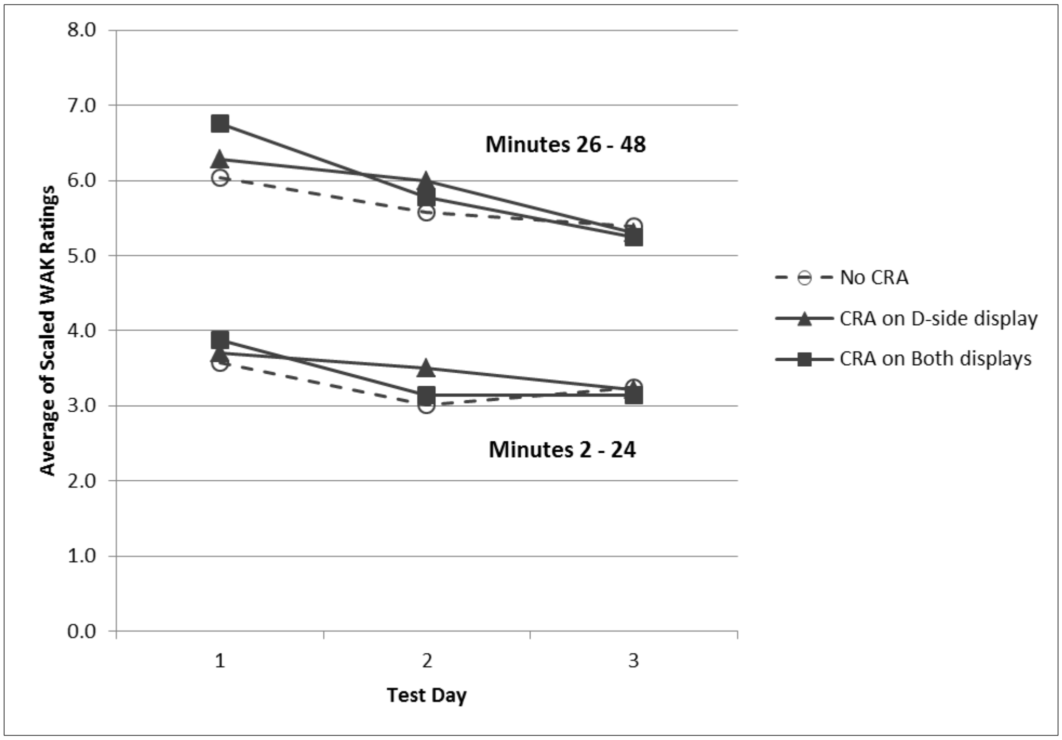


Figure 20. Average of scaled WAK ratings by day.

Average scaled ratings of subjective workload declined from the first to the second test day for all CRA-presentation conditions. Workload ratings in the later, busier half of the simulation run continued to show declines from the second to the third test day—suggesting that controllers were still expending some effort on learning, which was enough to be noticeable in busy conditions.

3.2 Measures: Time and Distance Flown

We used time and distance flown by aircraft inside sector boundaries as one measure of controller efficiency. In each experimental run, the same aircraft types flew the same routes on the same schedule. Only those controller actions that changed aircraft ground tracks were likely to lead to substantive differences in distance flown and/or time in the physical sector. Other aircraft time and distance data might show small variations due to changes of altitude, but these would be quite small. We filtered the time and distance data set to create a subset of aircraft (identified by beacon code) that were given a clearance that altered their ground tracks in at least one simulation run. There were 48 aircraft in the subset that flew in Sector ZKC_20 and 42 aircraft in ZKC_22 (16 of these flew in both sectors). We compared this subset of beacon codes across all runs.

Figure 21 shows the total time flown in controlled sectors by the subset of flights that flew varied ground-tracks. All measures are shown, although any effect should be most distinguishable in the Geographic Bounds measure. Figure 22 shows the total distance flown in controlled sectors by the subset of flights that flew varied ground-tracks.

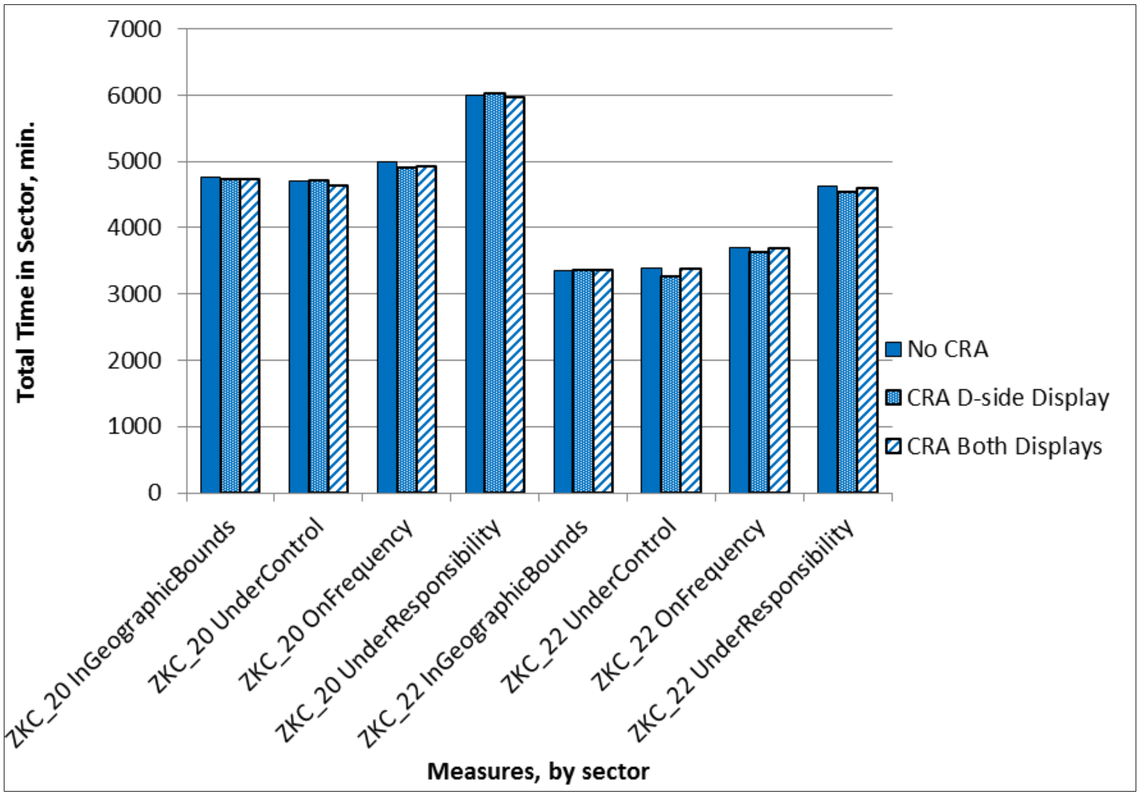


Figure 21. The durations aircraft were in the sectors.

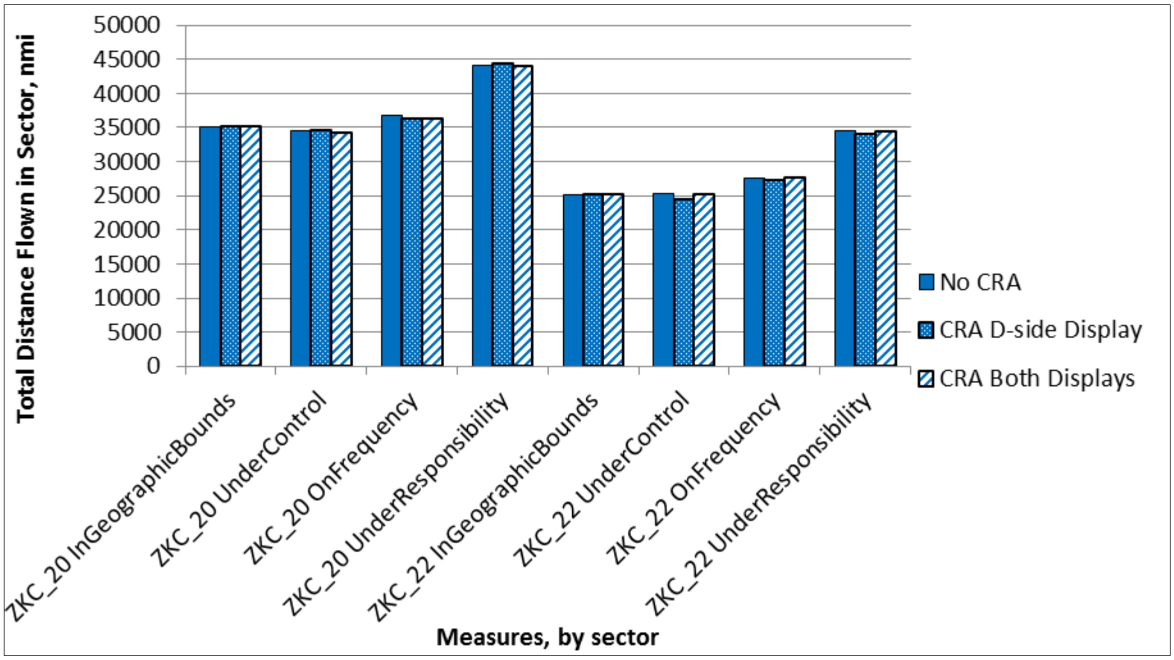


Figure 22. The distance aircraft flown in the sectors.

We used a repeated measures ANOVA to analyze the subset. The unit of analysis for the time and distance data was the team-pair of one R-side and one D-side controller. We analyzed the three methods of CRA display (None, D-side only, R- and D-sides), and the sector controlled (ZKC_20 or ZKC_22) as within-teams effects (The effect of sector was not of interest, per se, but it was included to account for the variation due to sector and route geometries). Table 3 shows the *F*-ratios and *p*-values for the time measure, and Table 4 shows the distance measure (nmi). Sectors ZKC_20 and ZKC_22 showed significant differences in distance flown, as expected. The effect of CRA presentation was not significant for any of the eight measures.

Table 3. Time Flying in Sector - Analysis

Time in Sector (min)	Within-Team Effect	*F*	*p*
In Geographic Bounds	Sector	1459.466	0.000
	CRA_location	0.118	0.889
	Sector*CRA_location	0.526	0.601

Table 4. Distance Flown in Sector - Analysis

Distance Flown (nmi)	Within-Team Effect	*F*	*p*
In Geographic Bounds	Sector	1503.993	0.000
	CRA_location	0.174	0.842
	Sector*CRA_location	0.459	0.640

Sector and route geometries may result in overall shorter origin-to-destination flight paths, but a longer path within any given sector. Conversely, a trajectory change that shortens the time/path within the sector could lengthen the overall path. Such effects could possibly obscure actual more-efficient routings. The simulation system projects time points along the trajectory of each aircraft, updating as needed. We extracted a data set of the final estimated time for arrival at the destination of each flight. Following the simulation, we also conducted simulation runs using the test scenario in two "baseline" conditions. Baseline 1 was "hands-off"; no inputs were made at either the controller or pilot workstations. In Baseline 2, a Subject Matter Expert—SME, a retired controller and air traffic control (ATC) supervisor—climbed departing flights, and descended arriving flights, but performed no tactical separation. We extracted the time-at-destination/last-fix for the baseline runs as well. Starting times of flights were nominally constant, so by comparing times for experimental runs to the times for the baseline runs we were able to identify flights that were delayed (predicted to arrive later than the baseline time) and those that had a saving (predicted to arrive earlier than baseline). Savings or delays calculated from times-at-destination are not affected by sector geometry. For this analysis, we included only overflights of the test sectors; not arrivals to or departures from local airfields. Because overflights flew their as-filed flight plans in both Baseline 1 and Baseline 2 simulation runs, we were able to average times for the two baseline runs to calculate time savings. Figure 23 shows the total time savings for overflights compared to the baseline average.

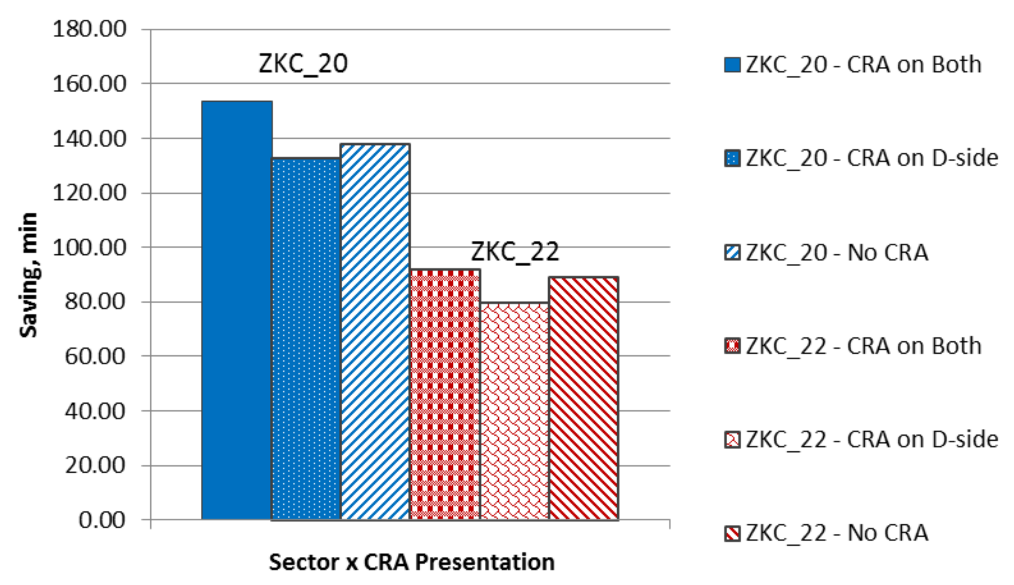


Figure 23. Total time saving.

Although sectors differed significantly in the time savings for overflights, $F(1, 8) = 8.12$, $p <= 0.021$, there was no significant effect of the CRA presentation condition, $F(2, 16) = 0.568$, $p > 0.578$. Figure 24 shows the total time delay for overflights. None of the apparent differences in delays across sectors or CRA presentation condition were statistically significant.

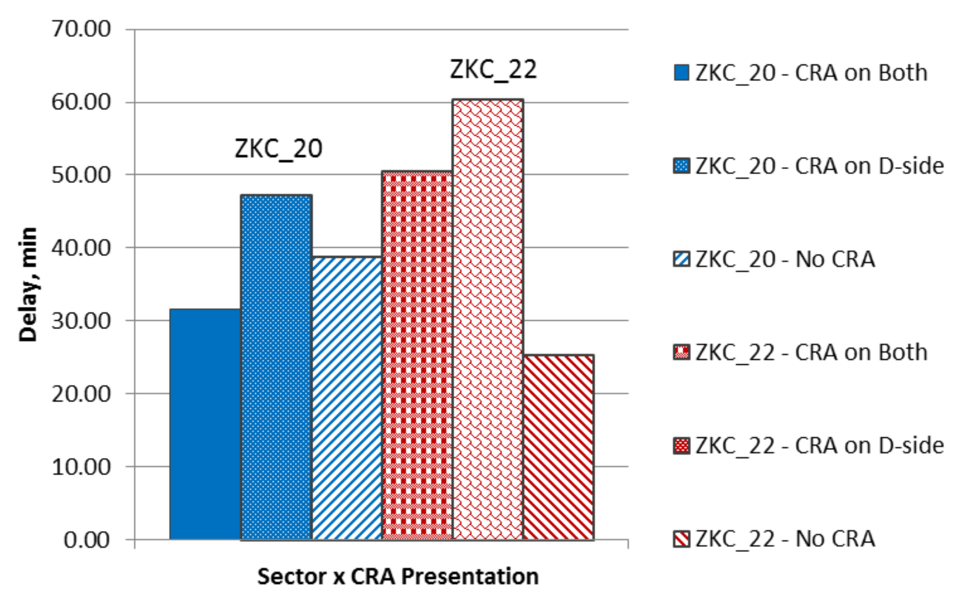


Figure 24. Total delay time.

3.3 Fight-Out Menus Interactions

There are many differences between the CRA use and the emulated ERAM (Baseline) use. First of all, most of the time controllers must use the Fly-Out menus to use the CRA. In the Baseline condition, controllers could type commands using the keyboard or generate commands by interacting with the flight-out menus as in the CRA condition. Because the CRA commands that were also used in the Baseline condition were altitude, heading, and speed commands, we present those frequencies in Figure 25. In the following, we compared the two conditions with the controller use of those commands only. In addition, because R- and D-side roles were quite different, we compared their activities at the R-side and D-side separately. It is obvious that controllers preferred using altitude clearances. Figure 25 clearly shows that the orders from the most frequent to least frequent were always Altitude first, Heading second, and Speed third in all control positions. In Figure 25, the x-axis abbreviations represent display condition (either CRA or Baseline), flight-menu field selected, and controller position (either R-side or D-side) as described in the figure's caption.

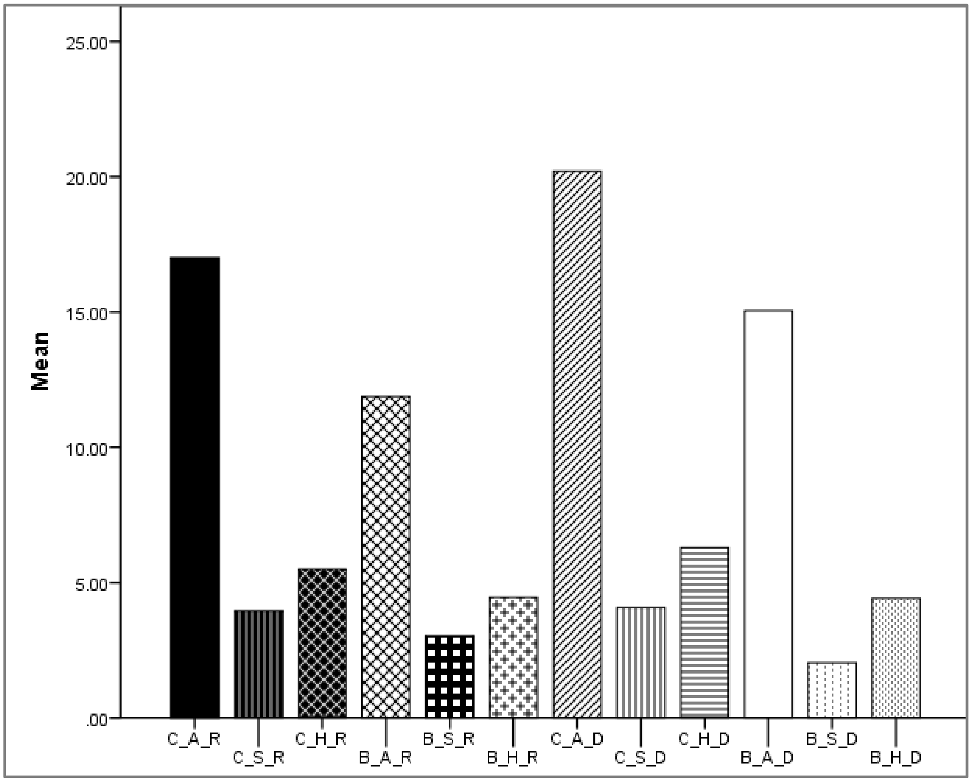


Figure 25. Number of clicks on datablock flight-parameter fields (Note. C_A_R: CRA_Altitude_Rside; C_S_R: CRA_Speed_Rside; C_H_R: CRA_Reading_Rside; B_A_R: Baseline_Altitude_Rside; B_S_R: Baseline_Speed_Rside; B_H_R: Baseline_Heading_Rside; C_A_D: CRA_Altitude_Dside; C_S_D: CRA_Speed_Dside; C_H_D: CRA_Reading_Dside; B_A_D: Baseline_Altitude_Dside; B_S_D: Baseline_Speed_Dside; B_H_D: Baseline_Heading_Dside).

3.4 Comparison of CRA and Baseline Flight-Out Menus Interactions

In our experiment, the CRA datablocks were available to the R-side controllers only when they were in the Both-display conditions. The Baseline datablocks were available only when they were in the Baseline condition. For the D-side position, we compared their activities during the D-side condition (where they had the CRA) with their activities when they were in the Baseline condition.

3.4.1 Frequency of the Flight-Out Menu Interactions in CRA and Baseline

We compared the frequencies of activities per controller because each controller had very different ways of interacting with the datablock fields (see Table 5). When participants had two experimental runs, we summed their frequencies and divided by two to obtain an average frequency per run.

Table 5. Average Frequencies of use at the R-side and D-side Positions per Run

	R-side		*D-side*	
Participant	**CRA**	**Baseline**	**CRA**	**Baseline**
1	63	58	52	53
2	39	3	16	15
3	57	12	54	24
4	7	35	41	26
5	11	8	28	12
6	5	10	12	5
7	9	8	30	14
8	37	31	45	34
9	23	16	28	28
10	17	13	22	11
11	34	12	39	9
12	27	10	33	18
Mean	**27**	**18**	**33**	**21**

For the R-side data, the test results showed that they used CRA significantly more often than Baseline (Wilcoxon Signed Ranks Test, $Z = 1.883$ ($p = .029$) which is larger than $Z = 1.645$ at $p = .05$ (one-tailed). The mean frequencies of CRA and Baseline datablock field uses were 17.88 and 27.25 per run, respectively. Only 2 out of 12 participants used Baseline flight-out menus more often than CRA's. For the D-side, we found only one participant used Baseline menus more often than CRA's, 52 vs. 53. The Wilcoxon Signed Ranks Test showed controllers used CRA menus significantly more often than the Baseline's, $Z = -2.759$, $p = .0048$.

3.4.2 Duration of the Flight-Out Menu Interactions in CRA and Baseline (in seconds)

Again, there were large individual differences in the duration of the CRA and Baseline use (see Table 6). We used *t*-test because duration is a continuous variable. The results of the paired *t*-tests showed that controllers spent more time when they used CRA than when they used Baseline at the D-side ($t = 2.585$, $df = 11$, $p = .025$) but not on the R-side ($t = 1.504$, $df = 11$, $p = .161$). The mean difference between CRA and Baseline for the R-side was 0.9 seconds and the difference for the D-side was 1.4 seconds. Most of them spent more time when they used CRA at the D-side ($M = 5.4$ seconds) than when they used it on the R-side ($M = 4.4$ seconds).

Table 6. Average Frequencies at the R-side and D-side Positions per Run

	R-side		*D-side*	
Participant	**CRA**	**Baseline**	**CRA**	**Baseline**
1	3.5	2.9	5.7	3.6
2	6.8	4.6	8.4	5.1
3	5.4	3.9	6.0	7.8
4	2.7	4.7	5.2	4.9
5	5.4	3.4	5.5	4.9
6	2.8	6.5	1.9	1.4
7	5.7	5.3	4.0	3.7
8	3.0	2.5	3.8	2.2
9	2.7	1.9	6.8	4.8
10	4.2	1.1	5.3	5.8
11	4.8	3.2	7.8	2.9
12	5.2	2.2	6.7	2.4
Mean	**4.4**	**3.5**	**5.4**	**4.0**

3.4.3 Wrong Altitude for Direction of Flight (WAFDOF)

In the field, the westbound over-flights should be on the even altitudes, and eastbound over-flights should be on the odd altitudes. If not, it is a deviation unless it was allowed by Approval for Request (APREQ). However, CRA did not consider it in presenting its resolutions. Whenever controllers encountered wrong altitude suggestions, they performed extra steps and used the altitude that was not WAFDOF (see Figure 26a, b, c, and d).

- **Picture (a):** ***Time - 0 min.*** As the controller clicks the red indicator with the left mouse button, the CRA flight-out menu pops out with the selected aircraft highlighted. The green colored numbers are the ranks by the CRA, and the best resolution was ranked as 1 and highlighted.
- **Picture (b):** ***Time - 13.5 sec.*** As the controller did not like the suggested altitude, he selected the ALT option at the bottom in the window by clicking the left mouse button to choose his own altitude.

- **Picture (c):** ***Time - 20.9 sec.*** He chooses 40,000 ft in green by clicking it with the middle mouse button. This action made the system to show the route of the aircraft, SWA576, automatically.
- **Picture (d):** ***Time - 32.1 sec.*** This action activates Data Comm, which sends it to the aircraft (blue square highlighted).

Figure 26. An example of controller's activities to avoid WAFDOF recommended by CRA.

The controller could have saved about 20 seconds if the CRA gave him the correct altitude (32.1 seconds − 12.5 seconds = 19.6 seconds). This faulty design of the CRA forced controllers to spend more time, unnecessarily, and might have given them a general impression that the CRA was not a good tool to use.

3.4.4 The Relationship Between Conflict Alert Indicators and CRA

A red conflict-alert indicator, showing a projected time to encounter a conflict, appeared at the zero line of the datablock (as previously shown in Figure 26a). We assumed that controllers might have used CRA flight menus comparatively more often when the indicators were on than when off when we compared the same two situations in the Baseline condition. As Table 7 and Table 8 show, that was not the case. The controllers' frequency of use of the CRA flight menus was similar to the Baseline whether the indicators were on or off. The Chi-square results for Tables 7 and 8 were $X^2(1, N = 12) = .3550$ and $X^2(1, N = 12) = .0096$, respectively. The value $X^2(1, 12)$ is 3.841 at $p = .005$.

Table 7. Average Frequencies of Flight-Out Menu use per Run for the R-side

Indicator	*CRA*	*Baseline*	*Total*
OFF	19.4	14.5	33.9
ON	7.0	4.9	11.9
Total	26.4	19.4	45.8

Table 8. Average Frequencies of Flight-Out Menu use per Run for the D-side

Indicator	*CRA*	*Baseline*	*Total*
OFF	21.0	16.4	37.4
ON	9.6	5.1	14.7
Total	30.6	21.5	52.1

3.5 Push-To-Talk

A multivariate, two-way (3 x 4) repeated measures analysis of covariance (ANCOVA) was carried out to determine whether using varying levels of CRA (Baseline, D-side Only, and R- and D-side) and Time interval (2-14 min, 14-26 min, 26-38 min, and 38-50 min of the scenario) had any effect on the total number of PTT communications and the average duration of each PTT communication using the number of aircraft each participant had averaged across all conditions as a covariate. Time interval was used as a second independent variable to determine whether the length of the scenario had any effect on the data. This analysis used the data from the R-sides since this is where the PTT's originated predominantly.

The multivariate ANCOVA showed a significant effect of CRA, $F(2, 15) = 4.36$, $p = .032$. The univariate test of CRA was examined and showed a nonsignificant $p = .052$. This nonsignificance led to no statistically significant pairwise comparisons between the levels of CRA. Figure 27 shows the data trends. There was no significant effect of PTT frequency, $p > .05$. There was no significant effect of CRA or Time interval on average PTT duration during any of the four time intervals, $p > .05$ (see Figure 28 for detailed data).

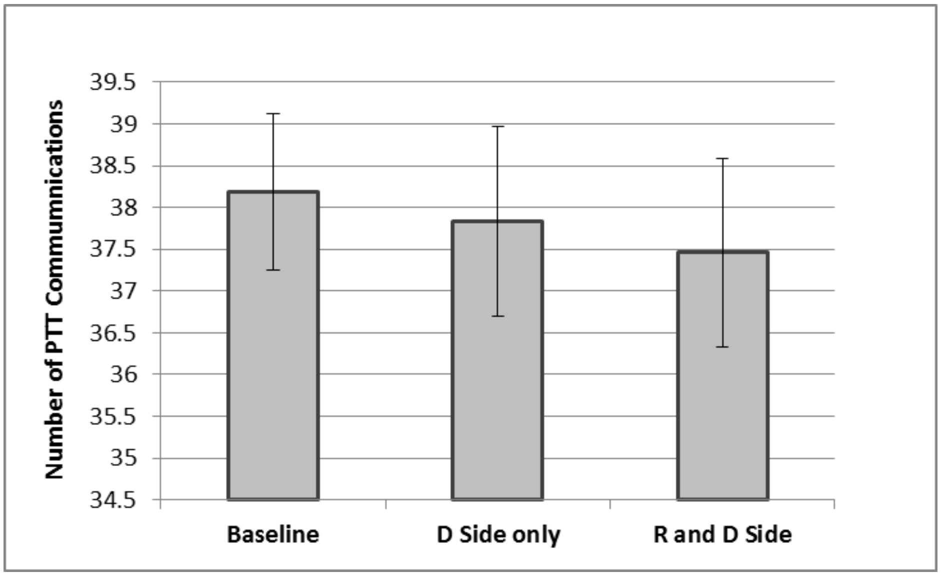


Figure 27. PTT counts by different CRA conditions.

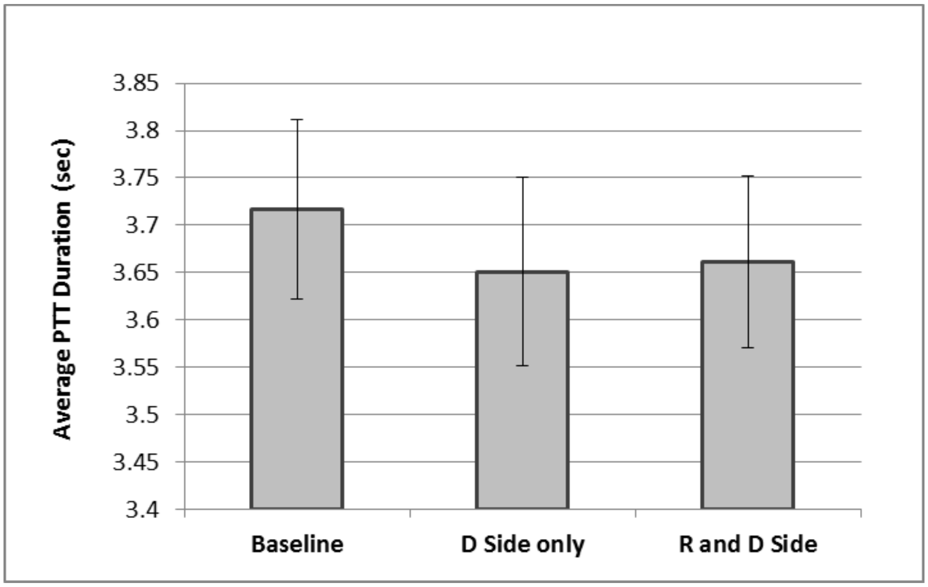


Figure 28. PTT durations by different CRA conditions.

3.6 Manual Handoffs

A multivariate, two-way (3 x 4) repeated measures ANCOVA was carried out to determine whether using varying levels of CRA (Baseline, D Side Only, and R- and D-side) and Time interval (2-14, 14-26, 26-38, and 38-50 minutes of the scenario) had any effect on the number of handoffs the R- and D-side controllers completed while using the number of aircraft each participant had averaged across all conditions as a covariate. Time interval was used as a second independent variable to determine whether the length of the scenario had any effect on the data. The R-side and D-side data were analyzed together as well as separately.

A multivariate ANCOVA showed a significant effect of CRA, *F*(2, 15), *p* = .030. The pairwise comparisons were examined and when CRA was available on the D-side only, there were significantly more initiated handoffs (*M* = 3.69, *SE* = .61) than when CRA was available on the R-side only (*M* = 2.92, *SE* = .47), *p* = .031. There was no significant effect of Time interval, *p* > .05 (see Figure 29 for detailed data).

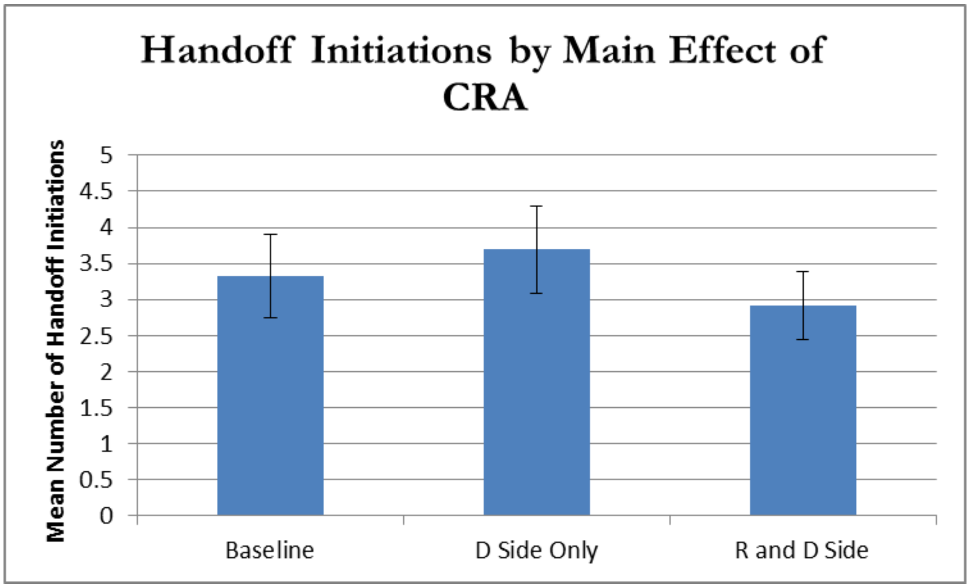


Figure 29. Manual handoff frequencies by the CRA conditions.

3.7 Controller Commands

For the analyses of controller entries, we included all entries—that is, independent of the input modality (mouse, keyboard, or a combination). We used a multivariate, two-way (3 x 4) repeated measures ANCOVA to determine whether using varying levels of CRA (Baseline, D-side Only, and R- and D-sides) and Time interval (2-14 min, 14-26 min, 26-38 min, and 38-50 min of the scenario) had any effect on the number of different types of commands issued by controllers using the number of aircraft averaged across all conditions as a covariate. We used time interval as a second independent variable to determine whether the length of the scenario had any effect on the data. We used total commands used in each experimental condition and found no significant effect of CRA or Time interval, *p* > .05 for each effect (see Figure 30).

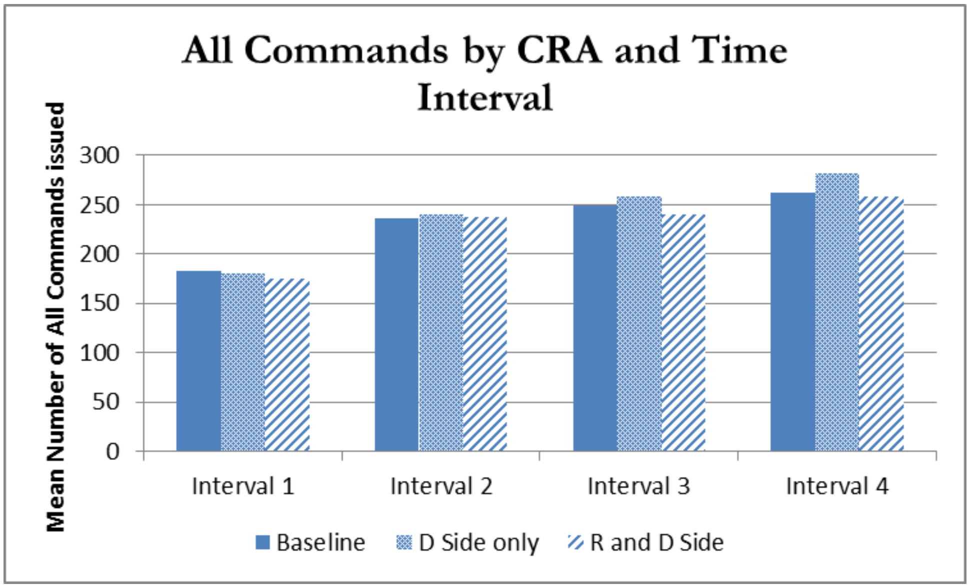


Figure 30. The frequency of total commands by CRA condition and time interval.

3.8 Individual Controller Commands

We also ran tests for individual commands separately. They were as follows:

- **Altitude_QQ** – The beginning of an Interim Altitude assignment message, which would be followed by said altitude, a space, and Computer Identification (CID).
- **Altitude QZ** – Assigned Altitude, which would be followed by an altitude, a space, and flight identification (FLID). In some facilities, the QZ command is equivalent to none or implied command QN. In that case controllers can enter anything that they would normally enter as an implied command after the QZ entry.
- **Confirm Assigned Level (example CAL DAL1590)** – Although the controller could use this, it is not a command that the controller would normally make. In the current implementation of the Data Comm prototype, the system automatically sends a CAL on initial contact (IC). DESIREE then verifies automatically that the aircraft is at the correct assigned altitude by testing if the response is AFFIRMATIVE.
- **Datablock Movement** (i.e., changing leading direction or length) – Accomplished by either dragging and dropping the datablock with a mouse (in the Future En Route Workstation Study prototype); selecting, moving, and then dropping it (as originally proposed in ERAM); or by selecting a cardinal position via the keyboard to move the datablock to (e.g., “9 FLID”).
- **Flight Plan Readout (FR N264)** – A command entered by the controller either through the keyboard by depressing the FP button (on the R-side) or the FR button (on the D-side) and then entering the FLID (call sign, beacon code, computer identification, or slew on the position symbol). An FR can also be

displayed in the Continuous Flight Plan Readout View by center clicking on the call sign of the datablock.

- **Halo (QP J 3 846)** – Displays a 5 nmi Halo around the selected aircraft. The 3 nmi Halo is not available due to the airspace configuration, so it defaults to 5 nmi.
- **Interim Altitude (QQ 240 JBU1570 or QQ JBU1570)** – Interim Altitude assignment message, which would be followed by an altitude, a space, and a FLID. If the controller omits the altitude value, this entry will remove an existing interim altitude from the aircraft representation.
- **PVD Select (QP 987 FWD or QP 019)** – Forces a Positive Visual Display (PVD) of an aircraft representation (datablock) to the selected sector. For example, depress PVD button, enter sector # to be displayed to (SP), followed by aircraft CID. The "QP 987 FWD" version is not available in ERAM, but it allows a controller to forward the command to the other controller's display. "QP 019" removes a datablock with the CID 019, previously pointed out to a controller from an adjacent sector, from her display. If a controller already owns the aircraft, the aircraft will disappear for a brief period before popping back on the display.
- **Route (QU ICT or QU 177 FWD)** – Entering the command "QU ICT" followed by a space then the CID updates the system with the controller intent to send the aircraft direct to ICT (Wichita, KS) provided ICT was in the route of flight, or another fix was added after ICT to join the previous route. Without a fix or set of fixes, the QU FLID entry toggles the display of the route for a FLID. Using the FWD suffix displays the route on the other controller's display. Currently, the use of a suffix to change the completion of the entry of a controller's intent does not exist in the en route automation.
- **Route Toggle (route display and route hide)** – Accomplished by depressing RTE key, (SP), then CID, to display the route of the selected aircraft, then repeating same to hide the route. Can also be achieved by clicking (center mouse button) in the space between the CID and SPEED areas in the third line of the datablock.
- **Select** – Controllers call the DESIREE SELECT entry a pick/enter or slew on an aircraft position symbol. The controller clicks on the position symbol with the middle button of the pointing device when she does not have track control for that the aircraft. In ERAM, this action will change the datablock into alternate datablock (ADB) if the aircraft was in a full datablock (FDB) state, a correlated limited datablock (CLDB) if the aircraft was in an ADB state, and an FDB if the aircraft was in an LDB state. In ERAM controllers refer to this as the cycle of life.
- **Uplink Held (UH JBU1570)** – Data Comm message from the en route automation telling the controller that a frequency change message is loaded and awaiting delivery to the aircraft by clicking on the brightened "H" in the FDB displayed to the left of the assigned altitude field.

None of the tests showed a significant difference between CRA display conditions (see Appendix G for the detailed test results). The analyses of all controller entries to update the system showed that the total number of entries per entry-type did not differ significantly between CRA conditions. The number of interactions increased with the 12-minute intervals, but this was likely the result of the increase of the number of aircraft over time (see Appendix G for the test results with graphs). See Appendix H for the pilot command analysis.

3.9 Access to Information in Fly-Out Menus

The CRA functions were only available to controllers through the interaction with the displayed items with the mouse. For example, to display the probed trajectories of a potential conflict, a controller clicked on the conflict pair in the Sector Queue; for access to the Aircraft Queue, a controller left-clicked on the Safety Portal; and to access the CRA Altitude Resolution menu, a controller left-clicked on the altitude field in the Full Datablock. We conducted repeated-measures ANOVAs using the multivariate approach on the number of times controllers accessed Altitude, Speed, Heading, and Safety Indictor Fly-Out menus to determine if controllers had changed their behavior and accessed functions and information through the datablock more or less often as a function of the availability of CRA.

3.9.1 Altitude Fly-Out Menus

We conducted a 2 x 3 x 4 (Position x CRA x Interval) MANOVA on the number of times controllers accessed Altitude Fly-Out menus within each 12-min interval, The multivariate tests showed significant main effects of CRA, $\Lambda = .553$, $F(1, 16) = 6.458$, $p = .009$, and Interval, $\Lambda = .275$, $F(1, 15) = 6.458$, $p < .001$. After Huynh-Feldt correction of sphericity, the main effects of CRA and Interval remained and the interaction between Position and CRA showed a trend, but did not reach significance, $F(1.9, 17)$, $p = .058$. We used Helmert contrasts to determine which levels of the dependent variables differed. At the .05 level, the Baseline was somewhat different from the conditions with CRA, but the difference did not quite reach statistical significance, $F(1, 17) = 3.188$, $p = .091$. When CRA was available on the D-side only, however, the number of times controllers accessed Altitude Fly-Out menus was significantly different than when CRA was available on the D- and R-sides, $F(1, 17) = 6.724$, $p = .019$. The first interval differed from the other intervals, and the second interval differed from the third and fourth intervals, $F(1, 17) = 27.949$, $p < .001$; $F(1, 17) = 26.674$, $p < .001$. The third and fourth intervals did not differ significantly from each another. The interaction between position and CRA seemed mostly due to differences between the R-side and D-side when comparing the baseline with the CRA conditions, $F(1, 17) = 4.058$, $p = .060$. However, as shown in the MANOVA results, the interaction did not reach statistical significance.

Figure 31 shows the effect of CRA and position on the number of times controllers accessed Altitude Fly-Out menus per 12-min interval. Under Baseline conditions, the R- and D-side controllers accessed the altitude menus the same amount of times. When we provided CRA on the D-side, the D-side controller accessed altitude menus more often than the R-side controller, resulting in an overall increase. Under conditions when CRA was available on both the R- and D-sides, the D-side controllers accessed the altitude menus about as frequently as when CRA was only available on the D-side. The R-side controllers accessed the altitude menus more often when CRA was available on both positions.

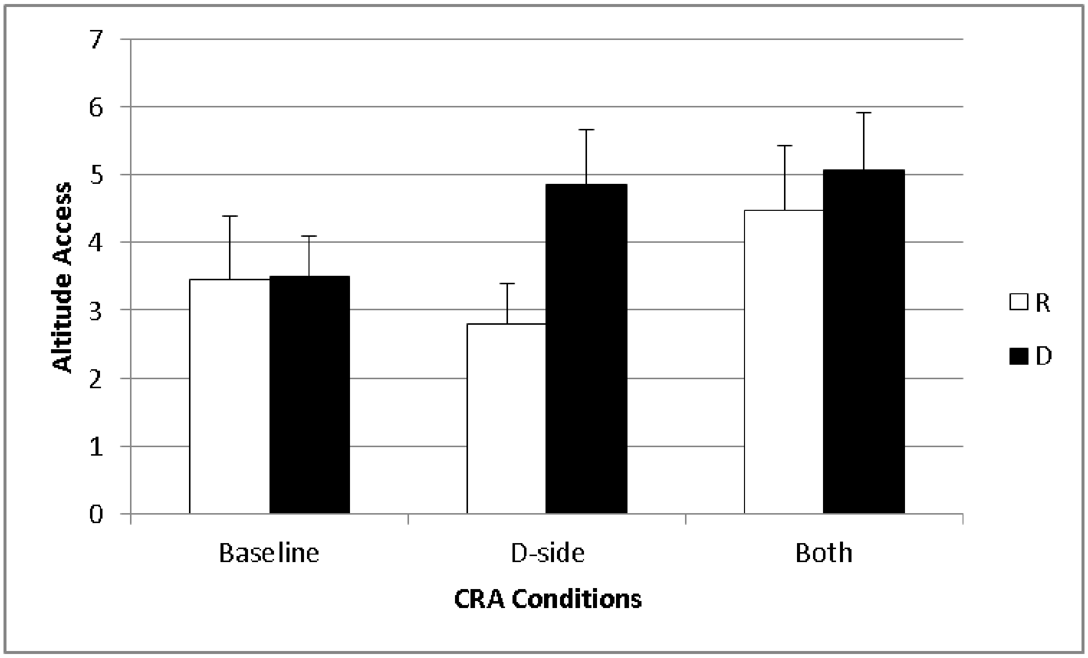


Figure 31. The number of times controllers access Altitude Fly-Out menus per 12-minute interval as a function of CRA condition and Position.

3.9.2 Access to Speed Fly-Out Menus

We conducted a 2 x 3 x 4 (Position x CRA x Interval) MANOVA on the number of times controllers accessed Speed Fly-Out menus within each 12-min interval. The multivariate tests showed a significant main effect of Interval, $\Lambda = .399$, $F(3, 15) = 7.523$, $p = .003$, and an interaction between the effects of CRA and Interval, $\Lambda = .332$, $F(6, 12) = 4.030$, $p < .019$. After Huynh-Feldt correction for sphericity, the main effect of Interval remained but the statistical significance for the interaction between CRA and Interval showed a trend, but it did not reach significance, $F(3.5, 59.4) = 0.302$, $p = .058$. The Huynh-Feldt correction for sphericity surfaced a three-way interaction between Position, CRA, and Interval, $F(4.0, 68.4) = .962$, $p = .009$. For the two-way interaction between the effects of CRA and Interval, we used a post-hoc Tukey HSD test to determine which pairs of conditions differed from one another. Controllers accessed the Speed Fly-Out menu significantly more often in the fourth interval under the CRA-D condition than in the first interval ($p =. 0.046$) and the second interval ($p = 0.023$) under the Baseline condition, respectively. We found the same pattern under the CRA-B condition—that is, in the fourth interval, controllers accessed the Speed Fly-Out menus significantly more often than in the first and second intervals under the Baseline condition ($p = 0.014$ and $p = 0.006$, respectively). Controllers accessed the Speed Fly-Out menus more often under the CRA-D in the third interval than under the CRA-B condition in the third interval ($p = 0.037$). Under the CRA-B condition in the fourth interval, controllers accessed the Speed Fly-Out menus more often than under the CRA-B condition in the third interval ($p = .011$). Although we could attempt to describe the pairwise differences within the three-way interaction between Position, CRA, and Interval, this would become rather confusing. Instead, we will discuss the interaction based on two figures that showed the interactions: one for the R-side and one for the D-side controllers (see Figure 32). The R-side controllers showed an increase in the number of times they accessed Speed FOMs with Interval. Because the overall number of speed entries did not change, the tendency for R-controllers must be to rely more on FOMs when traffic levels increased. Interval 3 does not fit that pattern—likely because the traffic levels during that time were too high and controllers reverted to keyboard-based entries. In Sector 22, the traffic level did not increase in Interval 4 and, because we used each team-sector combination as a separate observation, this might have led to the three-way interaction we found. When we focus on the impact of CRA as a function of Interval, we see that the R-side exhibited small changes during Interval 1, but the D-side shows an increase compared to the Baseline when we introduce CRA on the D-position. For the D-side, this remains true for all Intervals albeit not as pronounced as during Interval 1 where the traffic level is low. When CRA is available on the R- and D-positions, we do not see a substantial difference during the first Interval for the R-side controllers, but during the second Interval when we introduce CRA on the R-position as well, we see an increase for R-side controllers.

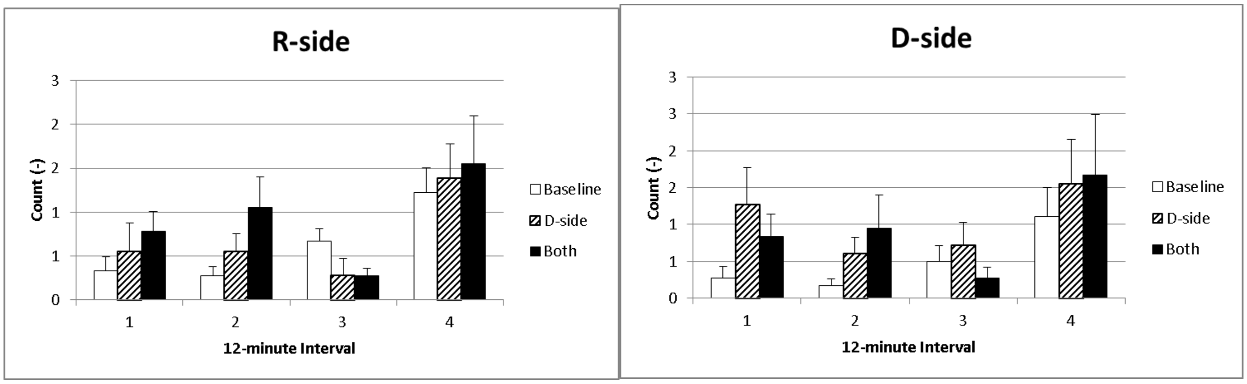


Figure 32. The number of times controllers access Speed Fly-Out menus per 12-minute interval as a function of Position, CRA condition and Interval.

3.9.3 Access to Heading Fly-Out Menus

We conducted a 2 x 3 x 4 (Position x CRA x Interval) MANOVA on the number of times controllers accessed Heading Fly-Out menus within each 12-minute interval. The multivariate tests showed an effect of Interval that just did not reach significance, Λ = .604, *F*(3, 15) = 3.281, *p* = .050, and an interaction between the effects of Position and CRA, Λ = .629, *F*(2, 16) = 4.726, *p* = .024. After Huynh-Feldt correction for sphericity, the main effect of Interval was significant, *F*(2.3, 39.4) = 3.978, *p* = .022, but the statistical significance for the interaction between Position and CRA did not reach significance, *F*(1.6, 2 6.4) = 2.342, *p* = .126. We used Helmert contrasts to determine which levels of the dependent variables differed. The effects in the three-way interaction are difficult to tease apart.

3.9.4 Access to Safety Portal Fly-Out Menus

We conducted a 2 x 3 x 4 (Position x CRA x Interval) MANOVA on the number of times controllers accessed Safety Portal Fly-Out menus within each 12-min interval (see Figure 33). The multivariate tests showed a main effects of Position, CRA, and Interval, Λ = .766, *F*(1, 17) = 5.195, *p* = .036; Λ = .554, *F*(2, 16) = 6.447, *p* = .009; and Λ = .462, *F*(3, 15) = 5.849, *p* = .008, respectively. We also found a significant interaction between Position and CRA, Λ = .311, *F*(2, 16) = 17.757, *p* < .001. After Huynh-Feldt correction for sphericity, the main effects and the interaction between Position and CRA were still statistically different although the statistics themselves and the degrees of freedom had changed. We conducted a post-hoc Tukey's HSD test to determine which pairs of conditions differed and found that the D-side controllers accessed the Safety Portal significantly more often when CRA was available than under the Baseline. For D-side controllers there was no difference between having CRA available on the D- or R-position. There was no difference across any of the conditions for the R-side. When CRA was available, however, the D-side always accessed the Safety Portal more often than the R-side.

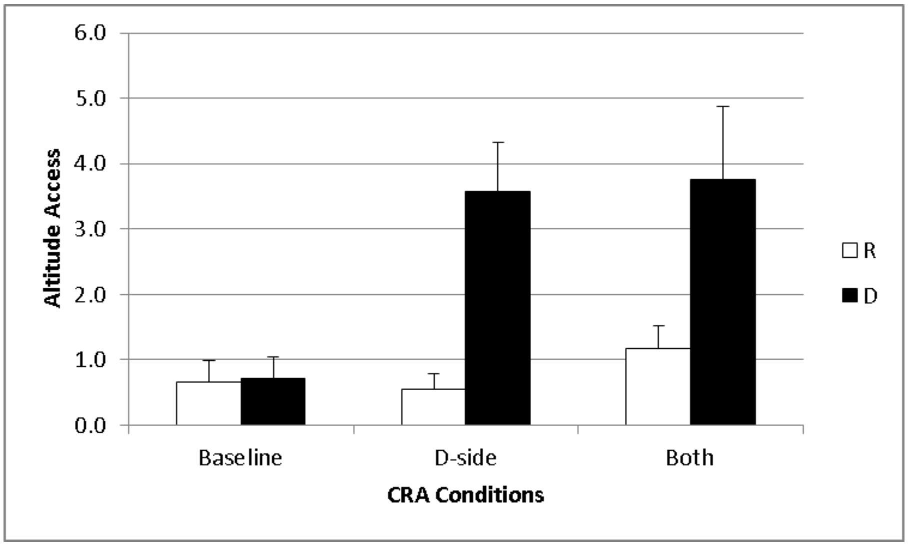


Figure 33. The number of times controllers accessed Safety Portal Fly-Out menus per 12-minute interval as a function of CRA condition and Position.

3.10 Tactical Conflict and Conflict Probe Alerts

In this study, controllers on the R-side and the D-side could experience both tactical conflict and conflict probe alerts. The tactical conflict alerts mimicked the behavior of the Conflict Alert as present in with DSR, HOST, and ERAM. In our data analysis, if conflict alerts occurred 20 seconds or longer between the same aircraft pair, then each conflict alert was included in the analyses as a separate event. If conflict alerts occurred less than 20 seconds between the same aircraft pair, then the conflict alerts were combined into a single event and duration was calculated by subtracting the last conflict alert end time from the first conflict alert start time. We did not include conflict alerts that occurred before 2 min and after 50 min. The conflict probe alerts consisted of an emulation modeled after the medium term conflict probe used by URET, but we reduced the look ahead for conflict detection from 20 min to 10 min and for trial planning from 20 min to 12 min. We did not display muted alerts to the R- or D-sides—that is, we suppressed potential conflicts on an uncleared portion of the trajectories. Finally, we combined red alerts (5-nmi between the center line of the trajectories) and yellow alerts (5-nmi between the conformance boundaries of the trajectories) and only displayed potential conflicts with a lateral separation of less than 10 nmi.

3.10.1 Tactical Conflict Alerts

Because traffic volume increased over time and the chance of having potential conflicts that set off conflict alerts increase with traffic volume, we included Interval and Number of Aircraft under Responsibility (NAR). To accommodate differences in the number of aircraft per CRA condition, interval, and team, we performed an ANCOVA with a varying covariate. The results of the ANCOVA include an estimate of the means and standard error after accounting for NAR. Figure 34 presents the results the estimated means and standard errors. Although the ANCOVA result for the effect of CRA was not significant, $F(2, 186.6) = 2.25, p = .108$, the estimated mean shows some interesting trends. First, the number of conflict alerts is quite low; less than one per 12-min interval. Second, the number of conflict alerts seems to increase with interval even after accounting for the number of aircraft under responsibility. Finally, the number of conflict alerts seems to decrease with the level of CRA available to controllers.

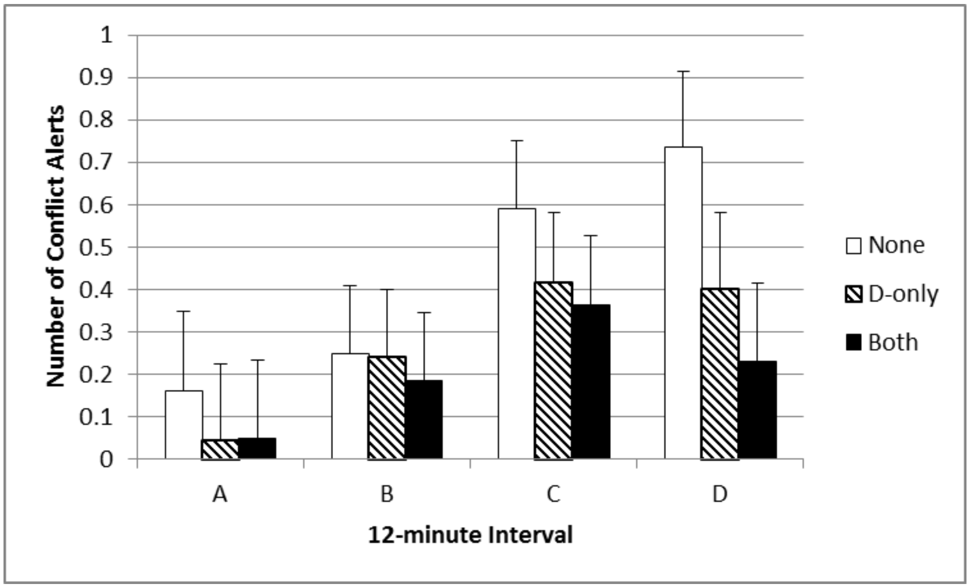


Figure 34. The number of Conflict Alerts as a function of CRA condition and 12-minute interval corrected for the effect of number of aircraft under control.

3.10.2 Conflict Probe Alerts

We identified the subset of aircraft pairs that would be in conflict based on the scenario traffic if controllers would not take action or only took action to meet restrictions or requested altitudes. To do so, we ran the scenario in two ways: one without any controller intervention and one with a SME climbing and descending aircraft to meet altitude restrictions and requested altitudes but otherwise not separating aircraft. When combining Sectors 20 and 22, there were 26 aircraft pairs that lost separation in one or both of these runs. Twenty-two conflicts occurred in both runs: 21 involved two level aircraft, and one involved a level aircraft and an aircraft that entered the sector climbing to a cleared altitude. The remaining four conflicts occurred in the climb/descend run only. Each of these involved a level aircraft and an aircraft that entered the sector at an interim altitude, requesting clearance to a higher altitude.

We analyzed the data from the simulation test runs for these specific aircraft pairs. We excluded data from one run in the non-CRA condition because it was incomplete. There were four total losses of separation; two in the non-CRA condition and two in the D-side CRA condition. We analyzed three categories of Conflict Probe alerts:

- **Red Alerts** – Aircraft predicted to be within 10 nmi on a cleared portion of their flight plans as defined by CRA.
- **5 Mile Alerts** – Red alerts with predicted separation less than 5 nmi as currently use in the NAS.

- **6-6 Alerts** – Red alerts with predicted separation less than 6 nmi in less than 6 min, one of the proposed changes to conflict probe under the Separation Management Program.

For each alert category, we calculated the proportion of expected conflicts alerted and the duration of the alert interval. We defined alert interval as the difference between the end of the alert and either the start of the alert or the first handoff request for an aircraft in the pair, whichever came later.

We analyzed the data with a Bayesian generalized linear model. We have included Figure 35 to explain how the Bayesian generalized linear model performs pairwise comparisons for the 6-6 rule. In the Bayesian approach, we calculate the distributions of the differences in a dependent variable (number of conflict probe alerts in this case) for each pair of conditions through Markov Chain Monte Carlo simulations. If the high density interval of the distribution does not include zero, there is strong evidence that the two conditions differ in terms of the dependent variable. We can see in Figure 35 that having CRA on the D-side, a lower percentage of the scripted conflict s resulted in conflict probe alerts than the Baseline as did having CRA on the R- and D-sides. There was, however, no difference in the percentage of conflict probe alerts between having CRA on the D-side only and having CRA on both the D- and R-sides.

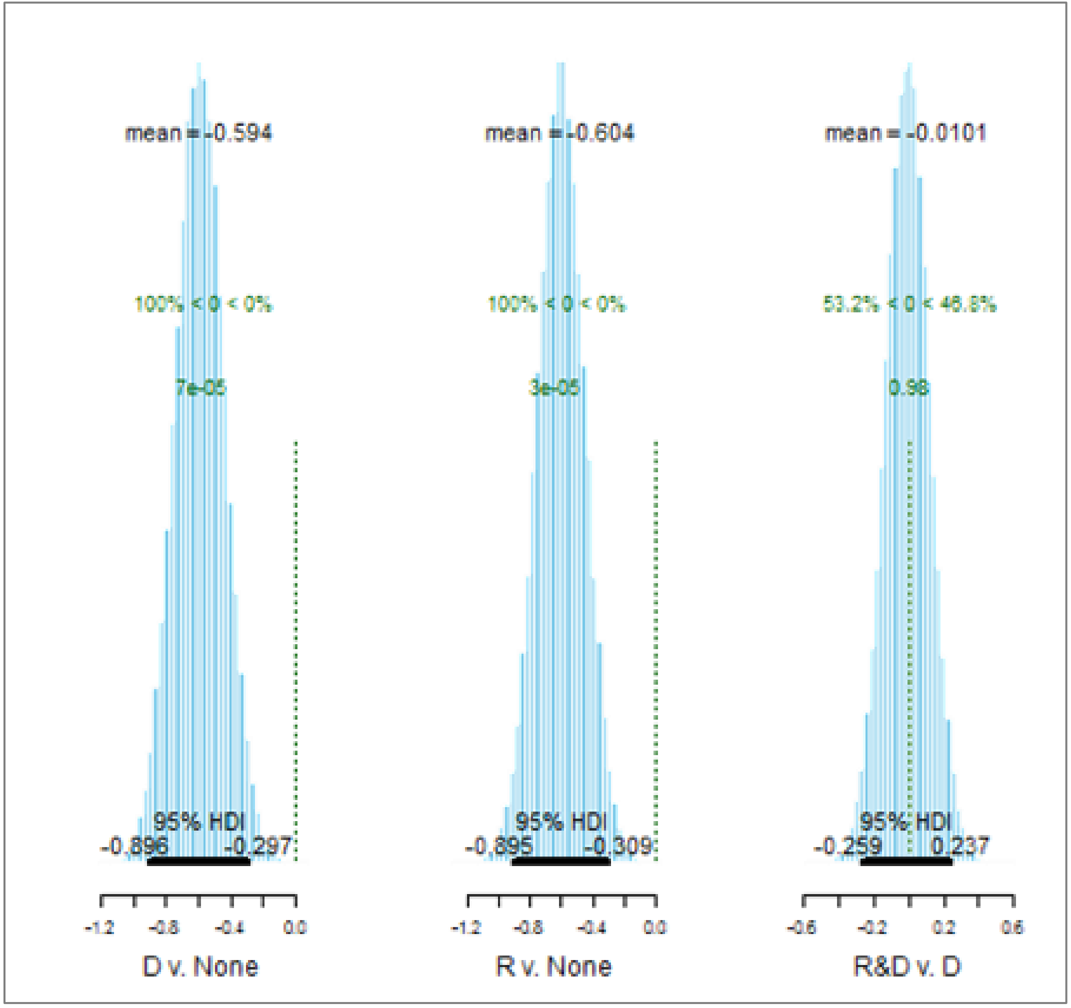


Figure 35. Distribution of mean differences used in the Bayesian generalized linear model comparison of the number of conflict probe alerts using the 6-6 alerting rule.

Table 9 contains the mean percentage of alerts and alert intervals for each condition, along with the *p* values for each pairwise comparison between conditions. CRA on the R- and D-sides reduced the number and duration of all three kinds of alerts significantly, as compared to the non-CRA condition. CRA on the D-side also significantly reduced the number of red and 6-6 alerts and reduced the interval of all three kinds of alerts. The decrease in 5 mile alerts was marginally significant. CRA on the R- and D-sides tended to result in fewer alerts and longer alert intervals than CRA on the D-side, but none of these comparisons were statistically reliable.

Table 9. Percentage of Alerts and Mean Alert Interval for the Expected Conflicts in each CRA Condition

	CRA Condition			Comparison *p*-values		
	None	**D-side**	**R- & D-sides**	**D v. None**	**R&D v. None**	**R&D v. D**
Red Alert %	81%	72%	70%	.037*	.008*	.50
5 Mile Alert %	72%	63%	56%	0.074	.001*	.11
6-6 Alert %	62%	24%	23%	< 10^$^{-5*}$	< 10^$^{-5*}$	.70
Red Alert Interval (min)	3.8	2.0	2.4	< 10^$^{-5*}$	.0003*	.08
5 Mile Alert Interval (min)	2.8	1.3	1.5	< 10^$^{-5*}$	< 10^$^{-5*}$	.48
6-6 Alert Interval (min)	1.2	0.5	0.5	< 10^$^{-5*}$	< 10^$^{-5*}$	.70

Note. * = Indicates statistical significance of $p < 0.05$.

3.10.2.1 Trajectory change timing

For each aircraft pair, we identified all of the altitude, heading, route, and speed changes made in response to clearances. We calculated two time measures:

- **Latency** – The difference in time between the first trajectory-changing command issued and the first handoff request.
- **Lead Time** – The difference in time between the expected loss of separation and the most recent trajectory-changing command issued while the aircraft remained in conflict status.

We restricted our analysis of these measures to the 22 conflicts that occurred in both of the non-separation runs. These are the conflicts that absolutely required controller intervention. Shorter latencies and longer lead times straightforwardly reflect better performance for this subset. The remaining four aircraft pairs do not enter the sector in a potential conflict situation, but are at risk of conflict if climbed or descended at the wrong moment. For these, quicker action does not necessarily reflect better performance.

We analyzed the data with a Bayesian generalized linear model. Table 10 contains the mean latency and lead times for each condition, along with the *p* values for each pairwise comparison between conditions. CRA on the R- and D-sides led to significantly improved latencies and lead times, as compared to the non-CRA condition. CRA on the R- and D-sides tended to result in longer latencies and shorter lead times than CRA on the D-side, but neither of these comparisons were statistically reliable.

Table 10. Trajectory Change Timing for the Expected Conflicts in each CRA Condition

	CRA Condition			Comparison *p*-values		
	None	**D-side**	**R- & D-sides**	**D v. None**	**R&D v. None**	**R&D v. D**
Latency (min)	10.3	7.7	8.3	< 10^$^{-5*}$	< 10^$^{-5*}$	0.26
Lead Time (min)	6.4	9.0	8.5	< 10^$^{-5*}$	< 10^$^{-5*}$	0.38

Note. * = Indicates statistical significance of $p < 0.05$.

3.10.2.2 Coordination between sectors

Controllers between two sectors can coordinate by using CRA. This coordination between two sectors is different from the coordination between R- and D-sides in the same sector, which Peterson, Bailey, and Willems (2001) studied (see Table 11). Their categories and percentages to the total frequency were Approval (1%), Handoff (3%), Point Out (6%), Traffic (39%), Altitude (8%), Route (15%), Speed (3%), Weather (6%), Frequency (5%), Flow Messages (5%), Flight Strips (5%), and Equipment (4%).

Table 11. Taxonomy of R- and D-side Controllers

Topic	**Definitions**	**Examples**
Approval	Communications about inter-sector control/approval requests.	"Get me control for descent on that aircraft." "APREQ N1234 climbing to FL330."
Handoff	Communications relating to the transfer of radar identification of a particular aircraft.	"Handoff N1234." "Did you handoff N1234?
Point Out	Communications relating to the transfer of radar identification of a particular aircraft when radio communications will be retained.	"Point out N1234 to 22."
Traffic	Communications about a traffic situation involving a specific aircraft. Includes conflict, spacing, other protected air space or terrain and the resolution of that situation.	"Are you watching that aircraft?"
Altitude	Communications about altitude not in relation to traffic.	"N1234 is requesting flight level 220."
Route	Communications regarding headings and/or amendments to route, not in relation to traffic situations.	"N1234 is on a 330 heading." "Next sector, 27, wants N1234 over WEVER."

Note. APREQ = Approval for Request. Table adapted from Peterson, Bailey, and Willems (2001).

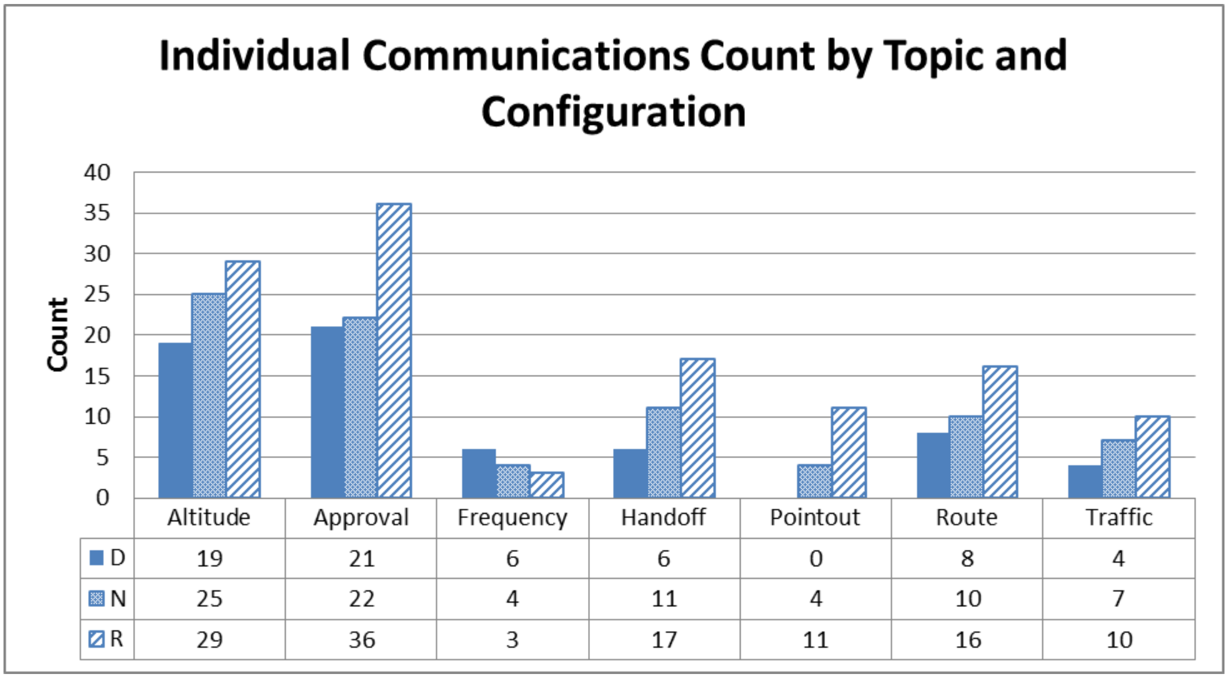
The verbal communication between the two sectors in our experiment would be about the traffic flow between one sector to the other. Thus, their communication and coordination taxonomy must be different from that of Peterson et al. (2001). However, in general, the controllers of the two sectors are still about controlling air traffic safely. Thus, we used Peterson et al.'s taxonomy as a baseline and modified it as necessary. Because we did not have a weather condition, and flight strips, we omitted those categories. We had equipment problems, but there was no need for them to coordinate because experimenters and engineers were standing by and solved the problems.

In our data, there was no coordination conversation about speed between the two sector controllers. Our categories are listed in Table 12. The results showed that there was no clear difference in coordination patterns between the CRA conditions in either communication categories and formats (see Figure 36 and Figure 37). Even with the CRA that had coordination features, controllers still coordinated using verbal communication as often. The most frequency coordination categories were about Altitude and Approval in all experimental conditions.

Table 12. Taxonomy of Verbal Coordination between Controllers of Two Adjoining Sectors

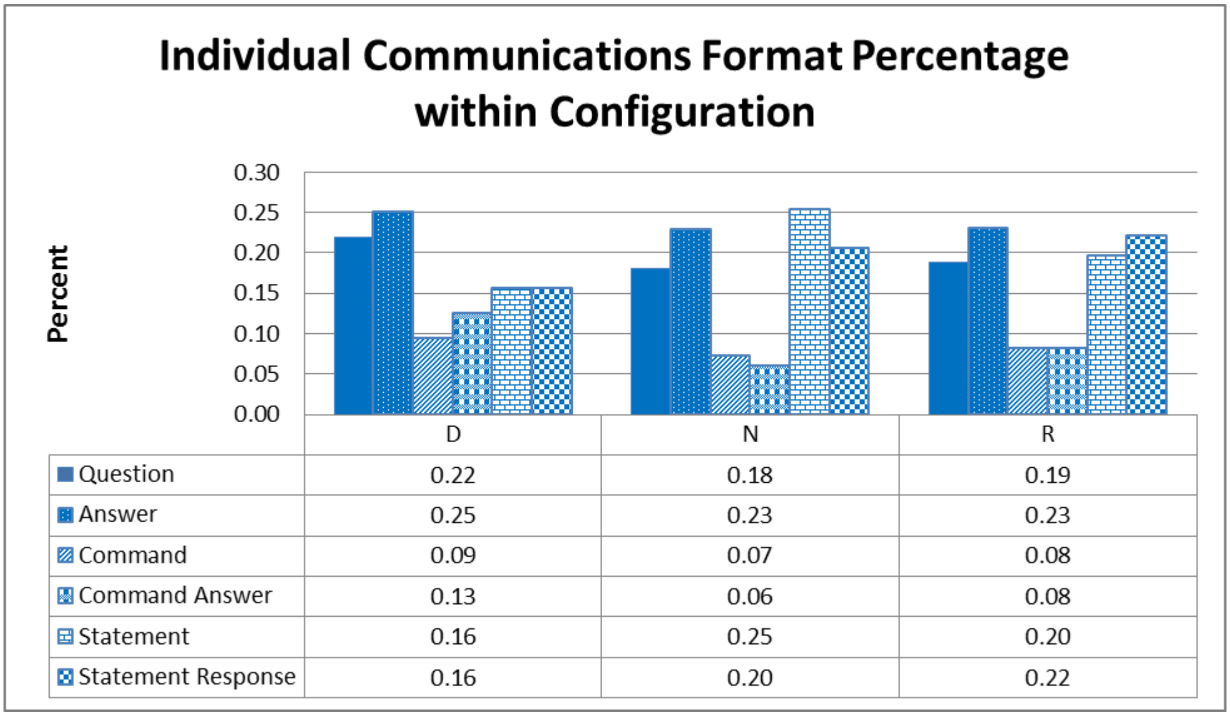
Topic	Definition	Example
Approval	Communications about inter-sector control/approval requests.	I'm gonna turn the Frontier eighteen sixty-one. (CRA D condition)
Handoff	Communications relating to the transfer of radar identification of a particular aircraft.	Handoff on two five four. (CRA R condition)
Pointout	Communications relating to the transfer of radar identification of a particular aircraft when radio communications will be retained.	You want to take three seven seven? (CRA R condition)
Traffic	Communications about a traffic situation involving a specific aircraft. Includes conflict, spacing, other protected air space or terrain and the resolution of that situation.	You see that coordination request with the Delta? With the green… (CRA D condition)
Altitude	Communications about altitude not in relation to traffic.	We're going to descend that Southwest when ya'll switch him… (CRA R condition)
Route	Communications regarding headings and/or amendments to route, not in relation to traffic situations.	You got that route... I gave that United? (CRA R condition)
Frequency	Communications about an aircraft's radio communications transfer or frequency assignment.	You still got United ten fourteen? (CRA D condition)

Note: CRA conditions are identified in the parentheses.



	Altitude	Approval	Frequency	Handoff	Pointout	Route	Traffic
■ D	19	21	6	6	0	8	4
▩ N	25	22	4	11	4	10	7
▨ R	29	36	3	17	11	16	10

Figure 36. The frequency of communication by categories and display configuration.



	D	N	R
■ Question	0.22	0.18	0.19
▩ Answer	0.25	0.23	0.23
▨ Command	0.09	0.07	0.08
▦ Command Answer	0.13	0.06	0.08
⊟ Statement	0.16	0.25	0.20
▩ Statement Response	0.16	0.20	0.22

Figure 37. The percentages of verbal communication types.

3.11 Functional Near-Infrared Data

3.11.1 Preprocessing of fNIR Data

When we first examined the data, we plotted all raw 16 channels separately before we applied any filters or calculations. We then went through each channel for each participant and created an Excel map of usable and unusable channels. No participant had good data from all 16 channels, and there was only one participant who had more than 12 channels valid for all 3 conditions. From this visual examination, we selected 15 runs out of 51 that had more than 8 valid channels. We filtered them further.

We used a sliding motion artifact removal (SMAR) algorithm to remove any detectable motion artifacts. We used a Butterworth filter after SMAR, and for that filtering we trimmed 10 seconds before and after the epochs. Then we filtered any channels that had less than 75% valid data points. SMAR and trimming removed many data points from the first minute of each run; so instead of using data from the first minute, we used data from the first 2 minutes as a baseline. There were a total of 46 runs out of 48 that had at least one "good" channel. Final oxygenation data used for analysis was an oxygenation level relative to the baseline averaged across all channels.

For data analysis, raw fNIR data was filtered through SMAR, linear phase filter with cut off frequency of 0.14 Hz. This attenuated the high-frequency noise, respiration, and cardiac cycle effects (Ayaz et al., 2010; Ayaz et al., 2012). We excluded saturated channels. We calculated blood oxygenation and volume changes relative to the first two-minute baseline for each of 16 channels using the modified Beer-Lambert Law.

In our analysis, we used the Jeffrey-Zellner-Siow (JZS) Bayes factor *t*-tests, via a Web-based program (at pcl.missouri.edu) developed by Rouder, Speckman, Sun, and Morey (2009). Compared to the traditional statistical method of Null Hypothesis Significance Testing (NHST), Bayesian model comparison allows researchers to state evidence for the null hypothesis. For example, in NHST, researchers can reject the null hypothesis, or fail to reject the null hypothesis (e.g., failed to find a difference between the means) but were not allowed to state the evidence for the null hypothesis (e.g., found evidence that means were not different). However, the Bayesian model compares probability of the null over alternative where Bayes Factor (*BF*) above 1 indicates evidence for the null while BF below 1 indicates evidence for the alternative. It gives the anecdotal, substantial, or strong evidence that two conditions are not different (see Table 13).

3.11.2 CRA Condition Comparison

Paired *t*-test and Bayes factor analysis showed no support for differences in mean oxygenation change between both R- and D-sides and D-side CRA conditions (see Table 13). Bayes Factor shows substantial evidence for null hypothesis (Bayes Factor between 3.00 and 10.00), providing evidence that there is no difference between CRA conditions within each position. In Table 14, No is the Baseline, D is the CRA D-side only, and Both is the CRA R- and D-side condition. For instance, 'No-D' is for the comparison between the Baseline and the D condition. Paired *t*-test Bayes factor analysis showed anecdotal evidence for no difference between mean oxygenation change for positions R-side and D-side (BF between 1 and 3) taking CRA conditions into account.

Table 13. Classification Scheme for the Bayes Factor

Bayes Factor	Interpretation
> 100	Extreme evidence for the null (no effect)
30 – 100	Very strong evidence for the null
10 – 30	Strong evidence for the null
3 – 10	Substantial evidence for the null
1 – 3	Anecdotal evidence for the null
1	No evidence
1/3 – 1	Anecdotal evidence for the alternative (effect)
1/10 – 1/3	Substantial evidence for the alternative
1/30 – 1/10	Strong evidence for the alternative
1/100 – 1/3	Very strong evidence for the alternative
< 1/100	Extreme evidence for the alternative

Table 14. Summary *t*-test and Bayesian Statistics for Condition Comparison

	D-side			R-side		
CRA condition	**No-D**	**No-Both**	**D-Both**	***No-D***	***No-Both***	***D-Both***
p-value	0.68	0.72	0.84	0.67	0.82	0.47
Sample size	7	7	7	7	6	7
t-value	0.44	0.38	0.21	0.45	0.24	0.77
Bayes Factor	3.38	3.46	3.63	3.37	3.39	2.84

3.11.3 Air Traffic Volume

Previous simulation studies have found correlations between traffic volume and oxygenation levels measured by fNIR (Ayaz et al., 2011). Similarly, we tested the relationship between oxygenation levels and traffic volume in our experiment. We defined low, medium, and high traffic by 7 to 13, 14 to 20, and 21 to 27 aircraft, respectively. We used the maximum number of aircraft by each participant in each of the traffic volume categories and the oxygenation change epoch as 10 seconds before and 10 seconds after the time aircraft count was reported. There was a significant effect for air traffic levels between low and medium traffic, $t(4) = 5.61$, $p < .005$, $BF = 0.07$; medium and high traffic, $t(4) = 5.54$, $p = .005$, $BF = 0.07$; and low and high traffic, $t(4)=8.76$, $p < .001$, $BF = .02$. When we differentiated the CRA conditions, the results showed no significant results (see Table 15).

Table 15. Summary *t*-test and Bayesian Statistics for CRA Condition Comparisons

	Low-med	Med-high	Low-high
		D-side No CRA	
p-value	0.005	0.068	0.011
Sample size	7	7	7
t-value	4.26	2.23	3.66
Bayes Factor	0.082**	0.68	0.15*
		D-side CRA on D	
p-value	0.063	0.025	0.030
Sample size	7	7	7
t-value	2.28	2.97	2.83
Bayes Factor	0.64	0.3	0.35
		D-side CRA on B	
p-value	0.002	0.051	0.003
Sample size	7	7	7
t-value	4.99	2.43	4.79
Bayes Factor	0.043**	0.54	0.051**
		R-side No CRA	
p-value	0.003	0.020	0.004
Sample size	7	7	7
t-value	4.75	3.14	4.48
Bayes Factor	0.053**	0.25*	0.067**
		R-side CRA on D	
p-value	0.010	0.134	0.015
Sample size	8	8	8
t-value	3.47	1.70	3.18
Bayes Factor	0.15*	1.25	0.21*
		R-side CRA on B	
p-value	0.001	0.004	0.000
Sample size	7	7	7
t-value	6.27	4.51	8.71
Bayes Factor	0.016***	0.065**	0.0036****

Note. *substantial evidence of difference, **strong evidence of difference, ***very strong evidence of difference, ****decisive evidence of difference.

Another interesting trend was that mean oxygenation changes were relatively flat when aircraft counts were below 10 but increased steadily after 10 aircraft. This trend was observed as a function of aircraft count under responsibility, aircraft count on the communication frequency, and the number of aircraft physically inside the sector. This provides some evidence that is no difference in effort between controlling 10 aircraft and controlling 7 aircraft.

We expected to find differences among CRA conditions when variance by aircraft traffic was accounted. However, we did not find differences among CRA conditions, and BF showed substantial evidence for null hypothesis (*BF* > 1).

3.11.4 Event Related Analysis

Unlike EEG, hemodynamic measures, such as fNIR and functional Magnetic Resonance Imaging (fMRI), have lower temporal resolution. Previous studies reported hemodynamic response delay periods to be about 10 seconds to 12 seconds (Bunce, Izzetoglu, Izzetoglu, Onaral, & Pourrezaei, 2006). Some researchers used fast NIR to examine the event-related optical signal (EROS). Unfortunately, EROS tended to have a low signal-to-noise ratio according to Bunce et al. (2006).

We were interested in how controllers' workload changed before and after they executed a clearance (i.e., command; see Figure 38 and Table 16 through Table 20). In our experiment, controllers could execute a clearance through the keyboard or through interaction with the Datablock (DB) click, using the mouse. We hypothesized that the oxygenation level would be lower 10 seconds after clearance issuance compared to 10 seconds before in both keyboard and DB modalities because controllers would evaluate the situation first, make decisions, and execute clearances. As soon as a controller issued a clearance, his/her oxygenation level would be lower because a decision was already made. However, we did not see any difference between before and after the clearance issuance.

Even though we did not see any difference in overall oxygenation levels across different CRA conditions, we assumed we might see a difference between them per command. In "Altitude-Keyboard" (see Figure 38), there was a significant difference among the oxygenation levels of the R-side controllers—that is, the baseline (No-CRA) condition level was higher than the CRA condition level. The R-side condition level was higher than the D-side condition level.

We also expected different cognitive processes to take place in keyboard command and DB command because the CRA pop-up was elicited only through DB click. However, as mentioned previously, we did not find a difference between oxygenation level for 10 seconds before and after command for both keyboard and DB command issuance.

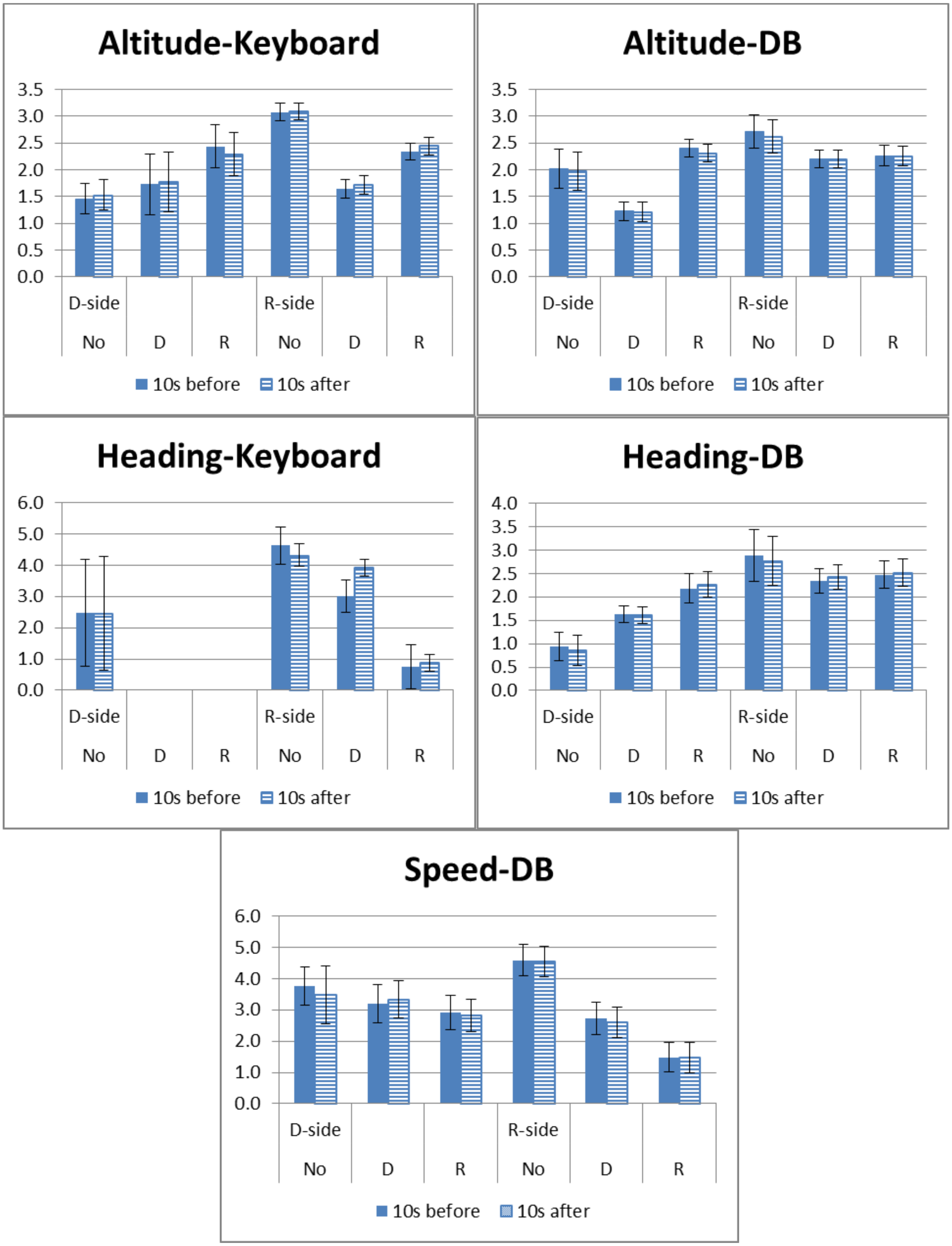


Figure 38. Oxygenation change for keyboard (left) and datablock (right) clearance. The error bars represent standard errors.

Table 16. Summary *t*-test and Bayesian Statistics for Keyboard-10s before Altitude Command

	D-side			**R-side**		
CRA Condition	***No-D***	***No-Both***	***D-Both***	***No-D***	***No-Both***	***D-Both***
p-value	0.675	0.054	0.319	0.000	0.001	0.003
Sample size	36	36	28	184	184	143
Sample size 2	26	28	26	181	143	181
t-value	0.42	2.02	1.02	6.30	3.24	3.07
Bayes factor	4.76	0.85	3.07	0.00	0.072	0.12
Difference				****	**	*
No Difference	*		*			

Note. *substantial evidence of difference, **strong evidence of difference, ****decisive evidence of difference.

Table 17. Summary *t*-test and Bayesian Statistics for Keyboard-10s before Heading Command

	D-side			**R-side**		
CRA Condition	***No-D***	***No-Both***	***D-Both***	***No-D***	***No-Both***	***D-Both***
p-value	-	-	-	0.085	0.036	0.116
Sample size	2	2	0	4	4	2
Sample size 2	0	0	0	4	2	4
t-value	-	-	-	2.35	17.81	5.41
Bayes factor			0.47	0.0002	0.11	
Difference	****	****	****		****	*
No Difference						

Note. *substantial evidence of difference, ****decisive evidence of difference.

Table 18. Summary *t*-test and Bayesian Statistics for Datablock-10s before Altitude Command

	D-side			**R-side**		
CRA Condition	***No-D***	***No-Both***	***D-Both***	***No-D***	***No-Both***	***D-Both***
p-value	0.055	0.329	0.000	0.154	0.227	0.786
Sample size	82	131	131	51	155	155
Sample size 2	107	82	107	79	51	79
t-value	1.95	0.98	5.12	1.45	1.22	0.27
Bayes factor	1.41	5.70	0.00	2.66	3.91	8.91
Difference			****			
No Difference		*			*	*

Note. *substantial evidence of difference, ****decisive evidence of difference.

Table 19. Summary *t*-test and Bayesian Statistics for Datablock-10s before Heading Command

	D-side			**R-side**		
CRA Condition	***No-D***	***No-Both***	***D-Both***	***No-D***	***No-Both***	***D-Both***
p-value	0.059	0.007	0.141	0.388	0.520	0.732
Sample size	24	49	49	18	57	57
Sample size 2	57	24	57	43	18	43
t-value	1.99	2.98	1.50	0.89	0.66	0.35
Bayes factor	0.91	0.11	2.33	3.35	4.06	6.08
Difference		*				
No Difference				*	*	*

Note. *substantial evidence of difference.

Table 20. Summary *t*-test and Bayesian Statistics for Datablock-10s before Speed Command

	D-side			R-side		
CRA Condition	***No-D***	***No-Both***	***D-Both***	***No-D***	***No-Both***	***D-Both***
p-value	0.536	0.332	0.724	0.015	0.000	0.083
Sample size	4	20	20	16	16	16
Sample size 2	27	4	27	20	16	20
t-value	0.70	1.15	0.36	2.75	5.22	1.85
Bayes factor	2.35	1.71	4.31	0.20	0.00	0.98
Difference				*	****	
No Difference			*			

Note. *substantial evidence of difference, ****decisive evidence of difference.

3.12 Eye Movements

Analysis of eye movements will either be provided as an update or in a separate document.

3.13 Electroencephalogram

Analysis of eye movements will either be provided as an update or in a separate document.

3.14 Losses of Separation

We had 12 losses of separation—among them, eight losses were caused by controllers, two losses were caused by simulation pilots, and one loss could have been caused by a simulation pilot or TGF. One loss was caused by an unreasonable climb rate that our simulation generated. Among the losses caused by controllers, four losses were in the baseline condition, three losses were in the CRA-D condition, and one loss was in the CRA-R condition (see Appendix I for the detailed descriptions). The following is one example of the loss:

- **At 28:40 min**, a controller issued a clearance via Data Comm to SQC364 to climb from FL370 to FL390.
- **At 30:57 min**, the controller issued a clearance via Data Comm to JBU2358 to climb from FL380 to FL400.
- **At 32:51 min**, SQC364 leveled at FL390. JBU2358 Mode-C indicated at FL385. The aircraft were separated by approximately 38 nmi.
- **At 33:59 min**, Conflict Alert activated when SQC364 was at FL390 and JBU2358 was at FL388 and climbing. They were separated by 23 nmi laterally.
- **At 35:15 min**, SQC364 was leveled at FL390 and JBU2358 was climbing thru FL391. At the moment, the loss of separation occurred as they were separated laterally less than 5nm laterally.

3.15 Post-Scenario Questionnaire Data

In this section, we present tables that show the participants' opinions about the CRA in the post-scenario questionnaire that we administered after each experimental run. For each of the questions, the average ratings are presented with a color code. The ratings in blue are positive, and it means the participants favored the feature or function mentioned in the question. The negative values are in red, and it means that the participants did not favor the feature or function.

There were two types of questionnaires. One was to compare the CRA conditions and the other was specific to the CRA or other features and functions in the simulation.

If the question was about comparing the CRA conditions, we tested the ratings using a nonparametric Friedman test that is the equivalent to a repeated measure ANOVA test. We used it because our questionnaire sample size was small and had missing data sometimes. If we tested ratings on a question that was about the CRA in general, or about a specific aspect of the simulation, we used one sample Wilcoxon Signed Ranks Test if the ratings were deviated significantly from 0 (zero). All significant averages were noted with * and *p* values in the tables. The averages without an asterisk (*) mean that they were not different statistically.

We did not have any weather conditions in the scenarios and did not include it in the following Table 21. The results showed that there were no significant differences among CRA conditions, and we do not see any distinct patterns. The rating for "Managing air traffic sequences" in the D-side CRA condition is negative. However, the rating value, -0.22, is a very small negative value that is close to 0 (zero). They also experienced difficulty in managing sector/position resources in all conditions.

Table 21. Post-Scenario Questionnaire Results

A. Air Traffic Control Tasks			
Question	Hindered Greatly **-5 to 5** Helped Greatly		
	Baseline	**D Side**	**R & D Side**
(1) Situation monitoring	**0.18**	**0.00**	**0.25**
(2) Detecting aircraft conflicts	**0.79**	**0.86**	**1.17**
(3) Resolving aircraft conflicts	**0.53**	**1.08**	**1.11**
(4) Managing air traffic sequences	**0.21**	**-0.22**	**0.19**
(5) Routing or planning flights	**0.26**	**0.14**	**0.03**
(6) Managing Sector/Position Resources	**-.18**	**-.14**	**-.22**

In Table 22, the controllers' ratings for *"B. Automation: a. How was the automation available in the scenario you just worked compared to DSR, HOST, and URET?"* showed different patterns to different questions. For *"Complexity of Fly-Out Menus," "Ease of Use of Fly-Out Menus,"* and *"Usefulness of Fly-Out Menu,"* controllers rated the CRA more negatively compared to DSR, HOST, and UERT. Controllers did not rate the coordination features in the CRA negatively.

Table 22. Ratings on the Comparative Effect of Automation compared to DSR, HOST, and URET

B. Automation: a. How was the automation available in the scenario you just worked compared to DSR, HOST, and URET?			
Question	Much Worse **-5 to 5** Much Better		
	Baseline	**D Side**	**R & D Side**
(1) Fly-out menus: Complexity	**-0.17**	**-0.58**	**-1.33**
(2) Fly-out menus: Ease of Use	**0.20**	**-0.39**	**-1.06**
(3) Fly-out menus: Usefulness	**0.80**	**0.67**	**-0.22**
(4) Coordination: Complexity	**0.24**	**0.82**	**0.17**
(5) Coordination: Ease of use	**0.48**	**0.94**	**0.61**
(6) Coordination: Usefulness	**0.58**	**1.38**	**0.89**

As Table 23 shows, it was evident that controllers liked Data Comm by the high values of their ratings of (5) Effect of Data Comm on workload with ratings of 2.26 (Baseline), 2.61 (D-side), and 2.28 (R-side). In addition, in all CRA conditions, controllers rated the electronic positively: (6) Effect of electronic coordination on workload, with ratings of 1.15 (Baseline), 2.08 (D-side), and 1.89 (R- and D-sides).

Table 23. Impacts of Automation on various Tasks

B. Automation: b. What was the impacts of the automation available in the scenario you just worked for the following items?			
Question	Very Negative **-5 to 5** Very Positive		
	Baseline	**D Side**	**R & D Side**
(1) Finding a solution to a conflict	**0.69**	**1.31**	**1.22**
(2) Situation awareness for potential conflicts	**0.31**	**0.92**	**0.72**
(3) Collaboration in solving conflicts within your team	**0.43**	**1.00**	**0.61**
(4) Collaboration in solving conflicts between sectors	**0.54**	**1.08**	**0.92**
(5) Effect of Data Comm on workload	**2.26**	**2.61**	**2.28**
(6) Effect of electronic coordination on workload	**1.15**	**2.08**	**1.89**
(7) Effect of CRA on workload	**-0.11**	**0.52**	**-0.03**
(8) Effect of Data Comm on conflict resolution	**1.06**	**1.56**	**1.25**
(9) Effect of electronic coordination on conflict resolution	**0.79**	**1.50**	**1.39**

For the question about Conflict Detection, the controllers approved the accuracy and timeliness of the CRA but thought it distracted them (see Table 24). They did not approve the CRA presentation format and interaction aspects when it was presented on the R-side.

Table 24. Comparison of the Automation to DSR, HOST, and URET

B. Automation: c. How does the Conflict Detection Available in the scenario you just worked compare to DSR, HOST, and URET?			
Question	Much Worse **-5 to 5** Much Better		
	Baseline	**D Side**	**R & D Side**
(1) Accuracy	**0.17**	**0.61**	**0.50**
(2) Timeliness	**-0.14**	**0.33**	**0.11**
(3) Distraction	**-0.29**	**-0.53**	**-0.61**
(4) Presentation format	**0.57**	**0.47**	**-0.06**
(5) Interaction	**0.40**	**0.14**	**-0.14**

As controllers responded to **B. Automation: d. Please rate the following about false alerts, missed alerts, nuisance alerts, and CRA solutions, if it was available,** they thought there were too many false and nuisance Conflict Probe Alerts (see Table 25). On the positive side, controllers felt that the number of missed Conflict Probe Alerts was acceptable. For the CRA conditions, the number of incorrect CRA solutions was not as many as in the baseline condition. They stated that there were many unacceptable CRA solutions.

Table 25. Automation in False Alerts, Missed Alerts, Nuisance Alerts, and CRA Solutions

B. Automation: d. Please rate the following about false alerts, missed alerts, nuisance alerts, and CRA solutions, if it was available.			
Question	Unacceptable **-5 to 5** Acceptable		
	Baseline	**D Side**	**R & D Side**
(1) The number of false Conflict Probe Alerts	**-0.06**	**-0.91**	**-1.08**
(2) The number of missed Conflict Probe Alerts	**0.18**	**0.20**	**0.14**
(3) The number of nuisance Conflict Probe Alerts	**-0.12**	**-0.74**	**-0.92**
(4) The number of incorrect CRA solutions if it was available in your previous scenario	**-0.18**	**0.06**	**0.14**
(5) The number of unacceptable CRA solutions it was available in your previous scenario	**0.18**	**-0.90**	**-0.77**

In their response to the quality of the CRA, they rated that its accuracy, timeliness, presentation format, and interaction positively (see Table 26). Again, they rated the CRA had distracted them.

Table 26. Quality of Advisories

B. Automation: e. If you just worked a scenario that included CRA, please rate the quality of the advisories.			
Question	Unacceptable **-5 to 5** Acceptable		
	Baseline	**D Side**	**R & D Side**
(1) Accuracy	**0.54**	**1.16**	**1.06**
(2) Timeliness	**-0.15**	**0.32**	**1.03**
(3) Distraction	**-0.69**	**-0.71**	**-0.53**
(4) Presentation format	**-0.08**	**0.58**	**0.14**
(5) Interaction	**-0.38**	**0.39**	**0.20**

3.16 Exit Questionnaire Data

In the first part of the Exit Questionnaire, we asked the participants about their experience with the CRA while controlling air traffic and asked about their justifications. The following is the summary of their responses. Table 27 through Table 30 show the results of the next part of the Exit Questionnaire. First, we asked about their opinions about the effect of the CRA on air traffic control (see Table 27). The controllers favored the CRA use overall. The majority of the participants (7 out of 12, 58%) favored having the CRA on both sides. The controllers approved the beneficial aspects of the CRA, overall, as shown in the tables in blue text; we summarize their detailed responses in the list that follows Table 27.

Table 27. The Effect of the CRA

1. The Effect of CRA	
Question	
(1) What was the effect of CRA on your air traffic control?	**1.08** (Hindered Greatly **-5 to 5** Helped Greatly)
(2) Where do you think CRA should be located? If you do not think it is needed, please mark 'No Need' below.	R-side only: 1 (8%) D-side only: 3 (25%) Both R- and D-side: 7 (58%) No Need: 1 (8%) -- Total: 12 (100%)
(3) What aspects of the current version of CRA you like?	**See the summary below in the text**
(4) What aspects of the current version of CRA did you not like?	**See the summary below in the text**
(5) When you coordinated between sectors, did you use CRA?	**See the summary below in the text**

What was the effect of CRA on your air traffic control?

- **Pros**: D-side more involved, liked ranks, a nice tool but needs work, resolutions to read easily and enter into the NAS.
- **Cons**: Wrong altitude for direction of flight (WAFDOF) solutions, focus off from the radar scope visually, too much information, hard-to-use menus.

Where do you think CRA should be located?

- **Both R- and D-sides**: Teamwork, sector queue with anticipated mileage, need for combined sectors, useful for high-volume traffic, availability of the conflict information to both,
- **R-side only**: Rarely needed D-side's help, sometimes D-side made decisions contrary to what I would do.
- **D-side only**: Help make decisions, D-side being responsible for coordination and scanning boundaries, easier for the D-side to use, finding solutions and scanning the display easier for the D-side with the CRA.
- **No need**: No longer team sector, too much distraction.

What aspects of the current version of CRA did you like?

- General idea of giving resolutions.
- I like all the resolution options.
- Electronic coordination and trial plans by time.
- R-side display of conflictions.
- A notation of the time left to encounter conflicts.
- Showing conflicts on the R-side
- Ranked options.
- When you get busy, you can use this tool to quickly evaluate and act.
- Showing the resolution ahead of time.
- The alert box.
- Eliminating phone calls.
- Alert character at the FDB.
- Timely ACL alerts.
- The "next fix" in fourth line for the off-course aircraft.
- Mouse-click entry for a full resolution.
- Altitude Fly-Out menu and conflict information tied to the FDB.

What aspects of the current version of CRA did you not like?

- WAFDOF.
- Needs to be less cumbersome, too many clicks.
- Sometimes there is no option presented.
- Too many false alerts.
- Too many solutions (needs to be two solutions per a/c).
- Ten minutes is too far ahead to be useful (between two and five minutes would be better).
- It hurts your scan because you rely on it too much.

- Sometimes it is complicated to use.
- Alerts were less timely than the current ACl product.
- Loss of fourth line unacceptable.
- No graphical trial plan available.
- Need for D-side to clear the fourth line data.
- "Situation Display" requires a great deal of work for D-side to manage.
- Menus distracting.
- The Fly-Out menus and conflict probe need work.
- Too many nuisance probe alerts.
- D-side being able to issue control instructions.
- D-side having a full radar screen. Why not having two sectors then?
- Sometimes the menus did not work.

When you coordinated between sectors, did you use CRA?

- **Yes:** 4 participants.
- **Sometimes:** 5 participants.
    - Sometimes alerts were not timely enough.
    - When busy, it is easier to call than looking at the CRT screen.
- **No:** 2 participants.
    - easier to lean over and talk instead of waiting for them to notice when busy, need to get used.
- **No Response:** 1 participant.

If you used it (CRA), was it helpful?

- **Yes:** 6 participants.
    - When busy, not used.
    - When it worked, it was great.
    - D-side and R-side need to work together. R-side cannot take hand-off on an A/C that D-side is about to make a coordination request on!
    - It makes coordination faster and more effective, cutting out the possibility for misunderstanding.
- **Sometimes (limited use):** 4 participants.
    - Needs to get used.
    - In some cases, such as when the aircraft was a significant distance from the sector boundary, it was easier to get an aircraft moved by the other (previous) sector than to wait to have the handoff and control of the a/c to do it yourself.
- **No responses**: 2 participants.

In general, controllers responded positive to the impact of conflict probe and CRA (see Table 28). Controllers rated the accuracy, timeliness, and usefulness as helpful with usefulness being statistically significantly better than 0 on a -5 to 5 scale (*p* = .010). The controllers rated the advisories provided by CRA as accurate, timely, and useful as well, with usefulness being significantly helpful. The controllers rated the Aircraft Queue and the reminder especially helpful (*p* = .046 and .008, respectively) and had a somewhat positive impression of the portals. The only feature that controllers rated somewhat negatively was the Sector Queue.

Table 28. Rating Each of the Conflict Probe and CRA Features on ATC

2. The Effect of CRA: Please rate each of the Conflict Probe and CRA features on air traffic control.	
Question	Detrimental **-5 to 5** Helpful
(1) Probe: Accuracy in Terms of Distance	**1.08**
(2) Probe: Timeliness	**0.75**
(3) Probe: Usefulness	**1.67* (*p* = .010)**
(4) Resolution: Accuracy	**0.58**
(5) Resolution: Timeliness	**1.08**
(6) Resolution: Usefulness	**1.33 (*p* = .059)**
(7) Sector Queue	**-0.08**
(8) Aircraft Queue	**0.92* (*p* = .046)**
(9) Reminder	**1.50* (*p* = .008)**
(10) Portals	**0.25**

The ratings of controllers of the impact of CRA on false, missed, and nuisance alerts and resolutions were somewhat negative with the exception of the number of missed conflict probe alerts (see Table 29). None of the ratings were significantly different from neutral point.

Table 29. Rating of the Effect of CRA on False, Missed, and Nuisance Alerts and CRA Solutions

3. The Effect of CRA: Please rate the following about false alerts, missed alerts, nuisance alerts, and CRA solutions.	
Question	Unacceptable **-5 to 5** Acceptable
(1) The Number of False Conflict Probe Alerts	**-0.08**
(2) The Number of Missed Conflict Probe Alerts	**0.33**
(3) The Number of Nuisance Conflict Probe Alerts	**-1.25**
(4) The Number of Incorrect CRA Solutions	**-0.25**
(5) The Number of Unacceptable CRA Solutions	**-1.25**

As Table 30 shows, the controllers were quite positive about having access to the CRA menus from the FDB and, especially, the CRA Altitude menu (*p* = .014). The controllers rated the availability of conflict resolutions (*p* = .035), coordination requests (*p* = .005), and time to loss of separation as positive as well. Although controllers did not find it as useful to have the system probe conflicts on the uncleared portion of the trajectory in their sector, they indicated that this was very useful when coordinating with other sectors (*p* = .025). Controllers rated having easy access to a display of an aircraft's trajectory by simply clicking on the call sign as very helpful (*p* = .025).

Table 30. Evaluation of CRA Features

4. Evaluation of CRA Features	
Question	Hindered Greatly **-5 to 5** Helped Greatly
(1, a) CRA menu accessible from the datablock: CRA altitude menu (click on the Altitude of Full Datablock [FDB] second line)	**2.00* (*p* = .014)**
(1, b) CRA menu accessible from the datablock: CRA heading menu (click on the CID field of FDB second line; for D-side, double-click on the Destination field of FDB third line or Heading field of FDB fourth line)	**0.58**
(1, c) CRA menu accessible from the datablock: CRA speed menu (click on FDB third line or the Speed field of FDB fourth line)	**0.08**
(2) Aircraft queue accessible from FDB third line	**0.73**
(3) Aircraft queue accessible from FDB fourth line	**0.27**
(4) Conflict trajectories and trial plans accessible for FDB third and fourth line	**0.25**
(5) Advisories of solutions in the CRA menus	**1.83* (*p* = .035)**
(6) Coordination requests from other sectors on FDB fourth line	**1.50* (*p* = .005)**
(7) Predicted conflict in Safety Portal	**1.33**
(8) The time until the highest alert in Safety Portal	**1.25**
(9) Notified Conflicts: A red color A for an AC to AC conflict that was more than nine minutes away in the Safety Portal	**0.42**
(10) No Notified Conflicts: A grayed number of minutes to an LOS for a downstream conflict in the Safety Portal	**0.25**
(11) Trial plans and uncleared and requested clearances on FDB fourth line	**-0.08**
(12) The concept of sharing the above (Item 11) with speed and free-text on FDB fourth line.	**0.83**
(13) Trial plans on FDB fourth line: Coordination request (CO or CL)	**1.17* (*p* = .025)**
(14) Trial plans on FDB fourth line: Uncleared altitude	**0.92**
(15) Trajectory shown by center-clicking AID (FDB first line)	**2.08* (*p* = .025)**

To test the simulation realism and research equipment ratings, we compared the controller ratings to 5.5, the center of rating scale (see Table 31). The controllers rated the simulations somewhat realistic and somewhat representative of a typical workday. However, they indicated that the simulated airspace was very realistic when compared with their actual airspace (*p* =.022). The controllers did not find that the online workload probe, WAK, interfered with their performance. When asked about the interference of the oculometer with their performance, the controllers indicated that it interfered significantly with their performance. The fNIR device, on the other hand, hardly interfered. The controllers felt that the simulation pilots responded fine, but not extremely well. Although the controllers may not have received enough training, they rated the training as effective.

Table 31. Simulation Realism and Research Equipment Ratings

5. Simulation Realism and Research Apparatus Ratings		
Question	**Rating Scale**	**Mean**
(1) How realistic was the overall simulation experience compared to actual operations?	Unrealistic **1 to 10** Realistic	**5.92**
(2) How representative were the scenarios of a typical workday?	Not Representative **1 to 10** Representative	**5.08**
(3) How realistic was the simulated airspace compared to your actual NAS airspace?	Unrealistic **1 to 10** Realistic	**7.08* (p = .022)**
(4) To what extent did the online workload rating interfere with your ATC performance?	Not At All **1 to 10** A Great Deal	**4.25**
(5) To what extent did the oculometer interfere with your ATC performance?	Not At All **1 to 10** A Great Deal	**6.70**
(6) To what extent did the fNIR interfere with your ATC performance?	Not At All **1 to 10** A Great Deal	**4.00**
(7) How well did the simulation pilots respond to your clearances in terms of traffic movement and callbacks?	Extremely Poorly **1 to 10** Extremely Well	**5.42**
(8) How effective was the training?	Not Effective **1 to 10** Extremely Effective	**6.42**

3.17 Summary of Exit Interviews

In general, controllers expressed some concerns in the current version of the CRA, such as the WAFDOF—even though they liked some features, such as the coordination function. Their detailed responses are presented as follows:

- The R-side had difficulty using the CRA when traffic volume was high and busy. So, controllers thought it would be useful on the D-side.
- The altitude resolutions did not take into account the current field practice; for example, odd altitudes for the East and even altitudes for the West (WAFDOF).
- The Coordination function of the CRA was useful.
- More training was needed. Controllers thought they should have received training for at least a week.

- The CRA took away the current function of the fourth line that is used in the field, which required adjustment because controllers use the fourth line quite often in the field.
- Controllers liked the capability of dragging the aircraft routes easily on the display, which enabled them to try out routes visually.
- The 12-minute look-ahead time was too short.
- Controllers felt that there was too much information in the CRA menu.
- Heading resolutions that required a large turn angle were unrealistic.
- Controllers claimed that there were wrong resolutions.
- Even if they selected a resolution, alerts remained red on the display.
- There were too many active information elements on the datablock area.

4. DISCUSSION

Controllers used CRA menus significantly more often than Baseline menus at both R-side and D-side. They also spent more time when they used CRA than when they used Baseline. The use duration difference between the two conditions was significantly longer when they used CRA at the D-side. When they used it on the R-side, the difference was not significant. This shows that the CRA effect was not as large on the R-side as on the D-side. Maybe controllers used CRA for planning rather than for solving immediate conflict problems, which is the R-side's main concern. However, our participant controllers differed in their opinions on which side to make the CRA available.

We assume that their use of Fly-Out menus is related to their use of the keyboard, and some controllers may have preferred using keyboard to using the Flight-Out menus on the datablock. If we relate their keyboard use to their Fly-Out menus use, we may be able to find individual differences and unique situations when controllers preferred one mode over the other. That information will be useful to modify or improve the design of CRA. It will also be useful for designing effective training materials for CRA.

With our data analysis, so far, we know neither why controllers used CRA menus more often than Baseline menus nor why they took longer with CRA menus than with Baseline menus. We did find that, sometimes, controllers need to execute extra steps in the CRA Flight-Out menus if the altitude recommended by the CRA was not an altitude they would choose (WAFDOF). Although CRA menus gave them more information than Baseline menus, without further data analysis, we do not know, definitively, whether the additional information in the CRA menus was actually more helpful to them. They might have used CRA menus just for collecting information or for actually avoiding conflicts. We did not have a chance to analyze how controllers dealt with conflict resolution selections offered in the CRA menus. We will need (a) to examine controllers' activities after selecting resolutions and (b) to evaluate whether CRA was more helpful to them than the Baseline menus. In the following section, we will discuss the limitations and results.

4.1 Limitations

Before we start the discussion of the results, we want to point out some of the limitations of this study and how it may have affected controller behavior and performance. We assumed that the training and performance challenges that the controllers faced during the experiment equally affected all conditions, because these challenges existed throughout the experiment. We counterbalanced experimental conditions to minimize learning effects, of course, but these challenges might have added variance to our dependent measures, which could have masked the effect of CRA. In the following paragraphs we want to reiterate the main issues that our participants overcame during the experiment.

The conditions that controllers encountered during this experiment were quite foreign to them for several reasons. First, we conducted this study in conjunction with the High Altitude experiment and used real sectors from the Kansas City ARTCC (ZKC, Sectors 20 and 22). We kept the airspace intact and used traffic samples extracted from field recordings. We then combined traffic from several days to increase traffic levels up to 150% of the MAP value for the sectors. To increase the pool of controllers from which we could recruit controllers, we excluded controllers from ZKC; therefore, controllers entered our experiment in unknown airspace and traffic levels that were beyond the levels that controllers would normally experience with these sectors.

Second, we exposed controllers to conditions used for our Baseline that NextGen envisioned during the time at which the FAA might implement CRA. Even the Baseline condition had capabilities that controllers did not encounter previously. The Baseline environment included Data Communications and Separation Management capabilities, as well as included changes to the CHI that are currently under development. Except for the ERAM interface, these capabilities do not currently exist in the field.

A third reason why the conditions during the experimental runs were foreign was that we were not able to provide the amount of training we would otherwise include in a study of an advanced concept like CRA. Because of our dependence on the High Altitude experiment, we had a limited amount of time available for training and relied on training transfer from the High Altitude experiment to the CRA experiment. That training transfer was not as great as we had expected. The controllers had seen ZKC airspace, but it had super-high sectors instead of the high-altitude sectors we used in our experiment. During the last day of the High Altitude experiment, the controllers participated in a Generic Airspace concept evaluation. The airspace used during that evaluation was also from ZKC, but it overlaid the sectors used in the CRA experiment. To ease learning the airspace, the High Altitude experiment had simplified adjacent sectors and frequencies. During the CRA experiment, we kept the airspace intact and required controllers to handoff to the correct sectors and switch aircraft to the correct frequencies—therefore, transfer of training of the airspace used during the Generic Airspace concept evaluation may not have occurred. The High Altitude experiment also provided controllers with a traffic mix that included Data Comm equipped aircraft, but we based its extension of the ERAM emulation on the CPDLC Build 1 and Build 1A interfaces. The CRA experiment used a more advanced interface prototyped for the Data Communications and Separation Management Program Office. Therefore, transfer of training of the Data Comm capabilities used in the High Altitude experiment may have occurred only for the general Data Comm concept and less for the automation features.

A fourth novelty that often occurs during concept research is that the simulations run on a system prototyped to mimic the concept and its features. Our team worked hard to create a stable system, but some of the features may need further refinement. In addition to comparing concepts, another reason to run these experiments is to determine whether we need to add or modify program requirements. We perform this in many of our studies, but for the CRA experiment this was more pronounced than usual. Our team had the challenge to integrate DESIREE with JEDI, and we created an interface that was a hybrid of DESIREE and JEDI features.

Finally, to assess how the changes in conditions affected our controller participants, we instrumented them with physiological equipment to measure oxygenation of the prefrontal cortex, to measure electrical activity of the brain, and to track eye movements. Controllers rotated their control positions (R- and D-sides) and sectors—our dependent measures for each of the conditions should be minimal but may show an effect, in general, thereby potentially reducing the range of values available to detect differences between conditions.

This experiment resulted in a large amount of data that enables us to look at changes in controller behavior and performance as well as the performance of the conflict probe and conflict resolution algorithms. We have reduced and analyzed most of the data sets that we collected.

In addition, when we focused on individual tasks, such as clearances, we found significant effects on some clearances. In this CRA experiment, controllers did not have to use the CRA menus if they decided not to.

4.1.1 Training

All participants indicated they did not receive enough training to learn the CRA features and functions. Although we counterbalanced our experimental design, we rotated controllers between positions and sectors as unique teams. Each controller ran the same condition, but at a different position or sector three times; once each day for three days. We counterbalanced conditions across days and controller groups. This provided us the opportunity to examine if workload ratings changed depending on how many times an individual controller had seen a given condition.

We applied a regression model to the workload ratings by test runs. The results showed that on the early runs, the workload ratings of the CRA conditions were higher than the baseline condition; but in the later runs, they were about the same as those of the baseline. We assume that the controllers had difficulty learning the CRA features at first, which contributed to their higher workload ratings. A higher reduction of workload occurred with continued use of CRA than with the Baseline—that is, the decreasing trend of workload ratings by days was steeper in the CRA condition than in the Baseline condition. We assume that this trend would have continued resulting in lower workload ratings in the CRA condition than in the Baseline condition if our simulation experiment had lasted longer.

4.1.2 Suggested Modifications to CRA

We had two types of questionnaires: a Post-Scenario Questionnaire and an Exit Questionnaire. Due to the limited amount of training, controllers' responses to post-scenario questions may not have been as definitive as their responses to the exit questions. In their exit questionnaire responses, controllers rated CRA as a useful tool. The majority of the controllers (58%) suggested using CRA on both R- and D-sides. A few controllers (25%) thought it should be located on the D-side only. One controller thought it was not needed at all. There were a few features and functions that did not receive positive ratings. One of the negative comments was that CRA provided altitude resolutions that were the wrong altitude for direction of flight, WAFDOF (i.e., CRA did not take into account that controllers use odd altitudes for the East and even altitudes for the West). Procedural separation of East and West bound traffic is an example of what controllers refer to as positive control (i.e., when a controller leaves an aircraft unattended for a short period, controls are in place that prevent a loss of separation or a deviation for several minutes). Another example of positive control is providing a temporary altitude assignment even though a controller knows that under normal circumstances the next controller will accept the handoff and continue the climb or descent before the aircraft levels off. A third example of positive control is the practice of preferring to turn behind a crossing aircraft rather than in front of that aircraft. During the discussion after the experiment, controllers also indicated that some of the resolutions turned aircraft in front of crossing traffic. Although the resolutions that turn aircraft in front of crossing traffic may be valid, controllers do not feel comfortable using them. This may be an indication that a controller's estimation of the longitudinal position (i.e., the position along the trajectory) is not as accurate as the lateral position. The trajectory model underlying the conflict probe has this limitation (i.e., *along track* error is larger than *cross track* error). Controllers also criticized that some of the CRA resolutions were incorrect. Another comment from several of the controllers was that it took too many steps to use the CRA features.

In a fast-paced environment, such as ATC, every step within a task represents valuable time. Reducing the number of steps will make a new automation feature more acceptable. Better integration of the CRA data in the ERAM interface may reduce some of these steps. For example, providing the optimal solution by itself with an option to explore more may allow us to keep more of the standard Fly-Out menus intact and reduce the amount of time spent on scanning the additional information in the CRA menus.

4.1.3 Workload

The analysis of the subjective workload data showed, as we expected, that workload increased with an increase in traffic. We did not find a difference in subjective workload between the Baseline and the CRA conditions. We had hypothesized that the workload would increase because of the limited amount of training and the large amount of new features that we introduced during the CRA conditions. Participants in the mini-evaluations and the HITL demonstration at MITRE/CAASD had the advantage because the CRA team had involved them from the very beginning. The advantage of that involvement was that it helped them to understand the CRA environment and the changes it underwent along the way. The participants in this study were active CPCs, from several ARTCCs, who did not have that advantage. Many of them had some training on ERAM, but they did not use ERAM at their sectors yet. Even the Baseline condition was a big change from the day-to-day operational environment. Unfortunately, our schedule did not permit us to include a true baseline—a condition that closely represents the current controller workstation environment. In this study, we found that the introduction of CRA did not result in an increase in workload. Despite the novelty of the airspace, the limited amount of training, and the new capabilities of the automation, controllers had the same overall workload in both Baseline and CRA conditions.

4.1.4 Conflict Detection and Resolution

Our results indicate that with CRA, controller seemed more strategic. We observed a trend in a reduction of tactical conflict alerts, which means controllers resolved conflicts earlier. The analysis of how controllers dealt with scripted conflicts shows this more strongly in a reduced number of the scripted conflicts, resulting in a conflict probe alert and shorter durations for those scripted conflicts that still resulted in conflict probe alerts.

5. RECOMMENDATIONS

The realities of schedules and budgets have severely constrained the scope of the CRA experiment. It reduced training time available to the controllers, the number of controllers, and the level of integration of the CRA functions in the controller workstation. Despite these constraints, we have found several positive indications that CRA provides benefits. Depending on scheduling and funding, we recommend the following:

1. Reduce remaining data and analyze the data as they become available.

   Potential benefits of CRA include that it may improve the quality of trajectories, thereby making the NAS more predictable and reducing the risk of loss of separation. We have shown that fewer tactical alerts took place and controllers resolved scripted conflicts earlier. We have not yet analyzed the data to determine the impact of CRA on the quality of trajectories.

   The data we collected also lends itself to determining the risk of loss of separation even under circumstances in which the system did not record a loss of separation. Analyses of that data would provide us with the insight of how CRA may affect the risk of losing separation and provide ATC researchers with a new measure to assess risk.

2. Analyze the available eye-movement data.

   We have collected and reduced the available eye-movement data and should determine whether visual-scanning characteristics have changed under CRA conditions. One concern we have is the complexity of the CRA Fly-Out menus reported by controllers. This complexity may force controllers to tunneling their vision to the menus at the expense of scanning for situation monitoring, scanning for conflict detection, and scanning for sequencing.

3. Conduct Bayesian statistical analysis of the existing data sets.

   When we use Bayesian statistics, some of the limitations of standard NHST do not apply. For example, when using NHST on a data set, we cannot take a look at the results (after a number of participants have completed the experiment) and decide to add a few more. Adding more observations after we have completed the NHST violates its main rule—that is, to set the significance level for the test and choose the number of observations prior to the experiment. Bayesian statistics do not have this limitation. Therefore, when using Bayesian statistics, we can collect data on additional observations (when funding becomes available).

4. Conduct regression analysis on workload as a function of the number of aircraft.

   The scenarios we used in this experiment used identical traffic samples (only the call signs changed between scenarios) and ramped traffic from 33% to 150% of the MAP value of the sectors. The data on controller workload and the number of aircraft are ideally suited for capacity benefits analysis. Data on previous studies have been used for this purpose even if the experiment did not directly address the benefits question (e.g., the use of data from Separation Management HITL-2 to determine capacity gains for Data Position display processors technical refresh). For CRA, we gradually increased the traffic level during each experimental run to conduct this type of analysis; however, because of time constraints, we have not completed them yet.

5. Collect data on additional observations.

   The small number of participants in this experiment did not give a good power to the tests we used. Increasing the sample size should improve it.

6. Fully integrate the CRA capabilities into a high-fidelity emulation of ERAM.

   Due to scheduling constraints, the integration of CRA on the high-fidelity emulation of ERAM in DESIREE was not optimal. JEDI views and menus augmented DESIREE capabilities. This resulted in a mixed interface that contains DESIREE and JEDI elements, instead of providing controllers with an integrated workstation display.

5.1 Ongoing and Future Activities

5.1.1 Detailed Analysis of Controller Interactions with CRA Fly-Out Menus

The analyses of controller interactions with the system provided evidence that controller behavior changed when CRA became available. The overall number of commands did not change, but controllers accessed CRA Fly-Out menus more often than the probed menus. The data we have collected additionally contains information about what controllers did when they accessed these Fly-Out menus. More detailed analyses of these activities will help explain if controllers accessed CRA Fly-Out menus to look at the information only, to implement proposed resolutions, or to create a resolution on their own. Combining the data from DESIREE and JEDI has been a challenge, and we did not complete reduction of that data set in time for this report. The reduced data is available, now, but requires further verification before our analyses can take place.

5.1.2 Analysis of EEG Data

The data we collected on electrical activity of the brain may show evidence for changes in workload with CRA conditions that we cannot capture with subjective workload ratings or fNIR assessment of cognitive workload in the prefrontal cortex. The analysis of EEG data sets is time and labor intensive, but it will serve several purposes. First, we will attempt to uncover how CRA may have changed brain activities to assess workload in a different manner. A secondary reason to further analyze this data set is to determine what additional measures we can derive.

References

Alexander, J. R., Alley, V. L., Ammerman, C. M., Hostetler, C. M., & Jones, G. W. (1988). *FAA Air Traffic Control operations concepts: Volume 3 – ISSS en route controllers* (DOT/ FAA/AP-87-01). Washington, DC: U.S. Department of Transportation, Federal Aviation Administration.

Ayaz, H., Cakir, M., Izzetoglu, K., Curtin, A., Shewokis, P., Bunce, S., & Onaral, B. (2012). Monitoring expertise development during simulated UAV piloting tasks using optical brain imaging. Proceedings from *Aerospace Conference, 2012 IEEE*. Big Sky, MT.

Ayaz, H., Willems, B., Bunce, S., Shewokis, P., Izzetoglu, K., Hah, S., … Onaral, B. (2010). Cognitive workload assessment of air traffic controllers using optical brain imaging sensors. In T. Marek, W. Karwowski and V. Rice (Eds.), *Advances in Understanding Human Performance: Neuroergonomics, Human Factors Design, and Special Populations* (pp. 21–32). Boca Raton, FL: CRC Press, Taylor & Francis Group.

Ayaz, H., Willems, B., Bunce, S., Shewokis, P., Izzetoglu, K., Hah, S., … Onaral, B. (2011). Estimation of cognitive workload during simulated air traffic control using optical brain imaging sensors. Foundations of Augmented Cognition. *Directing the Future of Adaptive Systems*, 549–558.

Bowen, K. (2000*). Concept of use for problem analysis, resolution, and ranking* (MTR00W0000049). McLean, VA: MITRE Center for Advanced Aviation System Development.

Bunce, S., Izzetoglu, M., Izzetoglu, K., Onaral, B., & Pourrezaei, K. (2006). Functional near-infrared spectroscopy: An emerging neuroimaging modality. *IEEE Engineering in Medicine and Biology Magazine 25*(4), 54–62.

Exum, M., Bolczak, R., Celio, J. C., & Poore, D. S. (2010). *Concept of operations for en route separation management enhancements* (MP090111R2). McLean, VA: MITRE Center for Advanced Aviation System Development.

Federal Aviation Administration. (2011). *FAA's NextGen implementation plan, March 2011*. Washington, DC. Retrieved from http://www.faa.gov/nextgen/media/ng2011_implementation_plan.pdf

Hah, S., Willems, B., & Phillips, R. (2006). The effect of air traffic increase on controller workload. In *Proceedings of the 50th Human Factors and Ergonomics Society Annual Meeting* (pp. 50–54). Santa Monica, CA: Human Factors and Ergonomics Society.

Mills, S. H., Kirk, D. B., & Rozen, N. E. (2011a). *Conflict resolution advisories heading menu Mini-Evaluation 2 - Briefs and preliminary results* (MP110081). McLean, VA: MITRE Center for Advanced Aviation System Development.

Mills, S. H., Kirk, D. B., & Rozen, N. E. (2011b). *Experiment plan for initial conflict resolution advisories, Build 1: Human-in-the-loop evaluation - Draft* (MTR110142). McLean, VA: MITRE Center for Advanced Aviation System Development.

Mills, S. H., Meyer, K. A., Syeda, S. H., Bowen, K. C., & Viets, K. J. (2012). *Concept of operations for initial conflict resolution advisories - Revision 1* (MTR120564). McLean, VA: MITRE Center for Advanced Aviation System Development.

Mills, S. H., & Rozen, N. (2010). *Conflict resolution advisories enhanced probed altitude menu: Mini-evaluation 1- Briefs and preliminary results* (MTR100043). McLean, VA: MITRE Center for Advanced Aviation System Development.

Peterson, L. M., Bailey, L. L., & Willems, B. F. (2001). *Controller-to-controller communication and coordination taxonomy* (DOT/FAA/AM-01/19). Washington, DC: Office of Aerospace Medicine, U.S. Department of Transportation, Federal Aviation Administration.

Radio Technical Commission for Aeronautics. (2013). *SC-214 Standards for Air Traffic Data Communication Services*. www.rtca.org

Rouder, J., Speckman, P., Sun, D., & Morey, R. (2009). Bayesian *t*-tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225–237.

Sollenberger, R. L., Willems, B., DiRico, J., Hale, M., & Deshmukh, A. (2010). *Human-in-the-loop investigation of automation requirements for separation management* (DOT/FAA/TC-10/07). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Stein, E. S. (1985). *Controller workload: An examination of workload probe* (DOT/FAA/CT- TN 84/24). Atlantic City International Airport, NJ: Federal Aviation Administration Technical Center.

Syeda, S. H., Bowen, K. C., Meyer, K. A., & Viets, K. J. (2011). *Concept of operations for initial conflict resolution advisories* (MTR110094). McLean, VA: MITRE Center for Advanced Aviation System Development.

Talotta, N. J. (1992). *Controller evaluation of initial data link air traffic control services: Mini Study 2, Volume 1* (DOT/FAA/CT-92/2,1). Atlantic City International Airport, NJ: U.S. Department of Transportation, Federal Aviation Administration, William J. Hughes Technical Center.

Willems, B., Hah, S., & Schulz, K. (2010). *En route data communications: Experimental human factors evaluation* (DOT/FAA/TC-10/06). Atlantic City International Airport, NJ: U.S. Department of Transportation, Federal Aviation Administration, William J. Hughes Technical Center.

Acronyms

ACID	Aircraft Computer Identification
AERA	Automated En Route Air Traffic Control
APREQ	Approval For Request
APRV	Approve
ARTCC	Air Route Traffic Control Center
ATC	Air Traffic Control
ATWIT	Air Traffic Workload Input Technique
CAASD	Center for Advanced Aviation System Development
CHI	Computer-Human Interface
CID	Computer Identification (of an aircraft)
CONOPS	Concepts and Operations
CPC	Certified Professional Controllers
CPDLC	Controller-Pilot Data Link Communications
CRA	Conflict Resolution Advisories
DB	Datablock
DESIREE	Distributed Environment for Simulation, Rapid Engineering, and Evaluation
D-side	Radar Associate Position
EEG	Electroencephalogram
ERAM	En Route Automation Modernization
EROS	Event-Related Optical Signal
FLID	Flight Identification
fMRI	functional Magnetic Resonance Imaging
fNIRS	functional Near-Infrared Spectroscopy
HITL	Human-In-The-Loop
JEDI	Joint En Route Decision Support System Infrastructure
MAP	Monitor Alert Parameter
NAS	National Airspace System
NextGen	Next Generation Air Transportation System
NHST	Null Hypothesis Significance Testing
PARR	Problem Analysis Resolution and Ranking
PTT	Push-To-Talk
PVD	Positive Visual Display

RDHFL	Research Development and Human Factors Laboratory
R-side	Radar Console Position
RTCA	Radio Technical Commission for Aeronautics
SC-214	Standards for Air Traffic Data Communication Services
SepMan	Separation Management
SMAR	Sliding Motion Artifact Removal
SME	Subject Matter Expert
TBO	Trajectory-Based Operations
TFM	Traffic Flow Management
TGF	Target Generation Facility
TLX	Task Load Index
URET	User Request Evaluation Tool
VSCS	Voice Switching and Communications System
WAFDOF	Wrong Altitude for Direction of Flight
WAK	Workload Assessment Keypad
WJHTC	William J. Hughes Technical Center
ZKC	Kansas Air Route Traffic Control Center
ZOB	Cleveland Air Route Traffic Control Center

Appendix A: Informed Consent Form

Informed Consent Statement

I, ______________________________, understand that this simulation, entitled "Evaluations of Conflict Resolution Advisories," is sponsored by the Federal Aviation Administration (FAA).

Nature and Purpose:

I have been recruited to volunteer as a participant in this simulation that will investigate: a) the availability of Conflict Resolution Advisories and b) the impact of having Conflict Resolution Advisories on the Radar Associate Position only. This simulation will evaluate these issues in traffic scenarios using a simulated En Route Automation Modernization (ERAM) system. I understand that the participants will be randomly assigned to work as either R-side or D-side controllers in some conditions. Depending on the condition, I will be asked to wear a head-mounted oculometer to record eye movements, Electroencephalogram (EEG) equipment, and a functional Near Infra Red Spectroscopy instrument to measure oxygenation of the brain. The results of the study will be used to determine the benefits and feasibility of integrating these components into the future en route environment.

Experimental Procedures:

Twelve en route Certified Professional Controllers (CPCs) from Level 11 and 12 facilities will participate in the simulation. Four participants will arrive at the lab at a time. They will spend 8 days at the lab over a 2-week period. They will travel in on a Monday and travel out on Friday of the following week. The first week, participants will work on the High Altitude experiment. During the second week, participants will join the Conflict Resolution Advisories Experiment II. At the start of the simulation, the participants will be randomly assigned to work as either an R-side or D-side controller for conditions. The airspace will consist of two adjacent sectors both staffed with an R-side and a D-side controller. Each participant will rotate through positions and sectors.

The participants will work from about 8:00 AM to about 4:30 PM every day with a lunch break and at least two rest breaks. The first morning of the Conflict Resolution Advisories Experiment will consist of an initial briefing to review project objectives and participant rights and responsibilities. It will include initial familiarization training on the simulated airspace, the system, and the procedures. The participants will then go the laboratory to begin hands-on training. They will complete practice scenarios prior to completing the test scenarios. All scenarios will be about 60-minute in duration.

During experimental scenarios the R-side participants will wear a head-mounted oculometer to record eye-movement data via infrared technology, a functional Near Infra Red Spectroscopy instrument, and an EEG equipment. The exposure to infrared illumination while wearing the oculometer is less than 4% of the intensity of that experienced when outside on a sunny day. The EEG equipment is a passive data collection equipment and collect electronic signals from the surface of the head. The data is for resarch purpose only and not for any clinical diagnosis purpose. The researchers are not qualified to diagnose any symptoms based on the EEG data collected.

The participants will provide workload ratings when prompted at designated intervals throughout each scenario. An automated data collection system will record system operations and generate a set of standard Air Traffic Control (ATC) simulation measures, including safety, capacity, efficiency, and communications. After each scenario, the participants will complete questionnaires to report their overall workload, situation awareness, and performance and to rate various aspects of the test condition. The simulation will be audio and video recorded.

After the participants have completed each of the simulation components, they will gather for a final debriefing session to provide final comments and feedback.

Anonymity and Confidentiality:

My participation in this simulation is strictly confidential. Any information I provide will remain anonymous: no individual names or identities will be associated with the data or released in any reports.

Benefits:

I understand that the only benefit to me is that I will be able to provide the researchers with valuable feedback and insight into the effectiveness of potential ATC tools and workstation configurations. My data will help the FAA to determine the benefits and feasibility of these modifications in this environment.

Participant Responsibilities:

I am aware that to participate in this study I must be a certified professional controller who is qualified at my facility and holds a current medical certificate. I must also have normal or corrected-to-normal (20/20) vision and do not wear bifocals, trifocals, or hard-contact lenses that are incompatible with the eye-tracking device used in this simulation. I will control traffic and answer the questions asked during the study to the best of my abilities. I will not discuss the content of the experiment with anyone until the study is completed.

Participant Assurances:

I understand that my participation in this study is completely voluntary and I can withdraw at any time without penalty. I also understand that the researchers in this study may terminate my participation if they believe this to be in my best interest. I understand that if new findings develop during the course of this research that may relate to my decision to continue participation, I will be informed. I have not given up any of my legal rights or released any individual or institution from liability for negligence.

The research team has adequately answered all the questions I have asked about this study, my participation, and the procedures involved. I understand that Ben Willems, Dr. Sehchang Hah, or another member of the research team will be available to answer any questions concerning procedures throughout this study. If I have questions about this study or need to report any adverse effects from the research procedures, I will contact Ben Willems at (609) 485-4191 or Dr. Sehchang Hah at (609) 485 5809.

Discomfort and Risks:

I understand that I will not be exposed to any foreseeable risks or intrusive measurement techniques. The only anticipated discomfort may be some discomfort from the oculometer head mount and EEG equipment. I agree to immediately report any injury or suspected adverse effect to Ben Willems at (609) 485-4191 or Dr. Sehchang Hah at (609) 485 5809.

Signature Lines:

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

Research Participant:____ Date:__________

Investigator:____ Date:__________

Witness:____ Date:__________

Appendix B: Biographical Questionnaire

Participant # ___________ **Date** ___________

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**

3. How long have you worked as **an Air Traffic Controller (include both FAA and military experience)**?	_____ **years** _____ **months**

4. How long have you worked as a **CPC for the FAA**?	_____ **years** _____ **months**

5. How long have you **actively controlled traffic** in the en route environment?	_____ **years** _____ **months**

6. How many of the **past 12 months** have you actively controlled traffic?	_____ **months**

7. Rate your current **skill as a CPC**.	Not Skilled	①②③④⑤⑥⑦⑧⑨⑩	Extremely Skilled

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

Appendix C: Post-Scenario Questionnaire

Post-Scenario Questionnaire

A. Air Traffic Control Tasks

For each of the following major tasks, please rate how well the situation you were in during the preceding scenario helped or hindered your ability to control air traffic. We are not asking any specific questions about CRA and any other features. We want to hear your general opinions on controlling air traffic in the previous run. There are six major air traffic control tasks which are explained on this page below. Please rate each task on the following page. The rating of -5 represents that you thought they limited or hindered your performance tremendously. The rating of 5 is the opposite: They helped you perform in a very positive manner. The rating of 0 means no effect.

We have listed the tasks for your reference below. Please consult them. Please circle the number that corresponds to your rating for each task on the next page.

A. Situation Monitoring: Checking and evaluating separation; Analyzing initial requests for clearances; Processing departure/En Route time information; Housekeeping.

B. Detecting Aircraft Conflicts: Detecting aircraft conflicts.

C. Resolving Aircraft Conflicts: Performing aircraft conflict resolution; Performing airspace conflict processing; Suppressing/Restoring alerts**.**

D. **Managing Air Traffic Sequences:** Responding to traffic management constraints/flow conflict; Processing deviations; Establishing arrival sequences; Managing departure flows; Monitoring non-controlled objects.

E. Routing or planning flights: Planning clearances; Responding to contingencies/emergencies; Responding to special operations; Reviewing flight plans; Processing flight plan amendments; Receiving transfer of control/radar identification; Initiating transfer of control/radar identification; Issuing pointouts; Responding to pointouts; Issuing clearances; Establishing, maintaining, and terminating radio communications; Establishing radar identification.

F. Assessing weather impact: Responding to significant weather information,

G. Managing sector/position resources: Assuming position responsibility; Executing backup procedures for communication failures/transient operation; Managing personal workload.

Participant # ___________ **Date** ___________

	Hindered greatly										**Helped greatly**
A. Situation Monitoring	-5	-4	-3	-2	-1	0	1	2	3	4	5
B. Detecting Aircraft Conflicts	-5	-4	-3	-2	-1	0	1	2	3	4	5
C. Resolving Aircraft Conflicts	-5	-4	-3	-2	-1	0	1	2	3	4	5
D. Managing Air Traffic Sequences	-5	-4	-3	-2	-1	0	1	2	3	4	5
E. Routing or Planning Flights	-5	-4	-3	-2	-1	0	1	2	3	4	5
F. Assessing Weather Impact	-5	-4	-3	-2	-1	0	1	2	3	4	5
G. Managing Sector/Position Resources	-5	-4	-3	-2	-1	0	1	2	3	4	5

If you have any additional comments about the positive or negative aspects of automation conditions during this scenario, please give us your feedback/opinions.

Participant # ___________ **Date ___________**

B. Automation

a. How was the automation available in the scenario you just worked compare to DSR, HOST, and URET?

Much Worse **Much Better**

A. Fly-out menus: Complexity	-5	-4	-3	-2	-1	0	1	2	3	4	5
B. Fly-out menus: Easiness of use	-5	-4	-3	-2	-1	0	1	2	3	4	5
C. Fly-out menus: Usefulness	-5	-4	-3	-2	-1	0	1	2	3	4	5
D. Coordination: Complexity	-5	-4	-3	-2	-1	0	1	2	3	4	5
E. Coordination: Easiness of use	-5	-4	-3	-2	-1	0	1	2	3	4	5
F. Coordination: Usefulness	-5	-4	-3	-2	-1	0	1	2	3	4	5

b. What was the impact of the automation available in the scenario you just worked for the following items?

Very Negative **Very Positive**

A. Finding a solution to a conflict	-5	-4	-3	-2	-1	0	1	2	3	4	5
B. Situational awareness for potential conflicts?	-5	-4	-3	-2	-1	0	1	2	3	4	5
C. Collaboration in solving conflicts within your team?	-5	-4	-3	-2	-1	0	1	2	3	4	5
D. Collaboration in solving conflicts between sectors?	-5	-4	-3	-2	-1	0	1	2	3	4	5
E. The effect of Data Comm on workload	-5	-4	-3	-2	-1	0	1	2	3	4	5
F. The effect of Electronic Coordination on workload	-5	-4	-3	-2	-1	0	1	2	3	4	5
G. The effect of CRA on workload	-5	-4	-3	-2	-1	0	1	2	3	4	5
H. The effect of Data Comm on conflict resolution	-5	-4	-3	-2	-1	0	1	2	3	4	5
I. The effect of electronic coordination on conflict resolution	-5	-4	-3	-2	-1	0	1	2	3	4	5

Participant # ___________ **Date** ___________

c. How does the Conflict Detection available in the scenario you just worked compare to DSR, HOST, and URET?

Much Worse **Much Better**

A. Accuracy	-5	-4	-3	-2	-1	0	1	2	3	4	5
B. Timeliness	-5	-4	-3	-2	-1	0	1	2	3	4	5
C. Distraction	-5	-4	-3	-2	-1	0	1	2	3	4	5
D. Presentation Format	-5	-4	-3	-2	-1	0	1	2	3	4	5
F. Interaction	-5	-4	-3	-2	-1	0	1	2	3	4	5

d. Please rate the following abut false alerts, missed alerts, and nuisance alerts. Also, please rate the following about CRA solutions if it was available in your previous scenario.

Unacceptable **Acceptable**

A. The number of false Conflict Probe Alerts	-5	-4	-3	-2	-1	0	1	2	3	4	5
B. The number of missed Conflict Probe Alerts	-5	-4	-3	-2	-1	0	1	2	3	4	5
C. The number of nuisance Conflict Probe Alerts	-5	-4	-3	-2	-1	0	1	2	3	4	5
D. The number of incorrect CRA solutions if it was available in your previous scenario	-5	-4	-3	-2	-1	0	1	2	3	4	5
E. The number of unacceptable CRA solutions if it was available in your previous scenario	-5	-4	-3	-2	-1	0	1	2	3	4	5

e. If you just worked a scenario that included CRA, please rate the quality of the advisories.

Unacceptable **Acceptable**

A. Accuracy	-5	-4	-3	-2	-1	0	1	2	3	4	5
B. Timeliness	-5	-4	-3	-2	-1	0	1	2	3	4	5
C. Distraction	-5	-4	-3	-2	-1	0	1	2	3	4	5
D. Presentation Format	-5	-4	-3	-2	-1	0	1	2	3	4	5
F. Interaction	-5	-4	-3	-2	-1	0	1	2	3	4	5

Appendix D: Exit Questionnaire

Participant # ___________ **Date** ___________

Exit Questionnaire

1. The Effect of CRA

The rating of -5 represents that you thought the system limited or hindered your performance very much. The rating of 5 represents that it helped your air traffic control very much.

Please note that your rating is **about the use of Conflict Probes and Conflict Resolution Advisories**.

A. What was the effect of CRA on your air traffic control?

Hindered greatly										**Helped greatly**
-5	-4	-3	-2	-1	0	1	2	3	4	5

Justification for your ratings:

B. Where do you think CRA should be located? If you do not think it is needed, please mark 'No Need' below.

R-Side only
D-side only
Both R- and D-side

No Need

Justification for your choice:

Participant # ___________ **Date** ___________

C. What aspects of the current version of CRA you like?

Justification for your response:

D. What aspects of the current version of CRA did you not like?

Justification for your response:

E. When you coordinated between sectors, did you use CRA?

If you used it, was it helpful?

Justification for your response:

F. Please rate each of the Conflict Probe and CRA features on air traffic control.

Some definitions:

False Alert: A false alert occurs when the automation indicates that it detected a potential loss of separation, but the aircraft will not lose separation.

Missed Alert: A missed alert occurs when the automation does not detect a potential loss of separation, but the aircraft will lose separation.

Nuisance Alert: A nuisance alert occurs when the automation correctly identifies a potential loss of separation after you have resolved the conflict situation.

Participant # ___________ **Date** ___________

Detrimental **Helpful**

a. Probe : Accuracy in Terms of Distance	-5	-4	-3	-2	-1	0	1	2	3	4	5
b. Probe : Timeliness	-5	-4	-3	-2	-1	0	1	2	3	4	5
c. Probe Usefulness	-5	-4	-3	-2	-1	0	1	2	3	4	5
d. Resolution: Accuracy	-5	-4	-3	-2	-1	0	1	2	3	4	5
e. Resolution: Timeliness	-5	-4	-3	-2	-1	0	1	2	3	4	5
f. Resolution: Usefulness	-5	-4	-3	-2	-1	0	1	2	3	4	5
g. Sector Queue	-5	-4	-3	-2	-1	0	1	2	3	4	5
h. Aircraft Queue	-5	-4	-3	-2	-1	0	1	2	3	4	5
i. Reminder	-5	-4	-3	-2	-1	0	1	2	3	4	5
j. Portals	-5	-4	-3	-2	-1	0	1	2	3	4	5

G. Please rate the following abut false alerts, missed alerts, nuisance alerts, and CRA solutions.

Unacceptable **Acceptable**

a. The number of false Conflict Probe Alerts	-5	-4	-3	-2	-1	0	1	2	3	4	5
b. The number of missed Conflict Probe Alerts	-5	-4	-3	-2	-1	0	1	2	3	4	5
c. The number of nuisance Conflict Probe Alerts	-5	-4	-3	-2	-1	0	1	2	3	4	5
d. The number of incorrect CRA solutions	-5	-4	-3	-2	-1	0	1	2	3	4	5
e. The number of unacceptable CRA solutions	-5	-4	-3	-2	-1	0	1	2	3	4	5

Participant # ___________ **Date ___________**

2. Evaluation of Conflict Resolution Advisory Features

We listed the Conflict Resolution Advisory features below. We will ask you about each one separately if it is a desirable feature for safer and more efficient air traffic control. Please rate one by one. The rating of -5 indicates that the particular feature limited and hindered your performance tremendously. The rating of 5 is the opposite, that is, it helped you perform in a very positive way. You can use a rating of 0 to indicate that the feature did not affect your performance.

If you want to comment further or explain your rating, please write it below or the back page with a corresponding feature number so that we can relate your responses to the particular features.

	Hindered greatly										Helped greatly
1. CRA menu accessible from the datablock:											
a. CRA altitude menu (click on the altitude field of FDB 2nd line)	-5	-4	-3	-2	-1	0	1	2	3	4	5
b. CRA heading menu (click on the CID field of FDB 2nd line; For D-side, double click on the destination field of FDB 3rd line or Heading filed of FDB 4th line)	-5	-4	-3	-2	-1	0	1	2	3	4	5
c. CRA speed menu (click on FDB 3rd line or the speed field of FDB 4th line)	-5	-4	-3	-2	-1	0	1	2	3	4	5
2. Aircraft Queue accessible from FDB 3rd line	-5	-4	-3	-2	-1	0	1	2	3	4	5
3. Aircraft Queue accessible from FDB 4th line	-5	-4	-3	-2	-1	0	1	2	3	4	5
4. Conflict trajectories and trial plans accessible from FDB 3rd and 4th line	-5	-4	-3	-2	-1	0	1	2	3	4	5
5. Advisories of solutions in the CRA menus	-5	-4	-3	-2	-1	0	1	2	3	4	5
6. Coordination requests from other sectors on FDB 4th line	-5	-4	-3	-2	-1	0	1	2	3	4	5
7. Predicted conflict in Safety Portal	-5	-4	-3	-2	-1	0	1	2	3	4	5
8. The time until the highest alert in Safety Portal	-5	-4	-3	-2	-1	0	1	2	3	4	5
9. Notified Conflicts: a red color A for an AC to AC conflict that was more than 9 minutes away in the Safety Portal	-5	-4	-3	-2	-1	0	1	2	3	4	5
10. No Notified Conflicts: a grayed number of minutes to a LOS for a downstream conflict in the Safety Portal	-5	-4	-3	-2	-1	0	1	2	3	4	5
11. Trial plans and uncleared and requested clearances on FDB 4th line	-5	-4	-3	-2	-1	0	1	2	3	4	5
12. The concept of sharing the above (Item 11) with speed and free-text on FDB 4th line	-5	-4	-3	-2	-1	0	1	2	3	4	5
13. Trial plans on FDB 4th line: coordination request (CO or CL)	-5	-4	-3	-2	-1	0	1	2	3	4	5
14. Trial plans on FDB 4th line: uncleared altitude	-5	-4	-3	-2	-1	0	1	2	3	4	5
15. Trajectory show by center clicking AID (FDB 1st line)	-5	-4	-3	-2	-1	0	1	2	3	4	5

Participant # ___________ **Date** ___________

3. Simulation Realism and Research Apparatus Ratings

1.	How **realistic** was the overall simulation experience compared to actual operations?	Unrealistic	1	2	3	4	5	6	7	8	9	10	Realistic
2.	How **representative** were the scenarios of a typical workday?	Not Representative	1	2	3	4	5	6	7	8	9	10	Representative
3.	How **realistic was the simulated airspace** compared to your actual NAS airspace?	Unrealistic	1	2	3	4	5	6	7	8	9	10	Realistic
4.	To what extent did the **online workload rating interfere** with your ATC performance?	Not At All	1	2	3	4	5	6	7	8	9	10	A Great Deal
5.	To what extent did the **oculometer interfere** with your ATC performance?	Not At All	1	2	3	4	5	6	7	8	9	10	A Great Deal
6.	To what extent did the **fNIR interfere** with your ATC performance?	Not At All	1	2	3	4	5	6	7	8	9	10	A Great Deal
7.	How well did the **simulation pilots** respond to your clearances in terms of traffic movement and callbacks?	Extremely Poorly	1	2	3	4	5	6	7	8	9	10	Extremely Well
8.	How effective was the **training**?	Not Effective	1	2	3	4	5	6	7	8	9	10	Extremely Effective

Please include any additional comments about the simulation that you would like us to know about.

Appendix E: Over-the-Shoulder Rating Form

Participant # ___________ **Date** ___________

Over-The-Shoulder Rating Form

Instructions for questions 1-37

Please evaluate the effectiveness of the controllers. Please write down observations and make preliminary ratings during the course of the scenario. However, please wait until the scenario is finished before making your final ratings. 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. Also, please write down any comments that may improve this evaluation form. Your identity will remain anonymous, so do not write your name on the form.

Rating	Label Description
1	Controller demonstrated extremely poor judgment in making control decisions and very frequently made errors.
2	Controller demonstrated poor judgment in making some control decisions and occasionally made errors.
3	Controller made questionable decisions using poor control techniques that led to restricting the normal traffic flow.
4	Controller demonstrated the ability to keep aircraft separated but used spacing and separation criteria that were excessive.
5	Controller demonstrated adequate judgment in making control decisions.
6	Controller demonstrated good judgment in making control decisions using efficient control techniques.
7	Controller frequently demonstrated excellent judgment in making control decisions using extremely good control techniques.
8	Controller always demonstrated excellent judgment in making even the most difficult control decisions while using outstanding control techniques.

Maintaining Safe and Efficient Traffic Flow

1. Maintaining Separation and Resolving Potential Conflicts - using control instructions that maintain safe aircraft separation - detecting and resolving impending conflicts early	1 2 3 4 5 6 7 8 9 10
2. Sequencing Arrival and Departure Aircraft Efficiently - using efficient and orderly spacing techniques for arrival and departure aircraft - maintaining safe arrival and departure intervals that minimize delays	1 2 3 4 5 6 7 8 9 10
3. Using Control Instructions Effectively - providing accurate navigational assistance to pilots - avoiding clearances that result in the need for additional instructions to handle aircraft completely - avoiding excessive vectoring or over-controlling	1 2 3 4 5 6 7 8 9 10
4. Overall Safe and Efficient Traffic Flow	1 2 3 4 5 6 7 8 9 10

Participant # ___________ **Date** ___________

Maintaining Attention and Situation Awareness

5. Maintaining Awareness of Aircraft Positions - 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	1 2 3 4 5 6 7 8 9 10
6. Ensuring Positive Control	1 2 3 4 5 6 7 8 9 10
7. Detecting Pilot Deviations from Control Instructions - ensuring that pilots follow assigned clearances correctly - correcting pilot deviations in a timely manner - avoiding excessive vectoring or over-controlling	1 2 3 4 5 6 7 8 9 10
8. Correcting Own Errors in a Timely Manner	1 2 3 4 5 6 7 8 9 10
9. Overall Attention and Situation Awareness	1 2 3 4 5 6 7 8 9 10

Prioritizing

10.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 mann	1 2 3 4 5 6 7 8 9 10
11. Preplanning Control Actions - scanning adjacent sectors to plan for inbound traffic	1 2 3 4 5 6 7 8 9 10
12. Handling Control Tasks for Several Aircraft - shifting control tasks between aircraft - avoiding delays in communications while thinking or planning control actions	1 2 3 4 5 6 7 8 9 10
13. Overall Prioritizing	1 2 3 4 5 6 7 8 9 10

Providing Control Information

14. Providing Essential ATC Information - providing mandatory services and advisories to pilots in a timely manner - exchanging essential information	1 2 3 4 5 6 7 8 9 10
15. Providing Additional ATC Information - providing additional services when workload is not a factor - exchanging additional information	1 2 3 4 5 6 7 8 9 10
16. Overall Providing Control Information	1 2 3 4 5 6 7 8 9 10

Participant # ___________ **Date** ___________

Technical Knowledge

17. Showing Knowledge of LOAs and SOPs - controlling traffic as depicted in current LOAs and SOPs - performing handoff procedures correctly	1 2 3 4 5 6 7 8 9 10
18. Showing Knowledge of Aircraft Capabilities and Limitations - avoiding clearances that are beyond aircraft performance parameters - recognizing the need for speed restrictions and wake turbulence separation	1 2 3 4 5 6 7 8 9 10
19. Overall Technical Knowledge	1 2 3 4 5 6 7 8 9 10

Voice Communications

20. Using Proper Phraseology - using words and phrases specified in JO 7110.65S - using proper phraseology that is appropriate for the situation - avoiding the use of excessive verbiage	1 2 3 4 5 6 7 8 9 10
21. Communicating Clearly and Efficiently - speaking at the proper volume and rate for pilots to understand - speaking fluently while scanning or performing other tasks - clearance delivery is complete, correct and timely - providing complete information in each clearance	1 2 3 4 5 6 7 8 9 10
22. Listening for pilot readbacks and requests - correcting pilot readback errors - processing requests correctly in a timely manner	1 2 3 4 5 6 7 8 9 10
23. Overall Voice Communication	1 2 3 4 5 6 7 8 9 10

Data Communications

24. Using Data Comm in appropriate situations - use in non-time-critical situations - equipped a/c only	1 2 3 4 5 6 7 8 9 10
25. Communicating Efficiently - choose efficient Data Comm message combinations - use complex Data Comm clearances appropriately - Data Comm clearance delivery is complete, correct, and timely - take full advantage of Data Comm message set	1 2 3 4 5 6 7 8 9 10

Participant # ___________ **Date** ___________

26. Monitoring Data Comm feedback and pilot replies and requests - Msg In, Msg Out, Msg Fail windows; DB field highlighting - processing requests correctly in a timely manner	1	2	3	4	5	6	7	8	9	10
27. Overall Data Comm	1	2	3	4	5	6	7	8	9	10

Conflict Probes

28. Using Conflict Probes in appropriate situations	1	2	3	4	5	6	7	8	9	10
29. Taking full advantage of Conflict Probes in a timely manner and efficiently	1	2	3	4	5	6	7	8	9	10
30. Overall Conflict Probes use	1	2	3	4	5	6	7	8	9	10

CRA

31. Using CRA in appropriate situations	1	2	3	4	5	6	7	8	9	10
32. Using CRA efficiently and correctly - taking full advantage of CRA	1	2	3	4	5	6	7	8	9	10
33. Overall CRA use	1	2	3	4	5	6	7	8	9	10

Teamwork

34. Task Allocation between R-and D-sides - Pre-coordinated plan (preferred); on-the-fly; none (least preferred) - Redundant vs. complementary action	1	2	3	4	5	6	7	8	9	10
35. Team Communication - Information sharing - Appropriate use: verbal, non-verbal (pointing, e.g.), graphical - Communicate working conventions (DB placement)	1	2	3	4	5	6	7	8	9	10
36. Team Situational Awareness - Anticipation of information need - Reminder of pending actions	1	2	3	4	5	6	7	8	9	10
37. Overall Team Work	1	2	3	4	5	6	7	8	9	10

Appendix F: Detailed Schedule

Detailed Schedule

Table F1 shows a 2-week schedule shared with the HA experiment using the later week for the CRA experiment. We had two sets of unique teams and rotated them across two sectors. We added our CRA specific information about air space managements such as coordination between sectors using the CRA interface. The participants received all necessary training on the first day, Monday. On the following days (Tuesday, Wednesday, and Thursday), they performed training scenarios and experimental scenarios.

Table F1. Available Team Configurations With a Swap of Teams Across Sectors

	Team Configuration					
	6 unique teams			Swapped across sectors		
Sector 1 – R-side	1	3	1	3	4	2
Sector 1 – D-side	2	1	4	4	2	3
Sector 2 – R-side	3	4	2	1	3	1
Sector 2 – D-side	4	2	3	2	1	4

We will assign each of the participants to an R- and D-side team. To increase our sample size we will create unique controller teams by rotating controllers among teams and also teams between sectors. Although it is possible to create six unique teams from four controllers and rotate them across the two sectors, our schedule can only accommodate twelve experimental scenarios. We therefore use four unique controller teams that will work each of the two sectors.

By swapping sectors but keeping the staffing experimental condition the same, that is, a team who worked on Sector 20 in the Baseline Condition will work on Sector 22 in the Baseline Condition, we can multiple the number of trials assuming the effect of the sector difference is minimal. However, we could not equip the EEG on the participants quickly enough to run all the position swaps. It took at least half an hour to equip the EEG on the participants. So, we decided to equip R and D-side controllers of Sector 22 who were close to the entrance of the simulation room. We used the position rotations of 1, 2, 3, and 4 for the first day, 3, 1, 4, and 2 for the second day, and 2, 3, 1, and 4 for the last day for all the participant groups.

In Table F2, the first number after P is the group number of 1 to 3 and the second number is the individual participant number of 1 to 4 because each group is composed of four participants, two teams of R- and D-sides. The resulting combinations of the team and sector for Group 1 is illustrated in Table F2. For Group 2, the participant number will be from P21, P22, P23, and P24.

Table F2. An Example of Experimental Conditions of Group 1

Position	Sector		Staff Condition	Test Run Numbers: Day
	20	22		
R	P11 (E & F)	P13 (EEG)	None (Baseline)	Test Run 1: Tuesday
D	P12 (E & F)	P14 (EEG)		
R	P11 (E & F)	P13 (EEG)	D-side	Test Run 2: Tuesday
D	P12 (E & F)	P14 (EEG)		
R	P11 (E & F)	P13 (EEG)	R- & D-sides	Test Run 3: Tuesday
D	P12 (E & F)	P14 (EEG)		
R	P13 (E & F)	P14 (EEG)	D-side	Test Run 4: Wednesday
D	P11 (E & F)	P12 (EEG)		
R	P13 (E & F)	P14 (EEG)	R- & D-sides	Test Run 5: Wednesday
D	P11 (E & F)	P12 (EEG)		
R	P13 (E & F)	P14 (EEG)	None (Baseline)	Test Run 6: Wednesday
D	P11 (E & F)	P12 (EEG)		
R	P12 (E & F)	P11 (EEG)	R- & D-sides	Test Run 7: Thursday
D	P13 (E & F)	P14 (EEG)		
R	P12 (E & F)	P11 (EEG)	None (Baseline)	Test Run 8: Thursday
D	P13 (E & F)	P14 (EEG)		
R	P12 (E & F)	P11 (EEG)	D-side	Test Run 9: Thursday
D	P13 (E & F)	P14 (EEG)		

Note. The characters in each parenthesis designate the equipment to wear: E = Eye movement, F = fNIR, and EEG = Electroencephalogram.

As shown above, not all combinations of experimental conditions could be randomized in a balanced way. Our first priority was to make sure that we randomized the conditions of CRA display conditions, that is, where the CRA would be displayed: no display (baseline), D-side, and both sides. All groups had three display conditions each test day, and these were counterbalanced across the groups (Table F3).

Table F3. Randomized Presentations of Three Display Modes on Different Test Days

Group	Day	First Run	Second Run	Third Run
1	**Tues**	None (Baseline)	None (Baseline)	None (Baseline)
	Wed	D-side	D-side	D-side
	Thurs	Both sides	Both sides	Both sides
2	**Tues**	D-side	D-side	D-side
	Wed	Both sides	Both sides	Both sides
	Thurs	None (Baseline)	None (Baseline)	None (Baseline)
3	**Tues**	Both sides	Both sides	Both sides
	Wed	None (Baseline)	None (Baseline)	None (Baseline)
	Thur	D-side	D-side	D-side

As we mentioned, we created two additional scenarios from each of the three training scenarios for test days. The following table (see Table F4) shows the randomized presentations of the nine training scenarios. The first digit of the number in each cell represents one of the three original scenarios. The second digit represents one of the three scenarios including the original scenario and two additional scenarios created from it. To balance out any possible effect due to the difference between the original three training scenarios on the test runs that will follow, we randomly assigned one of three scenarios from each set of the original scenarios to each test day. Table F5 shows the randomization of the test scenarios for the three groups of participants we had.

Table F4. Randomized Presentations of Nine Training Scenarios

Group	Tuesday			Wednesday			Thursday		
1	11	23	31	13	22	32	12	21	33
2	12	21	33	11	22	31	13	23	32
3	11	21	33	12	23	31	13	22	32

Table F5. Randomized Presentations of Nine Test Scenarios

Group 1	3	1	2	6	9	8	7	5	4
Group 2	4	6	1	9	8	5	7	3	2
Group 3	9	7	6	4	1	5	8	2	3

From the experimental condition we created the experimental run labels. Table F6 shows the run labels. For example, in G11234NP11R01, G1 means Group 1, the next four numbers correspond to the position numbers by sector in Table 4, N means No CRA displays used (Baseline condition), two numbers after P (practice) mean the practice scenario number shown in Table F4, and two numbers after R means the sequential run numbers starting Tuesday. T instead of P in other table cells means it is a Test scenario.

Table F6. Run Labels of Group 1, Group 2, and Group 3

G1, Tuesday	G1, Wednesday	G1, Thursday
G11234NP11R0 1	G13142DP13R07	G12314BP12R13
G11234NT03R02	G13142DT06R08	G12314BT07R14
G13142NP23R0 3	G12314DP22R09	G11234BP21R15
G13142NT01R04	G12314DT09R10	G11234BT05R16
G12314NP31R0 5	G11234DP32R11	G13142BP33R17
G12314NT02R06	G11234DT08R12	G13142BT04R18

G2, Tuesday	G2, Wednesday	G2, Thursday
G21234DP12R01	G23142BP11R07	G22314NP13R13
G21234DT04R02	G23142BT09R08	G22314NT07R14
G23142DP21R03	G22314BP22R09	G21234NP23R15
G23142DT06R04	G22314BT08R10	G21234NT03R16
G22314DP33R05	G21234BP31R11	G23142NP32R17
G22314DT01R06	G21234BT05R12	G23142NT02R18

G3, Tuesday	G3, Wednesday	G3, Thursday
G31234BP11R01	G33142NP12R07	G32314DP13R13
G31234BT09R02	G33142NT04R08	G32314DT08R14
G33142BP21R03	G32314NP23R09	G31234DP22R15
G33142BT07R04	G32314NT01R10	G31234DT02R16
G32314BP33R05	G31234NP31R11	G33142DP32R17
G32314BT06R06	G31234NT05R12	G33142DT03R18

Appendix G: Detailed Analysis of Data on Command Use

Flight Plan Readout commands were analyzed using a multivariate ANCOVA. There was no significant effect of CRA, *p* > .05. There was a significant effect of Time interval, *F*(3, 14) = 6.23, *p* = .007. Pairwise comparisons revealed that both Interval 1 (*M* = 9.50, *SE* = .89) and Interval 2 (*M* = 9.06, *SE* = .46) had more Flight plan readout commands than Interval 4 (*M* = 7.24, *SE* = .640), *p* = .007 and .014, respectively (see Figure G1).

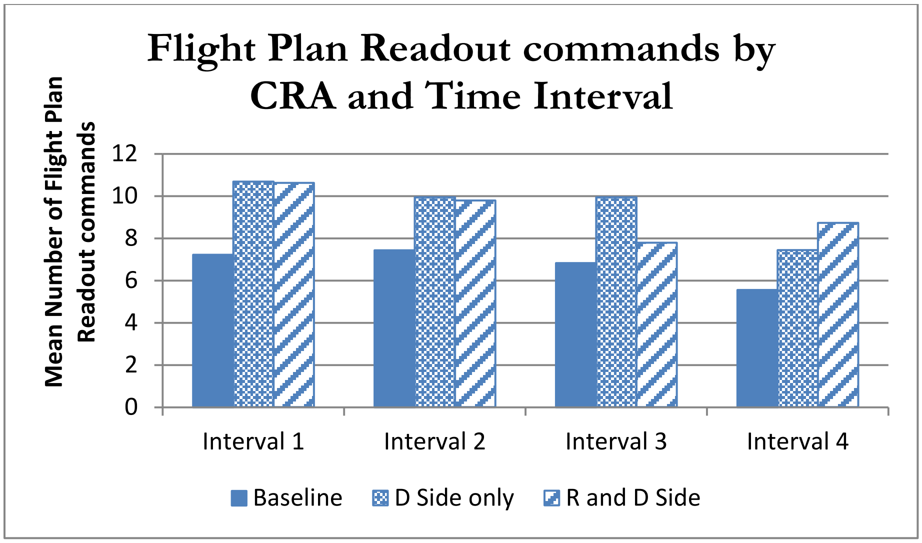


Figure G1. Flight Plan Readout Command frequencies by the CRA conditions and time intervals.

Confirm assigned level commands were analyzed using a multivariate ANCOVA. There was no significant effect of CRA, *p* > .05. There was a significant effect of Time interval, *F*(3, 14) = 4.11, *p* = .028. Pairwise comparisons revealed that both Interval 2 (*M* = 4.09, *SE* = .33) and Interval 3 (*M* = 4.24.06, *SE* = .36) had less commands than Interval 4 (*M* = 6.41, *SE* = .429), *p* = .007 and .014, respectively (see Figure G2).

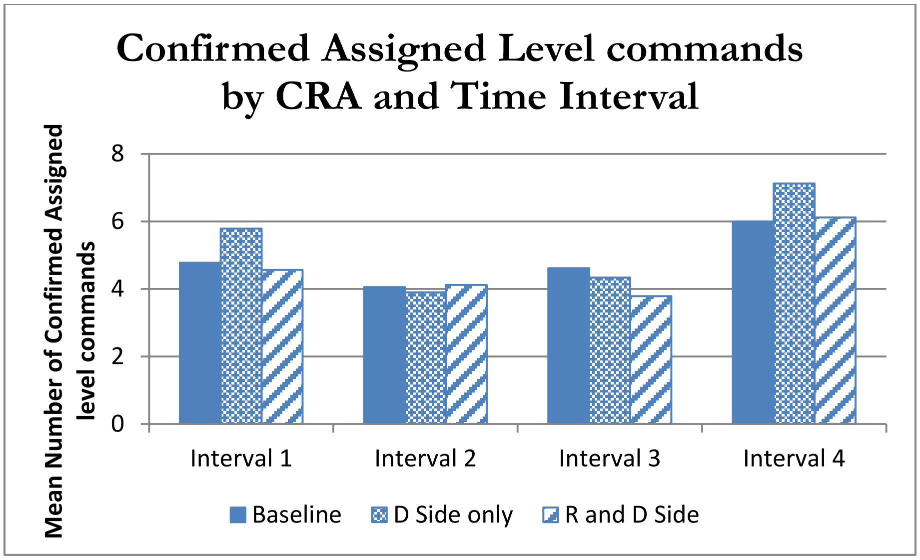


Figure G2. Confirmed Assigned Level Command frequencies by the CRA conditions and time intervals.

Route commands were analyzed using a multivariate ANCOVA. There was no significant effect of CRA, $p > .05$. There was a significant effect of Time interval, $F(3, 14) = 9.56$, $p = .001$. Pairwise comparisons revealed that Interval 4 ($M = 1.63$, $SE = .26$) had less commands than Interval 1 ($M = 3.94$, $SE = .61$), Interval 2 ($M = 3.35$, $SE = .63$) and Interval 3 ($M = 2.93$, $SE = .47$), $p = .000$, .009 and .002, respectively. Interval 3 ($M = 2.93$, $SE = .47$) had less commands given than Interval 1 ($M = 3.94$, $SE = .61$), $p = .008$. See Figure G3 for detailed data.

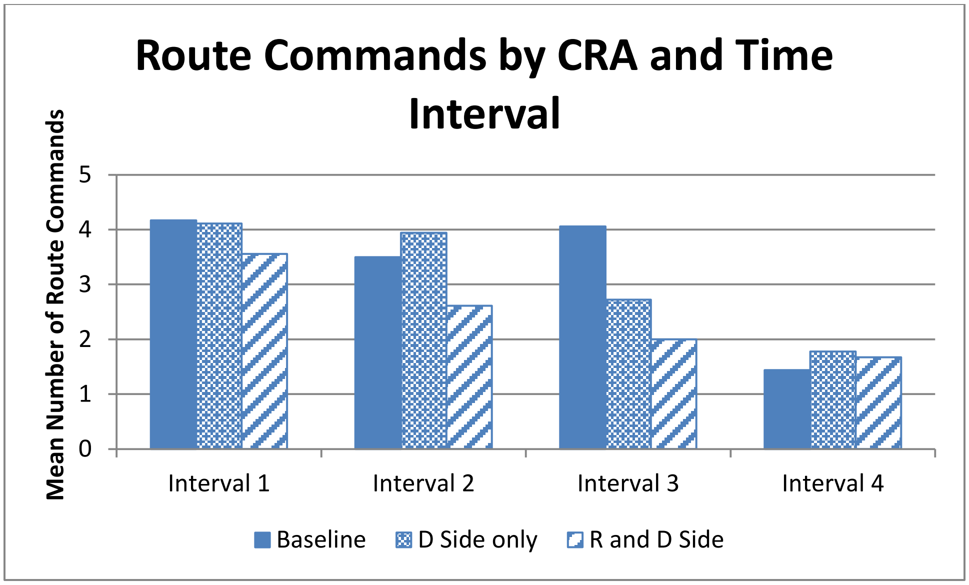


Figure G3. Route Command frequencies by the CRA conditions and time intervals.

Datablock movement commands were analyzed using a multivariate ANCOVA. There was no significant effect of CRA or Time interval, $p > .05$ for each effect (see Figure G4).

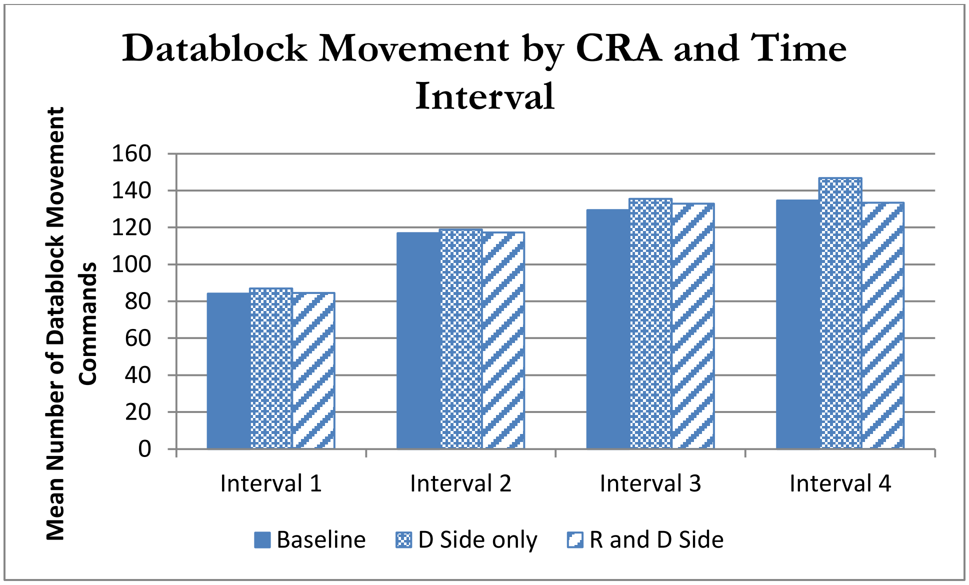


Figure G4. Datablock movement frequencies by the CRA conditions and time intervals.

Interim Altitude commands were analyzed using a multivariate ANCOVA. There was no significant effect of CRA, *p* > .05. There was a significant effect of Time interval, *F*(3, 14) = 3.39, *p* = .048. Pairwise comparisons revealed that Interval 1 (*M* = 2.46, *SE* = .26) had less commands than Interval 2 (*M* = 6.92, *SE* = .53), Interval 3 (*M* = 4.61, *SE* = .64) and Interval 4 (*M* = 7.15, *SE* = .55), *p* = .000, .002 and .000, respectively. Interval 3 (*M* = 4.61, *SE* = .64) had less commands given than Interval 2 (*M* = 6.92, *SE* = .53), *p* = .000. Interval 3 (*M* = 4.61, *SE* = .64) had less commands given than interval 4 (*M* = 7.15, *SE* = .55) (see Figure G5).

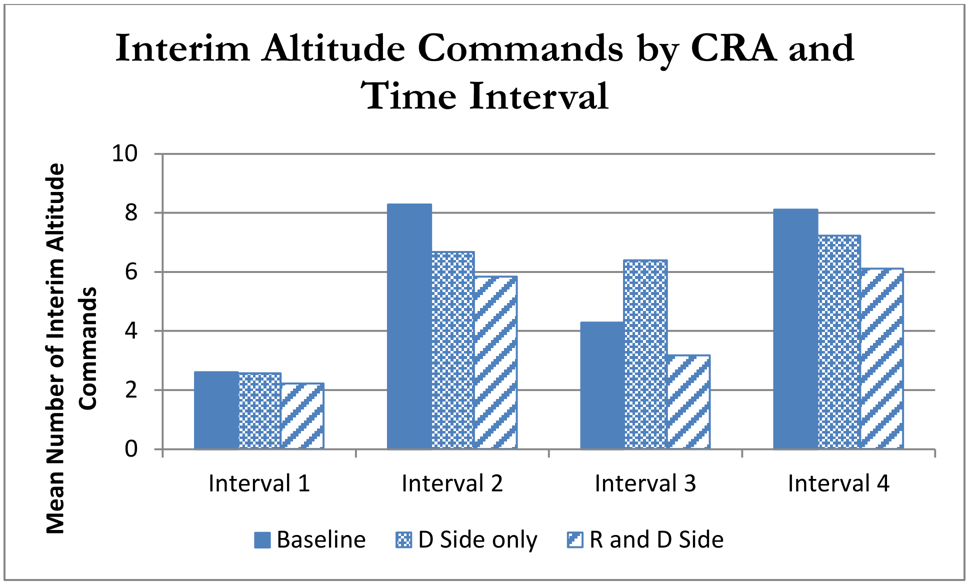


Figure G5. Interim Altitude Command frequencies by the CRA conditions and time intervals.

Assigned Altitude commands were analyzed using a multivariate ANCOVA. There was no significant effect of CRA or Time interval, *p* > .05 (see Figure G6).

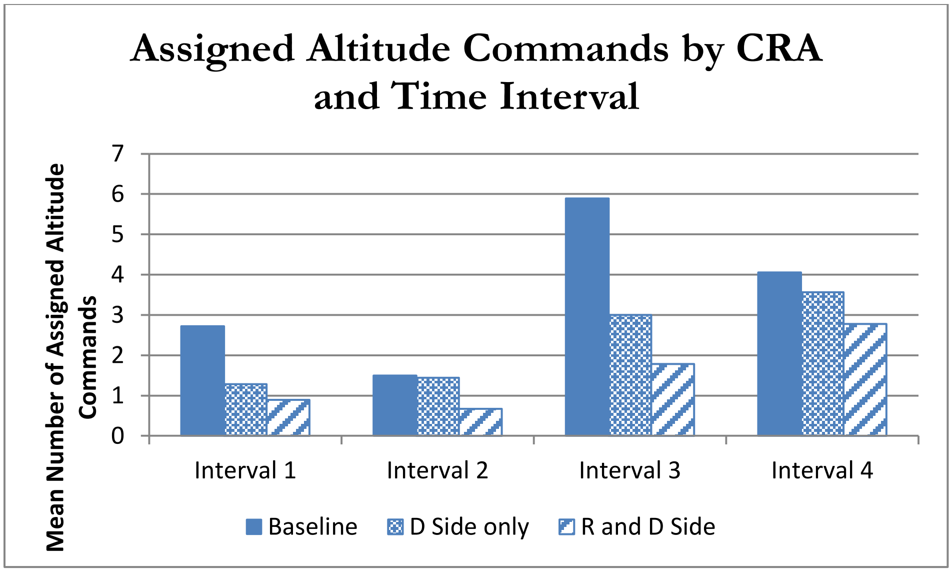


Figure G6. Assigned Altitude Command frequencies by the CRA conditions and time intervals.

Halo commands were analyzed using a multivariate ANCOVA. There was no significant effect of CRA, $p > .05$. There was a significant effect of Time interval, $F(3, 14) = 7.54$, $p = .003$. Pairwise comparisons revealed that Interval 1 ($M = 1.61$, $SE = .31$) had less commands than Interval 2 ($M = 2.56$, $SE = .43$), $p = .032$ (see Figure G7).

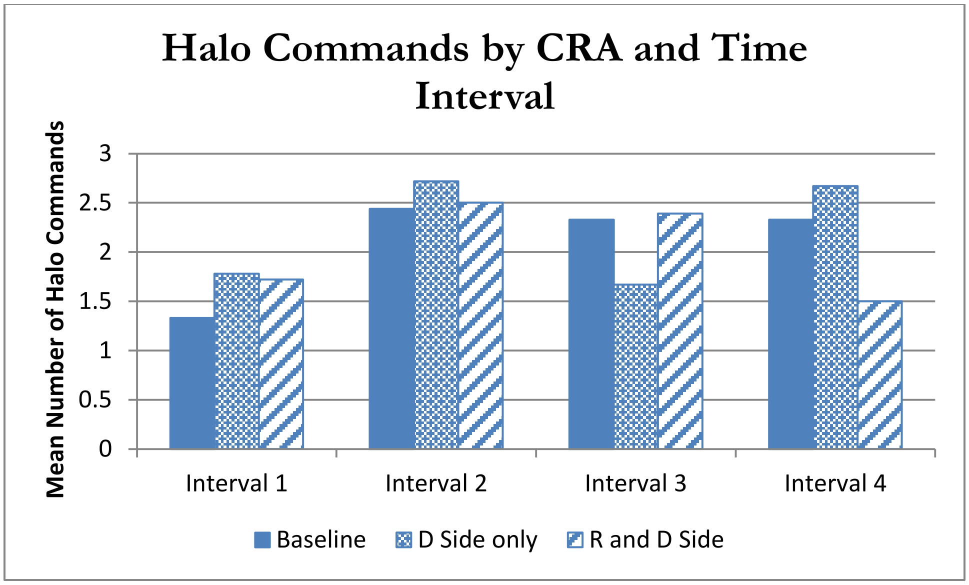


Figure G7. Halo Command frequencies by the CRA conditions and time intervals.

We tested the frequencies of Select command with a multivariate ANCOVA. The Select command is generated when a controller clicks the aircraft (i.e., target) symbol by the middle button to drop a full datablock or bring up the full datablock from a limited datablock. There was no significant effect of CRA or Time interval, $p > .05$ (see Figure G8).

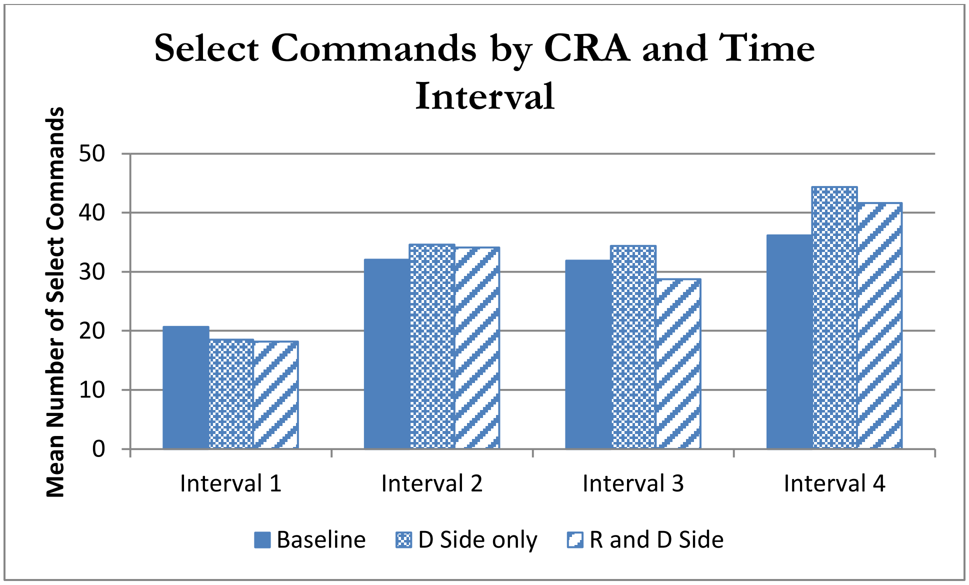


Figure G8. Select Command frequencies by the CRA conditions and time intervals.

Route toggle commands were analyzed using a multivariate ANCOVA. There was no significant effect of CRA or Time interval, *p* > .05 (see Figure G9).

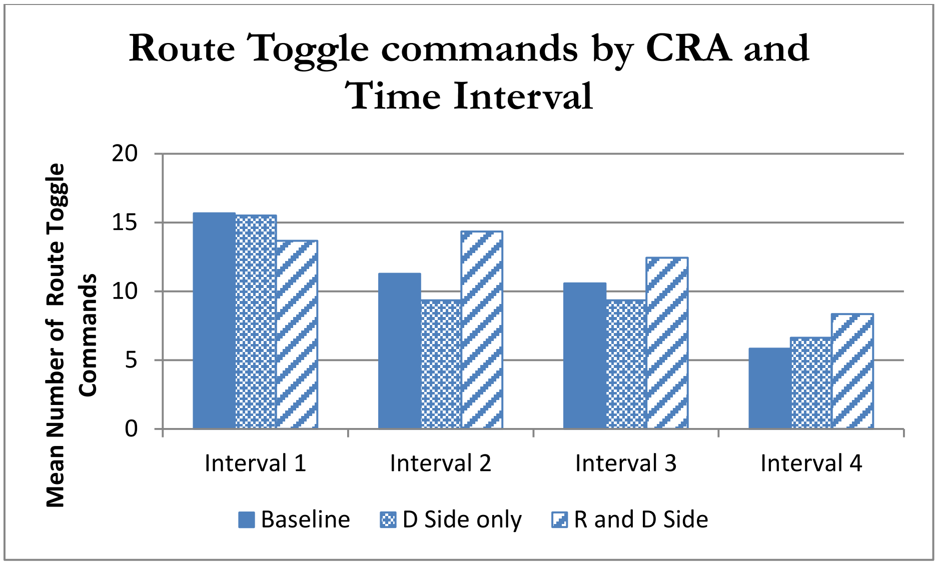


Figure G9. Route Toggle Command frequencies by the CRA conditions and time intervals.

Halo commands were analyzed using a multivariate ANCOVA. There was no significant effect of CRA, *p* > .05 (see Figure G10). There was a significant effect of Time interval, *F*(3, 14) = 5.74, *p* = .009. Pairwise comparisons revealed that Interval 1 (*M* = 1.83, *SE* = .14) had less commands than Interval 2 (*M* = 4.46, *SE* = .16), Interval 3 (*M* = 3.22, *SE* = .15) and Interval 4 (*M* = 3.94, *SE* = .13), *p* = .000 for each comparison. Interval 3 (*M* = 3.22, *SE* = .15) and Interval 4 (*M* = 3.94, *SE* = .13) had less commands than Interval 2 (*M* = 4.46, *SE* = .16), *p* = .000 and .033, respectively. Interval 3 (*M* = 3.22, *SE* = .15) had less commands than Interval 4 (*M* = 3.94, *SE* = .13), *p* = .013.

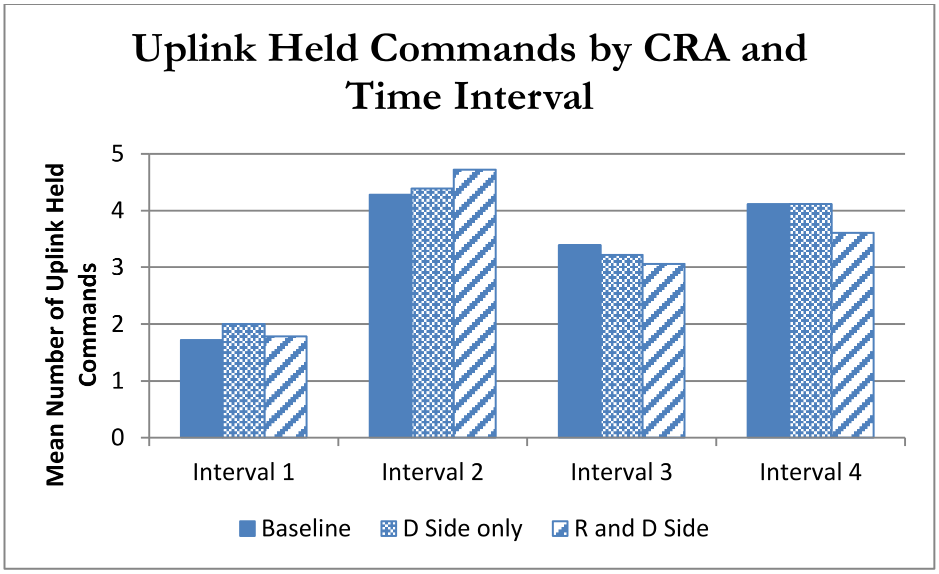


Figure G10. Uplink Held Command frequencies by the CRA conditions and time intervals.

Appendix H: Results of Pilot Commands

As you can see in the graphs below, for the most part, the data show similar trends between levels of CRA. If it looks like there are big differences in the graphs, it's because of the scaling of the graph, and not real statistical differences (see Figures H1 through H5).

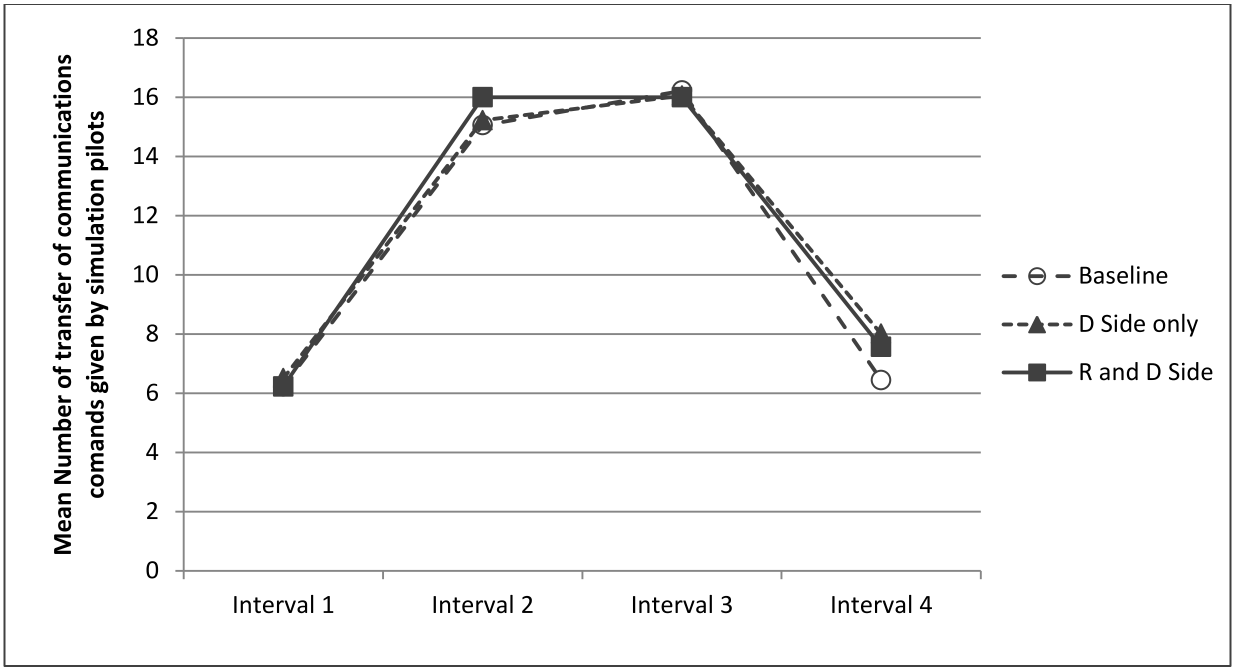


Figure H1. Transfer Communication frequencies by the CRA conditions and time intervals.

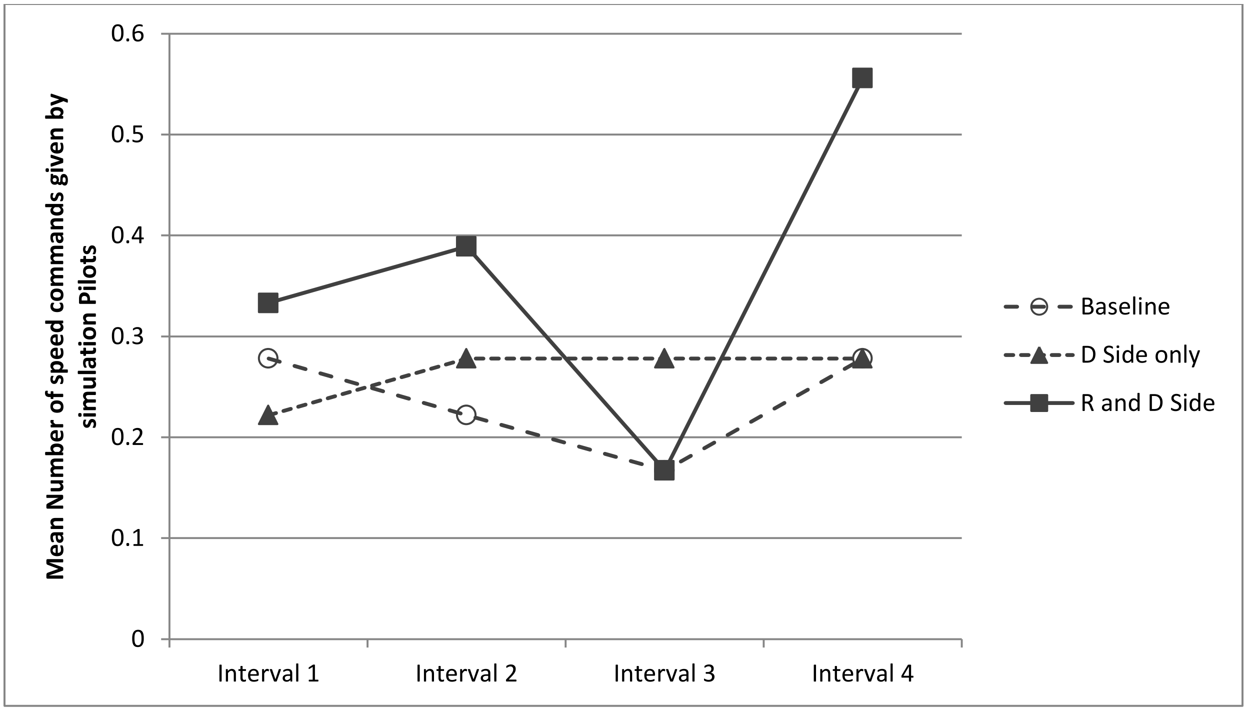


Figure H2. Speed Command frequencies by the CRA conditions and time intervals.

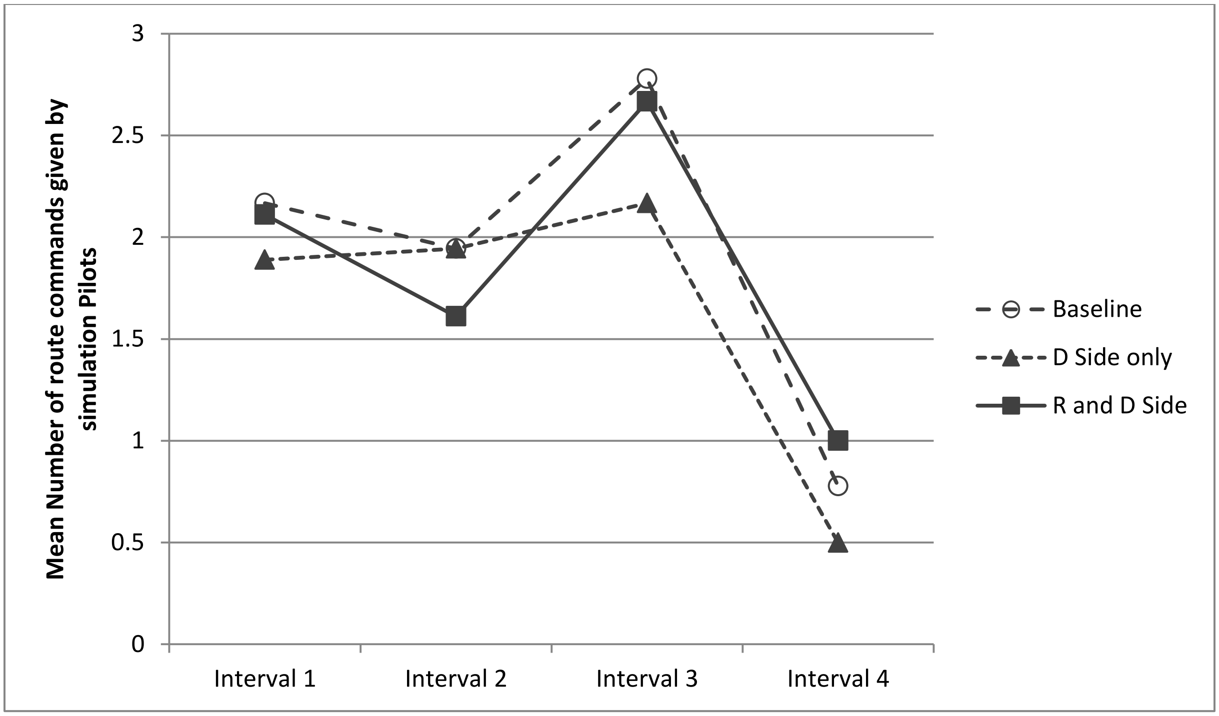


Figure H3. Route Commands frequencies by the CRA conditions and time intervals.

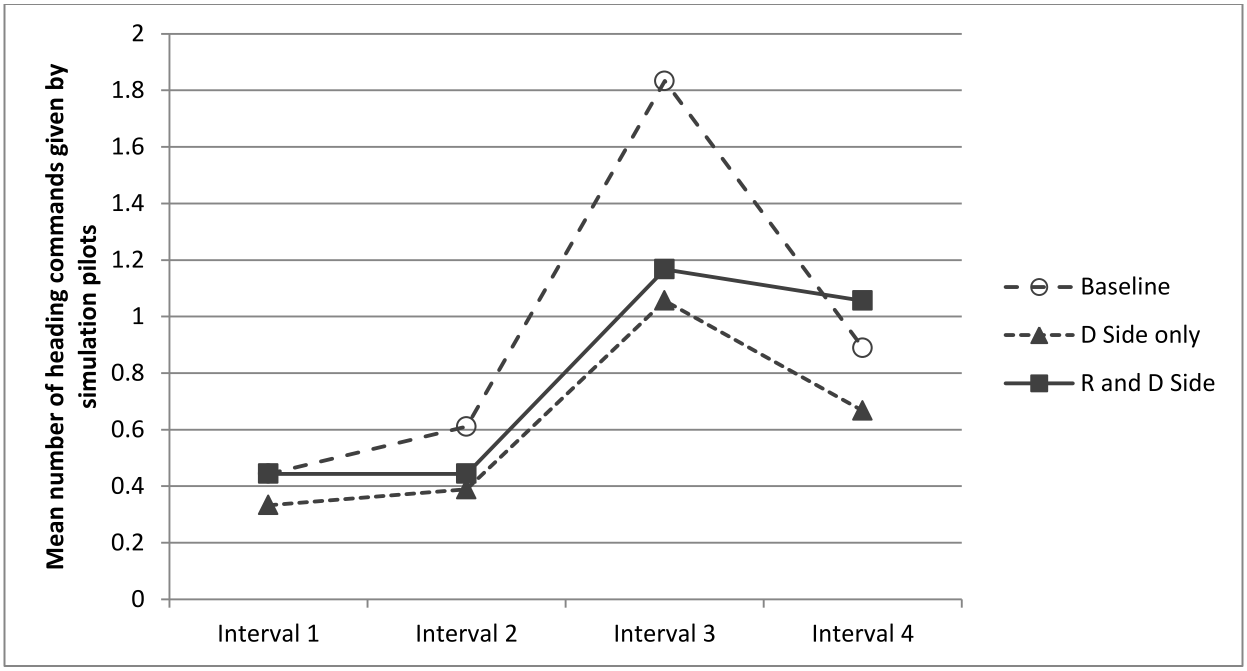


Figure H4. Heading Command frequencies by the CRA conditions and time intervals.

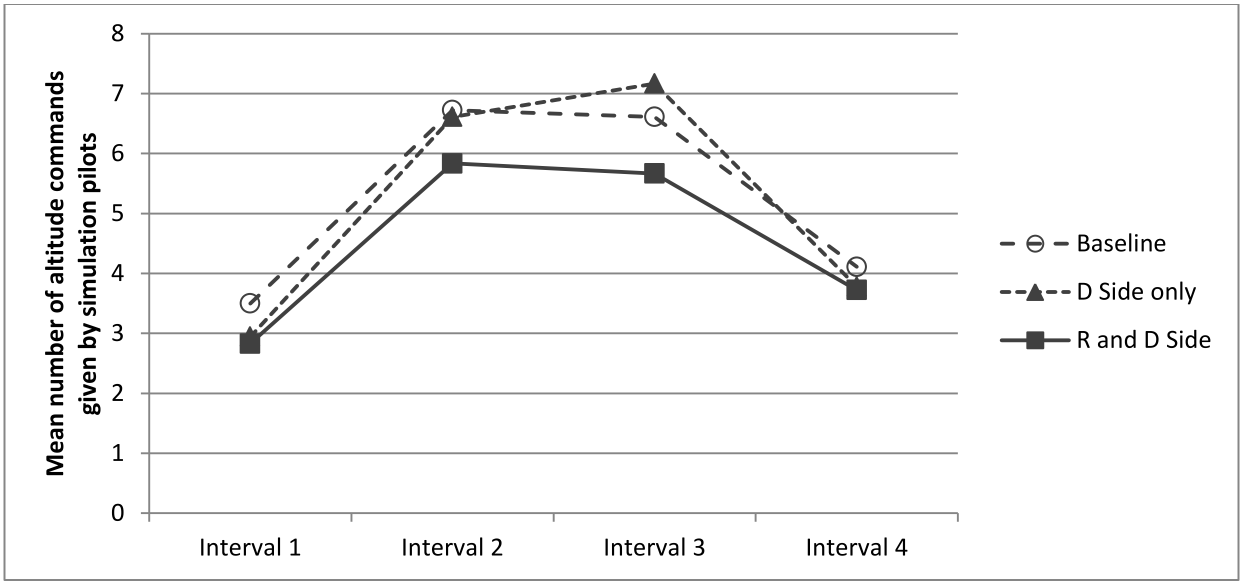


Figure H5. Altitude Command frequencies by the CRA conditions and time intervals.

Appendix I: Narratives of Operational Errors

SimTime	RunLabel	Sector	# ACFT	ACID’s	Narrative
34:12.3	TG12314NTt2.030.filteredR16	20	26	FFT1861 / SQC364	At 29:38, SQC364 was issued, via Data Comm, a climb from FL370 to FL390. At 30:36, approximately 1-minute later, SQC364 begins climb, Mode-C indicating FL372. At 32:36, SQC364 Mode-C indicates FL383, rate of climb approximately 650 feet per minute. At 33:00, SQC364 (@ FL384) and FFT1861 (@ FL380) are head-on and separated by 10 nmi. Up to this point, neither CRA nor Conflict Alert has given any indication that these two aircraft are in conflict. At 33:24, Loss of Separation occurs. At 33:32, Conflict Alert activates. **Controller error, with an alert system failure, and unreasonable climb rate by SQC364 as contributing factors.**
35:48.3	TG12314NTt2.030.filteredR16	20	27	JBU2358 / SQC364	At 31:23, JBU2358 was issued, via Data Comm, a climb from FL380 to FL400. At 32:03, approximately 30 seconds later, JBU2358 begins climb, Mode-C indicating FL381. At 33:56, Conflict Alert activates, SQC364 at FL390 and JBU2358 at FL386. At 34:09, controller turns JBU2358 twenty degrees left. At 35:15, separation is lost, SQC364 at FL390 and JBU2358 at FL389, JBU2358 climb rate at less than 300 feet per minute. Controller failed to issue a turn(s) that would have maintained lateral separation. **Controller error.**
33:00.3	TG13142DTt6.030.filteredR08	20	27	FFT1299 / N2696	At 22:17, controller issues climb to FL330 to N2696. At 22:44, controller issues a climb to N2906 to FL430. The clearance is rogered by similar sounding call sign N2696. At 25:03, controller re-issues climb to FL430 to N2906. N2696 again responds to the clearance “…out of FL255 on way up to FL430.” At 31:11, N2906 Mode-C indicates FL331. At 31:25, Conflict Alert activates between N2696 and FFT1299. At 31:55, separation lost between N2696 (@ FL337) and FFT1299 (@FL340), and less than 5 nmi separation. **Controller error.**
33:12.3	TG13142DTt6.030.filteredR08	20	26	FDX2323 / N2696	See narrative in previous scenario. At 31:34, Conflict Alert activates between N2696 and FDX2323. At 32:35, separation is lost between N2696 (@ FL345) and FDX2323 (@ FL350), and less than 5 nmi lateral separation. **Controller Error.**
34:48.0	TG13142DTt6.030.filteredR08	20	28	RPA1359 / VRD532	At 32:48, controller dropped datablock on VRD532, still within the protected airspace of Sector 20. At 33:05, controller issues descent to FL240 to RPA1359. At 33:57, Conflict Alert activates between RPA1359 and VRD532. At 34:13, targets merge, RPA1359 @ FL347 and VRD532 @ FL340, separation is lost. **Controller error.**

SimTime	RunLabel	Sector	# ACFT	ACID's	Narrative
20:24.4	TG22314RTt8.030R18	20	18	N3021 / SWA2100	At 18:36, N3021 issued descent to FL330. At 19:27 Sim Pilot requests say altitude for N3021. At 19:36, controller re-iterates FL330. At 19:59, N3021 Mode-C indicates FL328. Conflict alert activates at 20:10, N3021 Mode-C indicates FL329. **SIM Pilot error.**
36:00.3	TG23142NTt2.030R10	20	20	JBU2358 / SQC364	At 28:40, SQC364 was issued a climb, via Data Comm, from FL370 to FL390.At 30:57, JBU2358 was issued a climb, via Data Comm, from FL380 to FL400. At 32:51, SQC364 levels at FL390. JBU2358 Mode-C indicates FL385, and the aircraft are separated by approximately 38 nmi. At 33:59, Conflict Alert activates, SQC364 at FL390 and JBU2358 indicates climbing thru FL388, and laterally separated by 23 nmi. At 35:15, SQC364 is level at FL390 and JBU2358 reports climbing thru FL391, separation is lost as aircraft are separated laterally by less than 5nmi. **Controller error.**
47:48.2	TG23142RTt9.030R08	20	28	DAL1416 / SWA286	At 46:16, DAL1416 is given descent to FL300. Sim Pilot erroneously climbs DAL1416. Conflict Alert activates at 46:51. **SIM Pilot error.**
12:12.3	TG31234NTt5.030R04	20	15	AWE1978 / SWA1340	Reviewed from Sector 20 D-side replay. At 07:19 CRA alerts between AWE1978 and SWA1340, both aircraft at FL320 on crossing courses. At 08:46, controller turns AWE1978 15 degrees left for traffic. At 10:11, Conflict Alert activates. At 11:33, separation is lost. Controller failed to apply sufficient radar vectors to maintain lateral separation. **Controller error.**
16:36.3	TG32314RTt6.030R16	20	16	GJS2619 / N1085	N1085 climbing to FL370. Controller recognizes traffic, GJS2619 at FL340, and turns GJS2619 ten degrees right. Ten degree turn was not enough to insure separation, separation was lost. **Controller error**.
41:12.6	TG33142DTt3.030R10	20	27	UAL2011 / UAL620	Controller recognized aircraft were on converging courses at the same altitude. At 31:13, controller issues UAL620 direct HEC (Hector) to achieve parallel or diverging courses. At 37:14, Conflict Alert activates. At 38:17, controller re-clears UAL620 direct HEC, turn 20 degrees right until able direct. At 40:01, controller issues frequency change to UAL620. At 40:25, UAL620 turns left in confliction with UAL2011. **Either SIM Pilot error or TGF error processing route of UAL620.**

SimTime	**RunLabel**	**Sector**	**# ACFT**	**ACID's**	**Narrative**
49:12.6	TG33142DTt3.030R10	20	30	DAL293 / UAL2978	At 46:29, Conflict Alert activates between UAL2978, altitude FL320, ground speed 575 kts, behind DAL293, altitude FL320, and ground speed 445 kts. At 46:46 controller issues UAL2978 descent to FL300. At 48:59, 2 minutes and 13 seconds later, UAL2978 Mode-C indicates descending through FL310, a descent rate of less than 500 feet per minute**. System error, unreasonable climb rate.**