DOT/FAA/TC-08/14, II

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

Future En Route Workstation Study (FEWS II): Part 1 – Automation Integration Research

Ben Willems, Human Factors Team – Atlantic City, ATO-P
Sehchang Hah, Human Factors Team – Atlantic City, ATO-P

October 2008

Technical Report

U.S. Department of Transportation
Federal Aviation Administration

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/TC-08/14, II		

4. Title and Subtitle	5. Report Date
Future En Route Workstation Study (FEWS II): Part 1 – Automation Integration Research	October 2008
	6. Performing Organization Code AJP-6110
7. Author(s) Ben Willems, Human Factors Team – Atlantic City, ATO-P Sehchang Hah, Human Factors Team – Atlantic City, ATO-P	8. Performing Organization Report No. DOT/FAA/TC- TC-08/14, II
9. Performing Organization Name and Address Federal Aviation Administration Human Factors Team – Atlantic City, ATO-P William J. Hughes Technical Center Atlantic City International Airport, NJ 08405	10. Work Unit No. (TRAIS) 11. Contract or Grant No.
12. Sponsoring Agency Name and Address Federal Aviation Administration Human Factors Research and Engineering Group 800 Independence Avenue, S.W. Washington, DC 20591	13. Type of Report and Period Covered Technical Report 14. Sponsoring Agency Code ATO-P

15. Supplementary Notes

16. Abstract

The expected increase in air traffic makes it essential to investigate how to develop tools to assist controllers best without sacrificing safety, efficiency, or workload. The Future En Route Workstation Study (FEWS I) concept takes an integrated approach to controller workstation development that applies best human factors principles to the design of the computer-human interface to promote more efficient and effective controller interaction. The current simulation assessed controller performance and workload using the Display System Replacement (DSR), the En Route Automation Modernization (ERAM), and the modified FEWS interface (FEWS II). We measured controller eye movements, situation awareness, ratings about system features and functions, workload, performance, and measures of aircraft efficiency at air traffic levels corresponding to levels anticipated for 2015 and beyond. The FEWS II configuration resulted in fewer data entries than DSR or ERAM, although it did not result in a reduction in workload. Controllers had a significantly lower number of controller deviations when using the FEWS II configuration with a two-person sector and Data Communications available. We showed a reduction in controller workload when Data Communications was available in a two-person sector. Regression analyses showed that at the same workload level, controllers could handle more aircraft when they worked as a team and Data Communications was available.

17. Key Words	18. Distribution Statement
Air Traffic Control Controller-Pilot Data Link Communications Display System Replacement En Route Automation Modernization Future En Route Workstation	This document is available to the public through the National Technical Information Service, Springfield, Virginia, 22161. A copy is retained for reference at the William J. Hughes Technical Center Library.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	64	

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

Table of Contents

List of Illustrations

Tables Page

Executive Summary

The Federal Aviation Administration (FAA) has established a Target System Description for the National Airspace System (NAS) in 2015. The expected increase in traffic makes it essential to investigate how to develop tools to assist controllers best without sacrificing safe and efficient throughput of aircraft or producing information overload or excessive workload.

The existing En Route Air Traffic Control System has added new automation tools through a stove-pipe approach, attaching new tools to existing hardware and software as separate entities that controllers access through objects added to existing or auxiliary displays. The Future En Route Workstation Study (FEWS) concept takes an integrated approach to controller workstation development that applies *best human factors principles* to the design of the computer-human interface to promote more efficient and effective controller interaction. We conducted an initial study of the FEWS concept (FEWS I) to evaluate the new integrated design concept relative to Display System Replacement (DSR) and to identify any additional interface modifications that may be necessary to optimize the new FEWS II design. This study is the second in a series of efforts investigating (a) the integration of automation functions into the controller workstation and (b) the effect on the controllers' workload and their ability to manage increasingly higher levels of air traffic.

The current simulation assessed controller performance and workload using the FEWS II, the DSR, and the En Route Automation Modernization (ERAM) system; the FAA plans to replace DSR with ERAM by the year 2010. We measured controller eye movements, situation awareness, ratings about system features and functions, workload, performance, and measures of aircraft efficiency under the three system configurations at air traffic levels corresponding to levels anticipated for 2015 and beyond.

Researchers conducted the study at the FAA William J. Hughes Technical Center, Research Development and Human Factors Laboratory. We used the Distributed Environment for Simulation, Rapid Engineering, and Experimentation to emulate features and functions of the alternative workstations, the FAA Target Generation Facility to transmit information from simulation pilot workstations to the controller display, and the Center TRACON Automation System to transmit and receive additional system data when needed. We also used generic Air Route Traffic Control Center sector airspace with instrument flight rules.

We report on the selected results of experiments conducted on high-fidelity emulations of DSR, ERAM, and FEWS II workstations configured for a one-person or a two-person sector control with or without Data Communications availability. The FEWS II workstation configuration resulted in fewer data entries, although it did not result in a reduction in workload. The results indicate that the FEWS II workstation with a two-person sector and Data Communications available had a significantly lower number of controller deviations. We showed a reduction in controller workload when Data Communications was available in a two-person sector, but this reduction was not present when controllers worked traffic under the one-person sector conditions. Regression analyses showed that at the same workload level, controllers could handle more aircraft when they worked as a team using Data Communications instead of Voice Communications only. When controllers worked the one-person sector, they were not able to handle more aircraft with the addition of Data Communications.

1. INTRODUCTION

Historically, changes in the National Airspace System (NAS) have occurred in an evolutionary manner. Currently, the NAS is undergoing more rapid changes to adapt to greatly increasing levels of air traffic. Within the last decade, we have witnessed the introduction of the en route Display System Replacement (DSR), the User Request Evaluation Tool (URET), the Standard Terminal Automation Replacement System (STARS) in the Terminal Radar Approach Control (TRACON) environment, and the Domestic Reduced Vertical Separation Minimum program as well as several other automation systems. Within the next 5 to 10 years, we will see reduced lateral separation and changes to existing Air Traffic Control (ATC) systems and procedures. As a performance-based organization that provides services to its customers (airlines, passengers, and companies that rely on air transportation), the Federal Aviation Administration (FAA) Air Traffic Organization will continue to see faster changes in the NAS to meet increasingly challenging objectives and customer needs.

As the FAA continues to modernize the NAS, we need to evaluate how this will affect the controller workforce to ensure the realization of anticipated system benefits. Many proposed system changes may significantly alter the role of the controller, potentially affecting system safety and efficiency. Some changes may result in a less active role for the controller. The controller will spend an increased proportion of his/her time monitoring systems rather than directly controlling traffic. Consequently, the change from a more active to a more passive level of involvement with traffic can impair controller situation awareness (SA) and reduce system safety (e.g., Endsley, Sollenberger, Nakata, & Stein, 2000).

Additional system changes may also alter basic user interaction behaviors. Proposed automation systems, such as Data Communications, involve a shift from strictly voice-based communication to more computer-based interaction. Data Communications will enable transfer of text messages between ground-based systems and aircraft. Overall, the NAS modernization efforts have the potential to change the nature of the controllers' job in a revolutionary manner. Therefore, it is critical that we design new systems and tools using *best human factors principles* to ensure that users will be able to work with these technologies as effectively as possible.

1.1 Background

The existing En Route ATC System replaced the previous generation ATC workstations that included 19-in. (48.26 cm) circular, monochrome monitors, with workstations using high-resolution 29-in. (73.66 cm) color displays. On-screen interface elements replaced the control knobs and switches from the prior system. As DSR evolved, it started to include new tools outside of its original system structure. These new tools included URET and the Traffic Management Advisor (TMA). URET is a tool that assists controllers to detect potential conflicts and to assess whether proposed aircraft trajectories will be conflict free. TMA is a tool that displays aircraft metering data provided by the Center TRACON Automation System (CTAS).

The Next Generation En Route System, the En Route Automation Modernization (ERAM), will integrate some of the automation functions directly into the system architecture. ERAM will incorporate advanced flight data processing capabilities, enabling controllers from adjacent sectors and Air Route Traffic Control Centers (ARTCCs) to have greater access to flight information for aircraft currently outside of their control. However, these capabilities may affect traffic flow in their sectors. The introduction of ERAM will also include a modified Computer-Human Interface (CHI) that will consist of more windows-based operability, with additional list view elements that the controller can detach from larger lists and then place elsewhere on the display. More data block options will also be available. To determine whether these interface changes adequately meet user and system performance requirements, we need to evaluate the effect of the interface changes on the accuracy and speed with which controllers access and manage information.

1.2 Evolution of NAS Operation

The evolving air traffic system incorporates some new tools and functions that have the potential to allow the NAS to operate more strategically, using trajectory-based rather than sector-based control. Trajectory-based control allows aircraft to fly along the most fuel-efficient flight paths and, as a result, is appealing from an economic standpoint. The airlines would incur lower fuel costs and aircraft would potentially be able to reach their destinations faster. Prior research has suggested that the current sector-based system would have to change drastically from tactical to strategic control (e.g., Graham, Marsden, Pichancourt, & Dowling, 2000). However, other research (Willems & Heiney, 2002) has shown that an Airspace Coordinator is able to maintain a sector-based system in which we leave the responsibility of carrying out tactical trajectory-based requests to the sector controllers. The NAS, as a whole, would benefit from the advantages offered by a trajectory-based system, which allows controllers to continue to work in a familiar, tactical manner. This would permit us to leverage the extensive experience that our workforce has with the sector concept, as controllers would not need to alter their perspective on the aircraft.

The introduction of URET and TMA indicates that some changes in the NAS have already moved towards a more trajectory-based approach. Presently, the TMA estimates the time that aircraft will arrive at a metering fix and airport. Based on the airport acceptance rates, TMA calculates how much time an aircraft must gain or lose to fit into a sequence to meet its scheduled time of arrival. The sector controller receives data via the TMA tool and implements control actions to bring the delay time within an acceptable level; the Downstream sectors and the TRACON can absorb any small remaining delays. TMA uses a trajectory-based approach that has the potential to increase efficiency by distributing aircraft delays across sectors instead of requiring the sector with the metering fix to handle the full delay absorption. Single-center TMA allows a center to use this approach across sectors. Multi-center TMA, once implemented, will use the trajectory-based approach across ARTCCs. Similarly, URET alerts controllers within a sector that a potential loss of separation exists and requests that the sector in which the conflict may occur attempt to resolve it. URET uses a trajectory-based approach, probing each aircraft trajectory for potential conflicts with other trajectories or with airspace. Aircraft involved in a potential conflict may not have entered the sector that will receive the potential conflict. In that case, controllers may need to request changes to aircraft in adjacent sectors.

The current automation systems in the NAS are mostly passive alerting systems, because they indicate status or a potential condition but leave the controller responsible for taking action to meet a constraint (e.g., reduce a delay) or to avoid a problem (e.g., prevent a predicted loss of separation), and so forth. The National Aeronautics and Space Administration (NASA) and the MITRE - Center for Advanced Aviation System Development (CAASD) have explored capabilities that would provide controllers with resolutions to the issues that the automation tools detect, rather than just alert them to a situation. NASA's En Route Descent Advisor is the active sibling of TMA, whereas CAASD's Problem Analysis Resolution and Ranking tool is the active sibling of URET. Although available, the FAA has not implemented these systems because concerns remain as to when, where, and how to present these active decision support system advisories. Even in the more passive counterpart systems, there have been questions about the data to present. In TMA, issues have surfaced about whether to display the full delay needed to meet the metering restriction at the airport or to display only the portion of the delay for which the controller is responsible. Similarly, for URET, implementation issues have risen about whether to display conflict information to the controller who will receive the conflict (as currently implemented) or to the controller who actually has the aircraft under control.

1.2.1 Future En Route Workstation I

The first Future En Route Workstation Study (FEWS I) concept (Willems, Hah, & Phillips, 2008) attempted to enable controllers to interact with the projected capabilities of the NAS in 2015 while maintaining or increasing existing levels of safety. The overall objective was (a) to provide aircraft and airspace data to the controller when and where needed and (b) to present the data in a format that was easily accessible and interpretable. In many respects, this approach is similar to that discussed by Andriole and Adelman (1995) who set out to identify, define, and validate a set of requirements to design, develop, prototype, and evaluate changes to the interface and interaction routines that weapons directors experience when managing military air traffic. They derived their requirements from conventional task analyses, user interviews, and *best human factors principles*. We have reviewed many of the existing task analyses of ATC (e.g., Ammerman et al., 1987; Dittmann, Kallus, & Van Damme, 2000) to create a workstation concept that will enable controllers to handle their varied responsibilities in a more efficient manner. The following list contains the main human factors principles (e.g., Mejdal, McCauley, & Beringer, 2001) that we used to develop the FEWS I concept:

- Limit or eliminate the number of windows and lists, or make them optional.
- Provide access to information within a minimal number of steps.
- Present information when and where needed.
- Prevent time sharing of information.
- Keep information presentation consistent across display windows.
- Connect information displays related to the same object.
- Place related information close together.
- Use consistent layout formats to assist in the development of automated human behaviors.

We also reviewed existing concepts related to the evolution of the en route sector and the data available in the NAS and existing automation tools. We will discuss a concept of operation for the future en route sector and the suggested changes to the workstation to support that concept. In our simulation, controllers will use different workstation concepts to control traffic in the environment anticipated for the NAS in 2015. The Target System Description describes a snapshot of air traffic management in 2015 in terms of proposed technological capabilities and the effect on airspace design and procedures. The outcome of this study will provide guidance on how to access and use these capabilities at the sector, most efficiently, from a human factors perspective.

The FEWS I project involved several phases. First, we created a CHI Redesign Team to discuss potential enhancements to the Future En Route Workstation and prototyped those enhancements as the initial interface. We then recruited controllers to participate in a simulation that applied formal experimental protocols and measurement techniques to determine whether, and *how*, the concepts implemented in the FEWS I experiment affected controller ability to handle traffic safely and efficiently, relative to DSR. The data from the FEWS I test phase, which researchers completed in July 2005, have guided further modifications to the current FEWS II interface. The current phase will investigate how these interface enhancements affect the controllers' ability to handle traffic safely and efficiently.

During the CHI Redesign Team meetings, we learned from controllers that we could improve the system by modifying routine tasks (e.g., moving data blocks) that have little to do with the direct management of traffic. Recent simulations conducted at the FAA William J. Hughes Technical Center (WJHTC) Research Development and Human Factors Laboratory (RDHFL) indicated that about 25% of controller interactions with the system involved moving data blocks (Willems & Heiney, 2002; Willems, Heiney, & Sollenberger, 2002). One focus of our efforts, therefore, was to help controllers conduct these "housekeeping" tasks more efficiently so that they could devote their resources to managing traffic.

We focused on implementing features that apply across the interface and enhance the efficiency of multiple controller tasks. Our Subject Matter Experts (SMEs) agreed that the NAS already provides the data needed to manage traffic efficiently and that some design features convey these data effectively. However, systems in the NAS do not always apply these features consistently. For example, the implementation of a highlighting feature can enable controllers to group a subset of aircraft quickly, helping to achieve the objective of connecting related information. CTAS incorporates this concept by simultaneously highlighting an aircraft in the metering timeline as well as on the situation display if the user emphasizes either one. The ARTCCs also use this concept as part of the quick-look feature to allow controllers to determine which aircraft are under control of a particular sector. Controllers also emphasize flow-sectors to temporarily indicate which aircraft are traveling to a particular airport. Despite its usefulness, the "emphasis by shared characteristic" principle is not available throughout the system. Our focus was, therefore, to derive a set of features and design concepts like this one, apply them across the interface, and test them empirically to evaluate their usefulness.

During the FEWS I experiment, one of the open issues involved the display changes: Controllers pointed out that they had a hard time finding voice only aircraft when a pilot called in on initial contact. During preparation of FEWS I, we had suggested to our SMEs that we should indicate which controller within a team took action and indicate as well that the system may have taken action for the sector. The SMEs in the FEWS I experiment were of the opinion that the system did not need that feature because the team knew what was going on in the sector. The feedback from the FEWS I participants suggested that the introduction of this feature was necessary, especially when introducing an automatic handoff acceptance feature (Willems et al., 2008).

The features introduced in the FEWS I simulation included: Focus, Emphasis, Three-tier Full Data Blocks (FDBs), Enhanced Data (D)-side (with radar display), FDB dragging capability, Computed Indicated Air Speed, Conflict Probe (CP) indication, Value scrolling, and others.

1.2.2 Future En Route Workstation II

In this study, we applied many of the same principles as in Willems et al. (2008), but instead of using the DSR system, we used the ERAM system as the baseline and introduced FEWS II functionality onto the ERAM platform. In a similar manner, we applied ERAM principles to the FEWS functions. For example, we applied the FEWS principle of connecting display representations of the same aircraft to the ERAM views. When controllers "touched" an aircraft in any of the ERAM views, the corresponding representations in other areas (i.e., Electronic Flight Progress Strips [eFPS], radar area, and so forth) would highlight, making it easier for controllers to find related data elements. An example application of ERAM principles to FEWS objects is the availability of fly-out menus for font, brightness, and transparency control of FEWS specific views identical to ERAM views.

1.2.2.1 Data Communications Failed View

When controllers make a Data Communications entry using the fly-out menus in DSR or ERAM using the Controller-Pilot Data Link Communications (CPDLC) Build IA CHI, controllers need to check the UPLINK button on the fly-out menu to uplink the message to the aircraft. Because this is an extra step, we have assumed that at 70% equipage of the fleet, the system will most likely check the UPLINK button by default. The controller will uncheck only the UPLINK button when one of the aircraft does not have Data Communications available. In that case, however, controllers could execute the same action sequence to accomplish two different things depending on whether an aircraft has Data Communications available. In FEWS I, we had observed that controllers mistakenly assumed that they had uplinked a message to an aircraft, although the aircraft did not have Data Communications available (Willems et al., 2008).

1.2.2.2 Interactive Status of the Voice Frequency

The introduction of URET has removed many of the controller's responsibilities for maintaining paper Flight Progress Strips (FPS). Many of the functions that paper FPSs had, however, have slowly moved onto the radar display. Examples include the Multiple Flight Plan Readout (MFR) in DSR that displays several flight plans in a table instead of displaying them in plain text in the Computer Readout Device (CRD). Aircraft representations now have four lines in an FDB because controllers can now indicate coordinated heading and speed in the fourth line where they used to do that by annotating the paper FPS. One function that has not appeared on the radar display is the Voice Communications status. On paper FPSs, controllers used symbology (that sometimes was local to a facility) to indicate whether a pilot had called in or had indicated that he or she would switch to the next frequency.

In the comments on Post-Scenario Questionnaires (PSQs) and during debriefings in the FEWS I experiment, controllers had indicated that when working with the FEWS interface, it was difficult to find aircraft that had called in because the system automatically accepted handoffs (Willems et al., 2008). As a result, we introduced a feature that showed the system had accepted the handoff, as described in the next section. We introduced an interactive voice frequency status indicator that helped controllers to keep track of and to indicate which voice-based aircraft had called in or had switched to the next sector's frequency.

Awareness of whether or not a controller is in radio contact with an aircraft is important from a safety point of view. To assist controllers in keeping track of who had called in and which aircraft they had already switched, we introduced a function that enabled them to annotate the FDB with reminders. The reminders also helped the system trigger actions such as automatic FDB drop-off. When a controller accepted a handoff, the radar control indicator showed that his or her sector had track control, whereas the voice frequency indicator showed that the previous sector still had control. When the previous sector controller indicated that the pilot had switched to the current sector's frequency, the voice frequency indicator showed an arrowhead pointing into the FDB. At that point, the current controller could see that the pilot of the aircraft would soon call to establish two-way radio communications. As soon as the pilot calls in, the current controller acknowledges the two-way radio communications (the arrowhead, the voice frequency, and the radar control indicators disappear), and then the FDB becomes a standard FDB.

When the aircraft leaves the sector, the controller instructs the pilot to switch frequencies and indicates that the instruction has taken place by clicking on the voice frequency indicator. The next sector then goes through the acceptance sequence. The system keeps track of the voice frequency status and will drop FDBs automatically when it detects that the next sector has established two-way radio communications. An aircraft that has left the sector and still has an FDB, therefore, alerts the controller that something out-of-the-ordinary may be happening with that aircraft.

1.2.2.3 Show Who Took Action

In this study, we introduced a feature that showed who had taken action. For this study, we have limited this feature to handoff acceptance, but it may be useful for other tasks as well. When the D-side controller accepts a handoff, the system displays a dashed box. The assumption is still that the Radar (R)-side controller is the Team Lead. When a handoff acceptance takes place without a solid or dashed box around the FDB, this implicitly shows that the R-side controller has accepted the handoff.

1.2.2.4 Route Display

Controllers use the route display function to get a quick representation of the current route of an aircraft. In DSR and ERAM, controllers request a route display through a combination of the keyboard and trackball or the keyboard only, because a route display always requires the use of a "QU" quick action command. To save time, the DSR keyboard has an "RTE" function key that places the QU<SPACE> string in the CRD. The controller then completes the route display request by entering a Flight Identification (FLID), by typing the Computer Identification (CID) beacon code or by using the trackball to select the aircraft, followed by the <ENTER> key. Controllers can request flight plan readouts in the Response Area (RA) in a similar fashion. In the field, however, since the introduction of interactive FDBs, controllers can also request a flight-plan readout by clicking on the call sign with the center trackball button. This enables controllers to complete the flight plan readout without taking their hands off the trackball. We implemented an analogous function for the route display function by allowing controllers to click on the CID in the FDB with the center mouse button. This function worked as a toggle and turned the route display on and off.

1.2.2.5 Preferred FDB Offset

Controllers often use leader line orientation or FDB offset as a memory aid. For example, to remember that an aircraft lands at a certain airport, the controllers may offset an FDB in the NNW direction. To create such reminders, controllers use simple rules. The NAS contains the data to determine whether a flight matches these rules. Instead of having controllers create these reminders manually, the FEWS II algorithm automatically places FDBs at the preferred orientation for a sector.

1.2.2.6 Automatic Data Block Offset

Data block movement entries account for approximately 20% of all controller entries in the field. Controllers use some of these entries as memory joggers, but many offset entries are to prevent FDBs from overlapping, thereby obscuring data in one or more FDBs or other relevant data. An Automatic Data Block Offset (ADBO) algorithm can therefore potentially reduce controller entries substantially.

1.2.2.6.1 Existing Data Block Offset Mechanisms

In the NAS, controllers can currently manually offset data blocks in a number of ways. In all ATC domains, controllers can use their keyboard to move data blocks to eight specific orientations and at least three different distances for the position symbol.

The introduction of advanced automation into the NAS has provided the system with displays that contain a multitude of new aircraft representations that often resemble the data blocks that controllers use on the radar display. Examples of such displays are the Airport Surface Detection Equipment-Version X, the URET Graphic Plan Display (GPD), the CTAS Plan View Display Graphical User Interface (PGUI), and the Enhanced Traffic Management System Traffic Situation Display.

The interaction with the data blocks on the automation displays often extends interaction capabilities beyond those offered on the standard ATC displays. For example, in the GPD on URET and the PGUI on CTAS, controllers can drag a data block to a new position. The GPD also has its own ADBO algorithm.

In the TRACON environment, controllers have an ADBO algorithm available that has been around since the Automated Radar Terminal System and is now available in the STARS. The algorithm uses the discrete orientations and lengths of the leader line, resulting in sudden jumps of the data block to a new position when the algorithm prevents two data blocks from overlapping. The sudden onset of movement of the data block can be quite distracting, most likely caused by a grasping reflex of the visual system. The grasp reflex is a result of the visual system attempting to fixate areas in the visual field where it detects motion (Giorgi, Soong, Woods, & Peli, 2004). By fixating that area, we are able to pick up more detailed data and to decide whether the information requires action on our part. To avoid the distraction, controllers often disable the ADBO function.

URET uses a similar approach that calls an algorithm sequentially for each aircraft that marks the position symbol, leader line, and data block on a 100 x 100 bitmap. To move a data block, URET checks its tentative position against the bitmap. If already occupied, URET will try to find another position. Once the system finds an unoccupied position, it marks the new position on the bitmap, and URET continues with the next aircraft. The algorithm first checks to see whether the controller previously, manually offset the data block. If the algorithm finds that to be the case, it will keep the existing offset position. If not, the algorithm will iterate through three leader lengths and eight ordinals until it finds an unoccupied position. If the algorithm cannot find an occupied position, it will use the default orientation and length.

1.2.2.6.2 Research and Development

EUROCONTROL has developed a requirements document (Dorbes, 1999) that specifies a data block management system that includes data block offset preferences or defaults as well as settings for the ADBO algorithm. In a research program at EUROCONTROL, researchers have implemented an interface and algorithms to try to meet the documented requirements of EUROCONTROL-Consolidation of HMI for Operations, Evaluations and Simulations (ECHOES, 2004). In the implementation of the ADBO management system in ECHOES (2004), controllers have full control over how the algorithm offsets FDBs, including whether offsets are relative to screen or world coordinates, based on flow of traffic or based on aircraft.

Duverger (2005) developed a demonstration software package that used a cost function to determine the optimal data block placement, taking into account the preferred data block position as well as the location of data blocks, speed vector lines, and position symbols of other aircraft. Although Duverger's algorithms divided the search space into a grid, his algorithms also calculated the cost of placing a data block at a particular location and decided, based on finding a local optimum, where to place the data block.

1.2.2.6.3 Implemented Data Block Offset Algorithm

We implemented an FDB offset algorithm (see Appendix) that smoothly moved FDBs away from one another if the algorithm predicted FDB overlap. During initial user evaluations of the algorithm, we determined preferred parameters for the algorithm. For this study, we used fixed parameters. Controllers can toggle the use of the algorithm on individual aircraft by locking the relative FDB position. The system reminds controllers that an FDB is locked by changing the leader line from solid to dashed.

1.2.2.7 Third Tier View

During the FEWS I experiment, controllers complained that the third-tier object was hard to maneuver and had obscured parts of the traffic area. In this study, we have adopted the third tier as a flight data object. For example, when a controller double-clicks on the call sign of an FDB, the system places a third-tier object in a separate bay on the radar display. Controllers could interact with the third-tier object in the same way as they could on the electronic flight strip bay. The objects displayed on the flight strip bays were identical to the third-tier objects on the radar display; interaction with both objects was identical as well.

1.2.2.8 Interaction with Electronic Flight Progress Strips

To interact with the bays and the FPSs, the controllers used their fingers. We calibrated the two touch panels available for each position such that sharp objects like pens or pencils did not result in activation of the display elements, but a press with a finger would. By touching a bay header, controllers could highlight all aircraft in that bay on the radar display. Figure 1 displays part of one of the eFPS display panels.

Figure 1. Example of two bay headers for the metering fixes ILL and SGF as depicted on the eFPS panels.

To make interaction with the eFPSs on the touch panels easier, we modified the FEWS I implementation to create larger fields. The eFPSs contained touch-sensitive areas for Data Communications, NAS updates, and altitude, heading, speed, and route changes. When a controller touched the area on the flight strip that represented the first tier of the FDB, it toggled a selection of the flight strip and highlighted the call sign of that aircraft on the radar display. Directly touching the altitude, heading, speed, or route field on the strip puts the strip in edit mode. Figure 2 shows an example of the interaction with the route field to send an aircraft direct to a fix.

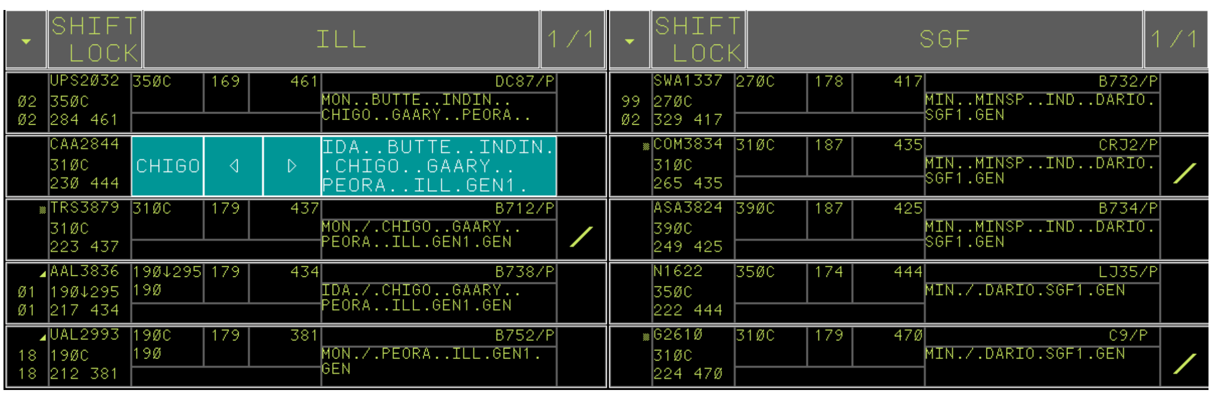


Figure 2. Example of interaction with eFPS on the eFPS panel to change a route.

Because we designed the strips so that the controllers could use their fingers to make changes to the aircraft, we used large buttons to change field values. Once the controllers put a flight strip in edit mode, they could change field values by stepping through field values using these buttons.

The *smart bays* also enabled controllers to move aircraft between our two metering fixes. When a controller selected an aircraft in the bay of one of the metering fixes and, subsequently, clicked on an “Insert” arrow on the other metering fix bay, the algorithm found a spot to join the arrival route over the other metering fix and inserted the FPS. After deselecting the strip and updating NAS or uplinking a route message to the aircraft, the route changed and the radar display showed the new route.

2. METHOD

The current study investigated three interface designs: DSR, ERAM, and FEWS II (see Table 1), at increasing traffic levels designed to reflect the anticipated volume of traffic in 2015 and beyond. We assessed the impact of Data Communications availability and sector staffing (SS) the three Workstation Configurations (WCs).

Table 1. Differences between DSR, ERAM, and FEWS II Simulation Configurations

System	**DSR**		**ERAM**		**FEWS II**	
Position	**R**	**D**	**R**	**D**	**R**	**D**
Hardware						
Input Device	Trackball	Trackball	Trackball	Trackball	Mouse	Mouse
	Keypad	Keypad	Keypad	Keypad		
Keyboard	R-side DSR Keyboard	D-side DSR Keyboard	R-side DSR Keyboard	D-side DSR Keyboard	DSR-R + Emphasis	DSR-D + Emphasis
Display	29 inches	19 inches	29 inches	19 inches	29 inches	29 inches
Computer-Human Interface (CHI): Aircraft Representation						
Track Data	Track and Position		Track and Position		Track and Position	Track and Position
Mode C Altitude	FDB		FDB		Integrated in Three-Tier FDB	Integrated in Three-Tier FDB
Assigned Altitude	FDB	Flight Plan Readout	FDB	Flight Plan Readout		
Indicated Airspeed		Though URET		Though URET		
Coordinated Heading	Line 4 of FDB		Line 4 of FDB			
Coordinated Speed	Line 4 of FDB		Line 4 of FDB			
Interaction with FDB	Flyout Windows		Flyout Windows		Edit/Scroll	Edit/Scroll
CHI: Windows and Lists						
Traffic Management Data	Present in TMA List + Range Data Block (RDB)		Present in TMA List + RDB		Integrated in three-tier FDB	Integrated in three-tier FDB
Conflict Alert (CA)	CA List		CA List			
Conflict Probe		URET		URET		
Trial Planning		URET		URET		
Data Communications	Build 1A	Build 1A	Build 1A	Build 1A		
Flight Plan Data (type, destination, etc.)	MFR Window or CRD	CRD	Continuous Flight Plan Readout (CFR) Window or CRD	CRD		
Emphasis	Multiple Dwell Lock/Fourth line indicators		Multiple Dwell Lock/Fourth line indicators		Emphasis Function	Emphasis Function
Multiple Flight Strip Readout	MFR Window		CFR Window		eFPS	eFPS
Flight Progress	Electronic Strips	URET Aircraft List (ACL)	Electronic Strips	URET ACL	eFPS	eFPS
Reminders	Electronic Strips	Electronic Strips	Electronic Strips	URET ACL	eFPS	eFPS
CHI: Other						
Route Display	Radar Display		Radar Display		Radar Display	Radar Display
Trajectory Display	Radar Display	URET GPD	Radar Display	URET GPD	Radar Display	Radar Display
Data Block Management					ADBO	ADBO

To increase the sector capacity further, we provided additional enhancements to the FEWS I concept. In FEWS II, we evaluated the effect of automatic data block management on controller performance and workload.

2.1 Participants

Twelve controllers from Level 11 and 12 ARTCCs participated in the study. All participants were non-supervisory Certified Professional Controllers qualified at their facility with a current medical certificate. The participants had visual acuity of not less than 20/30 corrected and could wear corrective lenses or soft contact lenses. Four of the 12 participants used corrective lenses. However, because we collected eye movement data, we excluded those who wore bifocals, trifocals, or hard contact lenses due to the design limitations of the oculometer. Our participants were seasoned controllers; mean age = 40 years (*SD* = 6), mean experience in ATC = 17.5 years (*SD* = 5.9). The participants gave their written consent to participate in the experiment by signing an informed consent agreement. The research team ensured the participants that their data were confidential; we used a coding system on the data forms to preserve anonymity. The FAA WJHTC Local Institutional Review Board reviewed routine ethical considerations for approval of this study.

2.2 Simulation Personnel

Simulation personnel included experimental staff and simulation pilots. The experimental staff included ATC SMEs, Engineering Research Psychologists (ERPs) and support engineers. The SMEs generated the traffic scenarios and conducted real-time Over-the-Shoulder (OTS) performance ratings. The ERPs conducted the data collection sessions, performed the data analyses, and prepared the technical reports. Support engineers ensured that the hardware and software were functioning appropriately.

2.2.1 Experimental Staff

ERPs and ATC SMEs conducted the simulation. The SMEs created the scenarios and provided performance ratings. The ERPs conducted the experiments, performed the data analyses, and wrote the final technical report. Support engineers ensured that the hardware and software were functioning properly to display and collect data. Additional support staff assisted in preparing, copying, and distributing forms and questionnaires.

2.2.2 Simulation Pilots

In this study, we used six simulation pilots. The simulation pilots either supported one sector or supported two sectors in two groups of three simulation pilots. The simulation pilots entered data at their workstations to maneuver aircraft, based on controller clearances given over a Voice Communications system.

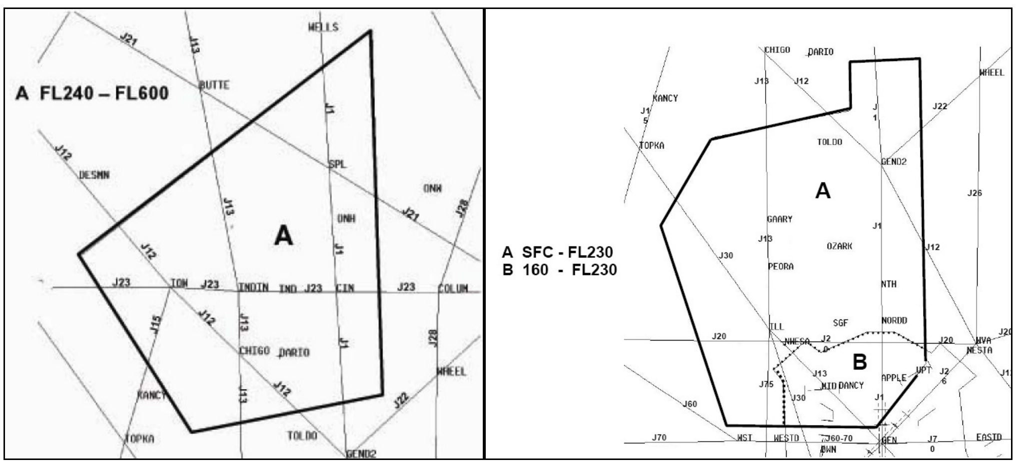
2.3 Simulation Equipment

The experiment took place at the RDHFL and used a high-fidelity ATC simulation environment that was fully reconfigurable to each workstation. We modeled airspace and scenarios for the training and test sessions in the ATC simulator. We used an integrated system, including the Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE), to emulate the three WCs and we used the Target Generation Facility to generate targets and airspace.

2.3.1 Airspace and Scenarios

The experiment included fifteen 60-minute test scenarios. The test scenarios were identical, except for their call signs. This allowed us to reduce the likelihood that participants perceived these scenarios as identical but enabled us to test each system under the most similar conditions possible.

During the simulation, the weather conditions required that Instrument Flight Rules be in effect. This study used generic airspace designed by researchers and SMEs to be realistic, yet relatively easy for controllers to learn (Guttman & Stein, 1997). A number of simulations, including the FEWS I simulations, have used the generic airspace (e.g., Sollenberger, Willems, Della Rocco, Koros, & Truitt, 2004). Initially, the generic airspace used in our simulations consisted of a single sector, but it has grown into a generic ARTCC containing several sectors, fix posting areas, TRACON facilities, and navigational aids, airways, STARs and Standard Instrumented Departures. Although we have not made an official name change yet, we think of it as Goldman Center or ZGN after the resident SME who implemented most of the facility. For this study, we used Sector ZGN08 and ZGN18 (see Figure 3).



Sector ZGN08 **Sector ZGN18**

Figure 3. ZGN08 and 18 airspace boundaries.

Sector frequencies are easy to learn because they all end with two digits that are identical to the sector number (e.g., 120.08 for ZGN08 and 120.18 for ZGN18). ZGN08 is a high altitude sector that meters traffic into Genera Airport using two meter-fixes located in ZGN18. ZGN18 is a low altitude sector located below and south of ZGN08. In this study, by LOA, controllers need to handoff aircraft to the low altitude sector 01 underneath at 24,000 feet (FL240) before the aircraft physically leave ZGN08. Controllers were responsible for feeding aircraft to several smaller airports, and controllers needed to handoff aircraft for these airports to the low sector at FL240. As illustrated in Figure 4, the traffic flow required controllers to merge two streams of traffic over Chicago (CHIGO), while absorbing delays before the Illinois (ILL) meter-fix, to meet arrival rates at Genera Airport.

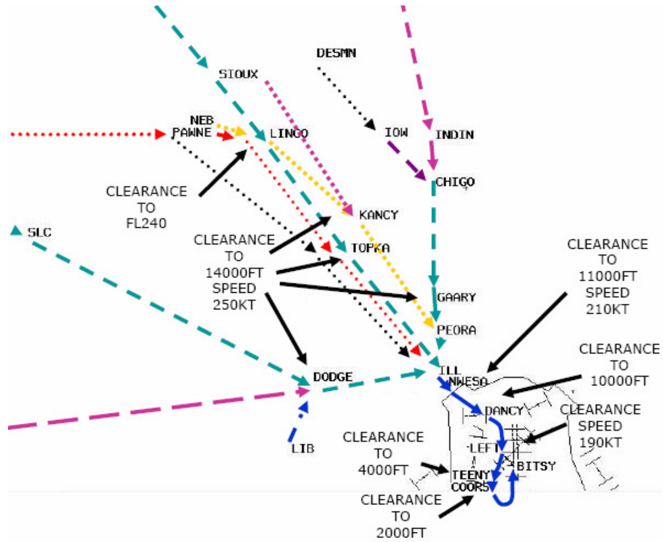


Figure 4. Metered traffic over ILL. Sector ZGN08 merges traffic over IOW and INDIN at CHIGO.

Controllers had a third stream of metered traffic over DARIO going to the Springfield (SGF) meter-fix into Genera Airport (see Figure 5). When necessary, controllers could balance traffic between the two meter-fixes.

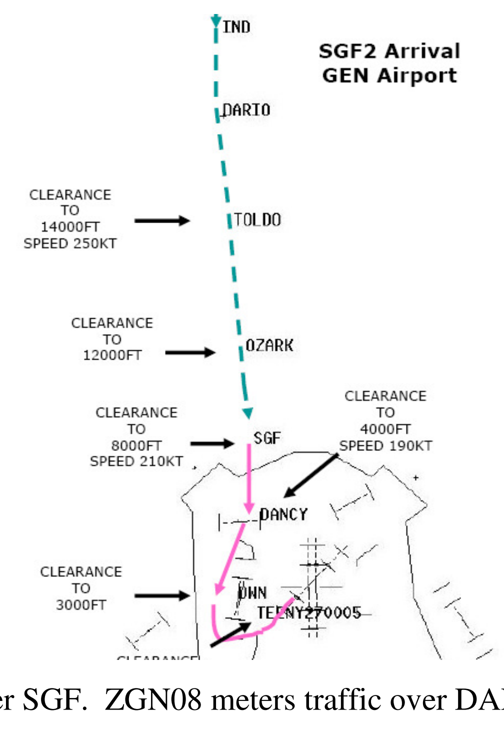


Figure 5. Metered traffic over SGF. ZGN08 meters traffic over DARIO.

We used traffic samples where the number of aircraft under control continuously increased. We changed our approach to scenario development because our operational stakeholder required that we expose our participants to the projected high traffic levels without the presence of Data Communications. From our FEWS I experiment, we had learned that, in many instances, controllers had struggled with 24 to 28 aircraft under control when digital Data Communications was not available. To expose controllers to a steady high traffic level (166% of current NAS levels or 35 aircraft for our simulated sector) would not provide much insight to when and how controllers will lose the "picture" and how they cope with that situation. Increasing the traffic from relatively low traffic levels to 166% levels and beyond provides us with a better opportunity to study what happens when controllers become saturated. The second reason for changing our approach was that in FEWS I, we used three relatively discrete levels of traffic. Although that provided us with enough data points to run a regression model to predict controller workload from the number of aircraft, the number of data points between the traffic levels was either low or non-existent. By increasing the traffic load steadily, we were able to better assess the relationship between taskload (e.g., number of aircraft) and controller workload. Each scenario started with the controller being responsible for 5 aircraft, and increased to as high as 50 aircraft by 50 minutes into the scenario (see Figure 6).

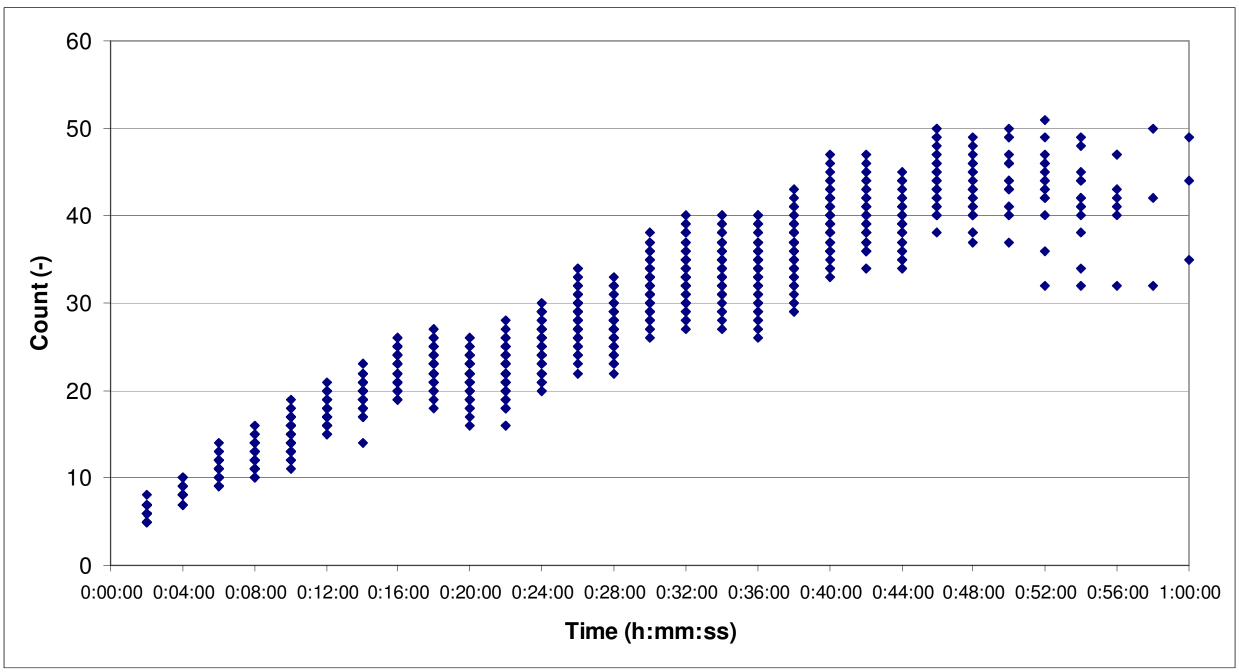


Figure 6. Number of aircraft in ZGN08 as a function of scenario time.

We generated additional scenarios for training and used a low traffic level scenario with 8 to 10 aircraft to introduce the participants to the Data Communications, ERAM, and FEWS II designs when training began on each system.

2.3.2 Controller Environment

The familiarization with the airspace, Letters of Agreement (LOAs), and Standard Operating Procedures (SOPs) used two adjacent controller stations equipped with a radarscope, a keyboard, and either a trackball or an alternative input device, depending on the system, for each participant team. One high-resolution (2,048 x 2,048 pixels) monitor displayed the radarscope while another displayed either a D-side CRD for the DSR and ERAM conditions or a second radarscope for the FEWS II interface conditions. The simulation pilots maneuvered the aircraft and issued ghost controller commands. A Workload Assessment Keypad (WAK) device (Stein, 1985) was available within easy reach of the participants to allow them to indicate workload ratings during the scenarios. An oculometer recorded the eye movements of the participants. A simulated radio system allowed for controller-pilot communications.

2.3.3 Simulation Pilot Terminal Configuration

A network linked the six simulation pilot operator displays with the controller workstation positions. Each simulation operator station allowed the entry of simulation pilot commands for up to 50 aircraft.

2.3.3.1 Display System Replacement Environment in a One-Sector Configuration

In the one-sector configuration (two-person sector), controllers had a 2,048 x 2,048 Barco Liquid Crystal Display (LCD) available on the R-side display that depicted the traffic situation (see Figure 7). Touch panels displayed eFPSs. Two touch displays showed incoming and current aircraft. The FPSs were not interactive but did update automatically when data in our simulated system received updates. Figure 7 shows two passive eFPS panels on the left-hand side. Next to the panels, the controller had a DSR radar display, keyboard, trackball, and Keypad Selection Device (KSD) available. To the right of the DSR radar display is the 2,048 x 2,048 display with a 1,024 x 1,024 inset, keyboard, trackball, and KSD for the D-side. The inset displayed the CRD and the URET Aircraft List (ACL). The eFPS panels on the right were identical to the ones on the left of the radar display. Above the displays were airspace maps for ZGN08 on the left and for Sector 18 on the right.

Figure 7. One-sector configuration in DSR.

To interact with the system, R-side controllers had the trackball, keyboard, and KSD available as currently used in the field. The keyboards on the R-side and the D-side were of the R-side type to make switching between configurations easier. The fielded R- and D-side keyboards are slightly different because the NAS contains certain data entries that are controller-position specific. For example, D-side controllers can enter a flight plan amendment command and have a dedicated function for that entry where the R-side does not have that capability. Although the KSD has six dedicated keys and two pairs of increment/decrement keys, we used only the increment/decrement keys for vector line increase/decrease and range increase/decrease in our study. We connected the keyboards and the trackball on the R-side display through a SunKeys device to the X386-based Linux computers that drive the displays.

The D-side controllers had a 2,048 x 2,048 Barco LCD available, but we used a window manager that made only a 1,024 x 1,024 area available to display the CRD and the URET windows. D-side controllers had a Logitech Optical Wheel Mouse available to interact with the display, but no KSD. The controllers could use Voice Communications to interact with simulation pilots through our communication system.

To key the microphone, they could use either a hand switch or a foot pedal. DESIREE emulated the DSR functionality in the R-side and the D-side. Most of the frequently used fielded functions were available, except for the annotation (drawing) function. The D-side controller had a D-side CRD available for data entry and feedback information.

DESIREE provided common message set (CMS) data to the Host Air Traffic Management Data Distribution System (i.e., HADDS) and the CTAS prototype, as if they had established a connection to a fielded Host Computer System (HCS). We ran the CTAS on 11 Sun Workstations under the Solaris Operating System (OS). The software included Radar Daemon that connected to DESIREE, an Input Source Manager, a Production Data Server, a Communication Manager, a Dynamic Planner, a Profile Selector, three Route Analyzers, a PGUI, and a Timeline Graphical User Interface. The CMS contains messages for system status, hand shaking, track updates, flight plans, and flight plan amendments. In return, the CTAS prototype provides estimated and scheduled times of arrival and delay times for aircraft that the system meters into an airport. Controllers could make entries to swap aircraft pairs of the same type or resequence aircraft in a way to meet metering constraints.

DESIREE also provided data through a Host URET Gateway (HUG) to the URET prototype as if it had a connection to a fielded HCS. The HUG ran on a Sun Workstation under the Solaris OS. The URET software ran on a DEC Alpha computer under the Open VMS OS. The URET software spawned several windows remotely onto the D-side controller display. The controllers had the ACL, GPD, and Plans Display available.

DESIREE has absorbed the CPDLC Build 1A functionality. The controllers could interact with an emulated Data Communications system built in support of the CPDLC program. All Data Communications services as envisioned for CPDLC Build 1A were available, except for the altitude downlink. The controllers had a message out, message in, and a menu text window available. The controllers could send Data Communications messages either through interaction with the data block (the UPLINK Button was active by default) or through a data entry in the Message Composition Area. The grammar used for Data Communications messages contained an “S” inserted before the FLID. During initial trials, controllers had asked us whether or not it was possible to replace the “<SPACE> S <SPACE>” sequence with a dedicated key to save data entry time. We implemented that change before the full experiment started.

To reduce the number of support staff to run a simulation, DESIREE had automated activities on the adjacent sectors. Adjacent sectors automatically initiated and accepted handoffs and switched frequencies when appropriate.

2.3.3.2 Display System Replacement Environment in a Two-Sector Configuration

In the two-sector configuration, the left 2,048 x 2,048 display continued to function as a radar display for ZGN08 (see Figure 8). Instead of displaying the URET windows for ZGN08 on a D-side position, they were available on an auxiliary display mounted over the radar display. The controllers could move their cursor from the radar display into the auxiliary display to make it possible to interact with URET without the need to have an extra keyboard and trackball. Two passive eFPS panels on the left-hand side are for ZGN08. Next to the panels, controllers had a DSR ZGN08 radar display, keyboard, trackball, and KSD available. To the right was a DSR ZGN18 radar display, keyboard, trackball, and KSD for the controller working sector ZGN18.

The eFPS panels on the right are for ZGN18. Above the radar displays were displays for URET that were accessible to the controllers by moving the cursor up from the radar displays and into the auxiliary displays. At the top of the photograph were the airspace maps.

Figure 8. Two-sector configuration in DSR.

On the D-side display, we removed the 1,024 x 1,024 inset and replaced it with a full 2,048 x 2,048 display. The controllers working ZGN18 had an identical setup to ZGN08. Each of the sectors used its radio frequencies to interact with the simulation pilots. ZGN08 and ZGN 18 shared a common airspace boundary and handed aircraft off to one another.

2.3.3.3 En Route Automation Modernization System Configuration

The hardware configuration used for the ERAM two-person sector emulation was identical to that used for the DSR emulation. The main differences between the ERAM system and the DSR/Host system are the hardware and software that process incoming data and enable controllers to interact with the system. The changes to the CHI are mostly part of an update scheduled for DSR but are now part of the ERAM deployment. These changes may affect the way controllers work with the system. The hardware and software that we used for the ERAM two-sector emulation were identical to the hardware and software that we used for the two-sector DSR emulation.

2.3.3.4 Future En Route Workstation System Configuration

The controllers had mostly the same hardware available as under the ERAM conditions, except for the pointing device. In the FEWS conditions, the controllers had a three-button Logitech Wheel Mouse instead of a trackball. The D-side controllers could still execute functions to assist the R-side controller but had the full 2,048 x 2,048 radar display available. The hardware and software that we used for the FEWS two-sector emulation were identical to the hardware and software that we used for the one-sector FEWS emulation, except that the two positions staffed separate sectors (see Figure 9).

Figure 9. Two-sector configuration for the FEWS configuration.

2.3.3.5 Communications Configuration

We used a communications system that was different from the Voice Switching and Control System used in the field. It had communication links between the controller, SME observer, simulation pilots, and experimenters and PTT recording capability. The equipment monitored and recorded times and frequencies of PTT activity for subsequent data reduction and analysis.

2.3.3.6 Oculometer

We used an oculometer consisting of an eye- and head-tracking system that records the Point of Gaze and pupil diameter of the participant by using near infrared reflection outlines from the pupil and cornea. For an extensive description of both the hardware and the software used for eye tracking, see Truitt and Willems (1999) and Willems, Allen, and Stein (1998). Willems et al. (1998) indicated that the exposure to the infrared illumination while wearing the oculometer is less than 4% of the intensity of that experienced when outside on a sunny day. Each participant wore the oculometer during the test scenarios as well as during the last two training scenarios for each system to become accustomed to the procedures and wearing the device.

2.3.3.7 Workload Assessment Keypad

The WAK is a portable version of the Air Traffic Workload Input Technique (ATWIT) and is a reliable and unobtrusive real-time, on-line measure of subjective workload (Stein, 1985). Each participant made a workload rating on the WAK device every 2 minutes throughout a scenario. The system prompted participants to respond by emitting several beeps and illuminating the keypad buttons. The participants pressed one of the keypad buttons labeled from 1 (*extremely low workload*) to 10 (*extremely high workload*). The participants had 30 seconds to respond; otherwise, the WAK records a code indicating that it received no response.

2.3.3.8 Video Camera and Video Recording Configuration

We recorded the video images of the controller (top and side views). At an observation station, video monitors provided a video display of all ATC positions. A video server compressed the audio and video, streamed it in real-time, and stored the data on a hard drive.

2.4 Experimental Design

The main experimental design is a 2 x 2 x 3 repeated measures design with three interface conditions (DSR, ERAM, and FEWS II); two Communications Configurations (CCs), Voice Communications or Data Communications; and two SS options (two-person sector or one-person sector), as illustrated in Table 2.

Table 2. FEWS II Simulation Design

Staffing Configuration		**2-Person Sector**		**1-Person Sector**	
Communication Configuration		VC	DC	VC	DC
Workstation	DSR				
	ERAM				
	FEWS II				

The participants worked in groups of two, with one working as an R-side and the other as a D-side within a sector or as two R-sides on adjacent sectors. We counterbalanced the order in which the groups completed the test scenarios for each system.

We presented the participants with a version of the NAS that had evolved into a system with a mixed fleet of 70-75% aircraft with state-of-the-art equipment expected for 2015, whereas the other 25-30% still used equipment from 2005, as used in FEWS I. Fully equipped aircraft had maximum automated flight capabilities (i.e., Data Communications), whereas the other aircraft continued to use 2005 capabilities.

2.4.1 Independent Variables

This experiment tested the effect of three Independent Variables (IVs) in two experimental designs. The IVs for the first design included the controller position: R08 or D08/R18. One of the controllers staffed the radar position at ZGN08 (i.e., R08) while the other controller worked either as a D-side at ZGN08 (i.e., D08) in the one-sector configuration or as an R-side controller at ZGN18 (i.e., R18) in the two-sector configuration. To determine the effect of staffing levels at the sector, we exposed controllers to traffic samples either as part of an R-side and D-side sector team or as a single R-side (Staffing Configuration). We provided controllers with a traffic sample that either had Voice Communications or had 70% of the aircraft equipped with Data Communications to determine the effect of Data Communications CC. Finally, to determine the effect of workstation, we used three WCs: One emulated DSR, the second emulated the expected operation and interface of ERAM, and the third emulated FEWS II. In the second design, we focused on the absence or presence of this automation function under the FEWS II display condition.

2.4.2 Dependent Variables

Table 3 summarizes the primary data measures we collected during the study. Each of the measures provides insight into aspects of controller performance and workload as well as aircraft efficiency.

Table 3. Primary Data Measures

Controller and pilot system entries
Aircraft data (e.g., heading, altitude, speed)
Eye movements of the R-side and the D-side controller (60 samples per second)
Workload ratings via Air Traffic Workload Input Technique (ATWIT)
Questionnaire ratings of workload, SA, performance, system usability
Communications between controllers and between controllers and simulation pilots
Push-to-Talk (PTT) data

The data collected during the simulations are scenario-based, interval-based or event-based. We calculated summary measures for each group in each test condition. These measures included eye movement data, workload data, WAK, and NASA Task Load Index (TLX) ratings, subjective measures of scenario difficulty, SA, controller and system performance obtained via the PSQ and SME OTS observations, and comparative ratings of the system configurations obtained via the Exit Questionnaire.

For each 2-minute interval, we collected WAK ratings and calculated interval-based or instantaneous variables to compare with workload ratings. For example, we calculated the number and average duration of Voice Communications for each interval prior to a workload rating, and we used the instantaneous count of the number of aircraft under controller responsibility at the time the system prompted the controller to rate his or her workload.

Event-based data sets included keystrokes, PTT events, and eye movement fixations. We used these data sets to establish a sequence of events to characterize controller tasks and subtasks to understand how changes to the SS, CC, and WC affected these tasks.

2.5 Procedures

Each participant pair spent 8 days completing training and testing. Each pair arrived at the RDHFL on a Tuesday morning and completed the simulation on Thursday of the following week. The simulation ran for a total of 12 weeks to accommodate the 6 participant pairs.

The first morning consisted of a briefing and familiarization period. The researchers explained the experiment, the oculometer, and the other equipment used in the study as well as the data collection tools. The researchers discussed the rights and responsibilities of the participants before asking them to sign the informed consent form. The participants then completed a Background Questionnaire that included demographic questions about age, experience level, need for corrective glasses, and other items pertinent to the experiment.

After instructing the participants about the LOAs and the SOPs, we began the training scenarios. Training always began with the DSR system to allow the participants to become familiar with CPDLC and the higher levels of traffic. We then counterbalanced the order in which the participants worked with the systems to train and test to minimize any order effects. The first day's scenario used 100% Data Communications-equipped aircraft. The first training scenarios had relatively lower traffic levels. DSR training included five training scenarios conducted over the course of one day. ERAM and FEWS II conditions received an additional day of training if that block of the simulation was the first time that controllers experienced the ERAM display features. The controllers completed up to five training scenarios each day.

Prior to the start of training, we instructed the participants on the use of the WAK tool and allowed them to become familiar with its operation during the training scenarios. The participants wore an oculometer during the final two training scenarios to allow them to become familiar with the apparatus and procedures. They wore the oculometer during all of the test scenarios.

The controllers then ran 15 test scenarios. They worked either in pairs, with one participant serving as an R-side and the other as a D-side within a sector, or as individual R-side controllers responsible for one sector.

Our SME observed the participant that functioned as the R-side on ZGN08 and provided OTS performance ratings. At the end of each scenario, the participants completed a PSQ to provide subjective ratings of performance, scenario difficulty, SA, workload, and reactions to the interface features. At the conclusion of the final test scenario, the participants completed an Exit Questionnaire to provide feedback about the simulation and comparisons of the test conditions.

3. RESULTS

During the presentation of the results of the data analyses, we set the alpha level to $p < .05$. Because of the small number of participants that we used in this experiment, we will occasionally present findings that did not reach significance at the $p < .05$ level but had a $p < .10$. We will mention these findings in the results as trends.

3.1 Scenario-Based Analyses

3.1.1 Post-Scenario Questionnaires

3.1.1.1 Perceived Performance

After each simulation scenario, we asked controllers to rate their overall level of ATC performance. The effects of CC and WC interacted, $\Lambda = .47$, $F(2,9) = 4.98$ (see Figure 10). Although the interaction was significant, the Tukey's Honestly Significant Difference (HSD) post hoc analysis did not reveal significant differences between conditions. The CC simple effect analysis under DSR revealed a trend that controllers perceived that they did not perform as well when Data Communications was available. A similar analysis showed that controllers did not perceive a difference in performance under ERAM. Finally, the simple effect analysis under FEWS showed a trend that controllers perceived that they performed better with Data Communications available. We conducted a WC simple effect analysis, which revealed that the controllers perceived their performance was significantly different between WCs when only Voice Communications was available, $\Lambda = .48$, $F(2,9) = 4.87$. However, the HSD post hoc analysis did not reveal differences between WCs. We found a trend that controllers perceived their performance was slightly less when they used ERAM instead of DSR and slightly less when they used FEWS instead of ERAM. The WC simple effect analysis under the Data Communications condition showed that controllers did not perceive a change in performance across the WCs (see Figure 10).

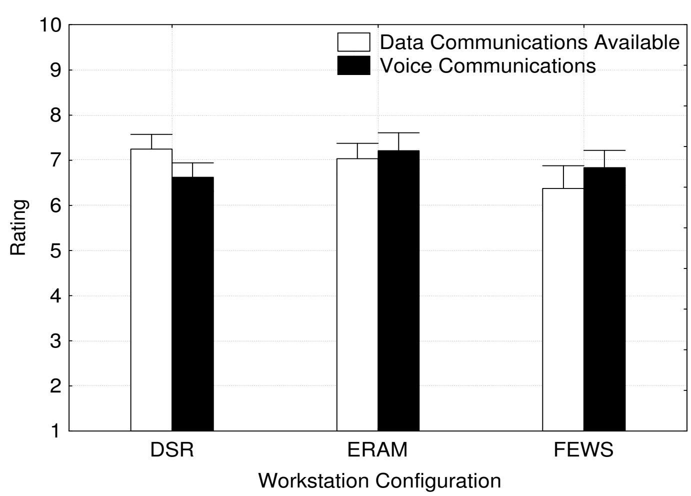


Figure 10. Performance ratings as a function of Communication Configuration and Workstation Configuration.

We also found an interaction between Staffing Configuration and Controller Position, $\Lambda = .38$, $F(1,10) = 16.01$ (see Figure 11). The controllers assigned to work traffic at the R08 position rated their performance as *better* when working in the two-person sector than in the one-person sector, whereas controllers working either as the D08 or the R18 position perceived that they had performed better when working as the R18 controller.

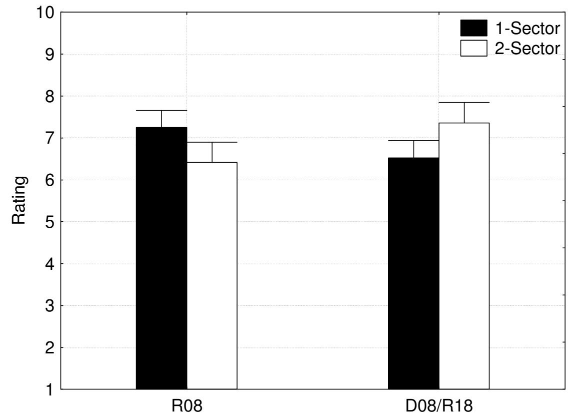


Figure 11. Performance ratings as a function of Staffing Configuration and Controller Position.

3.1.1.2 Communication Workload

After each simulation scenario, we asked controllers to rate their communication workload. When Data Communications was available, the controllers rated their communication workload significantly lower ($M = 5.3$, $SE = 0.5$) than when they could use only Voice Communications ($M = 7.4$, $SE = 0.3$), $\Lambda = .26$, $F(1,10) = 26.97$.

3.1.1.3 Situation Awareness

We asked controllers after each simulation scenario to rate their SA for current aircraft locations. We found a significant interaction between Staffing Configuration and Controller Position, $\Lambda = .55$, $F(1,10) = 8.02$. The R-side controllers perceived a reduction in SA for current aircraft positions when they worked the sector by themselves instead of with the assistance of the D-side controllers (see Figure 12).

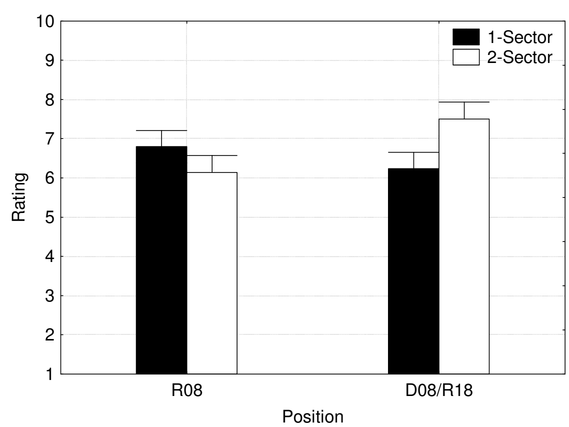


Figure 12. Perceived situation awareness for current aircraft locations as a function of Controller Position and Staffing Configuration.

We found that controllers had better perceived SA for current aircraft position when Data Communications was available than when they controlled traffic using only Voice Communications ($M = 6.5$, $SE = 0.2$ and $M = 6.8$, $SE = 0.3$, respectively), but the difference was very small and not statistically significant, $\Lambda = .70$, $F(1,10) = 4.22$, $p = <.07$. A small difference was also present in the interaction between the effects of CC and WC, $\Lambda = .54$, $F(2,9) = 3.83$, $p = .06$. The controllers indicated that when Data Communications was available, there was little change in their perceived awareness for current aircraft positions between WCs ($M = 6.8$, $SE = 0.3$). When using only Voice Communications, however, controllers rated their SA for current aircraft positions less for ERAM than for DSR and less for FEWS than for ERAM ($M = 6.9$, $SE = 0.24$; $M = 6.5$, $SE = 0.30$; and $M = 6.1$, $SE = 0.40$ for DSR, ERAM, and FEWS, respectively).

To determine how controllers perceived that our experimental manipulations had affected their ability to project aircraft positions, we asked them to rate their SA for projected aircraft locations. We found a four-way interaction between Controller Position, CC, Staffing Configuration, and WC, $\Lambda = .50$, $F(2,9) = 4.56$. We therefore followed up and analyzed the simple effects. Within the two-person sector and the one-person sector, individually, Controller Position, CC, and WC did not affect perceived SA for projected aircraft position. The controllers that always worked the R-side for ZGN08 did not perceive that Staffing Configuration, CC, or WC affected SA for projected aircraft positions. The controllers that worked either as a D-side on ZGN08 or as an R-side on ZGN18, however, perceived that the experimental conditions did affect their SA for projected aircraft positions, as indicated by the significant three-way interaction between CC, Staffing Configuration, and WC, $\Lambda = .13$, $F(2,4) = 13.56$. Further investigation of simple effects showed that controllers perceived that when they worked as the R-side on ZGN18 under the FEWS workstation condition, they had better SA for projected aircraft positions than when they functioned as a D-side on ZGN08, $\Lambda = .39$, $F(1,5) = 7.69$.

We asked controllers to rate their SA for potential loss of separation. We found a three-way interaction between CC, Staffing Configuration, and WC, $\Lambda = .51$, $F(2,9) = 7.24$ (see Figure 13). Further analysis showed that controllers rated their awareness for potential loss of separation in the two-person sector with Voice Communications using the FEWS workstation as significantly lower than in a one-person sector with Data Communications when using the ERAM workstation.

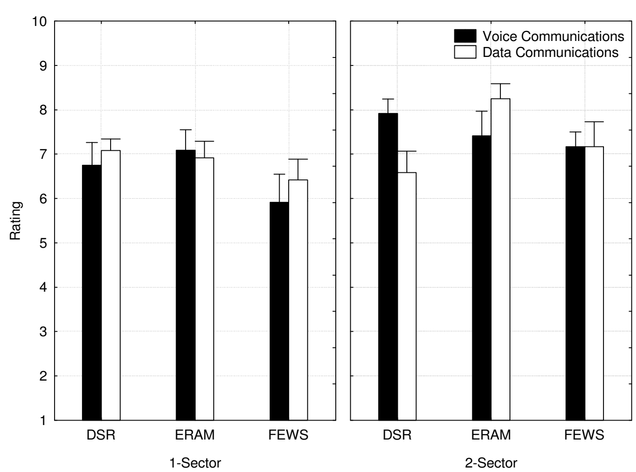


Figure 13. Perceived situation awareness as a function of Communication Configuration, Staffing Configuration, and Workstation Configuration.

To determine how controllers perceived how our experimental conditions affected their awareness for potential operation deviations, we asked them to rate their SA for potential operation deviations. We found a three-way interaction between the effects of Staffing Configuration, WC, and Position, and we found a two-way interaction between Staffing Configuration and Position, $\Lambda = .51$, $F(2,9) = 4.39$ and $\Lambda = .51$, $F(1,10) = 9.63$, respectively. Because the three-way interaction had an impact on the two-way interaction, we analyzed the three-way interaction by investigating simple effects. The R-side controllers did not indicate that our experimental conditions affected their awareness for potential operational deviations. The controllers indicated that their awareness for potential operational deviations was higher when they worked as the R-side controller on ZGN18 than as the D-side on ZGN08, $\Lambda = .22$, $F(1,5) = 17.47$, ($M = 7.30$, $SE = 0.43$ vs. $M = 6.35$, $SE = 0.52$). When the controllers worked with the DSR WC, they indicated that the effects of Staffing Configuration and Controller Position interacted, $\Lambda = .66$, $F(1,10) = 5.20$. Further analysis revealed that in the one-person sector, the controllers working as the R-side on ZGN18 perceived that they were more aware of potential operational deviations than controllers working as the R-side on ZGN08 ($M = 7.58$, $SE = 0.65$ and $M = 5.25$, $SE = 0.67$, respectively). When working ERAM, controllers did not perceive that CC, Staffing Configuration, or Controller Position affected their awareness of potential operational deviations. Under the FEWS condition, controllers indicated that their SA for potential operational deviations was significantly better when they worked as the R-side on ZGN18 than when they worked as the D-side on ZGN08 ($M = 7.83$, $SE = 0.75$ and $M = 5.67$, $SE = 0.75$, respectively).

3.1.1.4 Mental Resources

When we asked controllers how much of their total available mental resources they used for particular tasks, our experimental conditions did not always affect their ratings. We asked controllers about the following tasks:

- Searching for potential aircraft conflicts.
- Searching for direct routes.
- Planning control actions.
- Ensuring that aircraft conformed to control instructions.
- Situation monitoring.
- Resolving potential conflicts.
- Managing air traffic sequences.
- Routing or planning flights.
- Managing air traffic resources.
- Managing sector/position resources.

Controller responses indicated that the effect of CC and Controller Position on searching for potential aircraft conflicts interacted, $\Lambda = .66$, $F(1,10) = 5.11$ (see Figure 14). Although our analysis pointed at this interaction, the differences were small and further investigation did not reveal differences that we could attribute to Controller Position or CC changes. Controllers indicated that they spent between 58-67% of their mental resources on searching for potential aircraft conflicts.

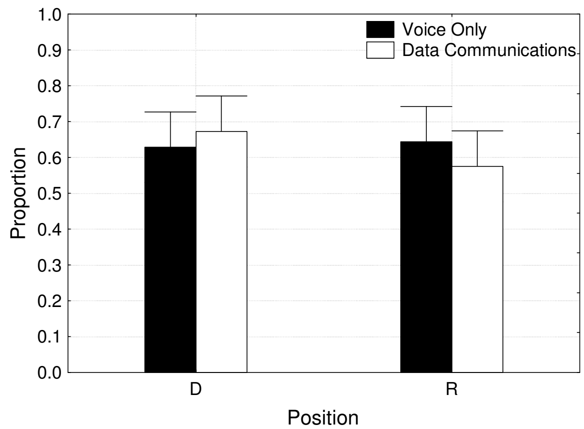


Figure 14. Perceived proportion of mental resources used for searching for potential aircraft conflicts.

Controllers spent less time searching for direct routes; 12-24% of total mental resources (see Figure 15). We analyzed how our experimental conditions may have affected how much controllers searched for direct routes and found 2 three-way interactions. First, CC, Staffing Configuration, and Controller Position effects interacted, $\Lambda = .62$, $F(1,10) = 6.20$. Although Figure 15 shows differences between sector positions, none of the simple effects showed significant results. Second, we found that Staffing Configuration, WC, and Controller Position effects interacted, $\Lambda = .50$, $F(1,10) = 4.42$, but none of the simple effects showed significant results.

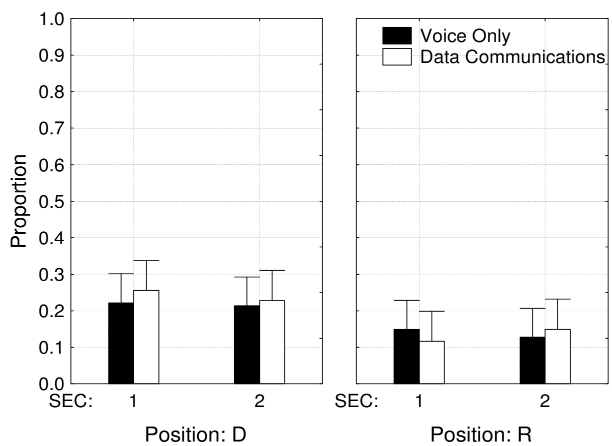


Figure 15. Perceived proportion of mental resources used for searching for direct routes.

The controllers indicated that they spent an average of 47% of their mental resources on planning control actions, but our experimental conditions did not affect that percentage. The percentage of mental resources used to ensure that aircraft conformed to control instructions varied from 23-45% and showed a four-way interaction between CC, WC, Staffing Configuration, and Controller Position, $\Lambda = .50$, $F(2,9) = 4.51$. Simple effects did not provide further insight into the interaction, except for a slight increase when Data Communications was available for the controllers that worked as an R-side on ZGN18. The controllers responded that they spent less mental resources on situation monitoring when they worked in a one-person sector than in a two-person sector, Λ = .41, $F(1,10) = 14.36$ ($M = 0.40$, $SE = 0.31$ vs. $M = 0.48$, $SE = 0.38$, respectively). Controllers spent about 48% of their mental resources on resolving potential conflicts, but our experimental conditions did not affect that percentage. Controllers spent more of their mental resources on managing aircraft sequences in the one-person sector, but the Staffing Configuration effect interacted with Controller Position, $\Lambda = .64$, $F(1,10) = 5.66$. The controllers that worked the R-side on ZGN18 under the one-person sector indicated that they had spent approximately 57% of their mental resources on managing aircraft sequences versus 26% when they had worked as a D-side on ZGN08 (see Figure 16). Our experimental manipulations did not affect how much of their mental resources controllers spent on routing and planning flights, air traffic resources, or sector.

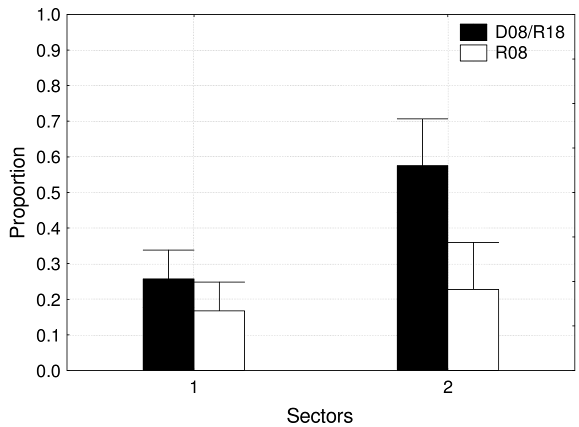


Figure 16. Perceived proportion of mental resources used to manage aircraft sequences.

3.1.1.5 Interface Effectiveness

To assess the effectiveness of the CHI, we asked controllers to rate how the interface affected safety, efficiency, sector operations, and control plan or strategy. The controllers indicated that the interface did not affect safety differently across our experimental conditions ($M = 6.41$, SE = 0.12). Controller responses indicated that the interface affected control efficiency more when Data Communications was available than when under Voice Communications conditions, Λ = .66, $F(1,10) = 5.22$ ($M = 5.73$, $SE = 0.28$ vs. $M = 6.47$, $SE = 0.43$), but WC did not affect perceived control efficiency. Our experimental manipulations did not change the extent to which the interface affected sector operations. Our experimental conditions affected the control strategy, but the effects of Staffing Configuration, WC, and Controller Position interacted, Λ = .27, $F(2,9) = 15.36$. We studied this three-way interaction, although we found a two-way interaction between Staffing Configuration and WC as well, because the three-way interaction also affected the two-way interaction. The R-side controller on ZGN08 did not perceive a change in control strategy. The controllers that worked either as D08 or R18 indicated that the effects of Staffing Configuration and WC affected their control strategy differently, $\Lambda = .12$, $F(2,4) = 14.98$ (see Figure 17).

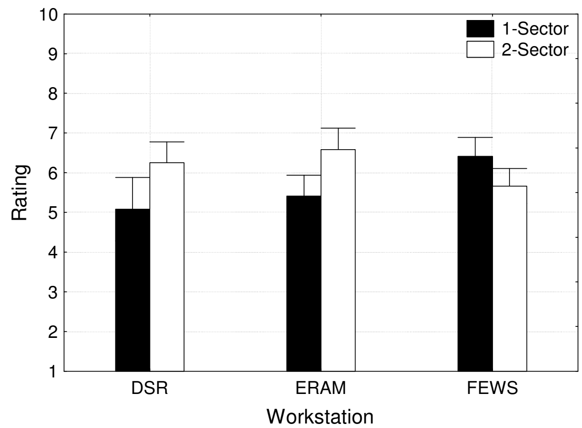


Figure 17. Perceived effect of the interface on control strategy for the D08/R18 controllers.

3.1.1.6 Interface Usability

To assess subjective interface usability, we asked the participants to rank the following statements on a 10-point scale that ranged from 1 (*complete disagreement*) to 10 (*complete agreement*).

- This system was simple and easy to use.
- The displays were clear and uncluttered.
- I could find the information I needed quickly and easily.
- I was able to prioritize information easily.
- Information was updated or changed on the display in an expected and predictable manner.
- The information was presented in an easy to understand format.
- It took only a few simple steps or actions to get information I needed, if it wasn't presented directly on the display.
- The colors and patterns used helped me locate what I needed quickly.
- The font and size of the text made it easy to read.
- The icons and graphics were easy to differentiate and interpret.

There was a four-way interaction of all IVs on the responses to our statement concerning simplicity and ease of use of the interface, $\Lambda = .30$, $F(2,9) = 10.35$. When controllers worked as a team in the two-person sector, the effects of CC and WC on the R-side responses interacted, $\Lambda = .19$, $F(2,4) = 8.63$. The R-side controllers rated that the system was simple and easy to use, with ratings ranging from 7 to 8 between conditions. The differences between conditions were only significant between the lowest and the highest agreement; they agreed most when they worked in the two-person sector with the ERAM configuration and Data Communications, and they agreed least when they worked with the FEWS configuration with Voice Communications only. The D-side controller ratings on the same statement ranged from 6 to 7, but CC and WC did not affect the D-side ratings. In the one-person sector where the controllers worked as R08 and R18, the R08 controllers agreed less with our statement when Data Communications was available, but it did not reach statistical significance.

The controllers agreed to some extent with our statement that the displays were clear and uncluttered. None of the effects of our experimental manipulations affected their ratings significantly, although controllers indicated that the FEWS displays were less clear and uncluttered than the other displays.

The controllers indicated that they *somewhat agreed* with our statement that they could find the information they needed quickly and easily, but the effects of Staffing Configuration and WC interacted, $\Lambda = .44$, $F(2,9) = 5.68$. Although the Tukey HSD post hoc analyses did not show significant differences, the controllers rated that information was easier to find in the one-person sector when using ERAM than in the two-person sector, but it did not differ under the other WCs.

Our experimental manipulations did not affect how controllers rated the ease with which they could prioritize information or the predictability of information updates on the displays. The controllers agreed that the system presented information in an easy to understand format; Data Communications increased slightly, but significantly, $\Lambda = .67$, $F(2,9) = 5.03$. Our experimental manipulations did not affect how much controllers agreed that colors and patterns helped them locate information or that font and font-size made text easier to ready.

3.1.1.7 NASA Task Load Index

The analyses of PSQ and OTS responses were somewhat problematic because of the protocol we used during the experiment. We instructed controllers to indicate when they reached a point where they could no longer control traffic safely and efficiently. At that point, we terminated the simulation. As a result, post-scenario items that probed the controllers about workload and difficulty were generally identical across experimental conditions, but the duration and traffic level at which the controllers terminated the simulation varied. To explore the impact of the presentation order and simulation duration, we conducted multiple regression analyses, where appropriate. The regression models often included participant teams as an IV to account for inter-team variability. To determine which variables warranted inclusion into the regression models, we used stepwise forward regressions.

The regression analyses conducted on the NASA TLX items included the following IVs: Controller Position, Staffing Configuration, Data Communications, WC, Team, presentation order, and simulation duration. We created dummy variables for the multi-level discrete IVs, WC, and Team. The responses for Team 2 and the DSR, one-sector, Voice Communications served as the baseline. For the 2 x 2 x 2 x 3 (Controller Position x Data Communications x Staffing Configuration x WC) design discussed in this report, we had 144 observations.

The Analysis of Variance (ANOVA) on controller ratings of the NASA TLX Mental Demand component revealed that, of the effects of Controller Position, Data Communications, Staffing Configuration, and WC, only Staffing Configuration had a main effect but it interacted with Controller Position (see Figure 18). The Tukey HSD post hoc analyses showed that the controllers that worked the R08 position did not experience a difference in Mental Demand, but the controllers that worked either the D08 (in the two-person sector) or R18 (in the one-person sector) experienced lower Mental Demand when working in the one-person sector. The controllers working the R08 and D08 positions (in the one-person sector) experienced similar Mental Demand levels; however, when working in the one-person sector, the R18 controllers experienced lower Mental Demand than the R08 controllers.

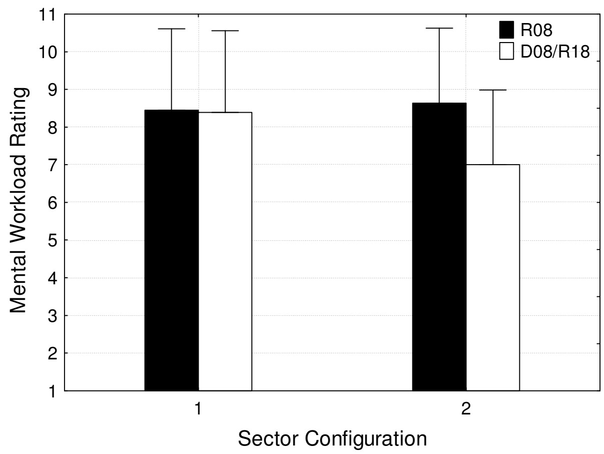


Figure 18. NASA-TLX Mental Demand Rating as a function of Staffing Configuration and Controller Position.

The regression model included an intercept plus five IVs (see Table 4). The model indicated that across conditions, the controllers perceived that the simulations required substantial mental resources (Intercept = 6.97, *SE* = 0.54, *t* = 12.98). The presentation order resulted in a significant increase in Mental Demand predictions; albeit, small and in the opposite direction of what one would expect as a result of a learning effect (*B* = 0.07, *t* = -4.28).

Table 4. Regression Results for NASA TLX Mental Demand Ratings

R= .62782786 R^2= .39 Adjusted R^2= .35 *F*(9,119)=8.60 *p* <.05						
	Beta	Standard Error	B	Standard Error	*t*(119)	*p*-level
Intercept			**6.97**	**0.54**	**12.98**	**.00**
Duration	0.19	0.10	0.00	0.00	1.97	.05
Position	**-0.31**	**0.07**	**-0.65**	**0.15**	**-4.28**	**.00**
Run	**0.27**	**0.08**	**0.07**	**0.02**	**3.15**	**.00**
Team 3	**0.30**	**0.08**	**0.84**	**0.23**	**3.58**	**.00**
Team 7	**0.31**	**0.08**	**0.90**	**0.24**	**3.79**	**.00**
DC	**-0.16**	**0.07**	**-0.34**	**0.16**	**-2.19**	**.03**
Team 6	0.17	0.09	0.46	0.24	1.96	.05
Team 5	0.11	0.09	0.31	0.23	1.34	.18
SC	-0.09	0.07	-0.18	0.16	-1.15	.25

Note. Statistically significant regression results are presented in bold.
DC = Data Communications; SC = Staffing Configuration.

As illustrated in Figure 19, Controller Position affected the controller ratings of Mental Demand, but our regression analysis showed that, in general, the D08 or R18 controller would result in a reduction of perceived Mental Demand; this is not obvious in Figure 19. The regression results indicated that the presence of Data Communications would result in a reduction of perceived

Mental Demand (*B* = -0.33, *t* = 2.19); a result that we did not find when analyzing the data using a repeated measures ANOVA. Finally, the inter-team differences that we eliminated in the repeated measures ANOVA resulted in two of our dummy variables having a significant contribution in the model. Teams 3 and 7 rated their Mental Demand higher than the Team 2 controls (*B* = 0.84, *t* = 3.5 and *B* = 0.90, *t* = 3.79, respectively). The full model explained 36% of the variance in the Mental Demand responses.

The ANOVA on the controller ratings of the NASA TLX Physical Demand component (see Figure 19) uncovered a three-way interaction between the effects of Staffing Configuration, Data Communications, and WC, Λ = .49, *F*(2,9) = 4.71.

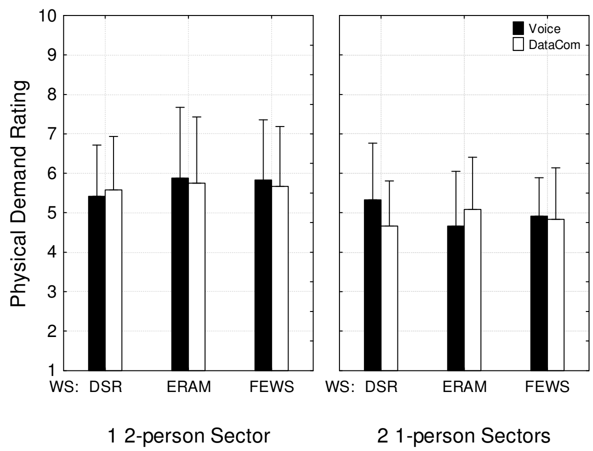


Figure 19. NASA-TLX Physical Demand as a function of Staffing Configuration, Data Communications, and Workstation Configuration.

3.1.2 Exit Questionnaire Controller Opinions

In the Exit Questionnaires, we asked the controllers to rate how well the WC supported some of their routine tasks. The controllers perceived that they managed traffic more efficiently when using the ERAM than when using the DSR configuration, but they perceived that they managed traffic most efficiently when using the FEWS configuration, Λ = .14, *F*(2,9) = 26.97. WC did not support locating information on the display differently, although the controllers that staffed the D-side position indicated that they perceived it was not as easy for them as for the R-side controllers. However, that difference did not reach statistical significance. Although the multivariate results indicated that there was no significant difference in the ease of avoiding potential conflicts across WC, the univariate results showed a trend towards easier detection when using ERAM than when using DSR and easiest under the FEWS WC. The controllers perceived that it was easier to resolve potential conflicts when using ERAM than when using DSR and easiest under the FEWS WC. There was no difference in the ease of maintaining SA between WC. The controllers indicated that it was easier to scan traffic effectively with the ERAM than with the DSR WC and easiest with the FEWS WC, but the result fell just outside of our criterion for statistical significance, Λ = .55, *F*(2,9) = 3.70, *p* = .07. Although there was a trend that indicated an increasing ease of providing timely control instruction from DSR to ERAM to FEWS, it did not reach statistical significance. Although the ANOVA results showed a significant difference between ease of managing traffic efficiently across WC, $\Lambda = .48$, $F(2,9) = 4.97$, a subsequent Tukey HSD post hoc test did not find significant differences between WC pairs. The controllers indicated that the support to maintain a manageable workload level was better when going from the DSR to the ERAM and then to the FEWS WC, but the ANOVA results perceived just outside of our criterion for statistical significance, $\Lambda = .56$, $F(2,9) = 3.60$, $p = .07$. Finally, we asked the controllers how well each of the WCs supported accomplishing all ATC tasks. Controllers responded that ERAM performed better than DSR, but that FEWS supported them best in accomplishing all ATC tasks, $\Lambda = .24$, $F(2,9) = 14.26$.

3.1.3 Operational Deviations

Controller deviations are incidents that do result in an operational error but are reportable and are an indicator of controller performance. The FAA has defined deviations and documented them for use during performance evaluation (FAA, 2007). In this study, we identified the following deviations:

- **Altitude Deviation:** The aircraft in ZGN08 with airport destinations in adjacent sectors did not leave the sector at an altitude of 24,000 feet or below as required by SOP.
- **Hand-off Deviation:** The current controller did not handoff an outgoing aircraft to the next sector controller prior to leaving the sector, unless pre-coordination had taken place.
- **Metering Delay Deviation:** A controller allowed an aircraft to leave the sector with a 1 minute or longer metering delay.
- **Frequency Deviation:** An aircraft left the sector without having the next sector frequency.
- **Wrong for Direction Deviation:** While at cruise altitude, west-bound aircraft were not on an even-numbered altitude or east-bound aircraft were not on an odd-numbered altitude.

Other types of deviations exist; for example, if an aircraft is close to the sector boundary within 2.50 nautical miles (4.63 km) horizontally or 1,000 feet (304.8 m) vertically, the controller commits a deviation if the controller does not point out the aircraft to the next sector controller. Another example is that when an aircraft leaves the sector, it should follow its current flight plan (i.e., aircraft should be FLAT tracking, where FLAT stands for FLight plan Assisted Tracking). We did not analyze the data of these two point-out related deviations. The shortest duration of the simulation runs for the one-sector condition in our primary design was 26 minutes. We used this duration for all runs to make each run equal in terms of traffic complexity.

The results showed large team differences. Groups 2, 3, 4, 5, 6, and 7 had 53, 43, 30, 18, 18, and 14 deviations total, respectively. For Group 4, we had missing data for one experimental run, which would have had ERAM, a two-person sector, and Voice Communications only. We substituted these missing data with the data calculated in the following way: The average number of deviations for other groups for this experimental condition was $30 \div 5 = 6$. After adding 6 to the total deviations of Group 4, which made 30 to 36, we ran chi-square tests to compare the groups; there was a significant difference among them, $\chi^2(5, N = 138) = 29.00$.

Figure 20 shows the total number of deviations per WC and CC combination when the R-side and D-side worked together on ZGN08 traffic. They had the fewest deviations when they controlled traffic in FEWS condition with Data Communications available. Most deviations occurred when controllers worked traffic in DSR with Voice Communications. Because we did not have many deviations and the distribution of deviations was non-normal, we used the Friedman Test as an overall test for significant differences between the six conditions. We used two-tailed Wilcoxon Signed Ranks Tests to compare pairs of conditions.

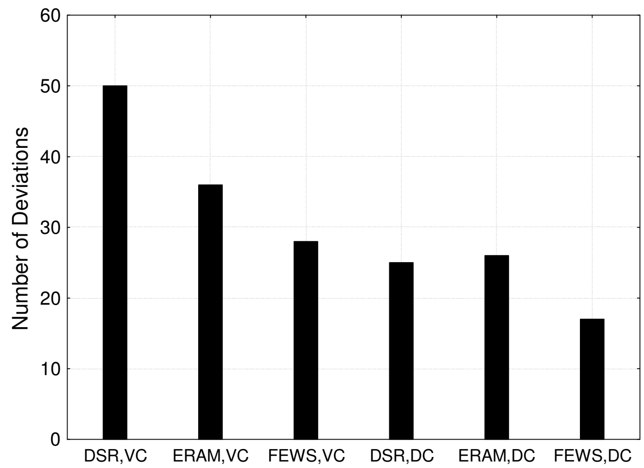


Figure 20. Total number of deviations for the Workstation Configuration and Communication Configuration combinations when R-side and D-side controllers worked on ZGN08 traffic together.

There was a significant difference among the six conditions, $\chi^2(5, N = 30) = 16.90$:

- A two-tailed Wilcoxon Signed Ranks Test that showed significant differences between DSR with Voice Communications and FEWS with Voice Communications ($z = -2.001$).
- DSR with Voice Communications and DSR with Data Communications ($z = -3.072$).
- DSR with Voice Communications and ERAM with Data Communications ($z = -2.866$).
- DSR with Voice Communications and FEWS with Data Communications ($z = -2.740$).

To compare deviation types, we created Team 4 data for the missing run by considering the relative deviation frequencies across deviation types by the other groups. We found no significant results when we compared WC for each deviation type, except Handoff Deviation in the Voice Communications mode, Friedman Test: $\chi^2(2, N = 6) = 6.12$. However, the Wilcoxon Signed Ranks Test did not show a significant difference between pairs of WCs for the Handoff Deviation.

3.2 Interval-Based Analyses

For the interval-based analysis, we used the same interval duration as we used for WAK data collection. Maintaining the same interval duration and timing makes it easier to compare the impact of Staffing, Communication, and WCs across data sets.

3.2.1 Keystroke and Button Press Data

The analysis of data entry performance focuses on the R-side controller of ZGN08. One of the promises of automation is to reduce the number of menial tasks. One way to measure a reduction in data entry-related activities is to measure the number of keystrokes and mouse or trackball clicks and their durations. We can then compare these four variables across experimental conditions. A reduction in data entry tasks should free resources that controllers can use for conflict detection and resolution and aircraft sequencing.

Although many controllers managed to work successfully through our traffic samples, longer than 14 minutes, in one of the simulations the team gave up after 14 minutes. We conducted two analyses – one with the team that gave up early, and an analysis without that team. We then compared our results between these two approaches. We conducted repeated measures Analyses of Covariance (ANCOVAs) to determine the number and durations of keystrokes and mouse/trackball clicks per 2-minute interval. The covariate used in the analysis is the interval end-time in seconds. The results of the analysis include mean values for the dependent variables after correction for covariate (time-interval).

When we included the first seven intervals (up to 14 minutes into the traffic sample), the ANCOVA revealed a significant change in the number of keystrokes as a function of WC, $\Lambda = .11$, $F(2,5) = 20.36$. The FEWS WC required significantly less keystrokes than either the DSR or ERAM WCs and reduced the number of keystrokes by more than 50% (see Figure 21). When we dropped the team (that had some of the shorter scenario durations) to look into the effect of our manipulations for higher Traffic Levels, we encountered a three-way interaction between Staffing Configuration, WC, and Team, $\Lambda = .04$, $F(8,3) = 9.90$. The main effect of WC was still present but interacted with the effects of Staffing Configuration and Team.

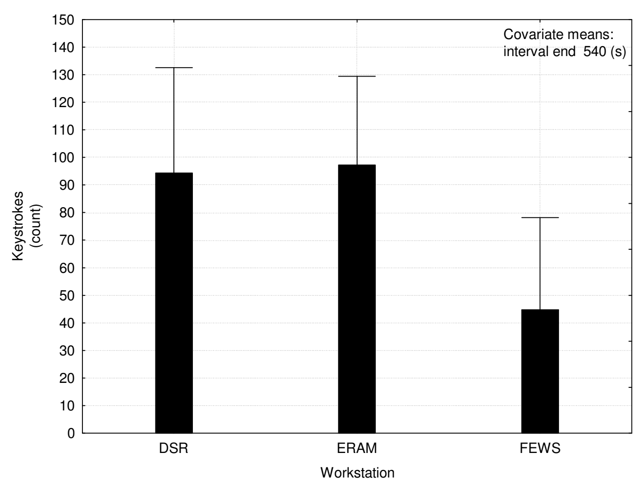


Figure 21. Adjusted mean number of keystrokes as a function of Workstation Configuration ($n = 6$, intervals through 14 minutes into the scenario).

In Figure 22, there is a large variability in the number of keystrokes between controllers but a consistently lower number for all teams when comparing the FEWS interface with either the DSR or the ERAM interface. When we compare the difference between two-person sectors and one-person sectors, there is a slight elevation in the number of keystrokes across all conditions reflecting the fact that the R-side controller for ZGN08 no longer has the assistance of the D-side controller and, consequently, needs to enter more data.

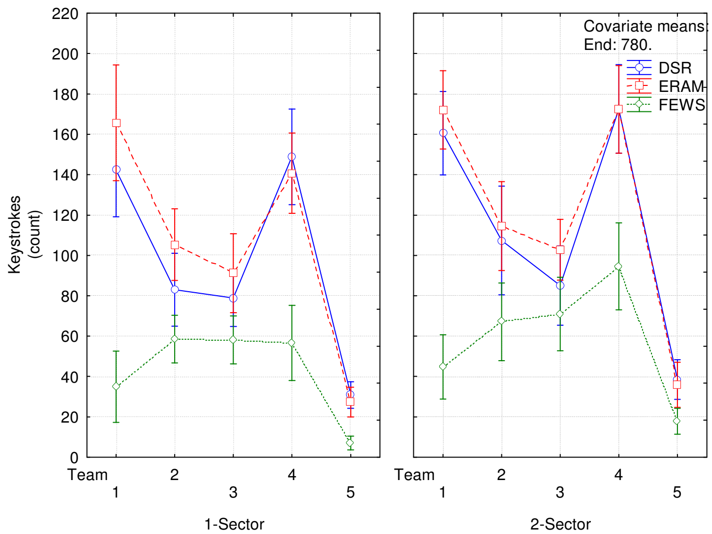


Figure 22. Adjusted mean number of keystrokes as a function of Workstation Configuration (n = 5, intervals through 24 minutes into the simulation).

The analysis of the use of the pointing device interactions under the different WCs revealed that controllers used the wheel-mouse significantly more often under the FEWS configuration, but that differed somewhat depending on Staffing Configuration, $\Lambda = .24$, $F(2,5) = 8.09$ (see Figure 23). Although Staffing Configuration itself did not have a significant impact on the number of pointing device clicks, the two-sector configuration resulted in a higher number. The difference between a two-person sector and a one-person sector was significant only under the FEWS condition. In the one-person sector, the ZGN08 controller no longer had the assistance of the D-side controller and, consequently, used the pointing device more. When using the FEWS configuration, controllers used the pointing device more because we created easy access to functions using the pointing device that normally required input that included the keyboard. We repeated the analysis on 5 participants and up to 22 minutes to determine whether the effect of WC and Staffing Configuration persisted, but found a three-way interaction between CC, Staffing Configuration, and WC, $\Lambda = .30$, $F(2,9) = 10.69$. This interaction showed that when only Voice Communications was available, the number of pointing device interactions of the R-side controller on sector ZGN08 did not change, depending on whether a single controller or a two-person team controlled sector ZGN08. However, when Data Communications was available, having a two-person team available resulted in a lower number of pointing device interactions for the R-side under the ERAM and the FEWS WCs. The availability of Data Communications provided the D-side the opportunity to alleviate some of the workload of the R-side by taking over some of the Data Communications clearances. This in turn resulted in fewer interactions through the pointing device for the R-side controller.

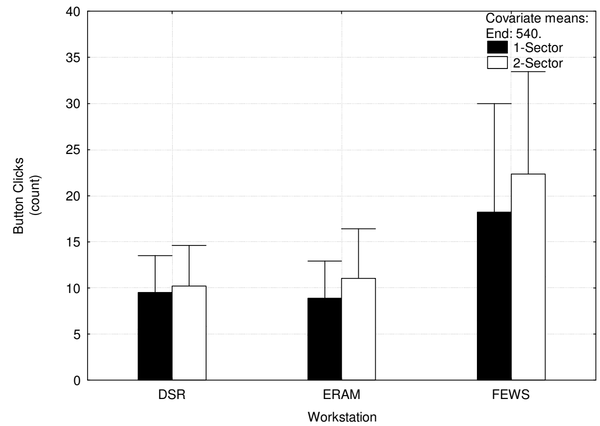


Figure 23. Number of button clicks on pointing devices as a function of Workstation Configuration and Staffing Configuration.

3.2.2 Subject Workload Assessment Keypad Data

For our statistical analysis, we treated the number of aircraft as a varying covariate. We compared WCs using controller workload to see which workstation would give less workload to controllers. Because we allowed controllers to stop the run when they perceived that they could no longer control air traffic safely, the lengths of the experimental runs varied. To compare WCs in the similar ATC complexity, we chose the shortest duration of experimental runs and analyzed all the experimental data up to that time. Team 2 had the shortest run when they stopped controlling air traffic at 14 minutes after the start of the experiment. The experimental condition for the run was FEWS workstation, Data Communications available, and a two-person sector. The maximum number of aircraft at 14 minutes was 23 aircraft. Because controllers rated their workload every 2 minutes, there were 7 data points for both workload ratings and aircraft count for each experimental run. The workload ratings in our graphs represent the projected ratings at an average number of aircraft (approximately 16 aircraft).

In comparing Staffing Configurations, we analyzed only the R-side data in the two-person sector condition and ZGN08 data in the one-person sector condition. The repeated measure ANOVA results showed that Mauchy's test of sphericity of the WC variable was significant (Mauchy's $W = .631$). Multivariate F-test results with the adjusted degrees of freedom showed that the variables we used did not contribute to the model significantly. The only meaningfully significant result was the interaction effect among Staffing Configuration, Data Communications, and Staffing Configuration, $\Lambda = .735$. Within-subjects tests showed significant interactions between Staffing Configuration and CC, $F(1,29) = 4.196$; and another among Staffing, Communication, and WC, $F(3,58) = 3.684$.

The results of the estimated marginal means of ratings show that in the two-person sector condition, R-side controllers rated their workload higher when they did not use Data Communications for all Staffing Configurations (see Figure 24). The order of their workload ratings from the lowest to the highest was DSR, ERAM, and FEWS when Data Communications was available. When it was not available, the order was DSR, FEWS, and ERAM (see Figure 24).

In the one-person sector condition, R-side controllers' ratings did not show large differences across WCs (see Figure 25). Data Communications seemed to help them when they used DSR and FEWS designs, but not when they used ERAM.

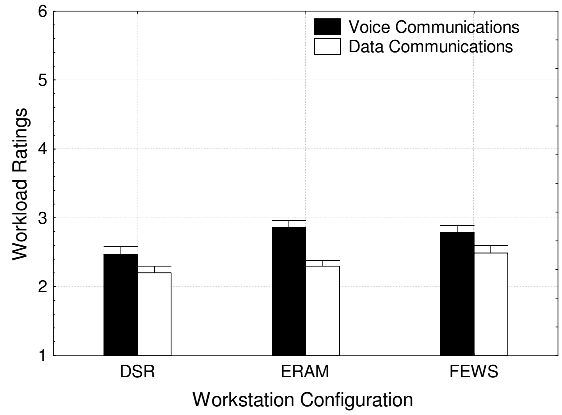


Figure 24. Marginal means of ratings for different Workstation Configurations by Communications Configurations for the two-person sector condition for data up to 14 minutes.

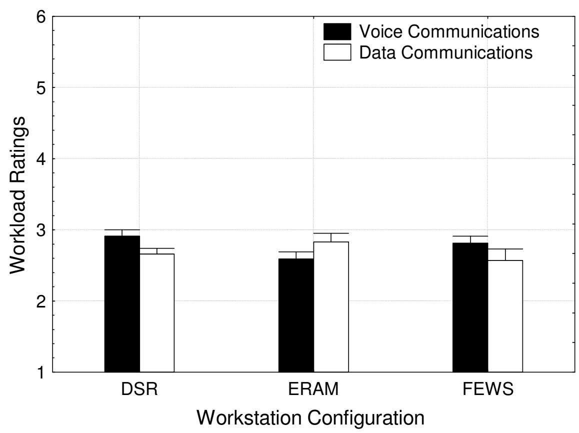


Figure 25. Marginal means of ratings for different Workstation Configurations by Communications Configurations for the one-person sector condition using data up to 14 minutes.

The length of the next shortest run was 24 minutes; therefore, the WAK response data for the probe at 24 minutes fell outside of our observations. When we extended the range of data up to the 22 minutes that excluded Team 2 data, the results were not much different (see Figure 26). The range of aircraft count at 22 minutes was between 16 and 28 aircraft. The major factor comparisons in both multivariate and Between-Subject analyses did not show any significant results. As evident with the two sets of Estimated Marginal Means, the patterns did not show a clear distinction between the 14 minutes and 22 minutes experimental run lengths (see Figure 27). Table 5 presents the results of the regression of workload ratings onto the number of aircraft for which the controllers were responsible. In Table 5, the final column shows the number of aircraft that we estimated using the regression models at a workload rating of 6.

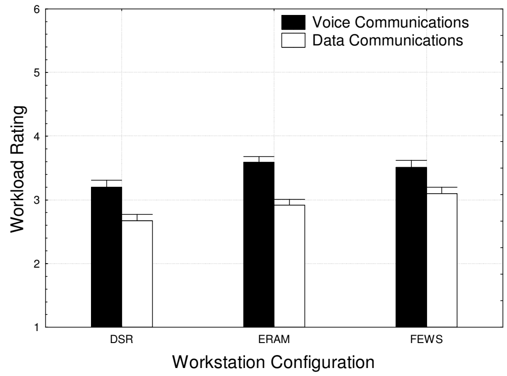


Figure 26. Marginal means of ratings for different Workstation Configurations by Data Communication Configurations for the one-sector condition for data up to 22 minutes, excluding Team 2 data.

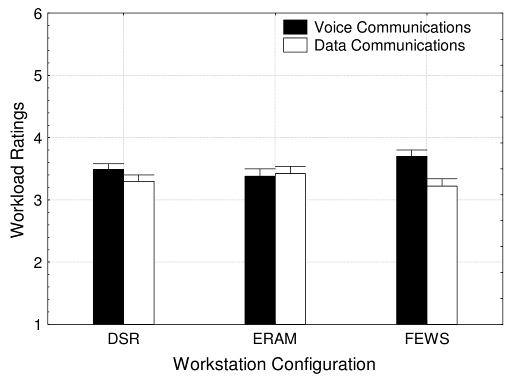


Figure 27. Marginal means of ratings for different Workstation Configurations by Data Communication Configurations for the two-sector condition for data up to 20 minutes, excluding Team 2 data.

Table 5. Intercepts and R^2 for all the Conditions

Staffing	Communication Configuration	Workstation Configuration	R^2	Intercept	Slope	Aircraft Count at Workload Rating 6
Two-person sector	VC	DSR	.589	7.438	3.047	25.7
		ERAM	.647	3.589	3.935	27.2
		FEWS	.655	6.721	3.162	25.7
	DC	DSR	.741	7.083	3.999	31.1
		ERAM	.692	5.140	4.296	30.9
		FEWS	.673	6.358	4.144	31.2
One-person sector	VC	DSR	.796	2.653	3.778	25.3
		ERAM	.712	6.211	3.132	25.0
		FEWS	.694	5.819	3.170	24.8
	DC	DSR	.770	5.207	3.244	24.7
		ERAM	.714	4.487	3.659	26.4
		FEWS	.653	6.449	3.219	25.8

Note. VC = Voice Communications; DC = Data Communications.

The availability of Data Communications increased the number of aircraft by about five (or 19%) over the number controllers could comfortably handle with Voice Communications only (see Figure 28). Note also that the number of aircraft controllers handled at that workload level without Data Communications is also five aircraft (or 24%) higher than the Monitor Alert Parameter (MAP) value we had assumed for this sector.

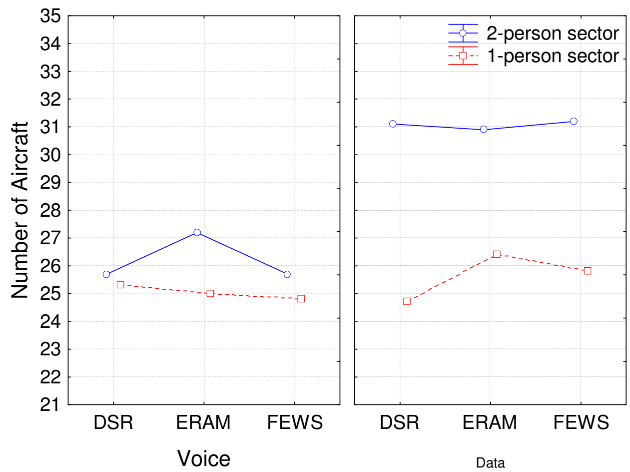


Figure 28. Aircraft count as a function of Communication Configuration and Workstation Configuration.

3.3 Event-Based Analyses

Some of the data used for this experiment are in a format that is very different from that in previous experiments conducted at the RDHFL. In this study, our interest is in making the CHI more efficient to reduce controller effort and increase the speed of data access and entry. To determine in what areas we have accomplished that goal, we measured performance with respect to individual event types. We have defined events based on cognitive job and task analyses conducted by Ammerman et al. (1987) as well as work done at EUROCONTROL by Dittmann et al. (2000). Ammerman identified four events that a controller can experience.

1. Clearance delivery
2. Clearance request
3. Visual flight rules
4. Amend altitude/route and destination

In the FEWS II scenarios, the controller encountered only the "amend" event. Ammerman et al. (1987) defined the amend altitude→route→destination event as follows:

> An amended altitude, route, or destination may be requested for any reason. It generally is used to obtain more direct routing, or to avoid weather/turbulence, or for fuel economy. The new route may be airway, NAVAID-direct-NAVAID, RNAV, radar vectors, or coordinates (latitude-longitude) to coordinates. Amended routes also may arise from a change in destination due to pilot request or contingencies. (p. A-9)

For our study, we included speed amendments as well as distinguished how controllers make the amendments. For example, an altitude amendment can consist of an interim altitude change, and the controller can enter such a change in several ways, such as using a keyboard or a pointing device, or through Voice Communications or Data Communications. We will use these events to determine which processes a controller has activated at a particular point in time.

3.3.1 Controller Commands and Entries

Although we have not completed the full formal analyses of the data sets that we have available, we will provide initial results in this section. With the data sets that we have available, we can analyze controller interactions with the system at many different levels. One of the research questions that we address in this study is how our experimental manipulations have affected controller task characteristics. Our approach has been to start at the subtask level and work our way up towards tasks that are more complex. To take advantage of the many observations that we have available per subtasks, we prefer to use each instance of a subtask as an observation instead of using mean values. Using mean values instead of the individual observation data would lose the detailed data available and would make some assumptions (e.g., that the mean represents the data for a given sample because the underlying assumption is that we have a normal distribution of values). The experiment had a repeated measures design, so we prefer to take advantage of the fact that each of our participants experienced the same conditions. Because of the dynamics of the simulations, however, the control of the traffic sample will be different between and even within controllers, resulting in an unequal number of observations between conditions. These observations themselves are not of a repeated nature (i.e., the fact that a given subtask executed on an aircraft under one condition does not correspond does not necessarily have an equivalent subtask on that same aircraft under another condition). Because we were unable to take full advantage of the data in a repeated measures design, we approached the data from three directions to circumvent the issue of having very large quantities of detailed data. First, we very conservatively looked at the data without taking advantage of the repeated measures design, knowing that we violated the normal distribution assumption. Second, we looked at the frequency distributions of our data sets to understand how our experimental manipulations changed the nature of the ATC subtasks. Third, we looked at the detailed observations to understand why subtasks may have changed.

For the first step, we looked at the duration of the data entry subtask (see Figure 29). Controllers interacted with the system either by requesting data from or by entering data into the NAS. Depending on the modality used for the data entry and the concurrent activities during data entry, the duration of the data entry varied. We calculated the average data entry duration for six subtasks: Assigned Altitude (QZ), Coordinated Speed and Heading (QS), Flight Plan Readout (QF or FR), Interim Altitude (QQ), Route (QU), and Reference Distance Indication (QP, J or HALO). We then treated these entry duration averages as separate observations for each of the teams, and we treated CC and WC as repeated measures. We found that with Data Communications available, the data entry durations increased significantly, $\Lambda = .51$, $F(1,35) = 33.75$ ($M = 2039$, $SE = 133$ vs. $M = 2493$, $SE = 155$ msec.). The data entry duration also varied with WC, $\Lambda = .67$, $F(2,34) = 8.42$. The data entry duration in the FEWS WC, $M = 2570$, $SE = 185$ msec, was longer than the entries in DSR or ERAM, $M = 2205$, $SE = 144$ vs. $M = 2022$, $SE = 132$ msec.

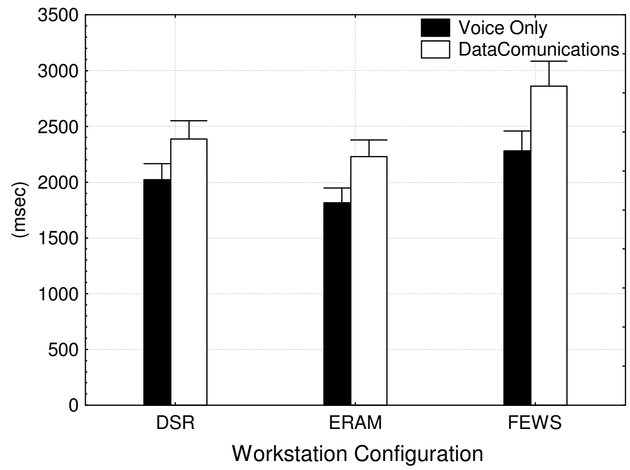


Figure 29. Data entry duration as a function of Communication Configuration and Workstation Configuration.

Given the strength of the effects, we analyzed individual data entry types. Our analyses revealed that the CC variable affected only data entries that involved a clearance delivery component (QU, QS, QQ, and QZ). For example, a route-related entry took longer when Data Communications was available, $\Lambda = .38$, $F(1,5) = 8.29$ ($M = 2458$, $SE = 144$ vs. $M = 2881$, $SE = 186$ msec). For some entries, an effect of WC either accompanied or interacted with the effect of CC (see Figure 30). For example, for interim altitude entries, the effects interacted, $\Lambda = .21$, $F(1,35) = 33.75$ ($M = 2039$, $SE = 133$ vs. $M = 2493$, $SE = 155$ msec, for Voice Communications and Data Communications, respectively).

The QU and QQ entries have shorter durations because controllers can use each of these entries with or without a new value, whereas the QS and QZ entries have to contain a newly assigned value. A QU entry without a value will toggle the display of the route for an aircraft, whereas a QQ entry without a value will remove an existing interim altitude value from the data block, exposing the assigned altitude.

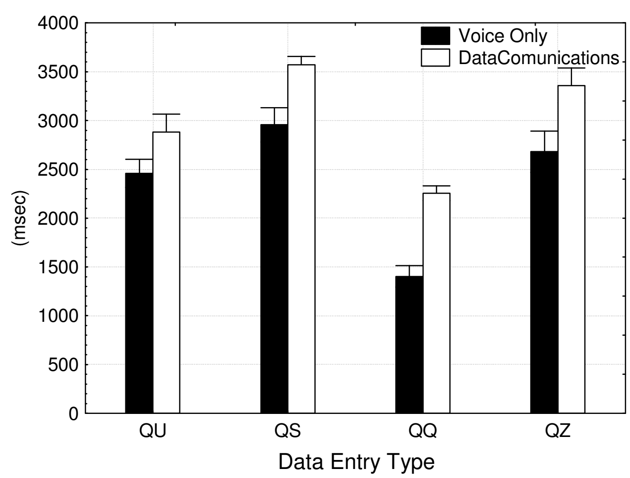


Figure 30. ATC clearance data entry durations as a function of Communication Configuration.

For non-clearance ATC data entries, such as QF, QP, Initiate Handoff (IH), Accept Handoff (AH), and Data Block Movement (MD), we did not find an effect of CC on data entry durations. We have plotted their means and standard errors in Figure 31. The durations for non-clearance data entries are much shorter than clearance-related entries. Most of the non-clearance data entries do not contain a new value for the entry. In Figure 31, the two non-clearance entries that require an additional value are the QP (a "J" to indicate a distance reference indicator request) and the IH (two digits to indicate the next sector identification) entries.

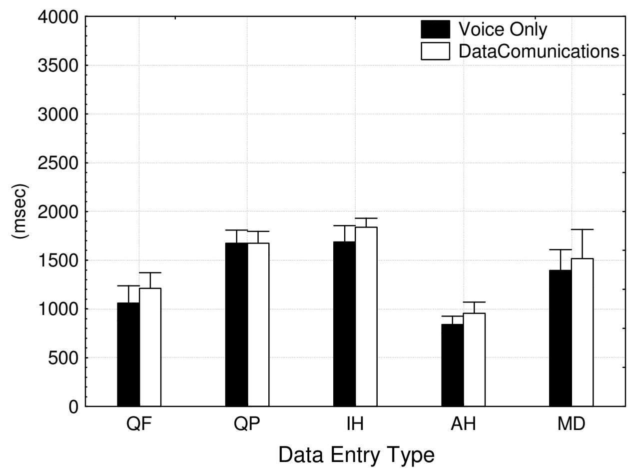


Figure 31. Non-clearance data entry durations as a function of Communication Configuration.

The use of different WCs had no effect on QF, QP or QU, and MD entries. For QZ, QQ, and QS entries, the data entry durations increased when controllers were using the FEWS workstation, whereas the data entry duration for AH and IH decreased. To understand why and how our changes to the workstation affected data entry duration, we looked at frequency distributions of the number and duration of these events. The distributions show that, although our analyses are quite robust with regard to violations of the normal distribution assumptions, our data sets do not have normal distributions. When we look at the distributions, we can easily identify what happens within the data entries both as a function of our experimental manipulations and within conditions. For example, in Figure 32, we plotted the relative distribution of data entry durations for flight plan readouts under Data Communications conditions. The distribution points out two different ways of obtaining a flight plan readout. First, we see a very short duration distribution below 0.4 seconds. Controllers that use a single click on the call sign in the data block can get a quick flight plan readout in the RA on their display. Other controllers use the keyboard only by using a quick-action key entry followed by the CID, but that takes longer and has a wider distribution from 1-2 seconds. It is of course possible to use a combination of keyboard and pointing device entries with some hesitation between keystrokes and button clicks accounting for the long tail in the distribution. We notice also that, mostly for the FEWS entries, there are a larger proportion of entries that last approximately 300-500 msec. Our data now enable us to further investigate why this has occurred. We use a flight plan readout entry here as an example, but we will look at an operationally more interesting example next.

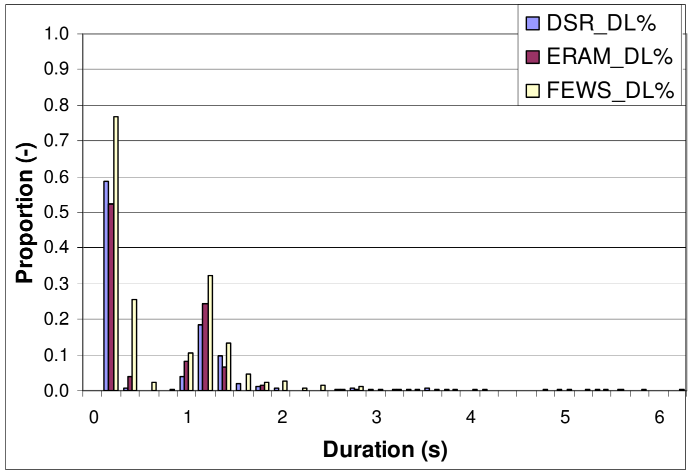


Figure 32. Distribution of flight plan readout entry durations for Data Communications conditions as a function of Workstation Configuration.

Using the results from the repeated measures ANOVA, we found that altitude entry durations took significantly longer when using the FEWS WC than either DSR or ERAM, $\Lambda = .11$, $F(2,4) = 16.80$. From Figure 33, we see that the distribution is not normal and that the distribution indeed has moved slightly over to the right-hand side, resulting in the longer durations that we had found for Data Communications over Voice Communications. We also see an increase in longer duration entries for the FEWS WC. The way we currently collect and process our data enables us to take a more detailed look at those instances where controllers took longer to make an altitude entry under the FEWS conditions.

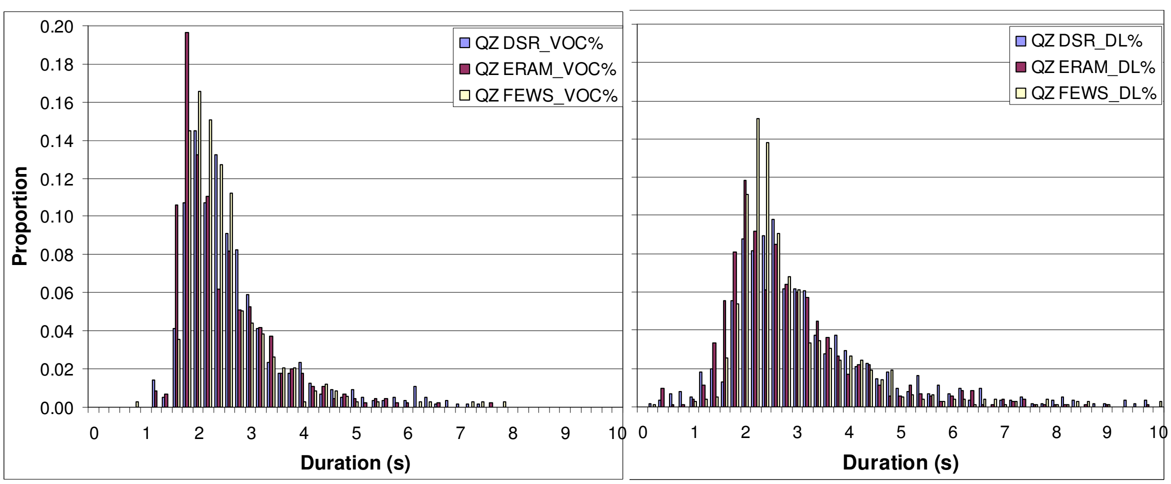


Figure 33. Distribution of altitude entry durations as a function of communications and Workstation Configuration.

4. DISCUSSION

During this complex Human-in-the-Loop simulation, we collected a wealth of data that we can use to look at many aspects of controller and system performance. This report can provide only a glimpse of the information we have obtained from the controllers that participated in the study. We have broken down the discussion of the results in the same way that we have conducted our analyses. We start with discussion of scenario-based observations, followed by a discussion of the interval-based results, and finally we discuss the event-based analysis we have conducted. We will end with our thoughts about future analyses and provide conclusions and recommendations.

4.1 Scenario-Based Observations

The reduction in perceived performance when Data Communications was available corresponded to the level of familiarity that controllers had with the different WCs (i.e., ERAM changed some functions but was much more like the familiar DSR display than the FEWS configuration).

The R08 controllers perceived a reduction in performance between the two-person and one-person Staffing Configurations because in the one-person sector, the controllers no longer had the assistance of the D08 controller and therefore worked the same traffic levels by themselves. The increase in perceived performance of the controllers that switched from the D08 to the R18 function when they moved from the two-person sector to the one-person sector occurred because they perceived that they were “only” assisting the R-side controller and therefore could not do as much for the R-side controller. As a result, when controllers worked the D08 function, they rated their performance lower than when they had worked as the R18 controller.

The availability of Data Communications reduced the perceived communication workload dramatically. We found a trend that indicated a different impact of the availability of Data Communications for the controller who functioned either as the D-side of sector S08 or as the R-side of ZGN18. The reduction in communication workload was less for that controller when Data Communications was available. In this experiment, the controller who worked the right-side position swapped functions, depending on the Staffing Configuration. When they operated ZGN18 as the R-side controller, they experienced lower traffic levels than the controller who operated ZGN08. When they operated as the D-side controller for ZGN08 and assisted the R-side, one would expect them to experience lower communication workload levels because D-side controllers normally do not communicate with pilots. So, when only Voice Communications is available in both functions, the right side controller would experience lower communication workload levels. However, when Data Communications is available, the right-side position will experience lower communication workload when operating as R18 because of the lower traffic levels in ZGN18. When Data Communications is available and working as D08, the right-side position will experience a communication workload that is still lower than the R08 position. Now that the D08 position can send Data Communications clearances to the aircraft, this will increase communication workload compared to Voice Communications conditions.

One would expect a reduction in SA for current aircraft position because of the increased number of tasks that the R08 controller needs to execute when working ZGN08 as a single-person sector. The increased cognitive difficulty of the ATC task will, for example, narrow the functional field of view (Williams, 1982, 1995), possibly resulting in a reduction of SA for current aircraft positions.

Although none of the simple effects had significant results, it is interesting to note that the D08/R18 controllers perceived that the effect of the interface on their control strategy changed between the one- and two-Staffing Configuration for DSR and ERAM than for FEWS. In the DSR and ERAM WCs, these controllers switched from a D-side display to a radar display in a similar fashion as they would in the field. The interface on these displays is very different in terms of display real estate but mostly in terms of the information and functionality available on the displays. As a result, the controllers have to change the control strategy from the controller in charge of the sector (as R18 in the one-person sector) to the D08 controller assisting the R08 controller in the two-person sector. When using the FEWS workstation, that effect of the interface on control strategy is not present in the Staffing Configuration because the controller has identical display capabilities independent of whether they work in the D08 or the R18 role.

Although we trained controllers on the changes in the ERAM and FEWS II environments, their participation in the study lasted only 8 days. During that time, they learned to use Data Communications, the ERAM changes to the interface, and the changes in the FEWS II concept of operations. We noticed that the controllers rated the FEWS workstation with Data Communications less easy to use than the other configurations.

We asked controllers to indicate how much they agreed with our statement that the displays were clear and uncluttered. The general ratings (between 6.5 and 8 on a 10-point scale) indicate that with the level of traffic that some of our controllers were working, none of the systems could maintain a clear and uncluttered display.

One of the focal points in this experiment was on the effect of one- vs. two-person staffing of S08. The R08 controllers worked scenarios until they reached a point where they perceived they could no longer safely and efficiently control traffic. The mental demand for the R08 controllers did not differ at the end of the simulation runs (independent of whether the controllers worked in a team of two controllers on a single sector or worked by themselves) because the R08 controllers continued to work until they perceived that the sector had become saturated. Therefore, the NASA TLX results on mental demand do not provide an answer on the effect of SS on sector productivity.

Compared to the baseline DSR condition with Voice Communications only, we saw a large reduction in the total number of deviations when controllers worked as a two-person team and had Data Communications and FEWS available. We expected to see deviation because controllers will, once they experience very high workload levels, start task shedding. The controllers' primary responsibility is to maintain separation. Safety is their number one concern. So, when their workload levels increase to a point where they can no longer complete all their tasks, the tasks that are not safety-critical will suffer first. For example, where controllers prefer to have FDBs separated to be able to see all information available about aircraft, during high workload levels they may leave overlapping FDBs and focus their resources on separating aircraft. When workload levels further increase, controllers will shed tasks that may lead to deviations, but not loss of separation. The changes incorporated in the FEWS II environment removed some of the housekeeping tasks and assisted in some of the tasks that could lead to deviations. As a result, in conjunction with the presence of Data Communications, controllers had fewer deviations. The changes in the FEWS II environment that we created specifically to help controllers to prevent deviations included automatic handoff acceptance, confirmation of transfer of communications on aircraft capable of Voice Communications only, and automatic transfer of communication on aircraft capable of Data Communications. To allow an aircraft to enter a sector's airspace without accepting the handoff of radar control from the previous sector constitutes a deviation. If a controller cannot accept a handoff from the previous sector, the controller must coordinate with the previous sector. In the FEWS II environment, we had instructed controllers that the automation would detect potential conflicts and only accept aircraft if no projected, potential conflict existed. The controllers sometimes accepted a handoff from the previous sector manually, but most of the time they let the system accept the handoffs. In a similar fashion, we also used the automatic handoff feature to transfer communications to the next sector. The feedback from the controllers indicated that the system may need to delay the transfer of communications closer to the sector boundary, but the combined automatic handoff and transfer of communications prevented controllers to have Data Communications-equipped aircraft to cross the sector boundaries without a transfer of communications. For aircraft that had only Voice Communications available, we introduced a feature that brought back the paper FPS annotation to indicate which aircraft were currently on the frequency. The FEWS II environment showed the controllers which sector had an aircraft on its frequency as soon as the next sector accepted radar control. So, if a controller scanned the edges of the sector and noticed an aircraft that still had the sector indicator in the voice frequency area, they realized that they needed to switch the aircraft to the next sector manually.

4.2 Interval-Based Results

ERAM holds the promise to reduce controller entries through its macro-functionality, but we did not find evidence of a reduction in the number of keystrokes for our ERAM emulation. Similar to our findings in FEWS I, we found a large reduction in the number of keystrokes when controllers used the FEWS workstation. The focus in our concept has been in reducing menial tasks or making them more efficient. Automating handoff acceptance and dropping FDBs, as well as the ability to move FDBs by dragging them to new positions, contributes to this reduction. Expressed in the number of keystrokes per minute, the FEWS environment reduced the data-entry rate from 50 keystrokes per minute to approximately 25 keystrokes per minute. Handoff acceptance and FDB drop-off automation removes many of the keyboard entries, where the more flexible ability to move FDBs shifts entries from the keyboard to the pointing device.

We corrected for the number of aircraft for which the controller was responsible, and we confirmed our results with the results of Willems, et al. (2008). We found that the availability of Data Communications reduced controller workload across WCs when two controllers worked a single sector. When we split the team and had only a single controller working ZGN08, there was an increase in workload under the FEWS condition when only Voice Communications was available.

To determine what the effect of Data Communications and workload had on the number of aircraft that controllers could handle, we conducted a regression analysis under each of the conditions and determined the number of aircraft that would result in a workload rating for ZGN08 of 6 on a 10-point scale. At a 5-6 workload level, the controllers had no spare time remaining, but still maintained the picture. The results of our analyses show that with the assistance of a D-side controller and only Voice Communications available, controllers had level 6 workload ratings at approximately 26 aircraft (a 24% increase over a MAP value of 21). When Data Communications was available, controllers had a level 6 workload rating at 31 aircraft (a 48% increase over a MAP value of 21). The number of aircraft at that workload level did not differ across Staffing Configurations. These results do not mean that we can expose controllers to 24% more traffic than the MAP value. At a workload level of six, controllers are at a turning point. They can still maintain the picture, but there is no spare capacity left. In the field, this happens on occasion as well and the MAP value protects controllers from having to work under these traffic loads for very long. We confirmed the results that we had obtained from the Willems et al. (2008) experiment, about the impact of Data Communications on a two-person sector. With Data Communications, the controller team can work more traffic than the MAP value. One of the reasons that we had created traffic scenarios that had an increasing number of aircraft over time was to fill in the gaps between the discrete levels we had tested in the FEWS I experiment (Willems et al., 2008). We have found that a two-person controller team reaches a saturation point at 48% above the MAP value. We also investigated how the CC and WC affected a one-person sector. Similar to a two-person team, controllers reached their saturation point beyond the MAP value. Contrary to our expectations, Data Communications did not increase the number of aircraft that a single controller could work at the workload threshold. Therefore, our positive findings for the increase in capacity when Data Communications is present for the two-person sector rely on the availability of the D-side controller. With Data Communications available, the D-side can offload the R-side more effectively, resulting in an increase in sector capacity that a one-person sector cannot achieve.

4.3 Event-Based Analyses

We found an increase in data entry duration when using Data Communications. In the current NAS, controllers must provide pilots with a voice clearance first, before updating the NAS. Controllers have coined failing to do so as "pre-loading the data block." Although controllers do not consider the incorrect sequence of events a deviation, it does pose a performance issue. The display must reflect what the controller expects the aircraft to do. As a result of the prescribed sequence of actions when using Voice Communications, controllers will mentally prepare a clearance, and then transmit the voice clearance to the pilot. The cognitive processing and preparation of the clearance has taken place even before the controller executes the voice clearance, and the data entry takes place in a fast and pre-programmed manner. When using Data Communications, however, the preparation and cognitive phase may spill into the data entry phase of the clearance, slowing down data entry enough to result in an observable increase in data entry duration. To test that this hypothesis is true, we had to find anomalies that we had seen in the behavior of our controllers. Some controllers will quite often violate the prescribed sequence of actions when using Voice Communications. If our hypothesis is true, we should be able to show those data entry times are longer for cases where controllers made the data entry first, followed by a voice clearance, because the preparation and cognitive phase may spill into the data entry phase as well.

The findings show that with Data Communications, the structure of some of the controller tasks has changed beyond a reduction in Voice Communications. The longer durations of the data entry tasks involved in providing ATC clearances are a clear indication of that change. This does not mean that the next higher-level task of clearance delivery has increased in duration. The sequential nature of the voice clearances makes it very likely that the overall clearance delivery by voice will take considerably longer than its Data Communications counterpart. The analysis of the higher-level tasks is beyond the scope of this initial report.

5. GENERAL DISCUSSION

The use of traffic samples that increased the number of aircraft over time made the data analyses more challenging. In our simulated environment, controllers worked traffic well beyond traffic counts that are operationally acceptable before they reached our subjective workload threshold. As Hah, Willems, and Phillips (2006) discussed, the subjective workload threshold is an average value; an unacceptable proportion of the controller population would experience higher workload at these traffic levels. The MAP values in use in the operational environment most likely reflect a buffer that will enable controllers to sustain traffic levels higher than the MAP value for a brief amount of time without losing the picture.

Some argue that controllers will never experience traffic levels as high as the ones encountered in our experiments. That would be true if the NAS continues to operate as it currently does. In the current environment, traffic levels would increase until predicted levels start to exceed the MAP value for a sector. At that point, the traffic management unit of the affected ARTCC will start to divert traffic around the congested sectors to spread out the traffic load across multiple sectors. When adjacent sectors reach their MAP values as well, national programs such as the Ground Delay Program take effect. When these temporary measures occur frequently, the FAA may have to limit the number of aircraft in or out of an airport.

To increase capacity between city-pairs, a given volume of airspace will have to be able to accommodate more aircraft. In the current sector-based operations, this is possible only through an increase in the MAP values. The MAP values originated in staffing models that accounted for the time controllers spent per aircraft and set limits to the number of aircraft that controllers could accommodate within a 15-minute interval. To accommodate an increased number of aircraft within a 15-minute interval while keeping the staffing levels constant, the NAS needs to reduce the amount of time controllers spend per aircraft.

The current experiment and Willems, et al. (2008) have shown that the controllers cannot support more than 31 aircraft under control. This traffic level corresponds with 47% more than the MAP value when we assume that the sector can accommodate (the maximum number of aircraft, 18, plus an additional 3 aircraft based on a well structured flow of traffic). Researchers that investigate future scenarios that have traffic levels of 3X or 63 aircraft in a similar volume of airspace have equated our results with controllers not being able to accommodate future demands. In their opinion, automation is necessary because the human operator is the *bottleneck* that stands in the way of increasing the NAS capacity. Proposed solutions to circumvent the human bottleneck in these future scenarios would include automated conflict detection, conflict resolution, and sequencing resolution. The danger when changing the human operator from an active ATC participant to a traffic manager, or monitor of the traffic situation, is that controllers will not have the level of SA needed to intervene when the system needs human assistance.

One of the arguments against placing controllers in an air traffic manager position and taking them out of the loop is that the NAS currently is as safe as it is because of the presence of human controllers. Proponents of the introduction of advanced automation systems often neglect the fact that complex systems often are resilient not despite of but because of the presence of human operators. Removing the human controllers from their active role in the NAS may remove the resilience of the NAS to recover from unanticipated circumstances.

Our interpretation of the results is that we need to find ways to extend the capabilities of human operators to accommodate increased traffic levels while maintaining a human-centered system. Therefore, our approach would create an environment that provides controllers with automation functions that support controllers to maintain SA at acceptable workload levels and enables controllers to step in when the system needs human assistance. The experiments so far have shown that workstation changes have reduced the use of controller resources and that Data Communications has increased sector capacity, but these changes alone are not enough to support the projected 3X capacity levels. Changes to the workstations (including automation functions) need to create an environment that reduces the amount of time spent per aircraft, supports SA, and enables recovery from non-nominal conditions.

Although controllers indicated that they did not perform as well when using the FEWS II configuration, their workload ratings and the amount of time they were able to work traffic did not change. Therefore, the unfamiliarity with the FEWS II workstation affected subjective performance but did not affect objective performance.

6. CONCLUSION

The data we collected during this experiment enabled us to look at many aspects of controller performance and behavior in detail. We have provided some of our results in this report with the caveat that to determine the impact of Staffing, Communications, and WCs on performance and behavior, we must look at their effect on several variables simultaneously. An example is the impact of WC. We expected that controller workload would diminish with the FEWS II interface, but we detected a slight increase instead. On the other hand, objective data shows that the FEWS II environment reduced the number of controller deviations and the number of keystrokes. If the increase in workload in the FEWS II environment is necessary to prevent the operational deviations, this may be acceptable as long as the workload is within acceptable levels. As an example related to the impact of CC, we found that a two-person sector was able to handle more aircraft when Data Communications was available than with Voice Communications only. However, Data Communications did not have that impact when we staffed the same sector with a single controller. Not only does the D-side controller seem to be necessary to take advantage of the presence of Data Communications, but our subjective data also shows that without the D-side controller, the R-side controller's workload is higher and SA is lower.

References

Ammerman, H. L., Bergen, L. J., Davies, D. K., Hostetler, C. M., Inman E. E., & Jones, G. W. (1987). *FAA air traffic control operations concepts − Vol. VI. ARTCC/Host en route controllers* (DOT/FAA/AP/86-01). Washington, DC: Federal Aviation Administration.

Andriole, S., & Adelman, L. (1995). *Cognitive systems engineering for user-computer interface design, prototyping, and evaluation.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Dittmann, A., Kallus, K. W., & Van Damme, D. (2000). *Integrated task and job analysis of air traffic controllers −Phase 3: Baseline reference of air traffic controller tasks and cognitive processes in the ECAC area* (HUM.ET1.ST01.1000-REP-05). Brussels, Belgium: EUROCONTROL.

Dorbes, A. (1999). *Radar label anti-overlap requirements, EUROCONTROL working paper* (ADW 99001). Brussels, Belgium: EUROCONTROL.

Duverger, A. (2005). *Development of a mathematical weighted formula to eliminate the overlapping of aircraft labels on the ATC radar display* (EEC Note No. 19/05). Brétigny-sur-Orge CEDEX, France: EUROCONTROL Experimental Centre.

ECHOES. (2004). *A human-machine interface for en route air traffic control.* CoRe HMI Specification. Bretigny-sur-Orge, France: EUROCONTROL Experimental Centre.

Endsley, M. R., Sollenberger, R. L., Nakata, A., & Stein, E. S. (2000). *Situation awareness in air traffic control: Enhanced displays for advanced operations* (DOT/FAA/CT-TN-00/01). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Federal Aviation Administration. (2007). *Operational data reporting requirements* (DOT/FAA/Order 7210.55D). Washington, DC: Author.

Giorgi, R. G., Soong, G. P., Woods, R. L., & Peli, E. (2004). Facilitation of contrast detection in near-peripheral vision. *Vision Research, 44*(27), 3193-3202.

Graham, R., Marsden, A., Pichancourt, I., & Dowling, F. (2000). Controller roles: Time to change. In *Proceedings of the 3rd USA/Europe Air Traffic Management R&D Seminar.* Napoli, Italy: EUROCONTROL.

Guttman, J. A., & Stein, E. S. (1997). *En route generic airspace evaluation* (DOT/FAA/CT-TN-97/7). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

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

Mejdal, S., McCauley, M. E., & Beringer, D. B. (2001). *Human factors design guidelines for multifunction displays* (DOT/FAA/AM-01/17). Washington, DC: Office of Aerospace Medicine.

Sollenberger, R. L., Willems, B., Della Rocco, P. S., Koros, A., & Truitt, T. (2004). *Human-in-the-loop simulation evaluating the collocation of the user request evaluation tool, traffic management advisor, and controller-pilot data link communications: Experiment I - Tool combinations* (DOT/FAA/CT-TN-04/28). 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.

Truitt, T. R., & Willems, B. F. (1999). *Implications of reduced involvement in en route air traffic control* (DOT/FAA/CT- TN-99/22). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Willems, B. F., Allen, R. C., & Stein, E. (1998). *Controller visual scanning II: Taskload, visual noise, and intrusions into controlled airspace* (DOT/FAA/CT-TN-99/23). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Willems, B. F., Hah, S., & Phillips, R. (2008). *Future En Route Workstation Study (FEWS I): Part 1 - Evaluation of workstation and traffic level effects* (DOT/FAA/TC-08/14, I). Atlantic City International Airport, NJ: FAA William J. Hughes Technical Center.

Willems, B. F., & Heiney, M. (2002). *Decision support automation research in the en route air traffic control environment* (DOT/FAA/CT-TN-02/10). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Willems, B. F., Heiney, M., & Sollenberger, R. (2002). *Study of an ATC baseline for the evaluation of team configuration: Information requirements* (DOT/FAA/CT-TN-02/17). Atlantic City International Airport, NJ: Federal Aviation Administration, William J. Hughes Technical Center.

Williams, L. J. (1982). Cognitive load and the functional field of view. *Human Factors, 24*(2), 683-692.

Williams, L. J. (1995). Peripheral target recognition and visual field narrowing in aviators and nonaviators. *International Journal of Aviation Pscyhology, 5*(2), 215-232.

Acronyms

ACL	Aircraft List
ADBO	Automatic Data Block Offset
ARTCC	Air Route Traffic Control Center
ATC	Air Traffic Control
ATWIT	Air Traffic Workload Input Technique
CA	Conflict Alert
CAASD	Center for Advanced Aviation System Development
CC	Communication Configuration
CFR	Continuous Flight Plan Readout
CHI	Computer-Human Interface
CID	Computer Identification
CMS	Common Message Set
CP	Conflict Probe
CPDLC	Controller-Pilot Data Link Communications
CRD	Computer Readout Device
CTAS	Center TRACON Automation System
DESIREE	Distributed Environment for Simulation, Rapid Engineering, and Experimentation
D-side	Data-side
DSR	Display System Replacement
ECHOES	EUROCONTROL-Consolidation of HMI for Operations, Evaluations and Simulations
eFPS	Electronic Flight Progress Strips
ERAM	En Route Automation Modernization
ERP	Engineering Research Psychologist
FAA	Federal Aviation Administration
FDB	Full Data Block
FEWS	Future En Route Workstation Study
FLID	Flight Identification
FPS	Flight Progress Strip
GPD	Graphic Plan Display
HCS	Host Computer System
HUG	Host URET Gateway

IV	Independent Variable
KSD	Keypad Selection Device
LCD	Liquid Crystal Display
LOA	Letter of Agreement
MAP	Monitor Alert Parameter
MFR	Multiple Flight-plan Readout
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
OS	Operating System
OTS	Over-the-Shoulder
PGUI	Plan View Display Graphical User Interface
PSQ	Post-Scenario Questionnaire
PTT	Push-to-Talk
RA	Response Area
RDB	Range Data Block
RDHFL	Research Development and Human Factors Laboratory
R-side	Radar-side
SA	Situation Awareness
SME	Subject Matter Expert
SOP	Standard Operating Procedure
SS	Sector Staffing
STARS	Standard Terminal Automation Replacement System
TLX	Task Load Index
TMA	Traffic Management Advisor
TRACON	Terminal Radar Approach Control
URET	User Request Evaluation Tool
WAK	Workload Assessment Keypad
WJHTC	William J. Hughes Technical Center

Appendix

Proposed FDB Offset Algorithm

Proposed FDB Offset Algorithm

The Distributed Environment for Simulation, Rapid Engineering, and Experimentation (DESIREE) group developed an algorithm to separate aircraft FDBs that uses a general point-charge force equation. In an aircraft's graphical representation, the center of the FDB and the position symbol serve as point charges. The FDB has a force pushing it away from other FDBs and from other aircrafts' position symbols while it has a force pulling it towards its own position symbol (see Figure). Therefore in the point-charge force equation F= (K*q1 *q2)/r2, F is the force of movement, which is equivalent to distance in my algorithm; K represents a constant; q1 and q2 represent the point charges of either two FDBs or a FDB and a position symbol; and r represents the distance between the two points.

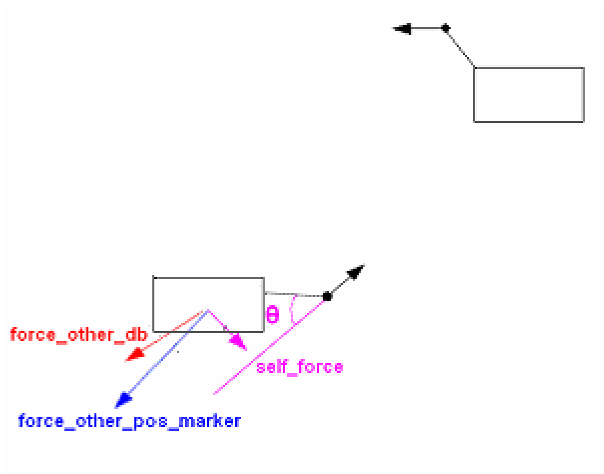


Figure. Forces separating data blocks and attracting data blocks to preferred orientations.

The forces from other FDBs and position symbols use the x and y components of the force vector. The force between the FDB and its own position symbol uses the vector component perpendicular to the direction of travel to try to move the FDB directly in front or behind its position symbol. The reason to have the other-force and the self-force is for balance and prevention of some overlaps between data blocks. Without self-force the FDBs tend to move wildly and generally will not move out of harm's way. The main advantage of the self-force, however, is to place the FDB in a relatively safe area; the area directly trailing or preceding a FDB is the spot the FDB seems to cause the least amount of overlaps.

The sum of all relevant forces determines how and where the FDB will move. The FDBs in the algorithm will only move away from FDBs or position symbols with which they currently overlap or may overlap in the future. If there are two FDBs heading towards each other, the algorithm will calculate only the force between the two FDBs and the self-force and not the force between the FDB and the position symbol. The reason to use only the forces that matter to the FDB is to prevent unnecessary movement.

An important aspect of the algorithm is the prediction of when two FDBs overlap or when a position symbol and a FDB intersect. To determine when two objects overlap, the algorithm looks at the parametric equations of x vs. t and y vs. t and compares points of overlap.

The algorithm will take the four endpoints where an overlap in the x and y axes occur and try to find a time where both the x and y overlap at the same time. In the case of parallel lines as in the y vs. t graph, the algorithm uses some predetermined points in the distant future as infinity. If the parametric equations find that two FDBs do overlap, the algorithm will put the two FDBs and the interval into a list. This list will determine which of the forces are relevant and subsequently calculate the total force. The algorithm regenerates the list at a previously set interval and the process repeats itself indefinitely. Intersections between a position symbol and a FDB follow the same logic.

Once a FDB is no longer overlapping with other FDBs and will not in the future, the FDB may move into a preferred position, but only if this option is set. The FDB will then move back into a preferred position and length at a preset rate. The program reads r and θ from a file.

Within the algorithm, many different settings could affect the outcome. The algorithmic settings include the amount of time to look ahead for overlap, the rate of the updates from the algorithm, a preferred polar speed, a maximum length for the leader line, and a preferred position for FDBs. A combination of these parameters enables us to move data blocks slowly and in small steps to prevent movement from distracting controllers. Most of these settings will not be available for change during the simulations unlike the FDB offset management system proposed by EUROCONTROL - Consolidation of HMI for Operations, Evaluations and Simulations (ECHOES, 2004).