ote technical note techn

Decision Support Automation Research in the En Route Air Traffic Control Environment

Ben Willems, ACB-220
Michele Heiney, Ph.D., Northrop Grumman Information Technology

April 2002

DOT/FAA/CT-TN02/10

Document is available to the public
through the National Technical Information
Service, Springfield, Virginia 22161

U.S. Department of Transportation
Federal Aviation Administration

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

NOTICE

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/CT-TN02/10		

4. Title and Subtitle
Decision Support Automation Research in the En Route Air Traffic Control Environment

5. Report Date
April 2002

6. Performing Organization Code
ACB-220

7. Author(s) Ben Willems, ACB-220 and Michele Heiney, Ph.D., Northrop Grumman Information Technology

8. Performing Organization Report No.
DOT/FAA/CT-TN02/10

9. Performing Organization Name and Address
Federal Aviation Administration
William J. Hughes Technical Center
Atlantic City International Airport, NJ 08405

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

12. Sponsoring Agency Name and Address
Federal Aviation Administration
Human Factors Division, Office of the Chief Scientist
800 Independence Ave., S.W.
Washington, DC 20591

13. Type of Report and Period Covered
Technical Note

14. Sponsoring Agency Code
AAR-100

15. Supplementary Notes

16. Abstract
This study examined the effect of automated decision support on Certified Professional Controller (CPC) behavior. Sixteen CPCs from Air Route Traffic Control Centers participated in human-in-the-loop simulations. CPCs controlled two levels of traffic, supported by either paper flight strips, electronic flight strips with conflict indication, or conflict probe and trial planning automation. Each CPC worked as either a radar or a data controller. The controller station included a display system replacement console and a prototype decision support tool (DST). The station consisted of operational hardware and software connected to an operational Host Computer System (HCS). The HCS received simulated radar from the William J. Hughes Technical Center Target Generation Facility. We assessed CPC behavior by using instruments that assessed situation awareness (SA), workload, visual scanning, trust, and performance. An increase in traffic or automation or working as a data controller lowered SA. An increase in traffic or automation or working as a radar controller increased workload. An increase in automation shifted the focus of the data controller from the radar display to the DST, whereas, an increase in traffic shifted the focus more to the radar display. With an increase in traffic or a decrease in automation, the trust in automation decreased. If subject matter experts indicated a difference between conditions, it involved better CPC performance under lower traffic and more automated conditions. The results have implications for controller training, distribution of responsibilities within controller teams, and air traffic control human-computer interface design.

17. Key Words
Air traffic control
Automation
Decision support
Performance
Simulation
Situation awareness
Visual scanning
Workload

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

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	265	

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

ACKNOWLEDGEMENTS

A project like the Decision Support Automation Research is impossible to conduct without the help of a large number of people. Many others within the William J. Hughes Technical Center had substantial contributions to the success of this project. There are three groups that we would like to mention because, without them, this project would not have survived. The staff of the Integration and Interoperability Facility supported the project by providing their expertise. They enabled us to do things on operational hardware and software that many thought impossible. Our support from the Target Generation Facility was superb. The Target Generation Facility and the Integration and Interoperability Facility made it possible to present the operational hardware and software with data that it took for live air traffic. Finally, we had continuous support from our Simulation Group that created an interface that our controller participants thought was the real Display Replacement System screen. Together, these three teams built an interface that picked up data from the operational network; displayed it on as a remote replication; and allowed us to use that replica to probe controller situation awareness.

We had the good fortune to collaborate with the Research and Development Laboratory Group ACB-840 of the Real and Virtual Environments division, NAS Human Factors Group of the NAS System Engineering Division (ACB-220), and the Civil Aeromedical Institute (CAMI). CAMI provided us with a taxonomy that captured intra-team communications. The data collected through this taxonomy provided further insight into how increased levels of automation affected controller teams. Our collaboration with CAMI was very productive and we are looking forward to continuing such efforts in the future.

We received support from several contract employees during this project. Employees from Titan Systems Corporation Air Traffic Systems Division and from Northrop Grumman Information Technology participated as experimental staff. The experimental staff worked side by side with the Simulation Group and kept our simulations up and running.

During this project, we received contract support as well from Mica Endsley. She has contributed an extensive amount of her knowledge to our introduction about situation awareness and automation. Mica Endsley also made her Situation Awareness Global Assessment Technique available to us to probe controllers for situation awareness.

Finally, we could not have done this experiment without our simulation pilots and controller participants. Our simulation pilots made the simulations so much more realistic. We have a long standing working relationship with our pilots, and we cannot emphasize enough how valuable they are to the success of our simulations. Our controller participants traveled from distant en route facilities to participate in this study. They learned to use an automation tool that they had never seen before and endured our ever-present data collection techniques. We learned a lot from them and understand better now how much there still is to learn.

TABLE OF CONTENTS

TABLE OF CONTENTS (Con't)

Page

LIST OF ILLUSTRATIONS

LIST OF ILLUSTRATIONS (Con't)

LIST OF ILLUSTRATIONS (Con't)

Figure Page

LIST OF ILLUSTRATIONS (Con't)

EXECUTIVE SUMMARY

Improving the current air traffic system with new technologies is necessary to reduce or manage current Air Traffic Control Specialists (ATCSs) workloads, while accommodating continued growth in air traffic. As the National Airspace System (NAS) continues to evolve, ATCSs' roles and responsibilities will change to adapt to the implementation of new, automated tools and procedures. The introduction of automated decision support tools (DSTs) will allow the ATCS to move from tactical control of traffic to more strategic control of traffic. DSTs identify potential conflicts and, in the future, will provide solutions to potential problems. Expected benefits from these new automated tools include increased safety, efficiency, and throughput.

Although research into the effects of automation on job performance exists, we know little about the effects of DSTs on performance of ATCSs. The introduction of DSTs into the Air Traffic Control (ATC) environment will change the way ATCSs work. We have only limited understanding of how much DSTs will affect ATCS strategies and procedures because of 1) limited applied research, 2) limited acquisition-funded studies that address specific risks, and 3) operational field trials without established baselines. Although we expect DSTs to improve safety and efficiency, it is important to investigate the effects of these anticipated changes on the ATCSs' behaviors. Through careful, controlled experimentation, researchers gain information about the possible consequences of changes in the ATC environment. The results obtained from experimentation can then guide development of the most appropriate and effective strategies to offset any detrimental consequences from these changes.

The current study examines the effects of automation and traffic volume on ATCSs in an en route ATC environment. To mimic the predicted increases in air traffic, we manipulated task load based on the volume of air traffic. ATCSs controlled both low and high traffic task loads. We also manipulated the level of automation ATCSs used in controlling traffic. This included three levels ranging from no automation to limited automation to full automation. Further, ATCSs rotated between the Radar (R-side) and Data (D-side) positions.

We anticipated changes in ATCSs' behaviors with the manipulation of automation, task load, and ATCS position. We examined ATCSs' behavior and cognitive processing through objective and subjective performance and behavioral measures. These measures examined ATCSs' situation awareness (SA), workload, visual scanning, communications, and trust.

The execution of this experiment was an achievement and required extensive cross-disciplinary teaming and new operational procedures. To pull together the different disciplines and organizations, we created a crosscutting team that involved test and evaluation, design and engineering, simulation, and human factors groups that spoke a common language and worked together to make things happen. For this experiment, we developed Generic Center Airspace on the Host Computer System and the DST. Although we had used our human factors instruments in other studies, we transferred them to operational hardware and software. Finally, to better assess SA, we developed a new SA assessment instrument called the SA Verification and ANalysis Tool (SAVANT) that used data from the operational systems to probe ATCS SA.

Sixteen ATCSs from Air Route Traffic Control Centers within the United States voluntarily participated in the experiment conducted at the William J. Hughes Technical Center in Atlantic City, NJ. ATCSs performed en route simulations under low and high task load; no, limited, and full automation; and in the R- and D-side positions. ATCSs controlled traffic in a generic airspace to make our findings easy to generalize and to increase the size of our potential participant pool. ATCSs received training on the generic airspace, the use of the DST (including hands-on training), and all equipment used in the simulation prior to experimental runs. We assessed SA using SAVANT, SA Global Assessment Technique, self-report measures, and over-the-shoulder (OTS) ratings by an ATC Subject Matter Expert. We obtained workload ratings from the Air Traffic Workload Input Technique (ATWIT), National Aeronautics and Space Administration's (NASA) Task load Index (TLX), and self-report measures. An eye tracking system collected visual scanning data. We examined ATCSs' landline and air-to-ground communications via push-to-talk software and intrateam communication via coding of digitized audio and video using the Federal Aviation Administration's Controller-to-Controller Communications/Coordination Taxonomy. ATCSs expressed their level of trust in the DST on self-report questionnaires. Post-Scenario Questionnaires provided self-report data from the ATCSs, and OTS ratings provided subjective performance data.

ATCS SA was lower with limited automation than it was with or without full automation. The subjective ratings on SA for potential violations and our objective SA measures showed that ATCSs displayed lower SA with full automation and under high traffic task load conditions. This finding shows that the benefit that automation provides under low traffic task load conditions no longer offsets the loss of D-side assistance on radar tasks. The introduction of automation may seem to have a benefit under low traffic task load conditions, but it may disrupt that balance between R and D positions under high traffic task load. We see this in the coordination between sectors. Under full automation and low task load, ATCSs spent a considerably longer amount of time communicating with the next sector than under full automation and high task load. Two explanations may account for this finding. First, when task load was low, ATCSs had enough time to conceptualize detailed plans to solve potential problems and discussed them in detail with the adjacent, controlling sector. In contrast, high task loads led the R-side ATCS to need more assistance from the D-side and pulled the D-side away from strategic planning using the DST. In the analyses of most of our data sets, we observed with automation, the D-side being pulled away from his or her traditional function of immediate assistance to the R-side. The ATC culture uses the D-side for immediate tactical assistance when traffic levels are high. Therefore, without changing the roles and responsibilities to include a strategic use of a DST, the D-side will have to choose between tasks and, under high traffic levels, the DST will likely go unused.

ATCSs indicated that the full automation conditions were less realistic and more difficult than the conditions without automation. Scenarios without automation mimic the current ATC field environment that utilizes flight strips, but no DSTs. The perceived difficulty of the DST may be because the ATCSs were novice DST users, and they needed to use conscious effort to use the DST. Other over learned behaviors (e.g., use of flight strips) are automated and effortless. Besides increasing the perceived difficulty ratings for the D-side ATCS, this also may have led to an indirect increase in the perceived difficulty for the R-side ATCS.

Task load had an impact on most of the dependent measures. ATCSs rated the low task load scenarios more representative and closer to their normal traffic levels. Although ATCSs did not view the high task load scenarios as representative of what they currently experience in the field, the results for the high task load scenarios may give us insight into what to expect when traffic levels increase in the field. Ground-to-air communications increased in number and decreased in duration under high task loads, whereas coordination using the landline increased. Within the team, ATCSs increased the number of comments regarding specific aircraft and altitude changes as task load increased. Trust in the DST fluctuated with task load levels. Although the algorithms in the DST did not change between conditions, ATCSs indicated that they trusted the DST the least under high task load conditions. It is under high task load conditions that the DST's presence is most needed; however, ATCSs are the least likely to trust and use it.

Workload ratings, as measured by the ATWIT, NASA TLX, and self-report ratings, indicated that as task load increased, workload increased. Interestingly, over time, workload ratings decreased under low task load, but they remained relatively constant under high task load. These results and our SA findings suggest that when task load was low, ATCSs were able to catch up and preplan control actions, decreasing workload over time. High task load conditions did not allow for this.

ATCSs and the OTS raters felt that an increase in task load resulted in decreased SA and performance in general. Under high task load conditions, the increased number of aircraft made it difficult for ATCSs to maintain the "picture," as reflected in each SA measure. Future sector-based SA was better, awareness for aircraft callsign letter and aircraft separation was higher, self-report SA was higher, and OTS (subjective performance) ratings were higher when task load was low. In contrast, SA was higher for awareness of aircraft clearances received when task load was high. Interestingly, our simulator data showed that ATCSs did attempt to setup the high task load scenarios to become more structured, potentially reducing the loss of SA.

We used our ATCS participants as their own controls, that is, they worked under the same conditions both as a R-side and a D-side. When working on the R-side, ATCSs' SA was better than when on the D-side was. Of course, the awareness for the tactical situation of a D-side can be less than that of an R-side because the D-side ATCS is less involved in tactical control of the sector.

Although the D-side ATCS is an integral part of the sector team, our participants indicated that they felt they controlled traffic less well when working as the D-side ATCS. These results reflect that the D-side does not directly control traffic, but, because our participants were all certified R-side ATCSs, they felt that they could have been of more assistance if they would have had R-side responsibilities. When working the D-side, our ATCS participants also indicated that high traffic task loads resulted in less of an increase in workload than when working the R-side. This may reflect spare capacity within the sector team on the part of the D-side.

1. INTRODUCTION

The introduction of new technologies into the National Airspace System (NAS) will significantly alter the role of the Air Traffic Control Specialist (ATCS). As the Federal Aviation Administration (FAA) moves forward with NAS modernization and the implementation of Free Flight concepts, the ATCS will move from a controlling role to a collaborating role (Office of Air Traffic System Development, 1997). With increases of up to 60% expected in air traffic over the next decade, automation will play a significant role in supporting the NAS and the ATCS (Office of Air Traffic System Development). The ATCS will be able to approve more user requested routes with the assistance of automated decision support tools (DSTs) (Federal Aviation Administration, 1999). As outlined by Kirk, Heagy, and Yablonski (2000), the anticipated benefits from the use of DSTs will encompass increases in safety, efficiency, and ATCS productivity.

We performed the current study to gain a better understanding of the impact of DSTs on ATCSs' cognitive and behavioral functioning. This study examined the effects of automation levels and increases in air traffic volume on ATCSs' situational awareness (SA), workload, visual scanning, communications, and the safety and efficiency with which they controlled traffic.

This study took place at the FAA Research Development and Human Factors Laboratory (RDHFL) and the Integration and Interoperability Facility (I^2F) in Atlantic City International Airport, NJ. The experiment involved engineers and researchers from I^2F Support Services of the En Route Branch (ACT-233), NAS Simulation Branch of the NAS System Engineering Division (ACT-510), NAS Human Factors Group of the NAS System Engineering Division (ACB-220), Civil Aeromedical Institute (CAMI), and their contractors.

1.1 Background

A primary goal of Air Traffic Control (ATC) is safety. ATC-related aircraft accidents are rare; however, at least one study has attributed over 90% of all errors to either ATCSs or supervisors (Kinney, Spahn, & Amato, 1977). Operating under high workloads, tight regulations, and challenging conditions, the ATCS job is both demanding and stressful (Finkelman & Kirschner, 1980). These conditions are likely to induce errors.

Improving the current ATC system with new technologies is necessary because of the projected increase in air traffic. These new technologies include the introduction of additional automated tools for all phases of flight. The FAA Office of Air Traffic System Development (1997) has laid out a plan that includes these efforts. Expected improvements from the new automation include increased efficiency and throughput. Table 1 lists these automation efforts in ATC (Gosling, 1992; Zweben, 1992).

Table 1. Active or Planned Automation Efforts

• Conflict detection and prediction	• Weather forecasting
• Conflict resolution	• Flight progress monitoring
• Automated flight planning & scheduling	• Situation classification
• Approach guidance	• Data fusion
• Departure guidance	• System function allocation
• Surface control management	• Modeling of the ATCS
• ATCS training	

The aviation industry has introduced automation into the working environment with the goal of increasing safety and efficiency in addition to supporting the continued growth of air traffic. Although extensive research on the general effects of automation on behavior exists, we know little about the effect of the introduction of automated DSTs on ATCSs. In their current capacity, ATCSs have developed strategies, procedures, and conscious, as well as automated behaviors to safely and efficiently direct traffic. These strategies include what has worked in the past for the ATCS, such as similarity of traffic flow through airspace. Facilities establish procedures that make traffic most efficient (e.g., silent coordination between Terminal Radar Approach Control (TRACON) and low altitude en route sector. This may include handoffs of aircraft that are suppose to climb to a certain altitude at a certain speed – ATCSs then do not need to use the landline (LL) to coordinate this). In new situations, ATCSs need to revert to consciously analyzing the situation to arrive at a solution that may work. In contrast, recognition-primed decision making (Klein, 1999) is an automated ATCS behavior where ATCSs recognize the solution to a problem more than the problem itself. The introduction of new tools will change the way they operate. Questions arise about the effect these new tools will have on the ATCSs. We expect that the introduction of automated DSTs in the en route environment will alter ATCSs' visual scanning, workload, and SA. These changes may in turn contribute to the overall safety and efficiency of the NAS.

1.1.1 Automation Issues

As the ATC system evolves to meet future increased traffic levels, automation will become increasingly more important. The automated tools presently in use include RADAR, Host, Display System Replacement (DSR), User Request Evaluation Tool (URET), Voice Switching and Communication System (VSCS), Conflict Alert (CA), and Minimum Safe Altitude Warning (MSAW). Ongoing projects that will result in new automated tools for future use include Direct To (D2), Traffic Management Advisor (TMA), En Route Descent Advisor (EDA), and Problem, Analysis, Resolution, and Ranking (PARR).

The literature widely discusses the technical and practical considerations of automation (Ignizio, 1991; Waterman, 1986) but does not address how it produces new problems for human performance (Endsley, 1993). As we continue to rely on the support of automation, it is essential to better understand its impact on human performance. As errors involving automated systems tend to be more cataclysmic and costly (Wickens, 1992), the human interface has become more important than ever.

1.1.1.1 Effects of Increased System Monitoring Task Load on ATCSs

Automation leaves the human with fewer functions but with a more complex system to monitor – a role in which people do not excel (Wickens, 1992). Out-of-the-loop performance problems are a major class of errors associated with automation. Operators who are out-of-the-loop are slower and less accurate at failure detection because they are passive decision-makers (Wickens & Kessel, 1979; Wickens & Kessel, 1980; Young, 1969). Human monitors have problems detecting system errors and performing tasks manually in the event of automation failure (Billings, 1988; Wickens, 1992; Wiener & Curry, 1980). In a review of automation problems, Billings traced six major aircraft accidents directly to failures to monitor the status of an automated system. In addition to delays in detecting that a problem has occurred, human operators may require a significant period of time to develop a sufficient understanding of the state of the system to act appropriately. This delay may prohibit them from carrying out their tasks or it may diminish their effectiveness (e.g., National Transportation Safety Board, 1990). We attribute these problems to lower levels of operator SA that can occur with automation approaches that place people in the role of passive monitor. In addition, a loss of skills needed to take over manually may result (Wiener & Curry).

1.1.1.2 Effects of System Complexity on ATCSs

Automation tends to increase system complexity. The increased probability of system failure related to the increased number of systems adds to the complexity of the user's job (Wickens, 1992). In addition, complex systems may increase workload, make the system more difficult to learn (Scerbo, 1996), and negatively affect SA (Endsley, 1999). The user must be aware of what mode is active and the procedures associated with all modes of the automated system (Scerbo). It may not be clear to the user what components are working, and the user may create his or her own model for the inner workings of the system. This makes the cause and effect relationship of the system less clear. The user may become “surprised” or “suspicious” of the automation when the user cannot account for an action taken by the automation that appears, to the user, to be illegitimate (Wickens & Hollands, 2000). This can lead to negative consequences if the user takes inappropriate actions to correct the perceived problem (Wickens & Hollands).

1.1.1.3 Effects on ATCS Decision Making

It is unclear whether automated systems will improve the quality of ATCS decisions. Increasing evidence suggests that ATCS performance does not benefit from the addition of traditional decision aids (Endsley & Kaber, 1996; Endsley & Kiris, 1995; Kibbe, 1988; Selcon, 1990). For example, the ATCS must make a comparison between his or her own way of doing the task and the way the system does the task and decide whether to accept the recommendations of the system. This can all take additional time, particularly if information is ambiguous. If the system advice is ambiguous or incorrect, the system is generally more likely to reduce human decision quality and speed.

Decision support advice represents a new source of information, adding to the decision-making problem as much as assisting it (Endsley & Selcon, 1997; Pritchett, 1997). With the presence of a decision aid, the user may experience problems stemming from limited knowledge or misperceptions of how the automated system works (Mosier & Skitka, 1996). Further, as noted by Mosier and Skitka, automated decision aids provide another type of decision heuristic from which a decision maker can act. In the context of ATC, this implies that an ATCS using a DST may accept the DST information without more in-depth analysis of the situation, may discount any conflicting information, or ignore other pertinent information. All of these have negative effects on ATCS SA and performance.

1.1.1.4 Effects on ATCS Workload and Performance

Wickens and Hollands (2000) noted that automation, when properly implemented, reduces workload involving muscular exertion, decision choice, and information acquisition and analysis. As applied to ATC, we would expect a DST to reduce ATCS workload initially through information acquisition and analysis and then, with future versions, decision choice. The purpose of one DST is to predict and identify future traffic conflicts with a 20-minute look-ahead time. The ATCS then formulates a decision to correct the conflict and can check it against the DST. Future DSTs will be more active and will compute a corrective action that the ATCS then implements.

Although automation has focused on reducing workload, Hart and Sheridan (1984) noted that automation often replaces workload involving physical activity with workload involving cognitive and perceptual activity. In ATC, workload will likely shift from one task to another, although the tasks will continue to be of a mostly cognitive and perceptual nature. Wiener's (1985) studies in commercial aviation found a significant number of pilots reporting that automation did not reduce their workload but increased it during critical portions of the flight.

Operator monitoring of the system, a component of automation, may induce high workload. When people use sustained attention to monitor, it induces considerable fatigue (Galinsky, Rosa, Warm, & Dember, 1993) and high-perceived workload (Becker, Warm, & Dember, 1991; Dittmar, Warm, Dember, & Ricks, 1993; Scerbo, Greenwald, & Sawin, 1993). Even when task load is not high, the requirement to vigilantly monitor automated systems imposes its own workload. In studies involving cockpit automation, pilots complained that automation required constant scanning, adding to workload (Wiener, 1985). For the ATC environment, these findings indicate that the addition of automation for many tasks may only shift the ATCS workload from monitoring traffic to monitoring both the automation and the traffic. This may result in anything from equal or higher workload under normal conditions to increases in workload under conditions where the automation has problems or traffic is unusual.

Harris, Goernert, Hancock, and Arthur (1994) found that when operators had to initiate task automation in response to an unanticipated increase in workload, they showed an increase in performance error on other manual tasks. This confirmed work by Parasuraman, Bhari, Deaton, Morrison, and Barnes (1992) indicating that operator initiation of automation was likely to increase demands when they were already high. This may explain why a paradox of automation exists. Operators indicate that under high workload they frequently turn the automation off

(Wiener, 1988). For the ATC environment, these findings indicate that under high traffic task loads, ATCSs may either disregard the automation or focus on it to the extent that their other tasks may suffer.

1.1.1.5 Effects on ATCS Situation Awareness

The effect of automation on ATCSs' SA is mixed. According to Wiener (1992) and Billings (1991), when appropriate integrated information is presented to the user, improvements in user SA may occur. However, negative effects of automation on user SA have been documented. Under automation, vigilance decrements, complacency, or a lack of trust in automation cause problems with SA. People may neglect monitoring tasks, attempt to monitor but do so poorly, or be aware of indicated problems but neglect them due to high false alarm rates. Passive processing of information under automation can make the dynamic update and integration of system information more difficult, decreasing SA. Changes in or a complete loss of feedback frequently occur either intentionally or inadvertently with many automated systems (Endsley & Kiris, 1995).

1.1.1.6 Effects on ATCS Trust in the Automation

The decision to rely on automation can be one of the most crucial decisions the operators of a complex system can make. Riley (1994) investigated factors that influenced when people would choose to initiate automation. In these studies, the choice to use automation did not relate to workload levels but rather to factors such as reliability, trust, and risk (Muir, 1988; Riley, 1994). An operator who does not trust a system may be more likely to commit errors (Danaher, 1980; Lee & Moray, 1992; Wiener & Curry, 1980). He or she may not use the system, even when it might be beneficial, and experience excessive workload because of this. This can lead to compromised overall system performance thereby undermining the reason for the automation. On the other hand, too much trust in an automated system can lead to errors (Lee & Moray, 1992; Stokes & Kite, 1994). The user

- may become complacent or overconfident in the automated decisions reached,
- may not properly monitor system state,
- may allow the system to perform functions best left to the user or to another aid, and
- may operate the system in ways for which it was not designed or when faulty.

1.1.2 Proposed Solutions for Automation Issues

The issues related to automation have led to proposed solutions such as "human-centered automation," support of SA, optimal level of automation, flexible function allocation, and relevant feedback provision.

Automation problems have led to human-centered automation to overcome the limitations of technology-centered approaches (Billings, 1988). With a human-centered approach, human roles have a higher priority (Hopkin, 1998) and give the user the greatest amount of satisfaction (Wickens & Hollands, 2000). In this context, automation is designed to assist the user in his or her tasks (Hopkin). In ATC, automation has often lacked a human-centered approach, resulting in rejection of the automation by the user (e.g., Standard Terminal Automation Replacement System Human Factors Team, 1997). ATCS-centered automation is necessary to ensure an ATC system that provides reliability, performance, safety, and acceptance. These efforts must augment the ATCS in handling the predicted increase in air traffic.

Wiener and Curry (1980) and Billings (1991) discussed the implementation of automation at various levels. Automation has an impact on the involvement of the ATCS in the decision process and thus, on performance in detecting system breakdowns and assuming control. Automation has typically resulted in decreased SA by removing the operator from involvement in system operation. An alternate approach to automation keeps the ATCS involved by determining a level of automation that minimizes negative impacts on SA (Endsley, 1987).

Endsley and Kiris (1995) found that by implementing functions at a lower level of automation, the operator remained involved in the active decision-making loop, SA remained at a higher level, and people were more able to perform the task manually when needed. To further explore the benefit and costs of intermediate levels of automation on overall human-machine performance, Endsley and Kaber (1999) developed a more detailed level of automation taxonomy (Table 2). They indicated that level of automation significantly affected both task performance and the out-of-the-loop performance problem. Automation aided the implementation of a task but hindered performance when involving joint human-automation option generation. Computer aiding at the action-selection (decision making) part of a task did not significantly have an impact on performance when compared to purely human decision making. Performance at the high end of the level of automation taxonomy was better than manual performance. It was, however, never as good as when the automation only assisted in the manual implementation aspects of the task, without becoming involved in the higher-level cognitive aspects.

Human ability to recover from and perform during automation failures significantly improved with levels of automation that required some human interaction in task implementation. Following an automation failure, time-to-recover and manual performance was worse with levels of automation that allowed advanced queuing of targets. Thus, automation strategies that allow operators to focus significantly in advance of current operations may contribute to out-of-the-loop performance decrements. Workload was lower and SA was better at some of the higher levels of automation.

Interface designs may need to support adaptive automation (AA) between the ATCS and the system with functions being passed back and forth as circumstances demand. AA allows for sharing of control between the user and the automated system (Scerbo, 1996). Past systems have not supported human operators in alternately assuming or delegating tasks. New interfaces need to provide the SA needed to support control transitions.

Table 2. Level of Automation Taxonomy (from Endsley & Kaber, 1999)

	ROLES			
LEVEL OF CONTROL	MONITORING the System	GENERATING Alternative Strategies	SELECTING a Strategy	IMPLEMENTING a Chosen Strategy
(1) Manual Control	Human	Human	Human	Human
(2) Action Support	Human/Computer	Human	Human	Human/Computer
(3) Batch Processing	Human/Computer	Human	Human	Computer
(4) Shared control	Human/Computer	Human/Computer	Human	Human/Computer
(5) Decision Support	Human/Computer	Human/Computer	Human	Computer
(6) Blended Decision Making	Human/Computer	Human/Computer	Human/Computer	Computer
(7) Rigid System	Human/Computer	Computer	Human	Computer
(8) Automated Decision Making	Human/Computer	Human/Computer	Computer	Computer
(9) Supervisory Control	Human/Computer	Computer	Computer	Computer
(10) Full Automation	Computer	Computer	Computer	Computer

The system design needs to prevent the erosion of skills. AA may aid in this goal, as demonstrated by Parasuraman (1992). AA acts to help maintain operator skills and performance levels by having the operator take over manual operations at regularly scheduled intervals.

Endsley (1996) points out numerous issues concerning AA. First, the characteristics of tasks that determine their optimal level of control and suitability for AA need investigation. Gluckman, Warm, Dember, and Rosa, (1993) and Carmody and Gluckman (1993) found different effects on workload, performance, and SA for AA involving a static (system monitoring) versus a dynamic (resource management) task. Durso, Gronlund, and Lewandowsky (1993) have proposed that if automation involves parts of an integrated task, more performance decrements will occur than if it involves the whole task. Second, examination of when AA should change its level of support is needed. Parasuraman, Molloy, and Singh (1993) examined whether manual control implemented at a pre-set periodic interval differed in effect from manual control implemented based on poor monitoring performance. They found no differences in the effect on subsequent human monitoring performance under automation. The insertion of a period of manual control was equally beneficial in both cases. Third, a major question lies in determining how to implement AA to provide the potential benefits without leading to loss of system awareness. Many systems have left it up to operators to invoke automation at their discretion. Leaving the system with the ability to turn itself on and off taxes the operator with keeping up with what the system is doing unless it provides the operator with a clear indication of the current status.

Research has also explored how much and when to use automation (Kaber, 1996; Kaber & Endsley, 1997). Kaber and Endsley found that level of automation had a large effect on operator SA, however, AA had an effect primarily on workload. Kaber found a significant interaction between AA and level of automation. The best strategy was when the human operator performed at a level of automation that allowed human strategizing and computer implementation (level of automation 3 - Batch Processing) during high automation cycle times (i.e., infrequent human manual control through AA). This combination was better than either fully automated performance or AA cycles that had high human-control times.

Norman (1989) cited fundamental changes in the amount and type of feedback provided by automated systems as crucial. We need to determine the necessary information to convey to the ATCS. For example, when the FAA automates flight strips, there may be loss of task load information associated with the number of flight strips or with the physical manipulation of flight strips (Garland, Stein, Blanchard, & Wise, 1992; Gronlund, 1992; Hopkin, 1991).

1.1.3 Summary

The successful implementation of automation will depend on the ATCSs' ability to acquire different types of skills, retain less frequently used skills, and adapt to different workload levels. The increased complexity of systems and the changes they induce in the degree of ATCS involvement requires human factors considerations in developing an effective ATC system.

Future ATC systems may impose a variety of challenges to ATCSs including information overload, non-integrated data, rapidly changing parameters, and various forms of automation. Evaluating the degree to which prospective system designs actually provide benefits to ATCSs is an important goal for ATC system evaluation.

1.2 Objectives

The introduction of DSTs into ATC will change ATCSs' behaviors and work procedures to accommodate these new tools. For example, ATCSs may no longer rely on flight strips, change Data (D-side) ATCSs' assistance to the Radar (R-side), and learn how to best use the DST to meet the ATCSs' primary goal of safety (i.e., moving from a novice DST user to an expert DST user). With the expected increases in air traffic, ATCSs will control higher volumes of traffic for longer durations and at more complex levels (i.e., more aircraft that require more control actions and may request user-preferred routes). The purpose of the current study was to explore in-depth the effects of levels of automation and traffic task load or volume on ATCSs in an ATC en route environment. It is important to investigate the effects of these anticipated changes on ATCSs' behaviors and cognitive functions. Through careful, controlled experimentation, researchers gain information about the possible consequences from expected and planned changes in the ATC environment. The results obtained from experimentation can then guide development of the most appropriate and effective strategies to offset any detrimental consequences from these changes.

We manipulated the level of automation that ATCSs used in controlling traffic. This included three levels ranging from no automation to limited automation to full automation. To simulate the expected increases in air traffic, we manipulated task load based on the volume of air traffic that ATCSs controlled. ATCSs controlled both low and high traffic task loads. Further, ATCSs rotated between the R-side and D-side positions.

We anticipated changes in ATCSs' behaviors and cognitive performance with the manipulation of automation, task load, and position. We examined their behavior and cognitive processing through objective and subjective performance and behavioral measures. These measures examined SA, workload, visual scanning, communications, and trust. We provide background information, the results, and the discussion for each of these measures under their respective headings in the Data Set Specific Analysis, Results, and Discussion section.

2. METHOD

In this study, 16 ATCSs performed en route ATC simulations at three levels of automation (No Automation, Limited Automation, and Full Automation) and two task load levels (High and Low). The ATCSs worked in teams of two, consisting of an R- and D-side.

2.1 Participants

Sixteen ATCSs from Air Route Traffic Control Centers (ARTCCs) within the United States served as voluntary participants. All participants were current, non-supervisory, full-time ATCSs and DSR certified. None of the participants was on medical waiver or in a staff position at the time of the experiment. Eleven participants had normal vision and five had corrected-to-normal vision. The oculometer design limitations excluded bifocals, trifocals, or hard contact lenses but allowed ATCSs to wear corrective lenses or soft contact lenses, if necessary. The mean age of the participants was 38.7 years (34–44). They had actively controlled traffic for a mean of 15.1 years (9-18.5). Six participants had worked at more than one facility during their ATC career. The participants worked air traffic for an average of 11.5 months in the preceding 12 months. None of the ATCSs had previously received training on a DST. Using a 10-point scale, participants rated their current skill level as an 8.7 (6-10) and their motivation to participate in the study as a 6.1 (3-9).

The Institutional Review Board of the William J. Hughes Technical Center approved the protocol. The ATCSs gave their written consent to participate in the experiment (Appendices A and B contain the participant recruiting form and the Informed Consent Form). The research team ensured them that their data would be completely confidential.

2.2 Experimental Staff

A research team composed of three Human Factors Engineers (HFEs) and two ATC Subject Matter Experts (SMEs) conducted the study. In preparation for the experiment, the HFEs designed the study, questionnaires, and procedures. The ATC SMEs designed the scenarios. After experiment completion, the HFEs performed the data analyses and wrote the final technical report. During the first week of training, one HFE and an ATC SME explained the study and the informed consent policy and trained the participants on the DST and the airspace.

For the experimental portion of the study, two HFEs and an ATC SME conducted the simulations. The HFEs managed the experiment, collected the data, and directed support staff. The ATC SME conducted the Over-The-Shoulder (OTS) ratings and completed SA Verification ANalysis Tool (SAVANT) forms with correct answers. Two system operators brought the Host and DST systems up for each simulation run, while another operator selected the appropriate scenario and prepared the Target Generation Facility (TGF) for each run. The study used six simulation pilots and one ghost ATCS. To allow rotation, researchers trained nine simulation pilots using procedures from past experiments. Support engineers ensured that the hardware and software functioned properly. Clerical staff assisted in preparing, copying, and distributing forms and questionnaires, in addition to preparing data sets, means and standard deviations (*SD*s), Multivariate Analysis of Variance (MANOVA), and Analysis of Variance (ANOVA) tables.

2.3 Materials

In the next sections, we describe the simulated airspace and scenario materials.

2.3.1 Airspace

The airspace used in this study was a generic en route ARTCC to make our findings easier to generalize and to increase the size of our participant pool (Guttman, Stein, & Gromelski, 1995). We created an airspace several sectors wide because the DST required a 20-minute look-ahead time. Only one of the sectors was active whereas the other sectors functioned as simulated ATCS sectors (i.e., ghosts). During the simulation, the weather conditions required instrument flight rules to be in effect. Figure 1 shows and Appendix C includes pictures of the airspace used in this study, called Genera Center. Within Genera Center, we used a high altitude sector 200 x 125 nms with boundaries from flight level (FL) 240 and above. Traffic flow consisted of arrivals handed off to intermediate sectors, departures climbing from intermediate sectors, and overflights through the sector.

For the staffing of Genera Center, ATCS participants controlled one sector (Sector 07 depicted in white in the center of Figure 1), while we staffed all adjacent sectors (depicted in gray in Figure 1) with one of our experimental staff. The R-side ATCS ensured separation, initiated control instructions, and operated radios. The D-side ATCS scanned the radar display for information and assisted the R-side ATCS when needed. The D-side ATCS also ensured separation, but on a separate position that shares the radar display with the R-side ATCS. The D-side ATCS operated the LL phones to adjacent sectors and facilities, accepted and initiated handoffs for the continued smooth operation of the sector, and ensured that the R-side was made immediately aware of any action taken. When working under a condition that included the DST, the D-side ATCS monitored the DST and integrated the information into the team's sector plan. The ghost accepted handoffs from our participant ATCSs and interacted with them when they requested control instructions from adjacent sectors. The ghosts confirmed the requested control instructions, contacted the simulated aircraft, and instructed the simulation pilots to fly the aircraft to comply with the ATCS request.

2.3.2 Scenarios

The experiment consisted of 18 different scenarios. We designed the scenarios with multiple conflicting traffic routes to necessitate the use of the DST for conflict resolution. Six scenarios contained moderate task loads for training sessions, and we used six low and six high task load scenarios for experimental sessions. Each training scenario lasted 30 minutes, whereas the experimental scenarios ranged in duration from 33 to 40 minutes to enable measures of SA with the SA Global Assessment Technique (SAGAT). We randomized the presentation order of the scenarios and rotated the scenarios through each automation level.

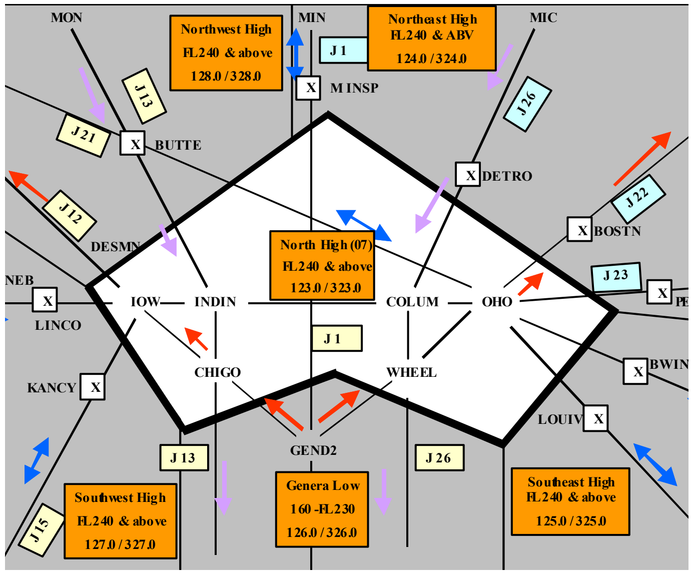

Figure 1. Illustration of Genera Center, Sector 7.

2.4 Location and Equipment

The following subsections describe the location where the study took place and the hardware and software we used during the study.

2.4.1 Location and Training and Experimental Setups

The experiment encompassed three separate locations and three different workstation platforms. Each location and operating system allowed for DST training, airspace training, and experimental sessions of the study to be conducted. We discuss each location and its platform in the following subsections.

2.4.2 DST Training Platform

Training on the DST and familiarization with the airspace, the Letters of Agreement (LOAs), and Standard Operating Procedures (SOPs) took place in a classroom in the RDHFL. The training hardware consisted of an instructor position and two student positions. These workstations displayed stand-alone versions of the DST. The DST itself ran on a DEC Alpha system (Digital Equipment Corporation, Boston, MA) in the I^2F with remote displays on the instructor and student positions that consisted of SUN (SUN, Palo Alto, CA) workstations with 19 in. displays, a keyboard, and a three-button mouse.

2.4.3 Airspace Training Platform

For training on the airspace, the RDHFL provided a high fidelity ATC simulation environment. We used two stand-alone ATCoach simulators (UFA, 1998), running on SUN machines. The ATCS stations included a radar display, full flight strip bay, an ARTCC keyboard, and a trackball. A high-resolution (2,000 by 2,000 pixel) monitor (Sony Corporation, Japan) displayed the radar display. Three PC-based simulation pilot workstations running on SUN operating systems were connected to each ATCoach (UFA) simulator to allow simulated pilot participation in the scenarios. The ATCoach (UFA) simulators allowed training of each participant on the airspace independently and simultaneously.

2.4.4 Experimental Platform

The experimental sessions took place in the en route lab in the I^2F. The I^2F provides areas and equipment for evaluation of ATC infrastructure enhancements and subsystems and integration and interoperation of ATC subsystems and functions and is configurable to support multiple users simultaneously.

For the experimental sessions, we used an integrated system including the TGF, the Host computer, and a full DSR workstation with all functions normally expected in an operational setting, and the DST. We used the TGF to generate targets and air space. A 2,000 by 2,000 pixel, 29 in. video display unit (Sony) presented the radar information. An Air Traffic Workload Input Technique (ATWIT) (Stein, 1985) was mounted immediately next to the DSR display within easy reach of the participant for input of workload ratings. The workstation had a DSR flight strip bay, an en route keyboard, and a trackball with three buttons. An LL allowed interfacility and intrafacility communications. The right trackball button served as a home key that returned the cursor to the center of the DSR when pressed. The D-side ATCS had access to the DST. Different from the configuration in the field, this DST ran on a 29 in. video display. A network linked six simulation pilot displays with the R- and D-side positions. Each simulation pilot station allowed control of several aircraft.

2.4.4.1 DST

The DST includes several display windows that capture various flight data and conflict information. One display window presents an aircraft list of all aircraft inbound to the sector (Figure 2). The list shows the controlling sector of a particular aircraft; flight data such as flight route, aircraft type, speed, altitude and beacon code; and conflict indication.

The conflict indication depicts red for predicted violations of less than 5 nms and yellow when less than 12 nms but more than 5 nms. As shown in Figure 2, TWA483 (the highlighted line) has a red conflict indicated, but the other aircraft(s) with which it is in conflict are not identified.

The graphic plan display window depicts aircraft and resembles the DSR display (Figure 3). It can show all conflicts, conflicts only in the ATCSs' sector, or conflicts of a specific type (e.g., red). As shown in Figure 3, DAL317 and DAL540 have a predicted conflict that is clearly shown on the graphic plans display.

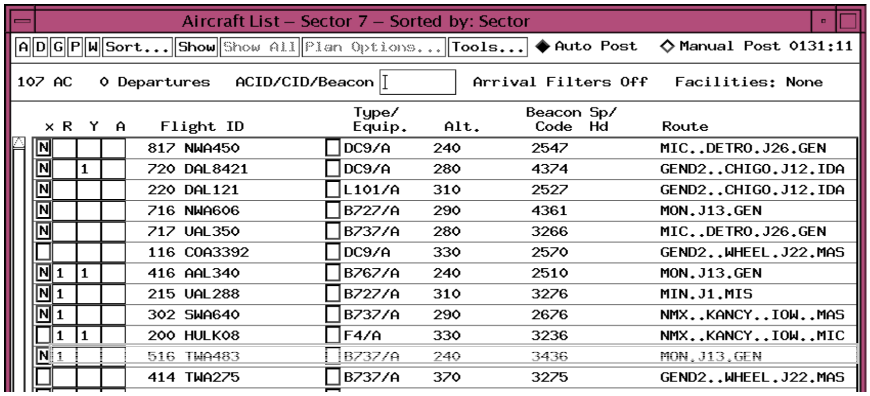

Figure 2. Aircraft list display window from DST.

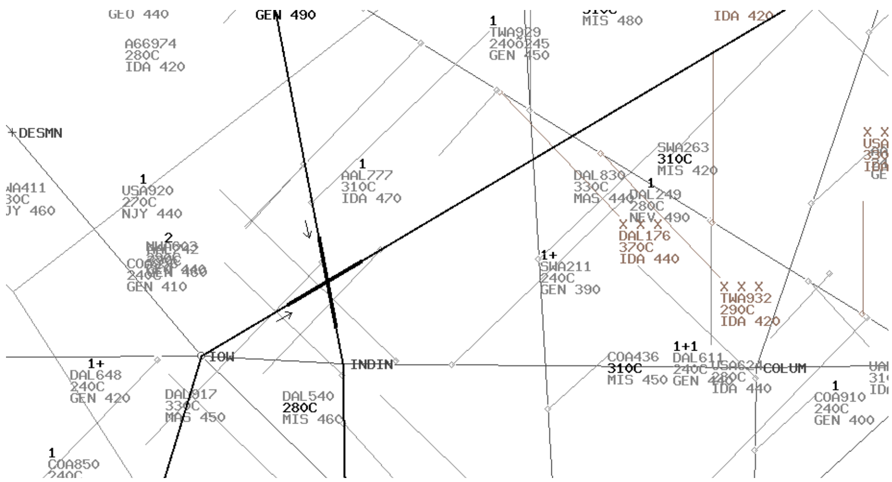

Figure 3. The graphic plans display window from the DST.

In Appendix D, a plans display textually depicts the computer and aircraft identifications, aircraft altitude, speed, and route along with any problem information (Figure D-1). The departure list shows flight information for proposed departure flights. It includes aircraft and computer identifications, departure time, aircraft type and equipment, filed altitude, beacon code, and route of flight. These plans are not probed for conflicts (Figure D-2). A response display shows system messages describing problems from an ATCS action (Figure D-3).

Trial flight planning allows the ATCS to enter in a revised altitude, speed, or route change to resolve a potential conflict. A trial plan can be created from the aircraft list, departure list, or graphic plans display. The DST then shows whether the change would no longer result in a conflict. This can be presented textually or graphically.

The DST also offers the capability to coordinate trial plans between sectors. However, the DST configuration used in the study did not allow coordination between sectors.

2.4.4.2 Video Camera and Video Tape Configuration

We taped the video images of both ATCS positions (Figure 4, top and side views of both D-side and R-side ATCSs) and streamed audio and video to disk. We copied the MPEG files onto a CD-ROM for further analysis.

Figure 4. Example of top and side views of ATCSs.

To record the intra-team communications, we developed a program, Orasis (Dauphin Technology, 1999), that ran on a portable device when conducting live coding. The software has a layout as depicted in Figure 5. The left column of buttons of the software indicated the message content. The second column indicated the message type. Finally, the third column indicated message modality. When conducting post hoc coding, we used a behavioral coding software program, Observer Video-Pro (Noldus, 2000). The software configuration used the Controller-to-Controller Communications/Coordination Taxonomy (C^4T) coding scheme developed by Peterson, Bailey, & Willems (in press). This allowed an SME to observe side and top views of both the R- and D-side and code behaviors in real-time. Using this matrix, we captured specific verbal and non-verbal communication behaviors as well as the sequence and timing of the behaviors.

2.4.4.3 Communications Configuration

We used communication systems that are functionally similar to those in use in the RDHFL. There were communication links between the ATC SME, OTS observer, simulation pilots, experimenters, and push-to-talk (PTT) recording. The equipment monitored communications and recorded times and frequencies.

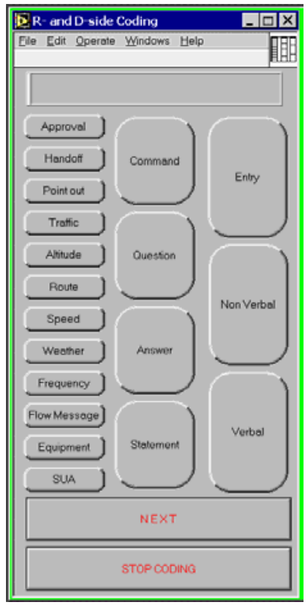

Figure 5. Coding software interface.

2.4.4.4 Situation Awareness Verification Analysis Tool

SAVANT ran on separate multiple network SUN workstations programmed in C++, Java, and ODS Toolbox (Orthogon, 1999), one for the D-side and one for the R-side ATCSs. Twenty seconds after ATWIT alerted the participants for a workload rating, both the R-side and D-side ATCSs' DSR monitors displayed the SAVANT query for 3 seconds; after that, the replication of the DSR screen appeared for 9 seconds while they responded to the query. For more detailed information on the implementation of SAVANT, see Appendix E.

The SAVANT display closely matched the DSR display. We presented the ATCS with several queries. The queries consisted of both Aircraft-Pair and Sector-Based questions. Aircraft-Pair questions were relational; they always asked about a pair of aircraft. The query asked about the relationship between one aircraft and either another aircraft or airspace for the current situation or a future situation. The questions assessed the comprehension and projection components of SA. All visual information was present except for the information about which the query was asking (Figure 6). It was a forced choice question. In contrast, Sector-Based questions were open-ended with multiple responses possible and assessed the ATCSs comprehension of the current situation only. The only information available for the question was the aircraft position (Figure 7). We recorded the ATCS responses by having them click on an aircraft.

2.4.4.5 Situation Awareness Global Assessment Technique

In addition to the SAVANT tool, we used the SAGAT to measure SA at the end of each scenario. SAGAT ran on Apple PowerBooks programmed in Hypercard. Section 3.1.2 provides detailed descriptions of the background, data reduction, analysis, and discussion of the SAGAT data.

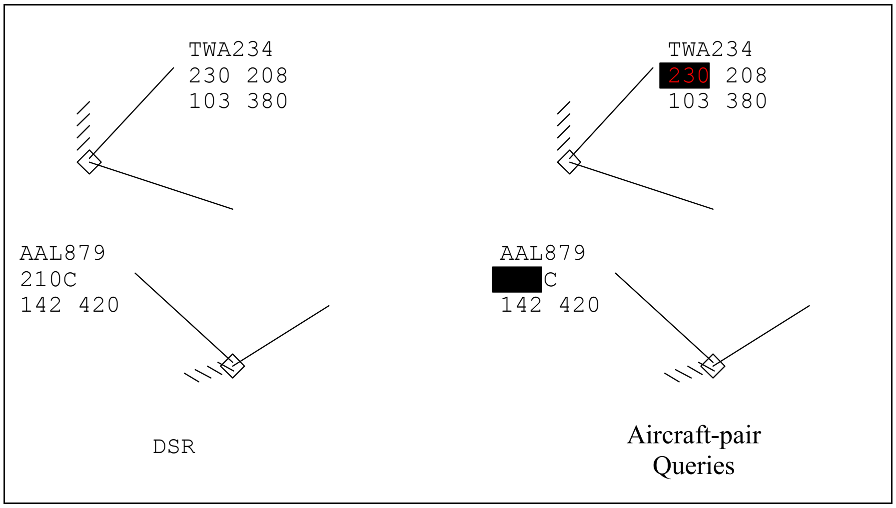

Figure 6. Comparison of DSR and SAVANT aircraft representation for aircraft-pair SA queries.

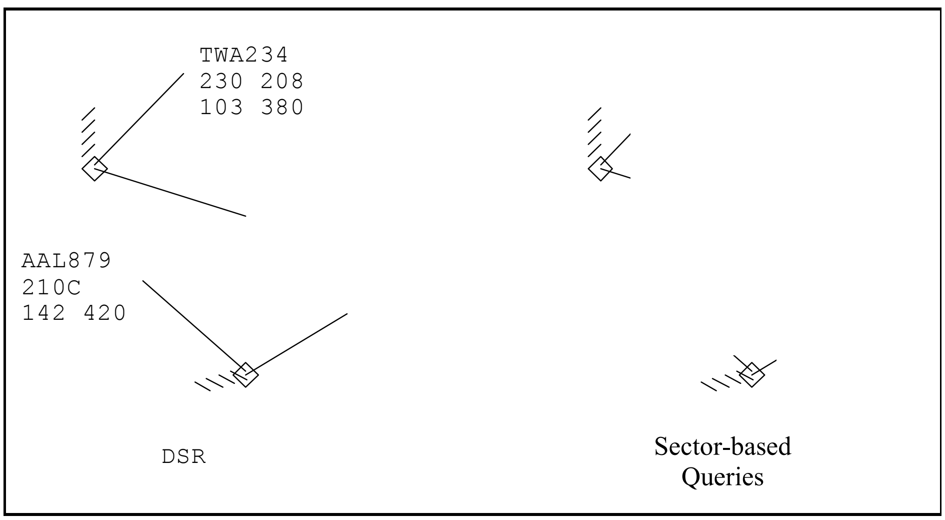

Figure 7. Comparison of DSR and SAVANT aircraft representation for sector-based SA queries.

2.4.4.6 Air Traffic Workload Input Technique

An ATWIT device (Stein, 1985) recorded response latencies (e.g., times to respond) and workload ratings during all conditions. The participants made a rating on the ATWIT device every 3 minutes. Before each rating, a tone alerted the participant who then had 20 seconds to make a workload rating. The participants used a scale of 1 (low workload) to 10 (high workload) (Appendix F contains the detailed ATWIT instructions). ATWIT is a reliable and relatively unobtrusive real-time, on-line measure of subjective workload.

2.4.4.7 Oculometer

We used an oculometer (Applied Science Laboratories, 1991) consisting of an eye and head tracking system. This system recorded the Point-of-Gaze (POG) and pupil diameter of a person 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, we refer the reader to previous reports (Willems, Allen, & Stein, 1999; Willems & Truitt, 1999). Willems et al. (1999) indicated that the exposure to the infrared illumination while wearing the oculometer is less than 4% of the intensity of that when walking outside on a sunny day.

To enable accurate calculation of the location of the POG, we determined the exact three-dimensional location of several surfaces (or scene planes) relative to the oculometer coordinate system. The procedures used for this initial calibration process measure distances of known points on the scene planes and determine the coordinates of each of these points relative to the oculometer three-dimensional coordinate system. The oculometer then used the position and orientation of the scene planes to determine the local coordinates (i.e., the coordinates relative to a two-dimensional coordinate system attached to each of the scene planes).

Once the oculometer software stored the exact position of the scene planes, one only needs a participant calibration before each of the simulations to correct for the way the head-mounted magnetic head tracker and optical eye tracker fit on the participant's head and for distortions in the optical system. We used a 17-point calibration grid displayed on 2000 x 2000 display similar to that depicted in Figure 8. During this final calibration, we instructed the participant to sit still and to focus his or her gaze on the numbered points as we called them out. We used the oculometer software to automatically enter the participant's POG for each of the 17 points. The software then used the known locations of these points to determine the adjustments it needed to make to fit POG to the exact location of the calibration points. At the end of the calibration procedure, the experimenter verified that the participant's POG coincided with the system's coordinates by having the participant look at several points of the calibration grid.

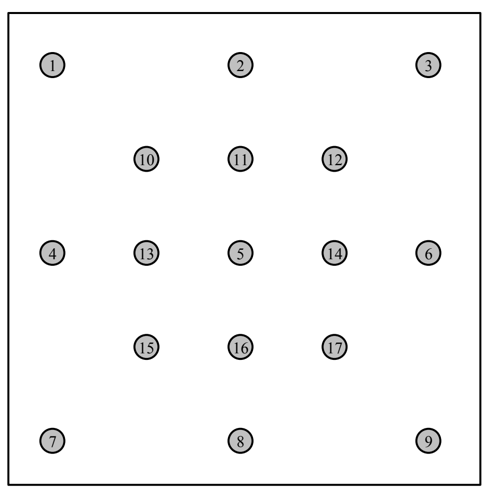

Figure 8. Example of the calibration screen used with the oculometer.

2.5 Design

Our study was a 2 x 2 x 3 design and contained the following IVs: three levels of automation, two levels of task load, and two levels of position. The DVs consisted of several measures to assess ATCSs' performance and behavior. We collected scores on all measures for all trials and compared them for each level of the experimental variables.

2.5.1 Automation

This study used three levels of automation. The no automation (baseline or first level) featured paper flight strips. The limited automation (second level) featured electronic flight strips with conflict indication but no trial flight planning (aircraft list from DST). The full automation or (third level) featured electronic flight strips with flight planning (aircraft list, graphic plan display, and trial plan display from DST).

2.5.2 Task load

We defined high and low task load as the difficulty of the scenario, operationalized by the number of aircraft and the complexity of the traffic situation. For this experiment, we defined a low task load scenario as having a level of complexity at which a first line supervisor would be about to remove the assistance of a D-side ATCS. We defined a high task load scenario as having a level of complexity at which a first line supervisor would be about to assign the extra assistance of tracker (i.e., move from a two- person team to a three-person team). Figure 9 shows the average number of aircraft for each 3-minute interval that ATCSs controlled traffic in both the low and high task load scenarios. The dashed lines represent the variability among the scenarios within each task load level.

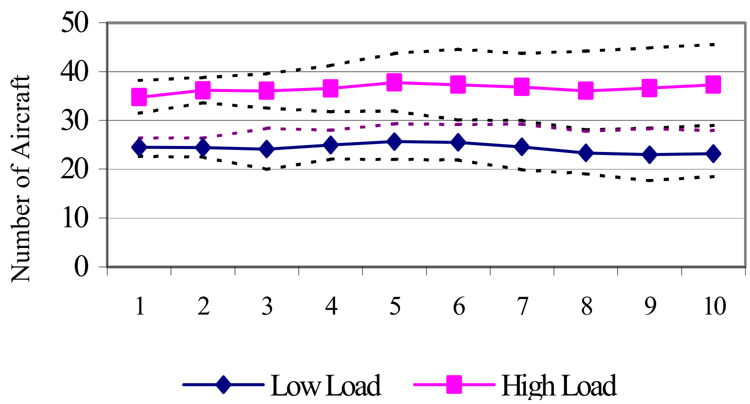

Figure 9. Average number of aircraft for low and high task load scenarios.

2.5.3 Position

We used two levels of position: R-side and D-side. Appendix G contains complete descriptions for R-side and D-side roles and responsibilities.

We calculated summary DVs for scenario and interval. After each scenario, we collected Post-Scenario Questionnaire (PSQ), trust, National Aeronautics and Space Agency (NASA) Task Load Index (TLX), OTS, and SAGAT ratings (Appendix H to Appendix J) and visual scanning and performance summary data. At intervals, we collected ATWIT and SAVANT ratings. Participants completed an entry questionnaire and, at the end of the study, an exit questionnaire (Appendix K and Appendix L, respectively). We used all questionnaire data in its raw form.

Table 3 summarizes the data sets we collected during the study. Each of the data sets provides insight into ATCS performance and behavior.

Table 3. The Data Sets Recorded during the Experimental Scenarios

- SA via SAVANT
- SA via SAGAT
- Workload via the ATWIT device
- Eye tracking of the D-side ATCS at 60 samples per second
- Continuous recording of communications between the R- and D-sides
- Recorded communications between ATCSs and simulation pilots on MPEG files, with a time stamp
- Unix-based PTT software identifying the speaker, and at what time and for how long the speaker keyed the microphone
- Aircraft data and pilot/ATCS entries into the system (TGF)
- System Analysis Recording (SAR) data (Host)
- Questionnaires

2.6 Procedure

ATCSs participated in the experiment for 2 consecutive work weeks. The first week consisted of familiarization and training, and the second week consisted of experimental sessions. The morning of the participants' first day consisted of a briefing and a familiarization period. The HFEs explained the experiment, the oculometer, differences between TGF and their own equipment, and the confidentiality of their identity. Researchers provided an informed consent briefing and assurance that participation was voluntary. Participants signed an informed consent. The ATCSs then completed an Entry Questionnaire that included demographic questions about age, experience level, need for corrective glasses, and so on.

The decision support automation training consisted of 2 days of classroom instruction and one day of hands-on training. After instructing the ATCSs about the LOAs and the SOPs, the ATCSs received background information on the DST. This included the functionality, human-computer interface, and operation of the DST. The ATCSs then received hands-on training with the DST. We replayed interactive scenarios during training. The next day, we briefed participants on the airspace, and they worked simulations to learn the airspace. The airspace modeled for the training sessions was identical to the airspace used during the experimental sessions. Our ATC SME trained each participant in the use of the airspace, scenario flow, and traffic type.

The first day of the second week consisted of hands-on training with all experimental equipment and procedures in place. Before beginning experimental trials, we trained each participant in the use of equipment, including the DSR workstation, communications, the oculometer, ATWIT, SAVANT, and SAGAT. At the end of training, participants mastered the airspace and all of the equipment used in the experiment. Finally, the last 3 days consisted of experimental scenarios. ATCSs had a 15 to 20 minute break between trials and 60 minutes for lunch. Appendix M presents a detailed schedule of activities.

3. RESULTS

We measured participant performance and behavior across five constructs: SA, workload, visual scanning, trust, and performance. To keep the background, results, and discussion for a specific data set in close proximity to one another, we report them per construct.

We categorized the data sets into three groups based on collection method: those collected (1) by scenario, (2) by interval, and (3) continuously (see Table 4). Data collected after each scenario include SAGAT and questionnaire ratings. Data collected by interval include ATWIT and SAVANT. Data collected continuously include visual scanning, communications, system entries, and System Analysis Recording (SAR) tapes. For analysis purposes, we calculated summary statistics on continuous data per interval and per scenario. We summarized the visual scanning data, the R- and D-side communications, ATCS/pilot communications, and pilot/ATCS system entries by scenario. We also calculated pilot/ATCS system entries by interval.

Table 4. Scenario, Interval, and Continuous Data Sets

Scenario Data	Interval Data	Continuous Data
• SA via SAGAT • Questionnaires	• SA via SAVANT • Responses to the ATWIT device	• Eye tracking of the D-side ATCS at 60 samples per second • Continuous recording of communications between the R- and D-sides • ATCS/pilot communications • Aircraft data and pilot/ATCS entries into the system • SAR data

Due to problems in initial counterbalancing, the first 10 participants did not experience each condition of the design. We replaced any missing data with the overall mean. The remaining six participants completed all condition levels. We performed analyses using their responses as a second data set. With the analysis of the second data set, we examined whether the replacement for missing data had removed effects of experimental conditions.

Unless otherwise noted, all results represent the findings for the data sets of 16 participants. When the results of the smaller, six participant data sets differed from the 16 participant data sets, we present these findings separately.

For a description of general statistical methods as well as detailed information about the statistical methods used in this study, see Appendix N. We computed MANOVAs to compare effects on multiple variables and ANOVAs for effects on single DV. We tested the Wilks' Λ statistic using a level of $p < .05$ and report the equivalent F statistic. We report the most commonly used alpha level closest to the actual p value obtained. If the results of the MANOVA were statistically significant ($p < .05$), we performed univariate ANOVAs to determine which of the DVs were significantly different across experimental conditions. We based the significance of an ANOVA result on an adjusted alpha level using the following formula:

alpha (overall) = 1-(1-alpha(individual))^n where n is the number of variables

or:

alpha (individual) = 1-(1-alpha(overall))$^{1/n}$

We report the adjusted alpha level with each analysis. If the result of an ANOVA was statistically significant, we performed appropriate post hoc tests to determine which conditions were responsible for the significance.

Other researchers have used a more lenient approach when investigating the effects of manipulation on DVs by not adjusting the alpha level. Such an approach may inflate the overall alpha level but allows researchers to investigate trends in the data. In the current study, we follow such an approach to investigate trends (Table 5). We use the term "trend" to indicate a primary trend. A primary trend indicates an effect that did reach significance at the multivariate level but had a *p*-value at the univariate level greater than the adjusted alpha but lower than .05. It also could indicate that it did not reach significance at the multivariate level but had a *p*-value less than the adjusted alpha at the univariate level. A secondary trend refers to an effect that did not reach significance at the multivariate level but was higher than an adjusted alpha but lower than .05 at the univariate level.

Table 5. Types of Trends

Trend	Multivariate	Univariate *p*-value
Primary	Significant	< .05, > adjusted alpha
Primary	Not significant	< adjusted alpha
Secondary	Not significant	< .05, > adjusted alpha

In the figures graphically representing the results, we provide means and SDs. The SDs indicate the between-subject variance.

3.1 Situation Awareness

ATCSs work in a dynamic environment in which they are continually assessing changing aircraft information in order to maintain aircraft separation while moving aircraft efficiently through the airspace. They must maintain "the picture" to fulfill these goals. As outlined by Endsley and Smolensky (1998), ATCSs base decisions on their mental model of the situation and refer to this as "the picture." In turn, what ATCSs refer to as the picture is what researchers investigate as SA (Endsley & Smolensky). Hence, the ATCS's ability to perform traffic management functions directly hinges on possessing comprehensive SA.

There is currently no agreed upon definition that captures the essence of SA (Endsley, 1989a, b; Fracker, 1989; Mogford & Tansley, 1991). Endsley's definition of SA is "...the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley, 1988a, p. 3). Most researchers agree that the perception and understanding of elements in the present situation and the use of this information to anticipate future events are important in maintaining SA.

SA is a cognitive construct separate from workload, performance, and decision making that highlights the importance of future perceptions and the anticipation of future events within a goal-directed dynamic environment (Durso & Gronlund, 1999). It forms input to decision-making and actions and is important for successful performance, and it has both temporal and spatial components that reflect the current goals of the ATCS (Endsley & Smolensky, 1998).

SA has been cited as crucial for effectively monitoring an automated system (Endsley & Smolensky, 1998) and, at decreasing levels, contributing to more severe operational errors (OEs) in ATC (Durso, Truitt, Hackworth, & Cruthfield, 1998; Rodgers & Nye, 1993). Its relationship with workload is more complex; high workload may lead to decreases in SA, whereas, decreases in SA may lead to higher workload levels (Endsley & Smolensky).

In the case of ATC, the R-side and D-side ATCSs form a team. Within this team, each person has roles and responsibilities. Task allocation defines the SA requirements of each team member. However, the team members have a need for a considerable amount of overlapping information. These are their shared SA requirements. The team can achieve effective joint performance when each member has the SA needed for his or her role, and the team members have the shared SA required for effective team coordination. Both R-side and D-side ATCSs need to be aware of the state of the aircraft within the airspace, but R-side ATCSs also need to be aware of the more action-oriented items of SA such as an aircraft conformance to clearances.

Measurement of SA

The process model developed by Endsley (1996) shown in Figure 10 depicts the issues that are involved in selecting measures of SA. This model shows the stages involved in the perception-action sequence. We have shown these closely coupled stages as separate for simplicity only. Moderating factors that may influence each stage appear on the left side. On the right, classes of measures that are appropriate to each stage appear. Some of these measures will be indirect indices of SA and others will be more direct.

Researchers have used many different methods to measure how operators develop and maintain SA. Previously used measures include physiological measures such as eye movements (Moray & Rötenberg, 1989; Wierwille & Eggemeier, 1993), verbal protocol analysis (Ohnemus & Biers, 1993; Sullivan & Blackman, 1991), retrospective recall (DeGroot, 1995; Kibbe, 1988), rating techniques (Reid & Nygren, 1988; Taylor, 1990), memory probes (Endsley, 1988a, b), and on-line queries (Durso et al., 1995). Most of these techniques have demonstrated some degree of validity and usefulness.

Vidulich (in press) completed an extensive meta-analysis of SA measures. This analysis provides an interesting comparison of the sensitivity of the various techniques for measuring SA, but the study does not address the issue of validity. In the following sections, we describe process indices and direct and indirect measures of SA.

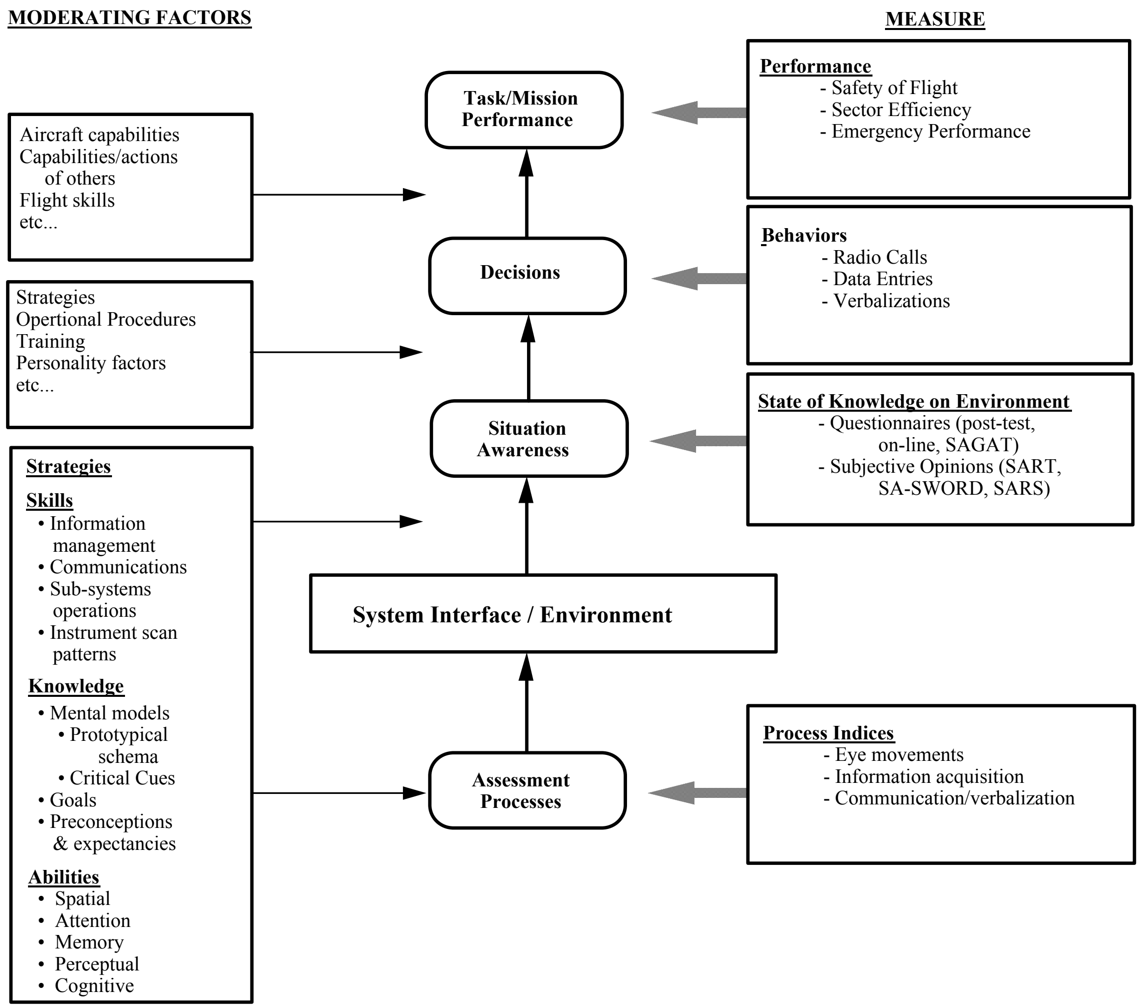

Figure 10. Measures of SA (from Endsley, 1996).

Process Indices

Individuals' characteristics will influence assessment processes used in acquiring information from the environment (Endsley & Bolstad, 1994). Differences in underlying abilities including spatial, attentional, memorial, perceptual, and cognitive contribute to this finding. Individuals will also form strategies, skills, and knowledge with experience and training that will contribute to their SA.

Process measures may provide an indirect indication of operator SA. Some may be useful in conjunction with each other. Process tracing tools used in the study of decision making may be applicable to the study of SA processes. Eye-trackers may help uncover how operators allocate attention in acquiring SA and their typical scan patterns. This information may provide insights into SA acquisition or into the mental models directing this process. Stein (1992) found that it takes about 2-3 minutes for scan patterns to stabilize after taking control of a sector. Smolensky (1993) found similar results for scan patterns and saccades that negatively correlated with a SAGAT measure of SA.

Studying the communications process between operators may provide useful information on the types of information that are lacking from displays, verbal techniques used for acquiring SA, and differences in SA strategies between individuals. Verbal protocols may help uncover what the operator attends to and may provide insight into how that information is integrated. Sullivan and Blackman (1991) used verbal protocols to investigate the relationship between working memory and long-term memory in maintaining SA.

These techniques provide useful information on SA processes. However, because verbal communications and verbal protocols take place in a limited period, they may not be complete representations to what ATCSs process or attend. Eye-trackers and information acquisition methods may capture the SA acquisition process.

Sarter and Woods (1991) have proposed using a scenario manipulation method. Here, researchers alter displayed information in some unpredicted way. Artificial manipulation during a simulation may provide useful insights into the SA process. Tenney, Adams, Pew, Huggins, and Rogers (1992) discuss using this technique to lead subjects “down the garden path” thus investigating factors that may lead directly to incorrect or missed assessments of situations.

Direct Measures of SA

The subjective assessment of SA is attractive because it is inexpensive and easy to administer. However, ATCSs may not know about their own inaccuracies or of what information they are unaware (Schroeder & Nye, 1993). However, self-ratings may be useful for providing an assessment of operators’ degree of confidence in their SA (Endsley & Smolensky, 1998).

One of the best-known subjective scales is the Situational Awareness Rating Technique (SART) developed by Taylor (1990). Participants rate a system design based on the amount of demand on attentional resources, supply of attentional resources, and understanding of the situation provided. Selcon and Taylor (1990) have found that SART correlated with performance measures. Selcon, Taylor, and Koritsas (1991) showed SART to be sensitive to changes in task demands, correlating with the NASA-TLX measure of workload. Crabtree, Marcelo, McCoy, and Vidulich (1993) found SART sensitive to most display manipulations, particularly the attentional demand scale. Researchers have compared the SART to the SAGAT. Endsley, Selcon, Hardiman, and Croft (1998) found no correlation between the SAGAT scores and the SART. However, the SART highly correlated with a subjective measure of performance and a subjective measure of pilot confidence level.

As another approach to developing a standardized subjective measure of SA, Vidulich and Hughes (1991) used a modified version of the Subjective Workload Dominance (SWORD) technique to obtain subjective evaluations of the SA provided by displays. SA-SWORD has individuals provide a comparative preference for displays on a 9-point scale based on their beliefs about the amount of SA provided by each. They found the technique discriminated between two display formats and had inter-rater reliability.

Outside observers can also assess subjective SA. An advantage of this method is that trained observers may have a more complete knowledge of reality than the rater. A shortcoming is that observers will have only limited knowledge about the ATCS’s concept of the situation. ATCS actions and speech may provide diagnostic information on explicit SA problems and indicate that the ATCS knows certain information, supporting observer judgments. Actions and verbalizations cannot provide a complete representation of ATCS SA. As pointed out by Durso and Gronlund (1999), SA is a cognitive construct. Outside observers cannot know what the participant knows mentally. Efforts to elicit more information may augment natural verbalizations, but this may alter the participant's distribution of attention, thus altering SA.

Questionnaires collect detailed information about ATCS perceptions, which, when evaluated against reality, provide a detailed objective assessment of SA. This type of assessment provides a direct measure of SA. We describe three ways of gathering this information: posttest, during simulations, or during interruptions in the simulation.

1. Posttest questionnaires. A detailed questionnaire after the completion of each simulated trial, allowing ample time to respond. Memories of dynamic SA will be less reliable with time. People tend to over-rationalize and over-generalize about past mental events (Nisbett & Wilson, 1977). Kibbe (1988) used this technique to evaluate SA as affected by automation of a threat recognition task.

2. Questionnaires during simulations. One way of overcoming the deficiencies of Post-Scenario Questionnaires is to ask ATCSs about their SA while they are carrying out their simulated tasks. It is possible to measure response time (RT) as an index of SA. Durso et al. (1998) recently investigated the use of this technique in an ATC task. They found that ATCS RT to probes about the current status of events in the simulation correlated with an SME's subjective ratings of ATCS performance. Willems and Truitt (1999) found an increase in RT with reduced involvement; however, workload increased.

3. Questionnaires given during simulator interruptions. To overcome the limitations of reporting on SA after the fact, one method freezes and suspends the simulation at randomly selected times, while ATCSs answer questions about the current situation. Researchers compare ATCS perceptions to the real situation as an objective SA measure. Collecting data in this manner provides an objective assessment of SA. The primary disadvantage of this technique is the temporary halt in the simulation. Several studies have used this technique to collect measures of SA on select parameters. Marshak, Kuperman, Ramsey, and Wilson (1987) administered queries on target location, altitude, and status in evaluating various map display formats. Fracker (1989, 1990) used queries to measure subject knowledge of target identification and location in several studies. Mogford and Tansley (1991) used queries regarding aircraft location and status in a study of ATCSs. One potential shortcoming of obtaining an indication of SA by using probes on a few predefined elements is that this may have an effect on subjects' attention during testing.

Indirect Measures of SA

Researchers may expect ATCSs to act in certain ways based on their SA. Some information about SA may be determined from examining behavior on specific subtasks of interest. Mosier and Chidester (1991) found that high-performing crews had fewer verbal communications than poorer performing crews. Rogers (1994) examined answers to on-line probes from the dispatcher and first officer. He found this measure sensitive to design issues surrounding implementation of an automated system. Other behavioral indices might include time to make a response (verbal or non-verbal), and correct or incorrect SA as identified from ATCS verbalizations and appropriateness of a given behavior for a particular situation.

Performance measures provide the advantage of being objective and are usually non-intrusive. Researchers have examined specific task performance as an indicator of SA. Hansman et al. (1992) used detection of errors as a measure of aircrew SA. They found that it was sensitive to differences between manual and automated programming modes but was not sensitive to the use of readback. Andre, Wickens, Moorman, and Boschelli (1991) measured navigation performance and aircraft control and found it was not sensitive to a display change as anticipated.

It is easy for subjects to bias their attention to the single issue under evaluation in a particular study through predictability and response priming. For example, Busquets, Parrish, Williams, and Nold (1994) measured the time for aircraft to respond to a runway incursion as a measure of pilot SA. This is a good example of a situation in which a single response outcome can be expected if the person has good SA.

One of the challenges in measuring SA is to probe ATCSs without them needing to transform knowledge into a format that we can measure. Experts are notorious for their inability to verbalize their knowledge or present a solution broken down in small steps. They will often refer to a "gut feeling" or indicate that they "just know" that they have chosen the right solution to a problem. To try to get around the pitfalls of assessing SA in expert ATCSs, we have created a technique that will leave the stimulus (i.e., the traffic situation displayed on the radar display) intact.

Another challenge is that the architecture of the NAS and Host prevents temporarily halting the simulation to assess SA. We, therefore, needed a technique that allowed us to probe controllers for SA while the simulation was running. Two well-known examples of techniques that use queries during simulations and ask operators about situation-specific information while they are carrying out their simulated tasks are the Situation Presence Assessment Method (SPAM) and the SAGAT.

In this experiment, we assessed SA while the simulation was running by using a method that incorporated the advantages of both the SPAM and the SAGAT. We refer to this method as the SAVANT. We also incorporated several other measures of SA - SAGAT, SME observer ratings, and self reported SA.

3.1.1 SAVANT

We developed an objective measure based on responses to queries that incorporated the advantages of both the SPAM and SAGAT called SAVANT. We used this technique to assess the effects of changes in automation, task load, and position on ATCS SA.

In SPAM, one of the experimenters calls the ATCS and asks a question about two aircraft. The experimenter then measures RT to queries as an index of SA (Durso et al., 1998; Willems & Truitt, 1999). The implementation of SPAM in a team situation is difficult because one would need to ask both R- and D-side ATCSs a question and measure the RT. Unless more than one experimenter administers the SPAM query, the manual determination of the RT is not possible. Aside from the difficulty of implementing SPAM in an ATCS team situation, SPAM has the drawback that the ATCS's task during the SPAM response changes to a visual search task. An underlying assumption in SPAM is that the ATCS uses the radar display as an external memory, and the time to find the information is an indication of SA (Durso et al., 1998). One way to use the same approach as SPAM while minimizing the search task is to use the radar display to query the ATCSs by highlighting the aircraft involved in the query. This has the advantage that ATCSs can use the radar display to respond to the query, eliminating the need to verbalize the answer. An additional advantage is that we can time the start of the query and the start and end of the response without the interference of the experimenter.

In SAGAT, the experimenter freezes and suspends the simulation at randomly selected times while operators answer questions about the current situation. The primary disadvantages of this technique are the temporary halt in the simulation and the use of a separate display that is different from the radar display. Further, one potential shortcoming of obtaining an indication of SA by using probes on a few predefined elements is that this may have an effect on participants' attention during testing. Queries given during simulations overcome the limitations of reporting on situation-specific information after the fact.

We refer to the queries that were similar to SPAM questions as Aircraft-pair queries. They assessed more relational information about aircraft pairs. We refer to the queries that were similar to SAGAT questions as Sector-Based queries. These queries assessed more global SA.

3.1.1.1 Data Analysis and Results

We divided the SAVANT data into several data sets (Figure 11). For the analyses, we separated the Aircraft-Pair queries from the Sector-Based queries. We then conducted analyses on a subset of the Aircraft-Pair data containing all responses and a subset containing correct responses only. These subsets contained queries about current and future situations. For the subset of data containing all responses, we conducted MANOVAs on RT and percent correct – one for current and one for future queries. For the subset of data containing correct responses only, we conducted two ANOVAs on the RT – one for current and one for future queries. For the Sector-Based questions, we calculated two variables. First, we calculated the percent of aircraft ATCSs answered. Secondly, we used the information on correct answers identified by the SME to calculate the percentage of correct answers. We then averaged the responses for the four questions that asked about which aircraft were not in communication, did not conform to clearances, needed a hand off to the next sector, and had not completed clearances yet. In all, we had five data sets. The MANOVA for the Aircraft-Pair questions included the percentage of responses and the percentage correctly answered relative to the SME's answers. We conducted subsequent ANOVAs when the MANOVA results were statistically significant. We provide the means, *SDs*, MANOVA, and ANOVA results in Appendix O.

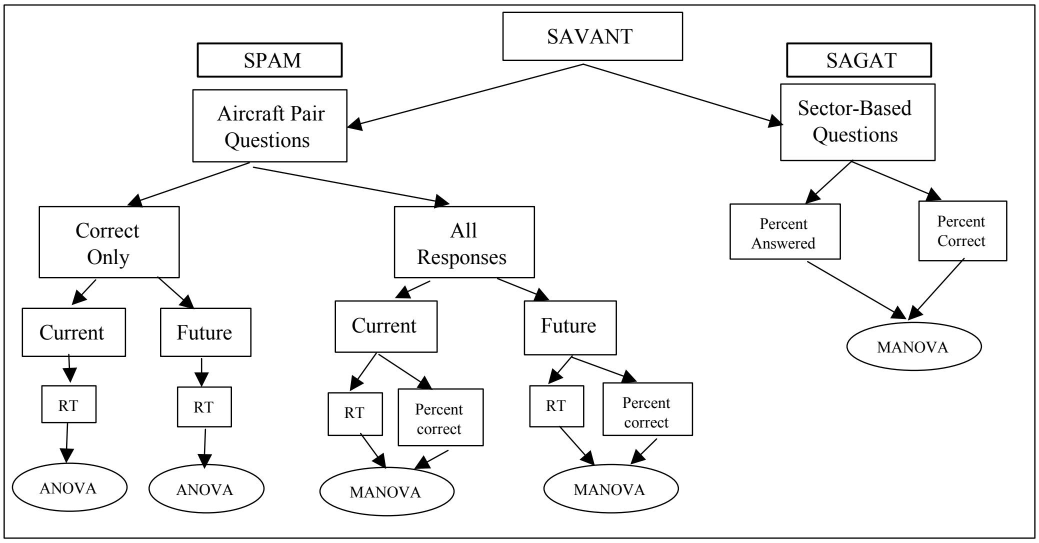

Figure 11. Flow chart for SAVANT data analysis.

3.1.1.1.1 Correct Responses Only – Aircraft-Pair Questions

We conducted two separate 2 x 2 x 3 (position x task load x automation) within-subjects ANOVAs. These were on the RT for correct answers only for the current and future SA questions, respectively. The ANOVA for current SA questions showed a significant effect for position [$F(1,15) = 5.86$, $p < .05$, Appendix O, Table O-5]. When sitting on the R-side, ATCSs responded faster to the current SAVANT questions than when on the D-side (Figure 12). The ANOVA for future SA showed a significant effect for task load [$F(1,15) = 10.81$, $p < .01$, Table O-6]. ATCSs responded faster to the future SA queries when task load was low (Figure 13).

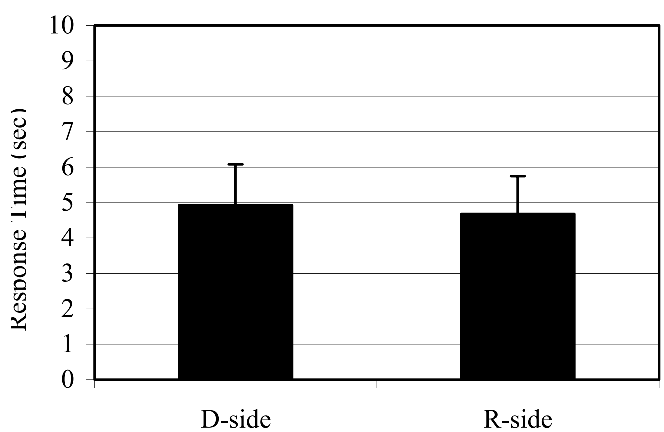

Figure 12. Response time for current correct answers by position.

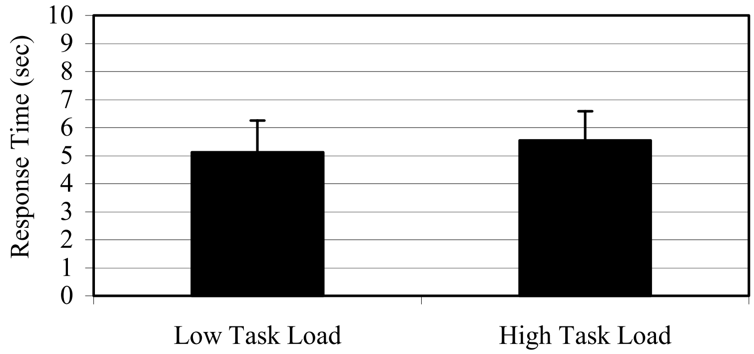

Figure 13. Response time for future correct answers by task load.

3.1.1.1.2 All Responses Answered – Aircraft-Pair Questions

We conducted separate 2 x 2 x 3 (position x task load x automation) within-subjects MANOVAs on the RT and percentage correct for current SA and future SA on all questions answered, respectively. The multivariate analysis for current SA showed significant results for position and task load [$\Lambda = .61$, $F(2,14) = 4.48$, $p < .05$; $\Lambda = .63$, $F(2,14) = 4.15$, $p < .05$, respectively, Table O-7]. The multivariate analysis for future SA showed significant results for task load [$\Lambda = .62$, $F(2,14) = 4.28$, $p < .05$, Table O-10]. Because of the significant multivariate findings, we conducted subsequent ANOVAs and used an adjusted alpha level of .025 to determine significance.

When examining ATCSs' RTs for current SA queries, we did not find significant effects for position using the adjusted alpha level; however, a trend showed that R-side ATCSs responded faster to the current SA queries than D-side ATCSs (Figure 14, Table O-8).

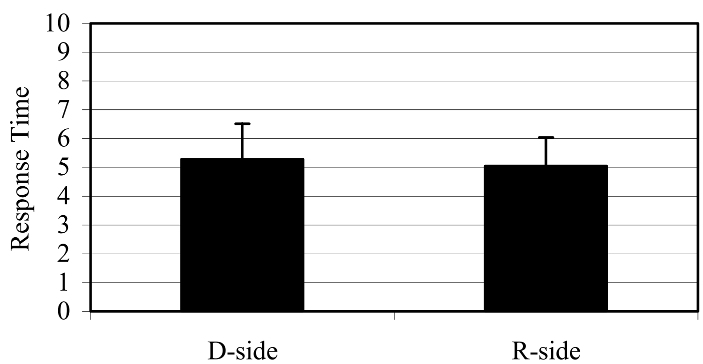

Figure 14. Response time for all responses of current SA by position.

In addition, a secondary trend for automation showed that ATCSs responded faster in the full automation condition than in the limited automation condition. There were no differences between full and no automation or limited and no automation (Figure 15, Table O-8). When examining the percentage of correct responses, the data did not reach statistical significance nor was a trend reached. However, when task load was high, the percentage of correct responses was higher (Figure 16, Table O-9).

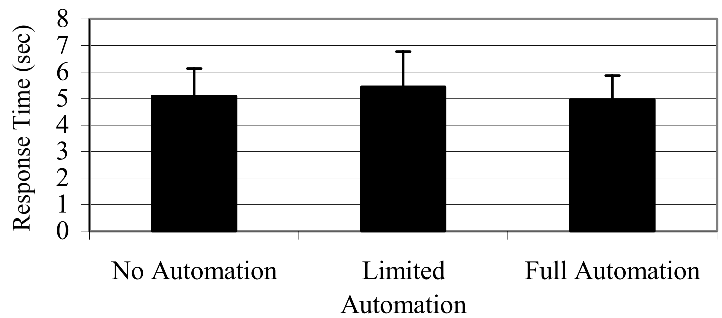

Figure 15. Response time for all responses of current SA by automation.

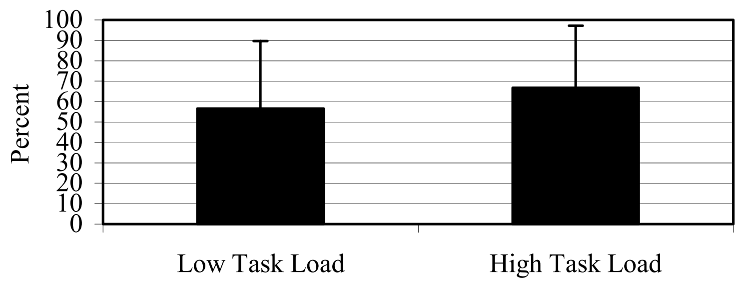

Figure 16. Percent correct for all responses of current SA by task load.

When examining ATCSs RTs for future SA queries, we found a significant effect for task load [$F(1,15) = 7.57$, $p < .05$, Table O-11]. ATCSs responded faster to the queries when task load was low (Figure 17). Analyses for the percentage of correct responses did not show any significant effects.

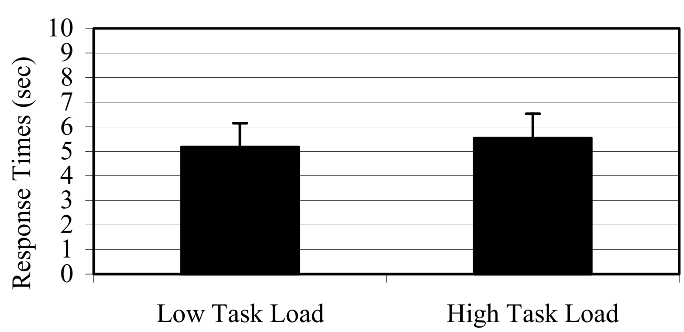

Figure 17. Response time for all responses of future SA by task load.

3.1.1.1.3 Sector-Based Questions

We conducted a 2 x 2 x 3 (position x task load x automation) within-subjects MANOVA for the percentage of responses and the percentage of correct responses relative to the SME's answers. The MANOVA showed a significant effect for position [$\Lambda = .44$, $F(2,14) = 8.70$, $p < .05$, Table O-13]. Because of the significant multivariate findings, we conducted subsequent ANOVAs and used an adjusted alpha level of .025 to determine significance.

The univariate analysis for the percentage of correct answers showed a significant effect for position [$F(1,15) = 13.40$, $p < .01$, Table O-15] and a trend for automation using an alpha of .05. When on the R-side, ATCSs answered a significantly higher percentage of questions correctly than when on the D-side (Figure 18). The trend for automation showed that, as automation increased, the percentage of correct responses decreased (Figure 19).

The univariate analysis for the percent of answers did not show any significant effects for the experimental manipulations. Analysis of the six-participant data set showed a significant three-way interaction [$\Lambda = .02$, $F(2,4) = 21.83$, $p < .05$, Table O-16] that we did not find in the full data set. However, the univariate follow-up analyses did not reach statistical significance. We did find a trend for the position x automation interaction for the percentage of correct responses. Automation had an effect for R-side ATCSs but not for D-side ATCSs. R-side ATCSs' percentage of correct responses was higher when they did not use any automation (Figure 20, Table O-18).

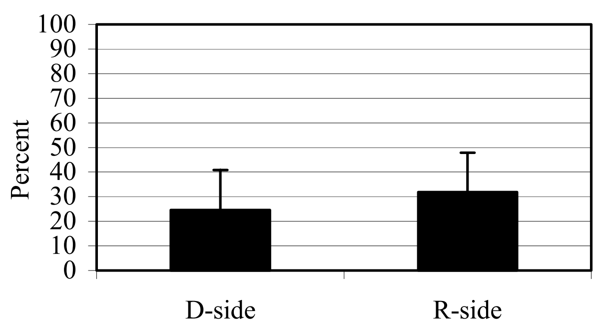

Figure 18. Percent of correct responses for Sector-Based questions by position.

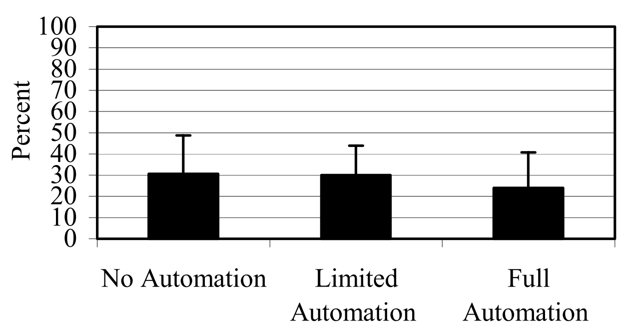

Figure 19. Percent of correct responses for Sector-Based questions by automation.

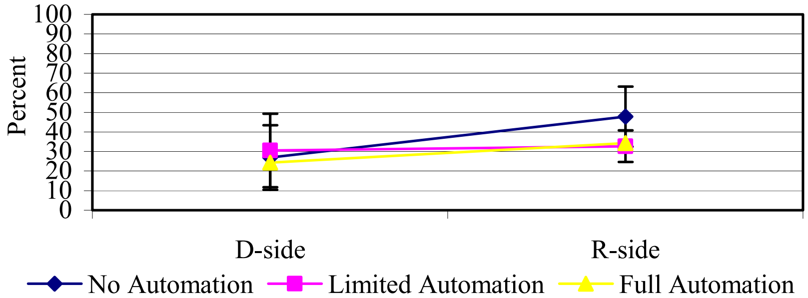

Figure 20. Percent of correct responses for Sector-Based questions by position and automation (N=6).

3.1.1.2 Discussion

The Aircraft-Pair queries assessed SA regarding the state of aircraft in the airspace. However, Sector-Based queries assessed more global SA for action-oriented items such as whether aircraft were conforming to clearances. This contrast in queries resulted in Aircraft-Pair questions being more applicable to both the R- and D-side ATCSs, but the Sector-based queries were more applicable to the R-side ATCSs. The Sector-Based questions focused more on tasks related to the tactical, R-side ATCS (i.e., control actions communicated to aircraft). The R-side ATCS needed to assess the overall comprehension of the actions issued to and performed by the aircraft, and the D-side ATCS gathered more second-hand information about the state of the aircraft location. Because of these differences, we expected to see the effect of experimental variables for both the R- and D-side ATCSs on the Aircraft-pair queries. Because Sector-Based queries focused more on issues related to tactical control of aircraft, we expected only the R-side to be affected by the IVs.

For the data set of all Aircraft-pair responses, a trend showed better SA under full automation compared to limited automation, although it was not significantly different compared to no automation. In the limited automation condition, ATCSs were less involved with the traffic. In contrast, full automation provided the graphic plan display that is a representation of aircraft positions and conflict information visually, allowing the D-side ATCS to stay involved. Even under no automation, the ATCSs had better SA than in the limited automation. In the field, ATCSs training includes building the “picture” from flight strips and maintaining that picture without a graphical representation. Therefore, by giving them flight progress strips, they may have been better able to stay involved than with the limited automation where ATCSs were not familiar with the format of displayed aircraft information.

ATCSs had better awareness of the current situation when working on the R-side. The R-side ATCS actively controls traffic, and the D-side ATCS has other responsibilities to attend to besides focusing on the radar screen. This gives R-side ATCSs an advantage because they are in a more active role related to the traffic compared to the D-side who is in more of a monitoring role. This compliments Willems and Truitt’s (1999) and Endsley and Kiris’ (1995) findings that SA was lower when passively processing information instead of actively processing it.

ATCSs had better awareness for future situations when task load was low. The fewer the aircraft, the easier it was to project them into the future, increasing SA. Because the Aircraft-pair questions assessed SA relevant to both positions, task load did not affect ATCS SA relative to their control position.

The higher percentage of correct responses in the high task load condition would seem to be counter to what we would expect. We believe that the replacement of missing data affected this result. We replaced more missing data in the high task load condition than in the low task load condition. However, ATCSs did worse than what guessing would account for.

Sector-based questions were really more for the tactical, R-side ATCS. When analyzing these questions, we focused on the percentage answered and the percentage correct. We did not use RT as an indicator of SA for these questions. We felt that RT was not appropriate because the first aircraft chosen would have a high RT, and subsequent aircraft tended to be selected in quick succession. Analysis of the mean for the percent of aircraft answered (the number of responses given by ATCSs) compared to the mean number of responses chosen by the SME and indicated that ATCSs did have enough time to click on all applicable aircraft to the chosen query.

SA was higher for these action-oriented items when ATCSs were on the R-side. The R-side needs to have an overall comprehension of whether aircraft are conforming to clearances; in contrast, the D-side ATCS needs to know the state of the aircraft. In addition, a trend showed lower SA under increasing levels of automation. This coincides to findings (Carmody & Gluckman, 1993; Endsley & Kiris, 1995) that SA decreases under automation. In the six participant data set, level of automation differentially affected SA through ATCS position. R-side ATCSs' SA was higher when no automation was present. These results indicate that the experimental manipulations affected R-side ATCSs more. When no automation was present, the D-side ATCS could devote full attention to the DSR screen and assist the R-side ATCS, which freed up some cognitive resources of the R-side ATCS. When no automation was present, the D-side ATCS could manipulate the data blocks allowing for easier viewing and no overlapping information.

The nature of the motor response to the queries resulted in slower RTs than in previous studies (e.g., Endsley, Sollenberger, Nakata, & Stein, 2000; Willems & Truitt, 1999). Before responding to the question, the ATCSs had to grasp the mouse and move it into position to answer the query, which slowed RT. Therefore, small differences may be more significant than the graphs seem to indicate. Another reason that may make it difficult to compare the current results to other studies is that ATCSs did not need to perform a search task because the SAVANT screen highlighted the aircraft in question for easy recognition.

3.1.2 SAGAT

Endsley and Kiris (1995) study contains full documentation on SAGAT, the ATC version. Normally, SAGAT measures three levels of SA by querying the ATCS during freezes in the trial. In the current study, we queried participants only at the end of each scenario. We terminated the scenarios between 33 and 40 minutes to prevent participants from creating a memorization strategy had they had knowledge about the time that SAGAT questions would occur. Each SAGAT query session asked ATCSs about the items in Table 6 (Appendix J).

Table 6. Levels of SA (from Endsley)

Level 1 SA – Perception of Traffic Situation	Levels 2 and 3 SA – Comprehension and Projection of Traffic Situation
aircraft position (+/- 7 miles)	aircraft separation
aircraft callsign	aircraft next fix
aircraft altitude (+/- 300 feet)	open clearances
aircraft ground speed (+/-20 knots)	clearance reception
aircraft heading (+/- 15 degrees)	clearance conformance
aircraft flight path change (vertical, turning)	aircraft communications
aircraft type	flight plan conformance
aircraft level of control	aircraft flight profile (inbound, en route, outbound
	aircraft hand-offs needed
	Special airspace separation

SAGAT contained questions about general traffic, specific aircraft inside and outside active airspace, future status of specific aircraft, and other questions designed to assess SA. We divided the SAGAT queries into several categories based on the level of SA the query assessed — perception, comprehension, or projection. Aircraft position, callsign letter, callsign number, altitude, speed, and heading are verbatim information from taken directly off the DSR display, and we categorized them as addressing Level 1 SA – Perception. Aircraft vertical change and turning are perceptual information gathered from the display over time. Aircraft type and level of control are also categorized as basic information regarding each aircraft that we categorized as part of Level 1 SA.

Aircraft separation problems, clearances not yet completed by the aircraft, clearances received correctly, clearance conformance, aircraft in communication with ATCS, handoffs needed, projected special airspace violations, next fix (projection of future flight path), and flight plan conformance composed the Level 2/3 SA category (comprehension of the status of the aircraft relevant to ATCS goals and projection of their future status).

3.1.2.1 Data Analysis and Results

We compared ATCSs' perceptions of the traffic situation as reported on the SAGAT queries to the actual state of the traffic situation at the end of the scenario. We scored ATCSs' answers as correct or incorrect (within tolerance levels specified in Table 6) and subjected them to an arcsine transformation to correct for non-normality of binomial data. The analyses included MANOVAs on each category of questions and subsequent ANOVAs when the MANOVA results were statistically significant and to assess any trends in the data. Appendix P provides the means, *SDs*, MANOVA, and ANOVA results. The graphs depict the mean from the transformed data set and the *SD* from the raw data set.

3.1.2.1.1 Level 1 SA

There were not enough degrees of freedom to perform a MANOVA on all the Level 1 SA query items. We performed 2 x 2 x 3 (position x task load x automation) within-subjects ANOVAs on each item and used an adjusted alpha level of .0056 to test for significance. We found several secondary trends among the Level 1 SA items. When working on the R-side position, ATCSs' awareness of aircraft location was higher than when on the D-side (Figure 21, Appendix P, Table P-4). A secondary trend of task load for callsign letter showed that ATCSs identified more callsign letters correctly under low task load conditions (Figure 22, Table P-5).

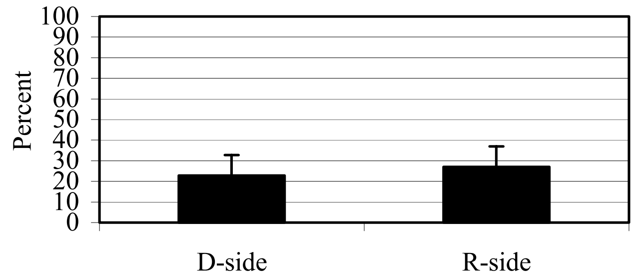

Figure 21. Percent correct for aircraft position by position.

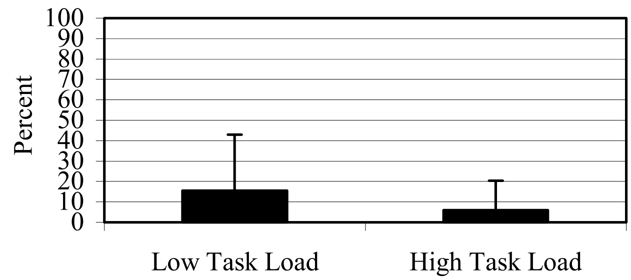

Figure 22. Percent correct for awareness of aircraft callsign letter by task load.

For headings, we found a secondary trend of the task load x automation interaction. Under full automation and high task load, ATCSs' awareness of headings was higher than full automation and low task load (Figure 23, Table P-9). Awareness for aircraft type was highest in the high task load, limited automation condition (Figure 24, Table P-12).

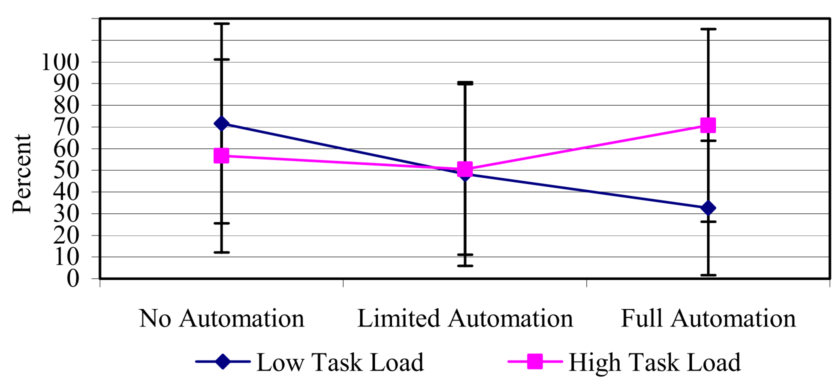

Figure 23. Percent correct for awareness for headings by task load and automation.

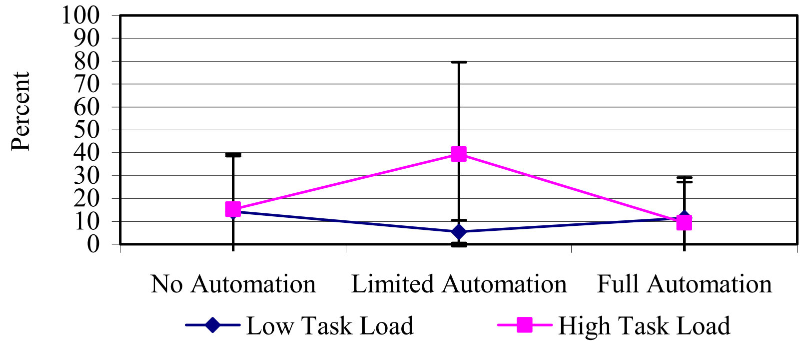

Figure 24. Percent correct for awareness of aircraft type by task load and automation.

3.1.2.1.2 Level 2 SA

There were not enough degrees of freedom to perform a MANOVA on all the Level 2 SA query items. We performed 2 x 2 x 3 (position x task load x automation) within-subjects ANOVAs on each item and used an adjusted alpha level of .006 to test for significance. Secondary trends in the data indicated that awareness of the next fix was lowest under limited automation (Figure 25, Table P-14). When on the R-side, ATCSs' awareness of aircraft flight profiles was higher than when on the D-side (Figure 26, Table P-15). A secondary trend of task load for awareness of separation showed that ATCSs had a higher awareness of separation issues under low task load conditions (Figure 27, Table P-16). Under high task load conditions, ATCSs were more aware of aircraft clearances received (secondary trend; Figure 28, Table P-18). Awareness for clearance conformance was lowest under the low task load, no automation condition (secondary trend; Figure 29, Table P-19).

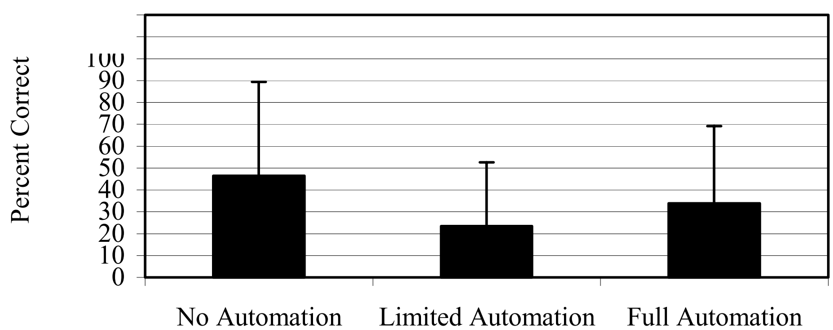

Figure 25. Percent Correct for awareness of next fix by automation.

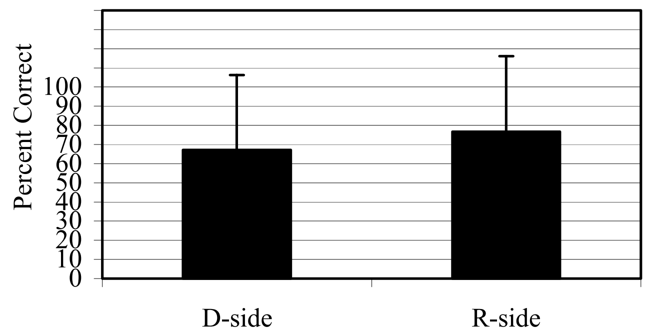

Figure 26. Percent Correct for awareness of flight profile by position.

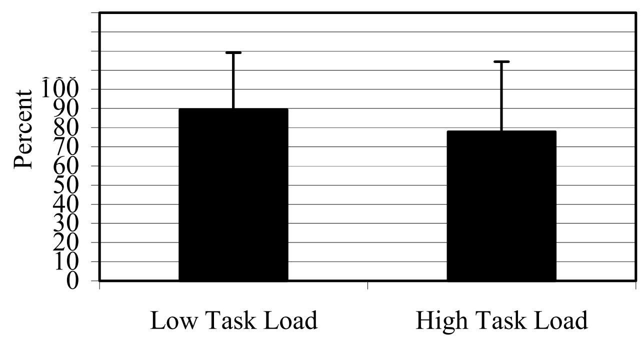

Figure 27. Percent correct for awareness of aircraft separation by task load.

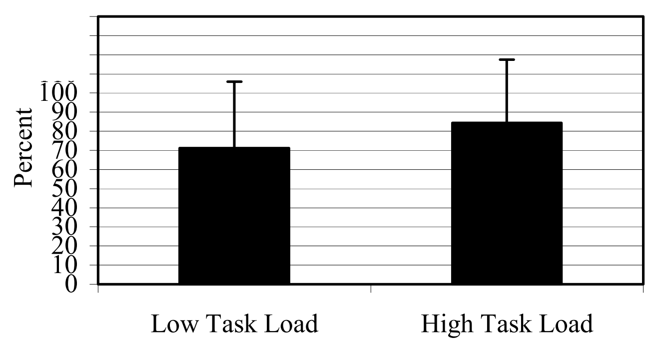

Figure 28. Percent correct for awareness of aircraft clearances received by task load.

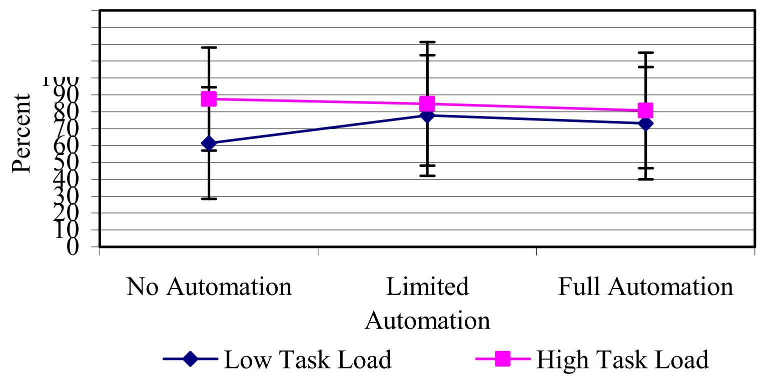

Figure 29. Percent correct for awareness of aircraft conformance by task load and automation.

3.1.2.1.3 Level 3 SA

We found no significant effects or trends for the variables assessing Level 3 SA.

3.1.2.2 Discussion

Differences in automation had an impact on ATCSs' awareness for the aircrafts' next fix. Awareness for the aircraft's next fix was lowest under limited automation conditions. In contrast to full automation or no automation, the limited automation did not allow ATCSs to maintain the "picture" as well. In limited automation, ATCSs only had electronic flight strips and could not pull up the graphical plans display resembling the DSR display. This appeared to hamper their ability to stay up with next fix information. ATCSs have traditionally trained to use paper flight strips to help build their picture, and, in the no automation condition, they may have used this to stay involved. Similarly, they may have used the graphical display provided in the full automation to maintain the picture.

Even though ATCS scores for aircraft positions were low, we did find that R-side ATCSs had better awareness of aircraft location than D-side ATCSs. R-side ATCSs spend more time viewing the radar display and actively controlling traffic, and the D-side ATCSs assist and are in a more monitoring role. The active control of traffic may increase SA as reflected in this item and supported by Willems and Truitt's (1999) findings that SA was higher for active ATCSs than monitoring ATCSs. We had a similar finding for awareness of flight profile. The R-side ATCSs were better aware of aircraft flight profiles – en route, outbound from an airport, or inbound to an airport in the sector – than the D-side ATCSs. This corresponds well with the notion that one has better memory for things that you do yourself (the R-side radioed the aircraft and discussed changes in flight profile with the pilots) than things that are done for you (the D-side received second-hand information after the R-side initiated the control action) (Slamecka & Graf, 1978, as cited in Albright, Truitt, Barile, Vortac, & Manning, 1994).

Under high task loads, ATCSs were less aware of aircraft callsign and aircraft separation; however, ATCSs' awareness for receipt of clearances was higher than under low task load conditions. In SAGAT, we queried ATCSs only on those aircraft that they initially placed on the SAGAT map. Under low task load, ATCSs may be able to recall more aircraft than the ones that required their attention. Under high task load, ATCSs most likely had so many aircraft that required their attention, that they could not recall all of them. Therefore, when queried about whether an aircraft that they recalled had received a clearance, under low task load, they may have remembered less information correctly because the recalled aircraft included unimportant aircraft as well. Under high task load, all recalled aircraft may have been important and whether or not one of these aircraft had received a clearance now was important to the ATCS, resulting in higher scores.

Task load also had an impact on several other variables; however, the level of automation altered that effect. For awareness of aircraft headings, task load had an impact for this item under full automation conditions; ATCSs were more aware of headings when task load was high. With full automation, ATCSs had conflict indications and reached out more to resolve those conflicts, bringing the headings for those aircraft to their attention. In addition, these aircraft may have been important because all recalled aircraft under high task load had the attention of ATCSs. ATCSs had higher SA for aircraft type under limited automation and high task load conditions. In limited automation, ATCSs could easily access aircraft type from the aircraft list. Under high task load, this information seems to have more importance to the ATCSs. Awareness of aircraft conformance to clearances was lowest under low task load, no automation. Under low task load, ATCSs may have remembered less information correctly because the recalled aircraft included unimportant aircraft as well important aircraft. Further, they did not have access to conflict indications and were not as concerned with possible conflicts because the likelihood of them was low. In contrast, under high task load, recalled aircraft included only important aircraft and the potential for possible conflicts was large, therefore, even under no automation conditions, ATCSs paid great attention to aircraft conformance.

Although Endsley (2001) validated aircraft placement from a larger screen to a smaller screen for SAGAT with pilots and found a 5.21 mile error within the tolerance band used in the current study, the overall percentage of correct responses for aircraft positions was quite low, approximately 20%. There are several explanations for the low percentage associated with awareness for aircraft position. First, the discrepancy in size between the DSR screen and the computer monitor may physically create a scaling problem, making it difficult for ATCSs to accurately place aircraft at positions coinciding with the DSR position. The large display used in Endsley's study was approximately 19 in.; in the current study, the DSR display was 29 in. On the other hand, the change in scale between the 29-in. DSR display and the SAGAT-probe display may have caused an additional cognitive challenge for ATCSs. We have not established that a change in scale does not distort ATCS recall for aircraft positions. The probe itself also may not access the relevant information within the ATCSs' knowledge. ATCSs may have the knowledge for aircraft positions relative to one another but cannot express them at the exact location as required by the format provided by the probe. Or, perhaps, the ATCSs really do not have the information stored in memory but instead use the radar display as an external memory. If this is the case, then a good indication of SA may be the speed at which they can find the information from the external memory. For instance, awareness of aircraft's next fix and aircraft type, although not available directly from the radar display, can be accessed from flight plan information. As pointed out by Durso et al. (1998), even though information is not in memory, knowing where to find it can imply good SA. Helbing (1997) suggested that ATCSs use aircraft positions as memory pegs (i.e., they store relevant information about aircraft by aircraft position and are not necessarily able to retrieve aircraft information when probed using cues such as callsign). Helbing's argument suggested that ATCSs store aircraft information in memory using positional memory pegs. The current research findings suggest that ATCSs do not have good memory for exact aircraft locations. Therefore, if Helbing's model is accurate, the spatial representation that ATCSs use to store aircraft information is not simply a scaled version of the actual displayed information.

Albright and Lewandowsky (1995) provide an additional explanation for the low percent correct on recall of aircraft positions. These researchers showed that in a low complexity ATC task, subjects displayed implicit momentum (i.e., when probed, subjects thought aircraft locations were slightly ahead of their actual location). Although tolerance bands used in SAGAT may have absorbed this effect, in this experiment, considering the level of traffic task load (necessary to justify the use of an R- and a D-side), such a forward projection may cause the low level of correct recall of aircraft positions.

Results for other Level 1 SA verbatim items showed that ATCSs' awareness for aircraft callsign was less than 20%, even lower than for aircraft position. In contrast, ATCSs' awareness for aircraft headings reached 60%. Implications point toward lower recall for alphanumeric information and higher for graphical information. Because ATCSs can derive awareness of aircraft heading by formatting the orientation of the vector lines, this gives them an advantage in recall. The results clearly show we cannot expect ATCSs to recall verbatim information. This does not mean that they do not have that information stored in memory; it simply means that the information is not accessible to us for probing. Therefore, in future studies, querying ATCSs for verbatim information will have no added value.

Awareness of receipt of aircraft clearances and conformance to clearances, Level 2 SA, was high. Although, in general, awareness for Level 1 SA items was relatively low, there did not seem to be an adverse effect for Level 2 SA. One could use the SAGAT results to infer that the information-processing model proposed by Endsley and Smolensky (1998) is not correct (i.e., an ATCS would not need Level 1 SA before achieving Level 2 SA). However, the results from this experiment do not indicate whether ATCSs have processed information from Level 1 SA into other variables or formats or that the raw information is still present but cannot be directly measured by Level 1 SA queries.

We did not have significant findings for Level 3 SA, the projection of elements into future situations. This clearly indicates that even with the DST that should have moved ATCSs to a strategic orientation, our ATCSs still controlled traffic in a mostly tactical fashion. This, of course, could also be because our ATCSs were novice DST users. It would be interesting to test ATCSs that have extensive experience using DST and compare those SA results with our current findings. One would hope to find that for the experienced DST users, level 3 SA has increased compared to our novice DST users.

Direct comparison across studies is difficult due to different traffic task loads and different numbers of ATCSs working a sector. In the present study, much higher traffic task loads were present (mean of 25 in low task load and mean of 35 in high task load) than in previous studies (e.g., mean of 19 in low traffic task load and 24 in high traffic task load scenarios) (unpublished data extracted from original data from Willems & Truitt, 1999). However, in the current study, we used both an R-side and D-side ATCS and previous studies used only an R-side ATCS. Previous research shows that there is a marked effect of traffic task load on SA as the number of aircraft present exceeds approximately 12 to 15 (Endsley & Rodgers, 1998). ATCSs will attend to only a subset of more important aircraft, thus lowering the percentage of aircraft attended to overall. The much lower absolute number correct in the present study (mean of 4.11 per trial), however, indicates far more than a workload effect. It is most likely reflecting a lowering of awareness for aircraft that had ceased to be important after the ATCS handed them off. Another explanation for this may be that with the time pressure of the current traffic, ATCSs' dynamic memory span reverted to approximately three items (Moray, 1986). Under lower traffic task loads and with only an R-side ATCS, Willems and Truitt found recall for 10 items. Therefore, the results presented here reflect lower SA than would be present for active traffic and perhaps the pattern of attention across elements of the situation would be different.

3.1.3 Post-Scenario Questionnaire

3.1.3.1 Data Analysis and Results

We conducted a 2 x 2 x 3 (position x task load x automation) within-subjects MANOVA on three items assessing participants' self-reported SA. These items included SA for current aircraft, projected aircraft locations, and potential violations. The multivariate analysis showed significant results for position and task load [$\Lambda = .52$, $F(3,13) = 3.95$, $p < .05$ and $\Lambda = .32$, $F(3,13) = 9.30$, $p < .001$], respectively, Appendix Q, Table Q-2. Because the effects of position and task load were significant at the multivariate level, we conducted subsequent ANOVAs for each SA item and used an adjusted alpha of .017 to determine significance. See Appendix Q for the means, *SDs*, MANOVA, and ANOVA tables.

When examining ATCSs' SA for current aircraft locations, we found significant effects for both position and task load [$F(1,15) = 11.33$, $p < .01$ and $F(1,15) = 11.91$, $p < .01$], respectively, Figure 30, Table Q-3. When in the R-side position, ATCSs rated their perceived SA for current aircraft locations higher than when they were in the D-side position. Further, when controlling traffic in a low task load scenario, ATCSs perceived they had higher SA for current aircraft locations than in high traffic task load scenarios.

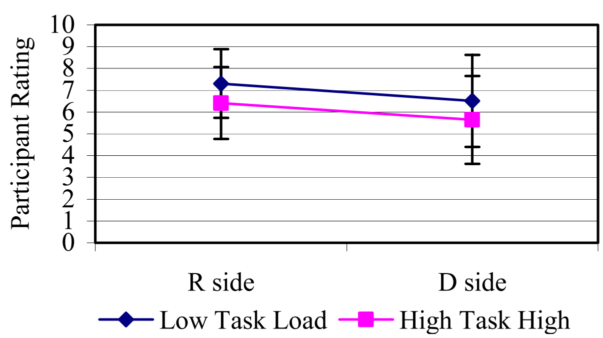

Figure 30. SA for current aircraft locations by task load.

For the item assessing SA for projected aircraft locations, only a significant effect of task load occurred [$F(1,15) = 12.76$, $p < .01$, Figure 31, Table Q-4]. Once again, ATCSs perceived they had higher SA for projected aircraft locations when controlling traffic in a low task load scenario compared to a high task load scenario.

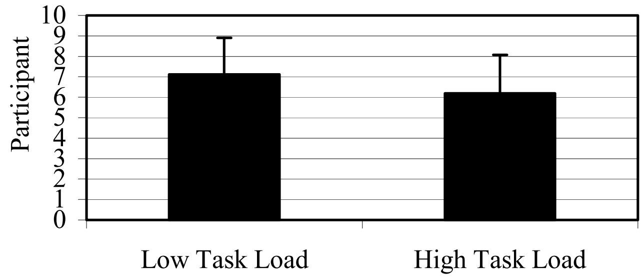

Figure 31. SA for projected aircraft locations by task load.

When asked to rate their SA for potential violations, ATCSs reported higher levels of SA under low task load scenarios in comparison to high task load scenarios [$F(1,15) = 22.21$, $p < .001$, Table Q-5]. The univariate results showed a secondary trend for the interaction between task load and automation (Table Q-5). As shown in Figure 32, the secondary trend indicated task load had the largest effect under the full automation condition. When ATCSs used full automation, their SA for potential violations was lower when the task load was high compared to low.

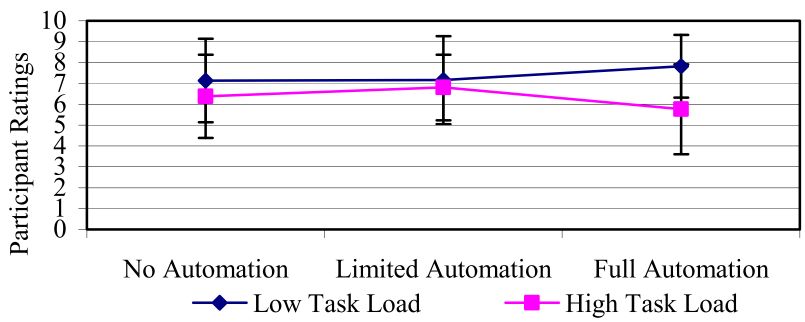

Figure 32. SA for potential violations by task load and automation.

3.1.3.2 Discussion

When assessing SA for potential violations, the degree of automation affected the impact of task load. The trend indicated that when automation was not present or was limited, the effects of high and low task load were not significantly different. In contrast, when ATCSs used full automation, high task loads led to ratings of lower SA for potential violations. This finding relates to other findings that indicated ATCSs felt the DST behavior was less predictable when task load was high and full automation was present. In the case of the DST, the ATCS had to constantly monitor it to fully utilize it. Constant monitoring of automation induces higher workload and can result in lower levels of SA. Under low task load conditions, the ATCS may be able to compensate for this, but, when task load is high, the ATCS is less able to maintain the picture and maintain SA. The ATCS's experience level using the automation may have contributed to lower SA for potential violations under high task loads. Our participants were novice DST users, and, as such, their skill level using the DST was not very high. Under low task load conditions, they had enough time to compensate for their lack of experience with the tool; however, under high task load conditions, their lower skill level with the DST became a factor influencing SA. We would expect that with additional training and experience using the DST, we would not see such a drastic drop in SA under high task load conditions. The amount of training and experience using the DST to offset this detriment in SA would need to be investigated.

Task load had a significant impact on all SA items. Increasing task load led to decreases in perceived SA for current aircraft locations, projected aircraft locations, and potential violations. Results imply that higher task load scenarios had an impact on the cognitive resources ATCSs allocated to SA. When task load was low, ATCSs were able to keep focus on the big picture and displayed higher SA.

Position affected SA for current aircraft locations. R-side ATCSs rated their SA higher along this dimension than D-side ATCSs. This item reflects the differences in the responsibilities of the R- and D-side ATCSs. The R-side ATCS actively scanned the DSR screen and directed traffic, but the D-side ATCS had other responsibilities that did not focus on the scope leading to lower SA.

3.1.4 Over-the-Shoulder

3.1.4.1 Data Analysis and Results

Maintaining awareness of aircraft positions, ensuring positive control, detecting pilot deviations from control instructions, and correcting own errors in a timely manner composed the maintaining attention and SA category. Appendix R contains the means, *SDs*, MANOVA, and ANOVA tables.

The MANOVA showed a significant effect of task load across the set of DVs [$\Lambda = .07$, $F(4,4) = 12.52$, $p < .05$, Table R-2]. The SME felt that ATCSs maintained awareness of aircraft positions, ensured positive control, and corrected their own errors in a timely manner better when the task load was low [$F(1,7) = 40.69$, $p < .001$, $F(1,7) = 56.52$, $p < .001$, and $F(1,7) = 20.05$, $p < .01$, respectively, Figure 33, Appendix R, Table R-3, Table R-4, and Table R-6]. A trend in the data suggests the ATCSs also detected pilot deviations from control instructions better under low task load conditions.

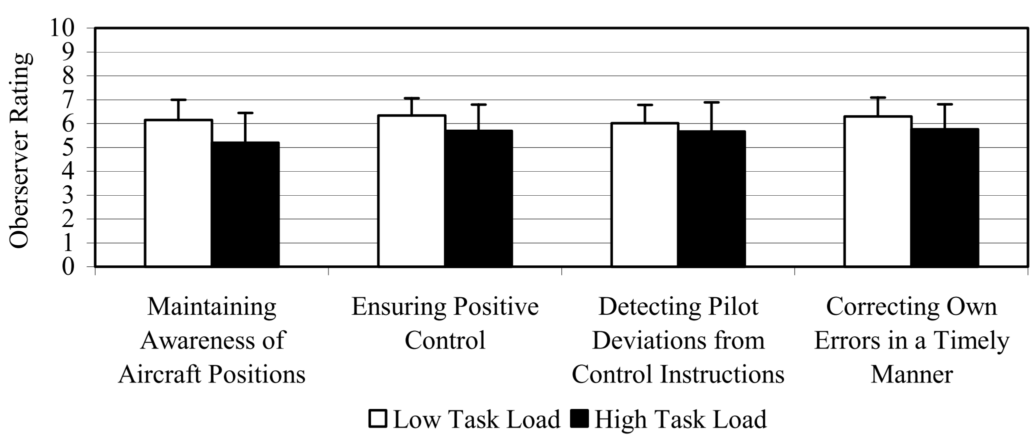

Figure 33. Maintaining attention and SA by task load.

3.1.4.2 Discussion

Interestingly, automation did not have an effect on the SME's ratings for SA. However, we expected that automation would have an effect on SA items from the OTS. These results imply that the change in automation levels did not affect ATCSs in any discernable manner along the maintaining attention and SA items, as observed by the SME. Due to limitations in resources, we used only one SME to complete OTS forms. The SME predominantly focused on the R-side ATCS. We wrongly assumed that any changes in the team should be shown on both sides. The R-side functioned as if everything was normal, but the D-side was using the automation. The automation did not directly affect the R-side as it did the D-side, as viewed by the SME. Therefore, an automation effect was not found.

Task load affected the SME's ratings of ATCSs' SA. High task loads contributed to lower SA ratings. Once again, high task loads made it difficult for the ATCS to maintain the big picture. This led to ATCSs loosing SA.

3.2 Workload

In this study, we used ATWIT, NASA TLX, and PSQs to assess participant workload.

3.2.1 Air Traffic Workload Input Technique

Stein (1985) first introduced ATWIT, which is an online measure that requires ATCSs to indicate, at set times, their perception of their current workload. ATWIT is an instantaneous probe that investigates overall perceived workload. Contrary to the NASA TLX (Hart & Staveland, 1988), the participants do not need to break down their workload by origin. Another advantage of the ATWIT over post-scenario ratings of workload is that ATWIT measures workload during the simulation instead of relying on participant's memory after the scenario. The ATWIT measure is a workload estimate based on a scale from 1 to 10 (Appendix F contains the detailed instructions that accompany the ATWIT device). The anchors used for the ATWIT scale relate directly to the task. The ATWIT device is a touch sensitive panel that displays a start button at the beginning of the simulation and a 10-point scale during the simulation. The ATCSs, prompted by a low tone, made a workload rating every 3 minutes. Each participant made 10 ATWIT ratings in a scenario allowing calculation of the mean and maximum rating for each scenario.

3.2.1.1 Data Analysis and Results

We created four workload-related data sets. The first two data sets contained 12 x 16 (scenarios x ATCSs) records that included the summary variables calculated per scenario for the full data set and the six-participant data set. The other data sets contained 12 x 10 x 16 (scenarios x intervals x ATCSs) records containing the summary variables calculated per 3-minute interval for the full and partial data sets.

To analyze the effect of the IVs on the subjective ratings, we used a MANOVA on the mean ATWIT ratings and mean response times (RTs). This MANOVA, structured as a 2 x 2 x 3 (position x task load x automation) repeated measures design, addressed the differences across scenarios. We also conducted a 2 x 2 x 3 (position x task load x automation) ANOVA on the maximum ATWIT ratings. To investigate the effect of time-on-task, we used ANOVAs on ATWIT ratings and RTs in a 2 x 2 x 3 x 10 (position x task load x automation x interval) repeated measures design because we did not have enough degrees of freedom for a MANOVA. See the tables in Appendix S for detailed results. We do not present redundant results from the scenario and interval data sets.

3.2.1.1.1 Mean ATWIT Ratings and Response Times

The 2 x 2 x 3 (position x task load x automation) within-subjects MANOVA showed significant effects for position, task load, automation and the task load x automation interaction [Λ = .36, $F(2,14) = 12.49$, $p < .01$, Λ = .11, $F(2,14) = 54.82$, $p < .001$, Λ = .43, $F(4,12) = 3.92$, $p < .05$, Λ = .62, $F(2,14) = 4.38$, $p < .05$, respectively, Appendix S, Table S-3]. Because of the significant MANOVA results, we performed subsequent ANOVAs. We used an adjusted alpha of .025.

The ANOVA for the mean ATWIT rating showed a significant effect for task load [$F(1,15) = 110.20$, $p < .0001$, Table S-4]. ATCSs' perceived workload increased with task load. Automation did not significantly affect perceived workload. The analysis did reveal a secondary trend that showed that perceived workload increased with increasing automation (Figure 34).

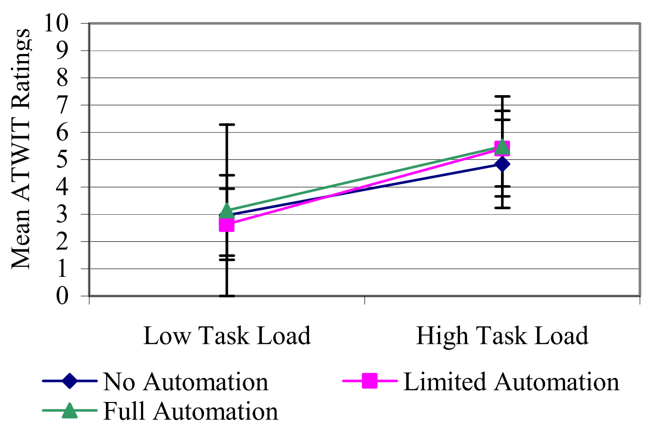

Figure 34. Mean ATWIT ratings by task load and automation.

The ANOVA examining RTs showed a significant effect for position [$F(1,15) = 24.81$, $p < .001$, Table S-5] and a significant interaction between position and task load [$F(1,15) = 7.03$, $p < .05$, Table S-5]. ATCSs responded to the ATWIT device faster when working on the R-side than the D-side. The level of task load affected this. The simple effect of position within low task load was significant [$F(1,15) = 30.94$, $p < .001$, Table S-5]. ATCSs responded faster when sitting on the R-side compared to the D-side; this effect was not significant when task load was high (Figure 35).

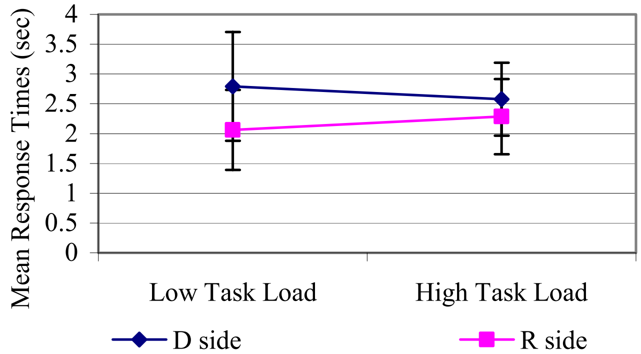

Figure 35. Mean ATWIT response time by task load and position.

3.2.1.1.2 Maximum ATWIT Ratings

The ANOVA examining the maximum ATWIT ratings showed significant effects for task load and automation [$F(1,15) = 145.18$, $p < .0001$ and $F(2,30) = 3.44$, $p < .05$, respectively, Figure 36, Table S-9]. Increasing task load caused an increase in maximum ATWIT ratings. When in the full automation condition, ATCSs rated their maximum workload higher than when they were in the limited automation condition. The full and no automation and limited and no automation conditions were not statistically different.

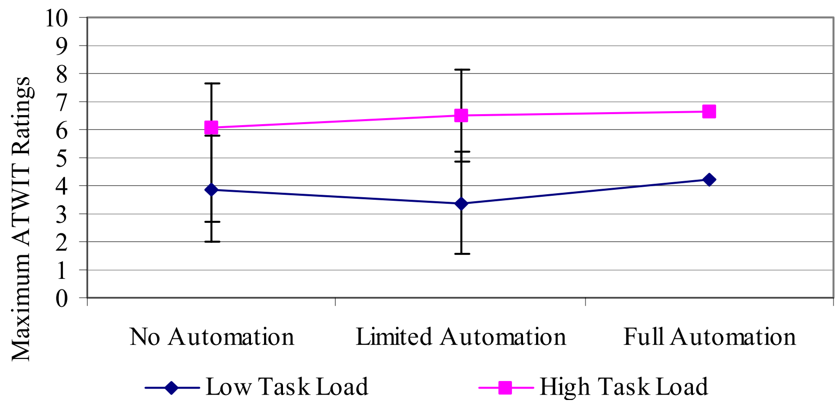

Figure 36. Maximum ATWIT rating by task load and automation.

3.2.1.1.3 Effect of Time-On-Task

We conducted separate 2 x 2 x 3 x 10 (position x task load x automation x interval) ANOVAs on ATCSs' ratings and RTs because there were not enough degrees of freedom to perform a MANOVA. We analyzed the effect of time-on-task using two separate ANOVAs and adjusted the alpha level to .025. We present results that are unique to only this data set and are not redundant with previously reported findings.

Examination of the simple effects of task load within interval showed significant results for task load within each interval [Fs(1,15) = 54.30, 57.68, 88.02, 72.76, 85.71, 130.94, 117.23, 94.85, 96.75, 130.04 , all ps < .001, respectively, Figure 37, Table S-11]. Over time, workload under low task load slowly decreased, but under high task load, it remained relatively constant. The

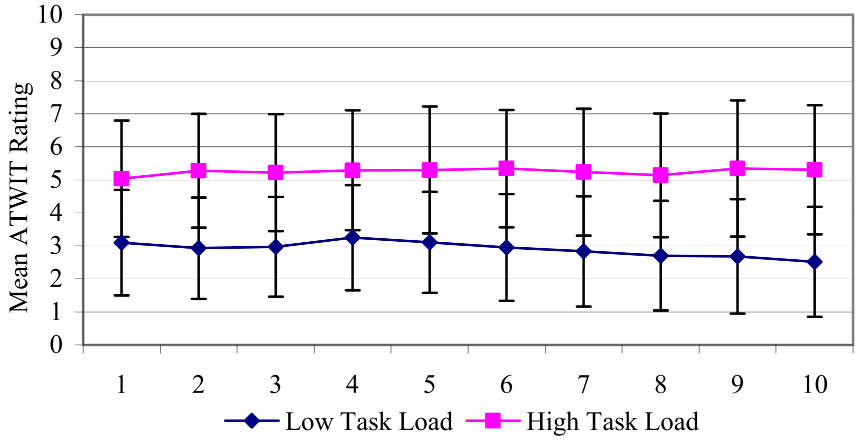

Figure 37. Mean ATWIT ratings by task load and interval.

results from the ANOVA using the last six participants showed slightly different results. We found significant position x task load interaction [$F(1,5) = 15.61$, $p < .05$, Figure 38, Table S-13]. The simple effects of task load were significant for both positions [Fs(1,5) = 105.71 and 31.00, ps < .01, respectively, Table S-13]. The simple effect of position within high task load was significant [$F(1,5) = 6.90$, $p < .05$, Table S-13], but the simple effect of position within low task load was not significant [$F(1,5) = .62$, $p > .05$, Table S-13]. The level of task load affected ATCSs' perceived workload when they were in both positions, although this effect was not as strong when they were in the D-side position (Figure 38).

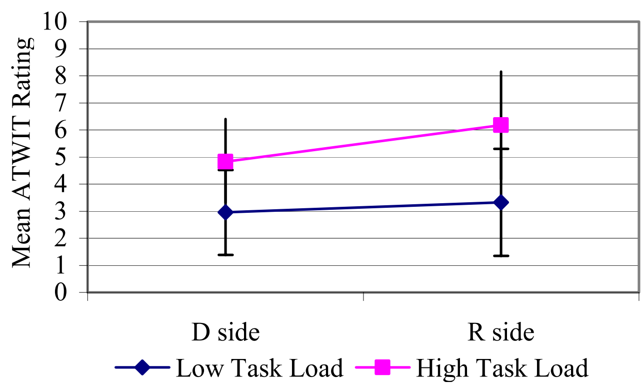

Figure 38. Mean ATWIT ratings by task load and position (N=6).

3.2.1.2 Discussion

Full automation did not statistically differ from the no automation conditions, although a trend did indicate an increase in perceived workload. According to Bressolle, Benhacene, Boudes, and Pari (2000), the introduction of automation tools in the ATC workplace will lead to an increase in ATCS effort while adjusting to the automation and attempting to incorporate it into the regular ATC tasks. During the introduction of automation tools, ATCSs will not yet possess automated behaviors that would make the use of the automation fast and effortless. When using the automation during that learning phase, its use will require extra effort, resulting in an increased workload. We remind the reader that we required that ATCSs did not have previous exposure to the DST. Although we gave them training and hands-on practice for several days, this was not enough to create automated behaviors. The increase in workload with full automation may be an indication of that. In practical terms, we expected one of two things to happen. Most likely, in high traffic task load situations, ATCSs would revert to their primary goal (safety) and leave the automation unused. During our experiment, we did observe that. However, even with ATCSs using the tool less in high task load situations, the perceived workload increased. The new format of presentation required ATCSs to consciously retrieve the information from the automation tool requiring increased effort. The alternative scenario was that ATCSs would attempt to use the automation tool to the best of their ability even in high task load situations. This would lead to larger increases in workload. Depending on the workload present, because of the traffic situation, this may distract (especially the D-side) from their regular duties. Questions arise as to how long this increase in workload lasts for novice DST users and whether workload eventually decreases once the ATCSs become expert DST users. The data from the current study do not address these issues.

Our manipulation of task load clearly resulted in an increase of perceived workload. ATCSs, however, rated their workload to be moderate under high task load conditions. During the experiment, we instructed ATCSs on how to rate the instantaneous subjective assessment of their workload. To help them rate their workload consistently, we provided operational anchors with several of the ATWIT values (Appendix F). At the low end of the scale, ATCSs should be able to accomplish all ATC tasks easily. However, at the high end of the scale, ATCSs would have to leave some of the ATC tasks unfinished. During the high task load scenarios, we observed ATCSs controlling traffic in a manner that reflects the high end of the ATWIT scale (i.e., some of the tasks remained unfinished). This could manifest itself in, for example, the late initiation or acceptance of assuming a strategy that reflects controlling for safety rather than elegance. The fact that ATCSs indicated that even under high task load scenarios, their perceived workload was only moderate shows an underestimation of their actual workload. The reader should keep in mind that our analyses focused on the differences between task load levels within each ATCS. The graphical representations show the average ratings between participants. We know from the raw data and our exit debriefing sessions that some of the individual ATCSs rated workload to be very high. The task load levels we implemented in the experimental scenarios were higher than what ATCSs currently experience in positions staffed with an R- and D-side team. In addition, the six-participant data set showed a stronger effect for task load for R-side ATCSs. The level of task load affected R-side ATCSs more directly than D-side ATCSs. The R-side ATCSs were actively controlling and communicating with traffic, thus increases in task load directly had an impact on the number of actions they needed to perform, which was then reflected in higher workload ratings. Willems and Truitt (1999) found lower workload ratings under monitoring conditions than active control. We would then expect that for the D-side ATCS, only a slight difference in workload ratings should be seen with increasing task load.

Workload ratings gradually decreased for the low task load scenarios; however, ratings for the high task load scenarios remained relatively constant over time. As the scenario progressed, ATCSs had an opportunity to "catch up" after the initial scenario start and, because task load was low, ATCSs had a chance to preplan better and stay ahead of the traffic, which led to perceptions of less workload. In contrast, the high task load scenarios did not give ATCSs as much of a chance to catch up and preplan.

When sitting on the R-side, ATCSs responded faster to the ATWIT device than they did sitting on the D-side; this effect was stronger in the low task load conditions than in the high task load conditions. Several reasons may account for this effect. Although the ATWIT device was located immediately to the left of the D-side position, when controlling traffic, the D-side ATCSs were physically sitting between their position and the R-side position. When ATWIT went off, the D-side ATCS then had to maneuver closer to the ATWIT to respond, but the R-side ATCS only needed to quickly reach up and hit it. In addition, R-side ATCSs may have responded faster to the ATWIT because they wanted to focus on their task of controlling traffic. By answering ATWIT quickly, they could continue with their task and not need to devote any cognitive resources to remember to hit it later (although, they were instructed to immediately respond to ATWIT when it went off). In contrast, D-side ATCSs, while performing their duties, may have given more cognitive resources to recall that ATWIT went off and finished their duties before responding. The gradual decrease for low task load and the constant level of workload for high task load clearly reflects this.

3.2.2 NASA Task Load Index

The NASA TLX consisted of six questions that asked for ratings of mental, physical, and temporal demands as well as performance, effort, and frustration levels (Hart & Staveland, 1988).

3.2.2.1 Data Analysis and Results

A 2 x 2 x 3 (position x task load x automation) within-subjects MANOVA examined the effects of the experimental variables across the set of NASA TLX items. We provide the means, *SD*s, MANOVA, and ANOVA tables in Appendix T. We found a significant effect for task load [Λ = .04, $F(6,10) = 42.14$, $p < .0001$, Appendix T, Table T-2] and a significant position x task load interaction [Λ = .14, $F(6,10) = 9.87$, $p < .01$, Table T-2]. Multivariate analyses of the simple effects of task load within position showed a stronger effect of task load for the D-side position [Λ = .01, $F(6,10) = 99.50$, $p < .0001$, Λ = .09, $F(6,10) = 15.93$, $p < .001$ for the R-side and D-side, respectively, Table T-2]. Position within task load indicated a stronger effect for position under low task load conditions than high task load conditions [Λ = .01, $F(6,10) = 124.45$, $p < .0001$; Λ = .12, $F(6,10) = 12.77$, $p < .001$, respectively, Table T-2]. Because of the significant task load effect and position x task load interaction, we performed subsequent ANOVAs for each item and used an alpha set at .0085 to determine significance.

ATCSs rated the mental, physical, and temporal demands of the scenarios and their levels of effort and frustration higher when controlling traffic from the high task load scenarios than the low task load scenarios [$F(1,15) = 148.82$, 30.20, 107.62, 33.89, and 49.60, respectively, all at

$p < .0001$, Figure 39, Table T-3, Table T-4, Table T-5, Table T-7, and Table T-8, respectively]. ATCSs rated performance lower when traffic was from the high task load condition than the low task load condition [$F(1,15) = 21.51$, $p < .001$, Table T-6]. The analyses also showed several secondary trends in the data. ATCSs experienced higher levels of temporal demand under full automation (Figure 40). ATCSs rated their performance better when on the R-side than the D-side (Figure 41). As the level of automation increased, the ratings for performance decreased. Further, the effect of automation was stronger under high task load conditions (Figure 42).

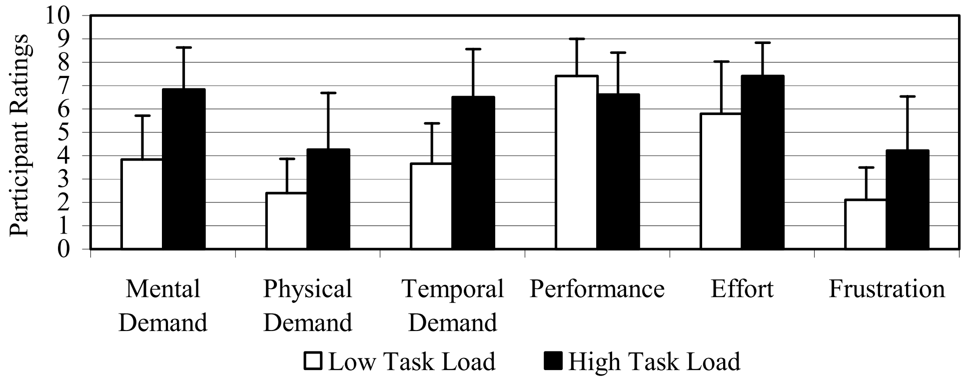

Figure 39. NASA TLX items by task load.

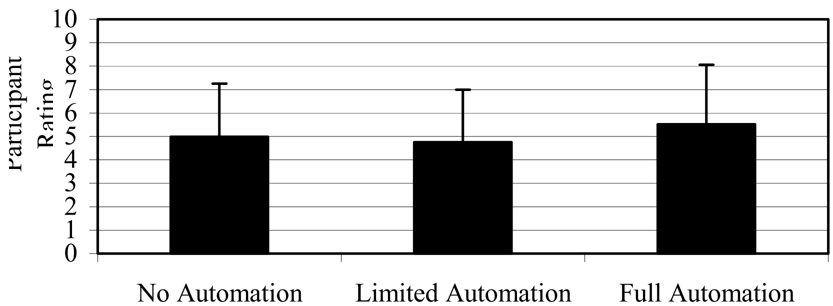

Figure 40. Temporal demand by automation.

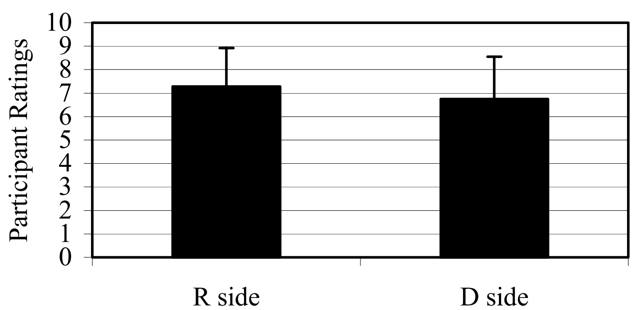

Figure 41. Performance by position.

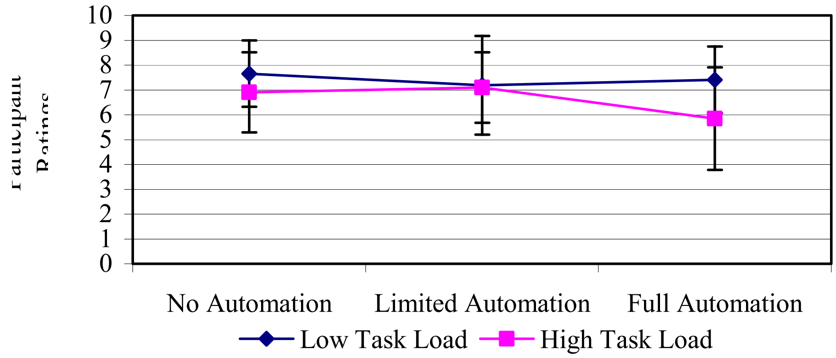

Figure 42. Performance by task load and automation.

3.2.2.2 Discussion

Task load modified the effect of automation. ATCSs rated their performance relatively equal in the low task load condition regardless of the degree of automation. In contrast, when working with high task loads of traffic, ATCSs felt their performance decreased when full automation was present. The ATCSs may have felt their performance decreased when full automation was present because, with full automation, the D-side ATCSs' responsibilities differed from their normal task requirements, and this may have altered their perception of their performance quality. In addition, the use of full automation changed the ATCSs' usual tasks and replaced them with slightly different tasks of monitoring the automation and interpreting the information. The ATCSs did not have a frame of reference for completing these new tasks and, when task load was high, believed they performed worse. During the debriefings, ATCSs mentioned that the use of the DST took them away from their D-side duties and out of the picture. It may also imply that ATCSs have a tendency to underestimate workload levels.

The manipulation of task load had a strong impact on the NASA TLX items. Results indicated that increases in task load led to higher ratings for mental, physical, and temporal demand, effort, and frustration. ATCSs stated their performance was worse when controlling traffic from the high task load scenarios. We expected higher ratings for workload measures with higher levels of traffic. At the volumes of traffic presented in the high task load scenarios, the ATCSs had a constant push of traffic that did not allow for them to "catch up" and preplan actions. This would affect perceived workload along the TLX items.

Besides task load and automation, position affected self-reported performance ratings. ATCSs felt they controlled traffic better when on the R-side in comparison to the D-side.

3.2.3 Post-Scenario Questionnaire

3.2.3.1 Data Analysis and Results

We conducted a 2 x 2 x 3 (position x task load x automation) within-subjects ANOVA on the item asking ATCSs how hard they were working during the scenario (Appendix U). Automation changed ATCS perception of how hard they worked [$F(2,30) = 3.51$, $p < .05$, Figure 43, Appendix U, Table U-1]. Tukey HSD post hoc tests showed that ATCSs stated they worked harder when full automation was present than when they had no automation. ATCSs' perceptions of working hard were not statistically different between the full and limited automation levels, or between the limited and no automation conditions. Task load or position did not influence their perceptions of working hard in the full data set. However, in the data set composed of the last 6 participants, ATCSs described working harder under high task load conditions than under low task load conditions [$F(1,5) = 29.91$, $p < .01$, Figure 44, Table U-2].

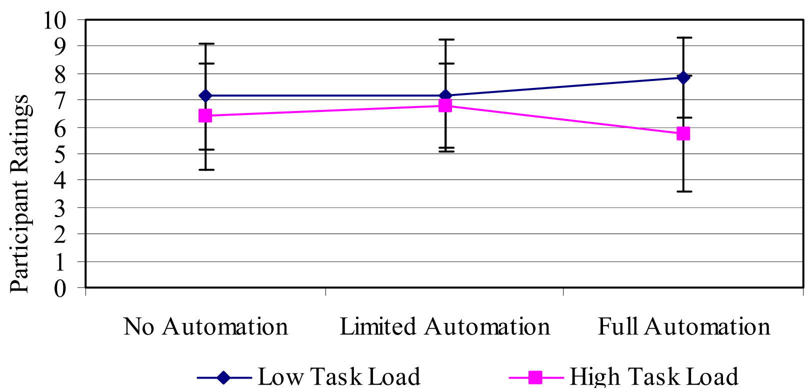

Figure 43. Working hard by automation.

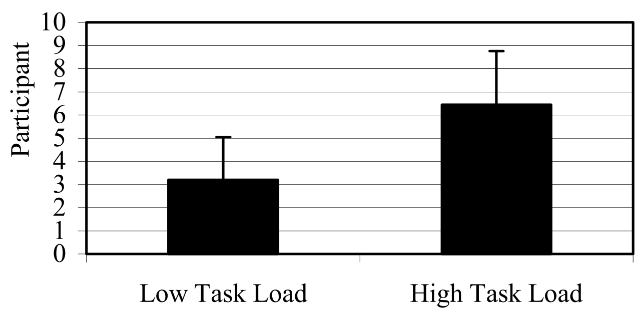

Figure 44. Working hard by task load (N=6).

3.2.3.2 Discussion

ATCSs described working harder when they used full automation relative to no automation. This coincides with other researchers findings (e.g., Wiener, 1985) that the use of automation often leads to perceptions of increased workload. The automation requires the user to monitor it and still perform their other tasks, in this case assisting the R-side ATCS. Researchers have documented that user monitoring of the system influences perceptions of higher workload (e.g., Becker et al., 1991). Of interest though, is that the self-reported workload ratings under full automation were only moderate. This would imply that ATCSs still have resources to devote to other tasks and that the introduction of automation did not overburden them.

Results from the six-participant data set showed higher ratings for working hard when ATCSs controlled traffic from the high task load scenarios. This would be expected. Higher volumes of traffic require more control actions, increasing workload. Position did not affect the perception of working hard.

3.3 Eye Movements

Air traffic control is a visually demanding task. ATCSs continuously monitor the radar display, flight progress strips, and other information displays for information that may require them to take action. The data gained from the visual scanning task (i.e., the monitoring and processing of visual information) forms the ATCS's main source of information. The ATCS integrates information obtained through radio communications, phone lines, his or her knowledge of the airspace, procedures, and other sources with that obtained from the radar display and other visual displays. To obtain data from visual displays, the ATCS scans them for new data or for situations that require examination that is more careful. The ATCS eye movements during the visual scanning task consist of stationary periods or fixations, jumps between fixations or saccades, and eye blinks (Willems et al., 1999).

Rötting (1999) reported fixation durations for reading tasks that averaged 225 ms. Fixation durations for image observations and search tasks averaged 330 ms and 250 ms, respectively. Although extensive research is available on eye movement characteristics in reading, the ATC task is more than reading alone. The long fixation durations of 500-600 ms in ATC (Willems et al., 1999) clearly reflects this.

In previous studies at the RDHFL, researchers used eye movement characteristics to measure changes in visual scanning behavior as a function of experimental conditions (Stein, 1992; Willems et al., 1999; Willems & Truitt, 1999). Willems et al. were the first to use head-mounted oculometry synchronized with dynamic ATC events. They were able to calculate the eye movement characteristics of participants looking at the radar display versus other data displays and aircraft versus other objects. Fixations on aircraft representations were substantially longer than on any other object. These researchers concluded that it is likely that more information processing takes place during the longer eye movement fixations on aircraft representations.

Willems and Truitt (1999) implemented several measures derived from the conditional information index, first introduced by Ellis (1986). Willems and Truitt adapted the conditional information index to investigate the randomness of the visual scanning distribution. Eye movement characteristics did not change as a function of involvement or task load. However, when the ATCS actively controlled traffic under high task load, the visual scan was less structured than when the ATCS passively monitored traffic. The results indicate that ATCSs maintain a better scan of traffic when they actively control it. At the outset, ATCSs rely more on relief briefing information, and, therefore, they can manage with shorter fixations. Once controlling traffic, the ATCS may need to fixate longer in order to evaluate the relevance of the displayed information.

The visual scanning measures in the present study focused on the D-side ATCS. In this experiment, we implemented the R-side and D-side ATCS positions. Our previous experiments have limited themselves to R-side ATCSs. The roles and responsibilities of the D-side ATCS are quite different from those of the R-side ATCS and will vary from ARTCC to ARTCC. The D-side ATCS functions as a second pair of eyes that assist the R-side ATCS in looking for information on the radar display. This would suggest that under normal circumstances, the D-side ATCS should display similar visual scanning behavior when looking at the radar display as we have previously found for R-side ATCSs (e.g., Willems & Truitt, 1999). The D-side ATCS, on the other hand, also has responsibilities that will move his/her attention away from the radar display. With the use of automation tools such as the one used in this experiment, one would expect the D-side ATCS to spend more time away from the radar display, attending to and resolving more strategic problems using the DST.

3.3.1 Data Analysis and Results

In this experiment, we used operational hardware and software to replicate the ATCS operational environment. Therefore, we developed software to record visual scanning related data. We collected two redundant data sources. The first data set consisted of the SAR tapes collected from the Host Computer System. The second data set consisted of Extensible Markup Language (XML) recorded by our middleware (Appendix E). These XML-files contained several messages normally available on the DSR token ring that we needed to determine the location and size of objects displayed on the radar display. Our programmers designed a conversion utility in JAVA that transformed the XML data into binary files that conformed to the format of our eye movement integration software (ITAP developed by programmers at the RDHFL). Our programmers also developed software that recorded which windows the DST displayed on the D-side display. The software collected the name, the location, and the size of each window and time stamped the data. We used the in-house developed ITAP software to integrate simulation and eye movement data. The ITAP program uses the converted XML files, the data recorded from X-event information from the DST display, and the eye movement data. ITAP determines with which scene plane the line of sight intersects and whether or not a sample belongs to a fixation, a saccade, or a blink (Figure 45).

The oculometer data formed the basis for the calculated variables related to visual scanning. Visual scanning targets included radar display, keyboard area, ATWIT device, flight strip bays, overhead sector maps, communication panels, Computer Readout Device (CRD), and the DST windows. Appendix V contains descriptions that are more detailed and information about the computation of the visual scanning DVs.

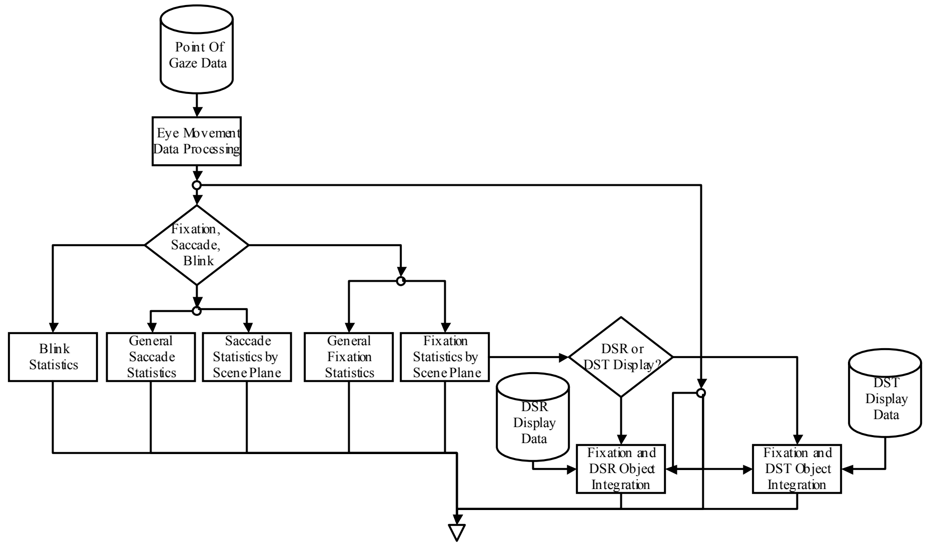

Figure 45. Eye movement data reduction and analysis process.

We reduced the raw visual scanning data. We expressed it as general, scene, object, or structure based eye movement characteristics and conducted appropriate analyses (Table 7). General eye-movement characteristics included fixations, saccades, and blinks. Fixation characteristics included time of onset, duration, the plane being looked at, the area covered by small eye movements within the fixations, and the coordinates relative to the plane. Saccade characteristics included information on the number and magnitude of the saccade and the average velocity during the saccade. Blink characteristics included number, mean duration, and mean distance. We integrated the eye movement data with the DST information for the scene based eye movement data. This included fixations and saccades for each scene plane within the experiment, as mentioned previously. Object based eye movement data included aircraft position symbol and Full Data Blocks (FDBs) from the radar display. Finally, structure of eye movement characteristics included conditional information, i.e., the predictability of the target or location of a fixation when we have information about the previous fixation. Object-based conditional information uses the probability that a fixation on object A (e.g., USA123) is followed by a fixation on object B (e.g., TWA46) and weighs that probability with the probability that a fixation fell on object A. Range-based conditional information extends the objects based principle. It divides the fixations into bins based on the distance from the next fixation and uses the probability that after a fixation landing in bin A, the next fixation falls in bin B. For box-based conditional information, we divide the radar screen into a grid of 10 x 10 and calculate probabilities for each cell of that grid. Finally, the ring-based conditional information is more applicable to terminal environments because it requires us to divide the radar screen into concentric rings around the center of the radar display. Appendix W contains the MANOVA, ANOVA, and means and *SDs* tables.

Table 7. Visual Scanning Analyses

Category	Visual Scanning Characteristics	Type of Analysis	New IVs and Levels
Scene based	**Radar Display and DST Display** **Fixations** (number, percent, mean duration) **Saccades** (number, percent, mean distance)	2 x 2 x 3 (task load x display x automation) MANOVAs and ANOVAs	Display: Radar display vs. DST display
Object based	**Aircraft Position Symbol and FDB** (number, duration)	2 x 2 x3 (task load x information x automation) MANOVAs and ANOVAs	Information: Aircraft Position Symbol vs. FDB
General	**Fixations** (number, mean duration, mean area, visual efficiency) **Saccades** (number, mean duration, mean distance, eye motion workload) **Blinks** (number, mean duration, mean distance)	2 x 3 (task load x automation) MANOVAs and ANOVAs	
Structure	**Object-based conditional information index** **Range-based conditional information index** **Box based (screen is divided into 10 x10 grid) conditional information index** **Ring based conditional information index**	2 x 3 (task load x automation) MANOVAs and ANOVAs	
Descriptive	**Visual scanning target** (D-side communications panel, R-side communications panel, D-side map, R-side map, CRD, ATWIT, flight strip bay 1, flight strip bay 2) **DST windows** (clock, aircraft list, graphic plans display, response display, trial plans, plans display, other displays)	Means and *SD*s	

We screened the visual scanning data based on fixation characteristics to remove any outlier data points. To do so, we used a multivariate approach that tested the Mahalanobis distance statistic. The Mahalanobis distance is the distance of a point from the centroid of a multidimensional space based on the IVs (StatSoft, 2000). We tested a given data point value against a critical value and plotted the values to determine whether any of the observations were outliers. In all, we had three observations that were outliers and removed them from further analysis.

3.3.1.1 Scene-Based Eye Movement Characteristics

The eye movement data integration software calculated fixations and saccades on each of the 10 scene planes examined in this study. Due to the manipulation of automation, we were most interested in the D-side ATCS's visual scanning of the DST display and the radar display. We created an additional IV (display type) and conducted a 2 x 2 x 3 (task load x display x automation) within-subjects MANOVA. For the other scene planes and DST windows, we only conducted descriptive statistics (Appendix W).

The MANOVA for the number, percent, and mean duration of fixations showed significant effects for display type, automation and the display type x automation interaction [$\Lambda = .39$, $F(3,13) = 6.72$, $p < .01$, $\Lambda = .14$, $F(6,10) = 10.28$, $p < .001$ and $\Lambda = .07$; $F(6,10) = 20.81$, $p < .0001$, Appendix W, Table W-38, respectively]. We used an adjusted alpha of .017 for subsequent follow-up ANOVAs.

The ANOVA for the number of fixations showed significant effects for display type and automation and the display type x automation interaction [$F(1,15) = 5.87$, $p < .05$; $F(2,30) = 24.54$, $p < .0001$; $F(2,30) = 50.07$, $p < .0001$, Table W-39, respectively]. Because there was a significant interaction between display type and automation, we explain the main effects of these variables within the interaction. The simple effect of display type was significant for the no and full automation conditions [$F(1,15) = 74.18$, $p < .0001$ and $F(1,15) = 19.14$, $p < .01$, Table W-43, respectively] but not for the limited automation condition. Under no automation, ATCSs fixated more on the radar display, whereas, under full automation, they fixated more on the DST display (Figure 46). Please note that for the purposes of the analyses, we did not remove any fixations or saccades on the SAVANT queries, which occurred on the same display as the DST. In other words, there were fixations and saccades on the DST even when the DST was not in use; these fixations and saccades are reflected by the none zero values in the no automation conditions. These fixations on SAVANT are consistent across all conditions and represent a systematic error or bias that is not dependent upon a particular condition.

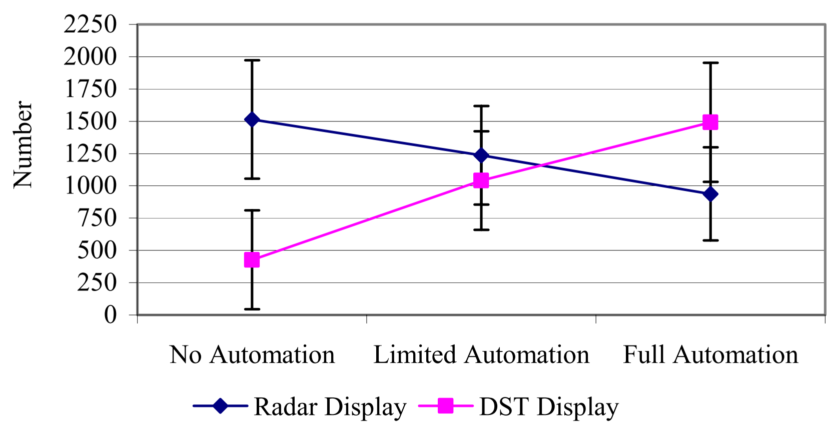

Figure 46. Number of scene plane fixations by display and automation.

For the percent of total number of fixations, we found a significant effect for automation and the display type x automation interaction [$F(2,30) = 29.09$, $p < .0001$ and $F(2,30) = 53.23$, $p < .0001$, Table W-40]. The simple effect of display type was significant under the no and full automation conditions [$F(1,15) = 70.34$, $p < .0001$ and $F(1,15) = 143.95$, $p < .001$, Table W-40, respectively], but not for the limited automation condition. As with the number of fixations, ATCSs spent a higher percentage of fixations on the radar display when no automation was present, but they spent a higher percentage of fixations on the DST display when automation was present (Figure 47). The percentage of fixations on the displays did not differ under limited automation conditions. We ask the reader to note that these percentages do not add up to 100 because we include only those fixations on the radar or DST display, and excluding fixations elsewhere (e.g., sector maps, CRD). If we included all fixations on all scene plane surfaces, we would get a percentage that adds to 100.

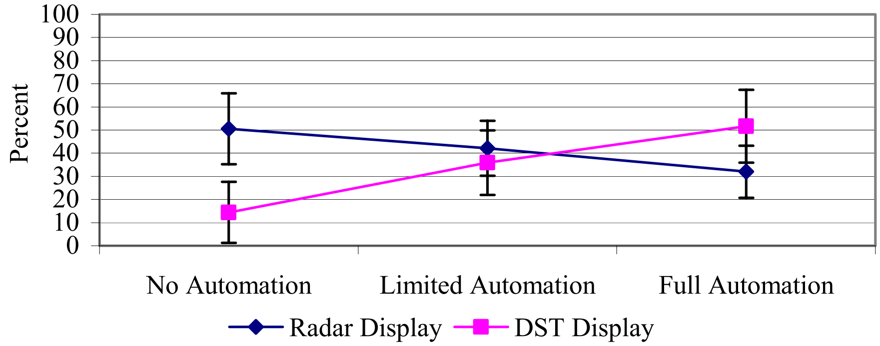

Figure 47. Percent of fixations by display and automation.

The effect of automation and the display type x automation interaction was significant for the mean duration of fixations [$F(2,30) = 11.34$, $p < .0001$ and $F(2,30) = 29.12$, $p < .0001$, Table W-41, respectively]. The type of display influences the main effect of automation. The two-way interaction between display type and automation showed that the simple effect of display type was significant for the no and full automation conditions [$F(1,15) = 27.02$, $p < .0001$, $F(1,15) = 4.54$, $p < .05$, respectively, Table W-41] but not for the limited automation condition. The mean duration of fixations was longer on the radar display than the DST display when no automation was present (Figure 48). In contrast, under full automation, the mean duration of fixations was longer on the DST display (Figure 48).

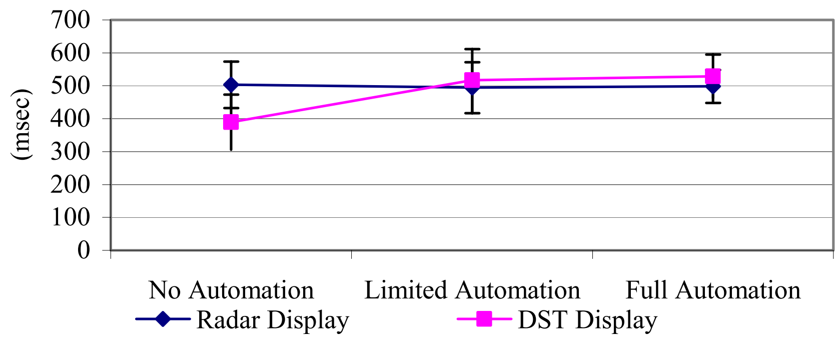

Figure 48. Mean duration of fixations by display and automation.

The MANOVA examining the number, percent, and distance of saccades showed significant effects for task load, display type, and automation, and the interactions between task load x automation, display type x automation were significant [$\Lambda = .26$, $F(3,13) = 12.44$, $p < .001$, $\Lambda = .14$; $F(3,13) = 25.87$, $p < .0001$; $\Lambda = .13$, $F(6,10) = 11.31$, $p < .001$; $\Lambda = .29$, $F(6,10) = 4.12$, $p < .05$, $\Lambda = .07$, and $F(6,10) = 20.66$, $p < .0001$, Table W-42, respectively]. We used an adjusted alpha of .017 for the follow-up univariate analyses.

The univariate results for the number of saccades showed a significant effect for automation and a significant interaction between display type and automation [$F(2,30) = 21.57$, $p < .001$ and $F(2,30) = 49.24$, $p < .0001$, Table W-43, respectively]. The display type influenced the effect of automation on number of saccades. The simple effect of display type within both the no and full automation conditions was significant [$F(1,15) = 61.72$, $p < .0001$ and $F(1,15) = 35.0903$, $p < .0001$, Table W-43, respectively]. The number of saccades on the radar display was higher under no automation conditions, while the number of saccades on the DST display was higher under full automation (Figure 49).

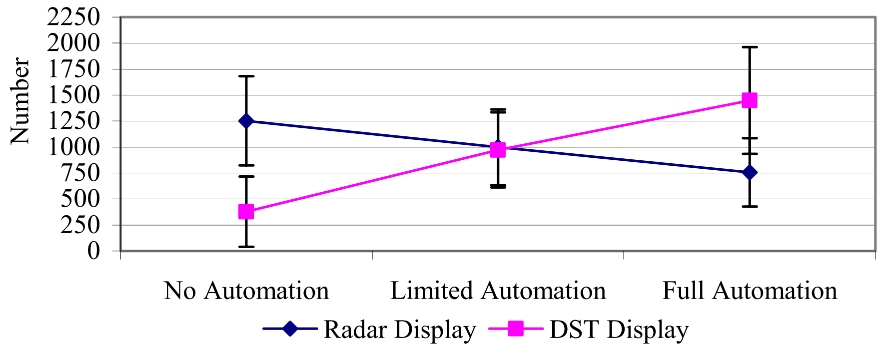

Figure 49. Number of saccades by display and automation.

The ANOVA for the percent of saccades spent on the two displays showed a significant effect for automation and the display type x automation interaction [$F(2,30) = 30.65$, $p < .0001$ and $F(2,30) = 57.96$, $p < .0001$, Table W-44, respectively]. The type of display qualified the effect of automation. The simple effect of display type was significant for both the no and full automation conditions [$F(1,15) = 60.84$, $p < .0001$, and $F(1,15) = 39.68$, $p < .0001$, Table W-44, respectively] but not for the limited automation condition. The percent of saccades on the radar display was higher than that on the DST display under no automation; in contrast, the percent of saccades on the DST display was higher under full automation conditions (similar trend to Figure 49).

For saccade distance, the ANOVA showed significant effects for task load, display type, task load x automation [$F(1,15) = 23.64$, $p < .001$; $F(1,15) = 79.06$, $p < .0001$; $F(2,30) = 6.37$, $p < .01$, Table W-45, respectively] and secondary trends for automation, display type x automation, and task load x display type x automation. We explain all effects within the context of the three-way interaction. The simple effects of display type, automation and the display type x automation interaction were significant under low task load [$F(1,15) = 36.34$, $p < .01$; $F(2,30) = 6.89$, $p < .01$; $F(2,30) = 6.98$, $p < .01$, Table W-45, respectively]. Under low task load conditions, mean saccade distance was higher for the DST display and no automation compared to limited automation. Mean saccade distance was longest under the no-automation condition and the DST display (Figure 50). Only the effect of display type was significant under high task load conditions [$F(1,15) = 77.27$, $p < .0001$, Table W-45]. When task load was high, saccade distance was longest on the DST display (Figure 51).

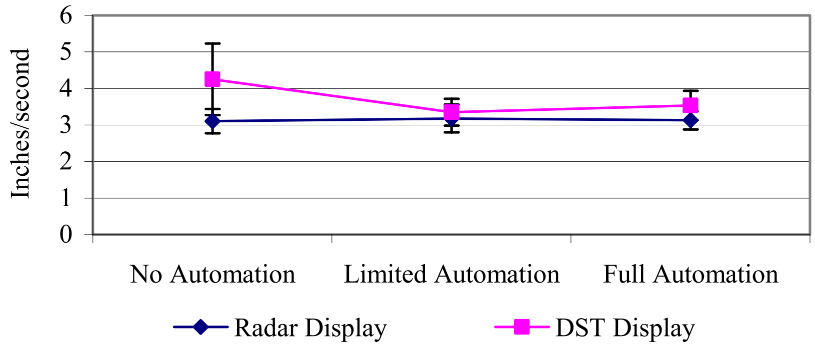

Figure 50. Mean saccade distance under low task load by display and automation.

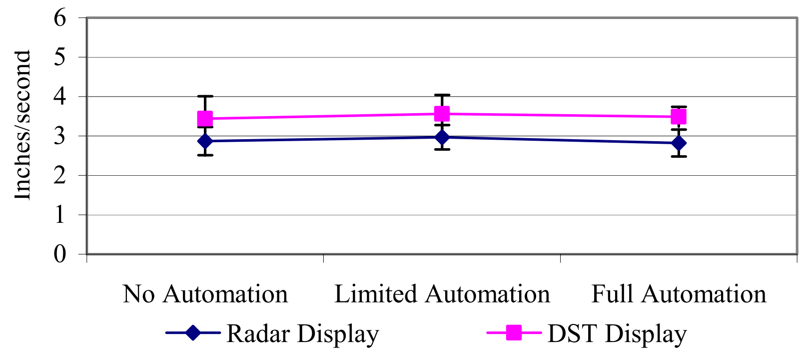

Figure 51. Mean saccade distance under high task load by display and automation.

For the analyses, we used DST display as one level of the display information variable. These analyses included any visual scanning hits on the DST display. However, the DST presented information to the D-side ATCS in several windows for which we have fixation characteristics. These windows include the aircraft list, graphic plans display, trial plans display, plans display, clock, and so on. Based on the level of automation, we instructed ATCSs to access certain windows only. In the no-automation condition, we did not display the DST windows. For limited automation, ATCSs could view the aircraft list, whereas, in full automation, all DST windows were available for viewing. Please note that the visual scanning values for the limited automation condition are somewhat inflated for three reasons. First, ATCSs occasionally and accidentally opened additional windows such as the graphic plans display during the limited automation conditions. We could not modify the DST to inhibit ATCSs from opening windows other than the aircraft list in the limited automation condition. Instead, the ATC SME reminded ATCSs to immediately close windows that they accidentally opened. In addition, the grand mean substitution we used for replacement of missing data inflates the values. The SAVANT responses are also included in these values, as they are for the full automation condition. We set all values for the no automation condition to zero because the DST was not running during these scenarios. Appendix W contains the means and *SD*s for these windows. Here, we only highlight those descriptive statistics that are of interest. As shown in Figure 52, ATCSs fixated on the aircraft list more than on the graphic plans display under full automation and did not take advantage of the trial plans capability.

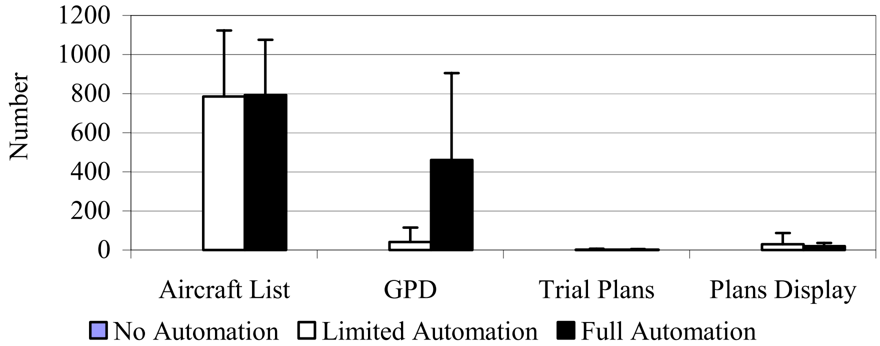

Figure 52. Means and *SD*s for fixations on DST windows.

3.3.1.2 Object-Based Eye Movement Characteristics

The object based eye movement characteristics included the aircraft position symbol and the FDB as an extra IV - information. We conducted a 2 x 2 x 3 (task load x information x automation) within-subjects MANOVA on the number and mean duration of fixations. The MANOVA showed significant effects for information, automation, the task load x information interaction, and the information x automation interaction [$\Lambda = .02$, $F(2,14) = 285.46$, $p < .0001$; $\Lambda = .22$, $F(4,12) = 10.86$, $p < .001$; $\Lambda = .21$, $F(2,14) = 3.75$, $p < .05$, $F(4,12) = 11.12$, $p < .001$, Table W-47, respectively]. We used an adjusted alpha of .025 for the follow-up univariate analyses.

The ANOVA for the number of object based fixations showed significant effects for information, automation, and the information x automation interaction [$F(1,15) = 209.58$, p <0001; $F(2,30) = 24.965$, $p < .0001$; $F(2,30) = 15.57$, $p < .0001$, Table W-48, respectively]. The two-way interaction between display and automation qualified both of the main effects, and we focused on this. The simple effect of information within the no-, limited-, and full-automation conditions was significant [$F(1,15) = 209.22$, $p < .0001$; $F(1,15) = 78.82$, $p < .0001$; $F(1,15) = 153.66$, $p < .0001$, respectively Table W-48] but not for the full automation condition. Although more fixations occurred on the FDB, automation attenuated this effect. As the degree of automation increased, the number of fixations on FDB decreased (Figure 53).

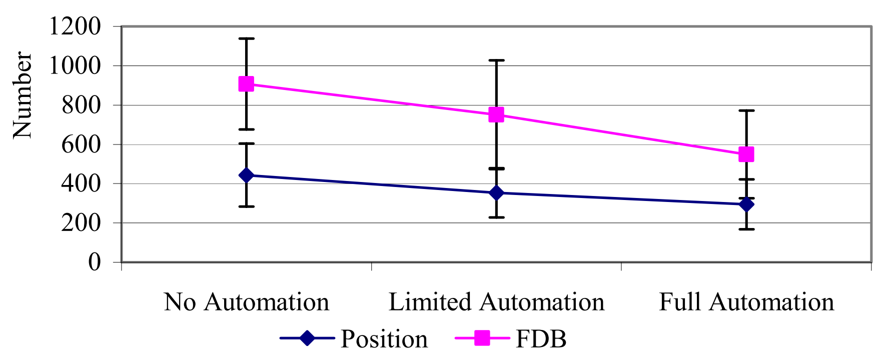

Figure 53. Number of object based fixations by information and automation.

The ANOVA for the duration of object based fixations showed significant effects for information and task load x information [$F(1,15) = 20.85$, $p < .0001$; $F(1,15) = 5.64$, $p < .05$, respectively Table W-49]. For the two-way interaction, the simple effect of information was significant for low and high task loads [$F(1,15) = 203.56$, $p < .0001$ and $F(1,15) = 123.50$, $p < .0001$, respectively Table W-49]. The fixation duration was longer for FDBs; however, low task load attenuated this (Figure 54).

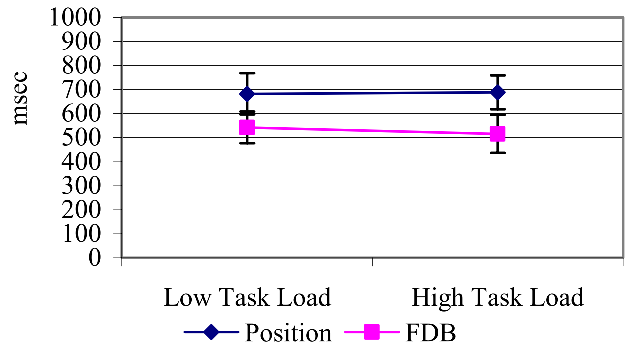

Figure 54. Fixation duration by task load and information.

3.3.1.3 General Eye Movement Characteristics

We conducted separate analyses for each of the general eye movement characteristics – fixations, saccades, and blinks. These MANOVAs, structured as 2 x 3 (task load x automation) repeated measures designs, addressed the differences across scenarios. The MANOVA for fixations included the number, mean and *SDs* for duration, area, and visual efficiency. We did not find significant differences at the multivariate level; however, at the univariate level, several trends existed. A trend for automation on the mean duration of fixations showed that as automation increased, the mean duration of the fixations increased (Figure 55, Table W-20). As automation increased, the *SD* of the duration of the fixations increased (Figure 56, Table W-21). Secondary trends in the data showed effects for task load and automation for visual efficiency. As task load increased, visual efficiency decreased (Figure 57, Table W-24). As automation increased, visual efficiency increased (Figure 57, Table W-24).

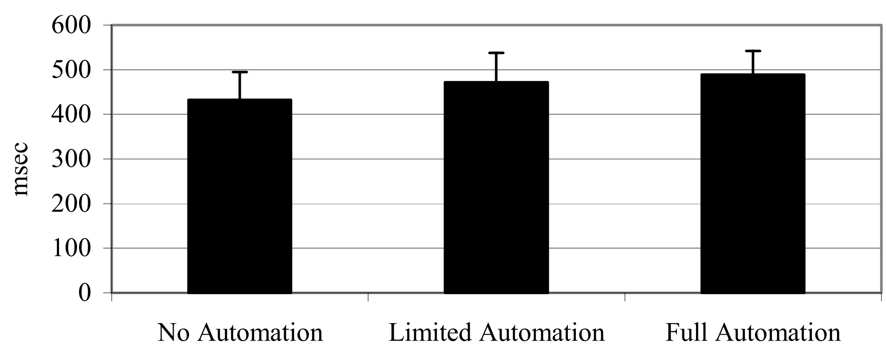

Figure 55. Mean duration of fixations by automation.

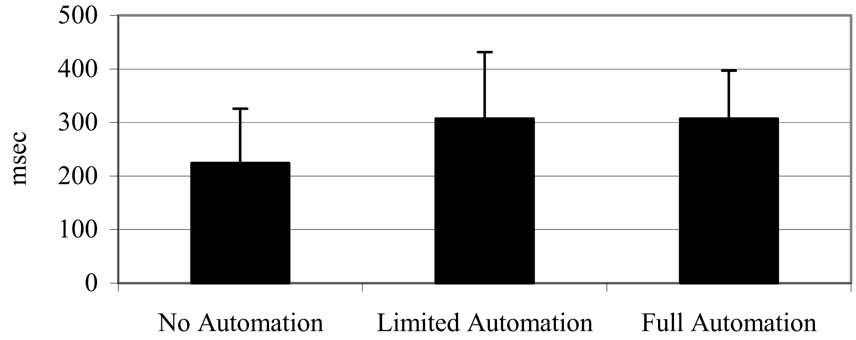

Figure 56. *SD* of fixation durations by automation.

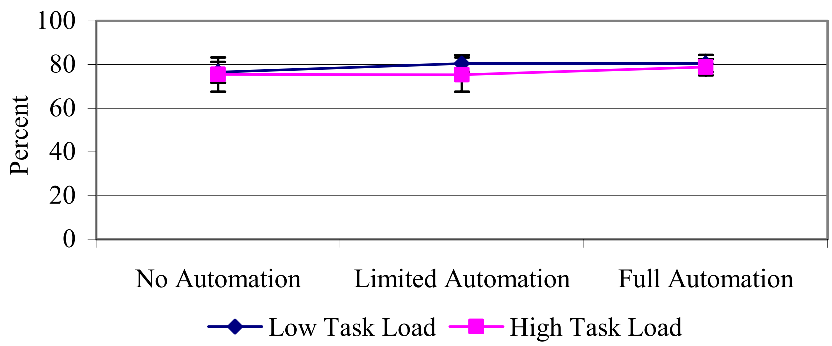

Figure 57. Visual efficiency by task load and automation.

The MANOVA for saccades examined the effects of task load and automation on the number, mean, and *SD* of durations, mean and *SD* of distance moved, and eye motion workload. The effects of task load and automation did not reach statistical significance at the multivariate level, however, univariate results showed several trends in the data. As automation increased, the mean distance traveled in the saccade decreased (Figure 58, Table W-27). A secondary trend showed that as task load increased, the number of saccades increased (Figure 59, Table W-26). As automation increased, the eye motion workload indicated that the eye moved less per second (Figure 60, Table W-38).

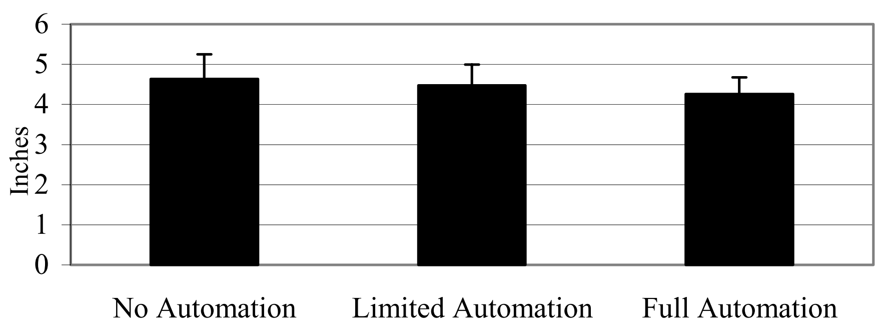

Figure 58. Mean distance traveled in saccade by automation.

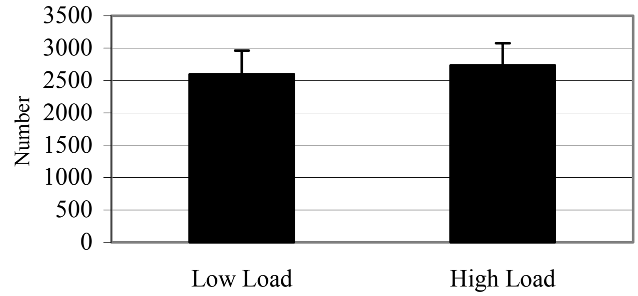

Figure 59. Number of saccades by task load.

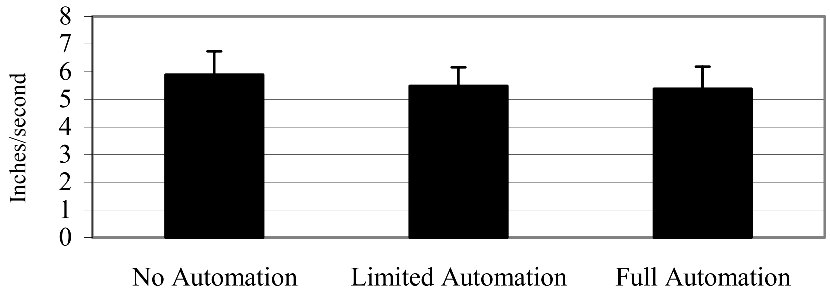

Figure 60. Eye motion workload by automation.

The MANOVA for blinks examined the effects of the experimental variables on the number of blinks, blink duration, and distance. No effects reached statistical significance at the multivariate level nor were there any trends at the univariate level of analysis.

3.3.1.4 Structure in Eye Movement Characteristics

The 2 x 3 (task load x automation) within-subjects MANOVA examining structure of eye movement characteristics showed significant effects for task load [$\Lambda = .16$, $F(4,12) = 26.52$, $p < .001$, Table W-50]. We used an adjusted alpha of .013 for subsequent univariate analyses.

The ANOVA for the object conditional index showed a significant effect for task load [$F(1,15) = 66.71$, $p < .001$, Table W-51]. The structure in the visual scan decreased under high task load conditions (Figure 61, Table W-51). The ANOVA for box conditional information showed a significant main effect for task load [$F(1,15) = 5.34$, $p < .05$, Table W-53] and a trend for automation. As automation and task load increased, the structure of the visual scan decreased (Figure 62).

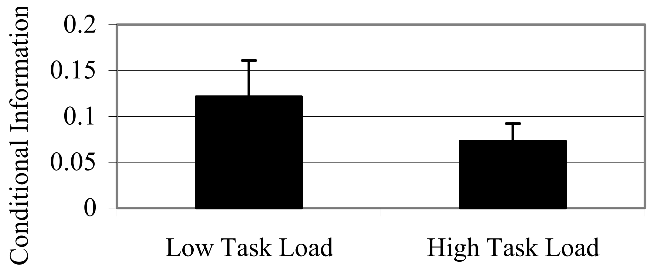

Figure 61. Object conditional index by task load.

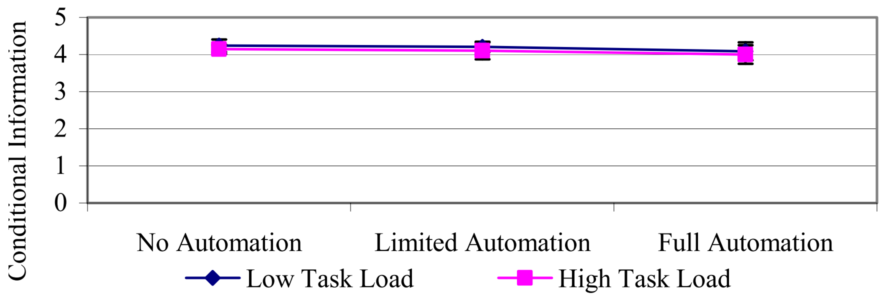

Figure 62. Box conditional information by task load and automation.

3.3.2 Discussion

The general eye-movement characteristics indicate that fixations were longer with increasing levels of automation. Based on the results from previous studies, we had expected that the characteristics of fixations on information on the radar display would not change because the stimulus did not change, and, therefore, the mechanism to pick up information from the stimulus would not change either. The information on the DST was either quite similar to that on the radar display (i.e., an aircraft representation in a graphical format) or similar to information on a flight progress strip in electronic format. If anything, we expected a decrease in fixation duration.

In previous studies, we tracked eye movements of an R-side ATCS. We had made the null hypothesis that the eye movement characteristics of an ATCS would not change between the R-side and the D-side ATCS position. We find, however, that the distribution of the number of fixations between the position symbols and the data blocks of aircraft is very different for D-side ATCSs compared to R-side ATCSs. In our experience with R-side ATCSs, they fixated on data blocks and aircraft positions an equal number of times. This may have occurred because we used a spotlight model that creates a 2-in. radius around the center of a fixation. We felt that the ATCS then absorbed information about the aircraft as a whole and not by a particular piece of information (i.e., FDB or position symbol). This, in turn, would result in an equal number of fixations on the data blocks and the position of the aircraft closest to the center of the fixation. In the current experiment, we find a very different distribution of fixations. The D-side ATCS consistently focused more on the FDBs. Overall, there were more fixations on the FDB and aircraft position symbols under conditions without automation because the D-side ATCS is more of an assistant for the R-side ATCS. With an increase in automation, the number of fixations on the radar display decreases (i.e., the R-side ATCS will have less assistance of the D-side ATCS in scanning the radar display for information).

The difference in fixation distribution for the D-side ATCS compared to the findings from our previous studies may also indicate a different strategy used by the D-side ATCS. The D-side ATCS focuses less on aircraft position symbols than the R-side ATCS in previous experiments. Cushing (1989), for example, suggests that ATCSs have a process that checks distances between aircraft. That suggestion, however, can still result in a pattern in which an ATCS would either fixate on two aircraft positions sequentially, fixate on a location close to both aircraft and process the position information, or separate the fixations on the positions by several other fixations but process the position information after having retrieved all information needed to check the distance. The R-side ATCS has a tactical function with separation as the primary goal. Fixating on aircraft positions allows the R-side ATCS to determine if the required separation standards are met. In contrast, the D-side ATCS looks less at aircraft positions but looks for information in the data blocks. Although processing information may occur in areas away from the fovea (i.e., further away from the center of the fixation), evidence exists that the visual system prefers processing information that is present in the foveal area. The shift in the scanning strategy between the R-side ATCS and the D-side ATCS seems to show that the D-side ATCS is less tactical and focuses on the data block information instead.

One of the concerns about the data analyses and results of our visual scanning work is that eye movements are the outcome of possibly several underlying cognitive processes. Viviani (1990) suggests that ATCSs, when exploring the ATC workstation for information, have a set of beliefs about the current situation, break that situation up in subsets, and execute sequences of fixations to verify or update the knowledge about the current situation. It is our belief that the scanning or monitoring process executes in parallel with these processes. Knowledge of the resulting output in the form of a visual scanning pattern may not tell us much about the architecture of those cognitive processes. From the literature, we have learned that "good scanners" will show little structure in the visual scan in an attempt to cover most of the information display, but we cannot assume that ATCS visual scanning solely consists of a monitoring process that needs to detect conspicuous events. Research in other areas such as scanning of x-rays by radiologists suggests that shorter fixations relate to the monitoring process, whereas the longer "evaluation" fixations directly relate to decision making and more detailed diagnosis (Carmody, Nodine, & Kundel, 1981). Eye movements in the ATCS task definitely have an action-oriented component to it, as well. Gross (1998) found, for example, that fixation on an aircraft that was about to be subject to a communication event increased in duration up to 40 seconds before the communication event took place. During the communication event itself, fixations durations were not longer or did not take place on the aircraft. This finding may indicate that an ATCS absorbed information from the aircraft presentations, planned the communication action, and then executed the action without visually obtaining further information about the aircraft. An R-side ATCS will have the results of these type of action-oriented visual scanning outcomes interspersed with the outcomes related to situation assessment (i.e., event detection). Therefore, the summary variables reported in previous studies did not distinguish monitoring from ATC-related activities involved in delivering a clearance. The eye movements of the D-side ATCS reported here, however, have removed most of the monitoring process because the D-side ATCS has a very different function than the R-side ATCS. The resulting eye movements may very well depict more of the action oriented eye movement behavior than we have previously found in R-side ATCSs.

For the scene-based eye-movement characteristics, the fixation and saccade results indicate that D-side ATCSs spent a great amount of time away from the radar display. Currently, D-side ATCSs act as a second pair of eyes for the R-side ATCS; however, the introduction of automation changed this and drew the visual scanning of the D-side to the automated tool. With the introduction of automation, the assistance the D-side can offer the R-side ATCS changes. Thus, automation tools have substantial implications for the future role of the D-side.

Descriptive statistics show that D-side ATCSs did not utilize the trial plans capability of the DST. This indicates that, most likely, the ATCSs identified conflicts from the DST. However, they conceptualized their own solutions to those conflicts without accessing any possible solutions from the DST. This may also indicate that the ATCSs trusted the DST to predict and identify conflicts but not to provide solutions to those conflicts. Further, the ATCSs were novice DST users, and this may have had an impact on using the trial plans capability less.

The object-based eye-movement characteristics showed that the number of fixations was higher for the FDB, although, as automation increased, the number of fixations on the FDBs decreased. ATCSs also spent more time fixating on the FDB, particularly under high task load conditions.

High task load and limited and full automation led to lower levels of structure in the ATCS's visual scan. The D-side ATCS seems able to follow events better under low task load conditions. Both automation conditions most likely reduce visual scan structure because the ATCS becomes increasingly strategic in focus. ATCSs' plans no longer come from the layout of the airspace but from the information provided by the DST. Under no automation, the radar display provides the feedback used in the scan.

In Figure 46, under conditions without automation, ATCSs still fixated 10-15 % of the time on the display on the D-side. For the purposes of the analyses, we did not remove any fixations or saccades on the SAVANT queries or during the first minute (i.e., during the countdown), which occurred on the same display as the DST. In other words, there were fixations and saccades on the DST even when the DST was not in use; these fixations and saccades are reflected by the none zero values in the no-automation condition and would account for approximately 250 fixations under high task load conditions. The remaining fixations on the D-side display under conditions without automation are most likely fixations during transitions to other scene planes such as the D-side keyboard, the maps, and the flight strip bays. The number of these fixations is a systematic error, that is, they exist during all automation levels. Comparisons between automation conditions, therefore, are still valid, although the absolute value of the number of fixations is lower than those used in the analyses.

3.4 Communications

During normal operations within an ATC team (R- and D-side ATCSs), the two ATCSs have predefined responsibilities as defined by FAA Order 7110.65L (Federal Aviation Administration, 1998) and local orders. Each member of the ATC team has specific communication responsibilities (e.g., R-side ATCS communicates with pilots). In addition, the team members develop communication patterns between them.

Changes in the way an ATCS team works may manifest itself in intrateam communications, ATCS to pilot and ATCS to ATCS in adjoining sector communications. Although considerable research has focused on the effects that decision aids will have on pilot-to-ATCS communication, relatively little research exists on how DSTs will affect ATCS to ATCS communication (Kanki & Prinzo, 1996). We know very little about the daily task-related communication exchanges between R-side and D-side ATCSs. This type of communication becomes especially important when the new DSTs attempt to enhance the tactical and strategic decision-making capabilities of ATCSs. What effect might these new technologies have on existing patterns of intrateam communications and communications between ATCSs within adjoining sectors? Without an understanding of the existing patterns of intrateam communications, there is no empirical way to answer that question.

Because the recordings required by the NAS do not capture the internal communication between ATC team members, we videotaped all task-related R-side/D-side communications. An ATC SME coded communications using the C^4T (Peterson et al., in press). The C^4T has three communication categories: the topic of communication, the grammatical form of communication, and the method of communication. Thus, the C^4T captures the “what” (topic) and “how” (form and method) of communication.

A frequency analysis of the communications between the D-side and R-side resulted in the following observations. Most of the intrateam communications were aircraft traffic (40%) and route of flight (15%). The least frequent communications were inter-sector coordination (1%). R-side and D-side ATCSs did not statistically differ in the topic of their communication. However, they did differ in the grammatical form of their communication. Whereas the D-side had a higher percentage of statements and observations (56% vs. 30%), the R-side had a higher percentage of answers (43% vs. 35%). It appeared that, compared to the R-side, the D-side ATCS was the initiator of more communication. Finally, there was no intrateam difference in the method of communication. The use of voice alone was the most frequent method of communicating (70%). The remaining 30% of communication contained a mixture of verbal and non-verbal expression. This latter finding suggested that any changes affecting the line of sight between R-side and D-side positions could disrupt the adaptive use of intrateam nonverbal communication.

One of the more consistent findings in the literature on ATCS-to-pilot voice communications is that workload affects the quality and quantity of communication exchanges (Prinzo & Britton, 1993). In this literature, workload is primarily measured by the number of aircraft at a given time (i.e., task load) that are under the control of an R-side ATCS. As task load increases, there is a corresponding trend toward an increasing number of communication errors (Morrison & Wright, 1989). Research suggests that as ATCSs and pilots become overburdened, the clarity of their communications begins to suffer. This, in turn, puts ATCSs and pilots at a higher risk of committing readback/hearback errors (Morrison & Wright). Standard ATC protocol requires pilots and ATCSs to repeat what they heard so that the sender knows that ATCSs received messages accurately. A readback/hearback error occurs when an incorrect pilot or ATCS readback of information goes uncorrected.

Changes in workload also affect the kinds of communication exchanges that occur. In a study of ATC communications in a simulated en route environment, as workload increased, ATCSs in top-rated teams issued more communication reports to pilots than did lower-rated teams. Compared to lower-rated teams, top-rated teams issued shorter messages as a means of insuring accuracy (Human Technology Incorporated, 1991). ATC SMEs conducted OTS ratings of team performance. However, as is the case with experiments, whether these results generalize to the broader ATC population or are specific to a given experimental manipulation is not yet clear.

Although related to workload, the effects of technology on communication are unclear. From the concept of monitoring, an operator has no need to communicate unless some event occurs that requires the actions of another. Whether an operator uses or trusts technology, one might argue that communications would increase during the transition period of adjusting to the new technology. This would be especially true if there were problems with the technology and team members had to decide whether to trust it. However, once the transition period passes, one would expect communications to return to a previous baseline.

3.4.1 Intrateam Communications

Appendix X contains descriptions of C^4T sub levels within each of the three categories. For information on the development and operational validation of the C^4T, see Peterson et al. (in press). The intrateam communications analyses represent a joint project between the RDHFL and CAMI. As part of their requirements at CAMI, Bailey, Willems, and Peterson (in review) wrote a paper on this information.

3.4.1.1 Data Analysis and Results

Because the communication component of the experiment was primarily descriptive in nature, we did not replace missing data with mean substitutions but, instead, dropped the cases from further analysis. In all, we dropped three teams from the analyses. Of the remaining five teams, we coded 3,194 communication events.

To determine if differences existed in the patterns of communication operating between the field study and the experiment, we analyzed the frequency data by topic, format, and expression. Although we did not model the experiment after a particular en route ARTCC, the scenarios reflected real world events. Thus, one might expect similarities between intrateam communications within a field and experimental setting.

Table 8 shows comparisons for the percentage of R-side and D-side communications related to the topic of communication, its grammatical form, and the method of expression. Although the percentages differed for the topics of communication presented in Table 8, both lab and field assessments identified the same top three topics. These included communications about traffic, route of flight, and aircraft altitude. Compared to the field, the most noticeable difference in the experiment was the lack of communications about the weather, point-outs, and traffic flow. There was also little communication about flight strips.

As Table 8 shows, the grammatical form of communications also differed between the two environments. The field results show a strong tendency for the D-side to make statements

(55.9%) and the R-side (42.8%) to provide answers. In contrast, for the experiment, the results show both R-side and D-side predominately making statements (58% vs. 77.3%, respectively).

Verbal communication is the method of choice for R-side and D-side ATCSs. However, as Table 8 shows, in the experiment, the D-side had a stronger tendency to use a mixture of verbal and nonverbal expressions than did the R-side (24.7% vs. 5%). For the field setting, the percentage of mixed messages was similar for both the R-side and D-side (14.6% vs. 16.8%). The two settings also differed in the percentage of nonverbal communications used. In the field, 13.9% of the communications was solely nonverbal for both the R-side and D-side. This is in contrast to the lower percentages recorded during the experiment (R-side 0.5%, D-side 2.8%).

Table 8. Comparisons of R-side and D-side communications in field and laboratory setting

	En route Center		Laboratory Setting	
	R-side%	D-side%	R-side%	D-side%
Communication Topic				
Traffic	41.0	37.9	53.7	51.2
Route of flight	14.2	15.6	13.1	11.7
Altitude	7.1	8.0	16.0	21.1
Weather	5.5	6.8	0.0	0.0
Point-out	5.0	6.1	0.0	0.0
Traffic flow	5.2	5.6	0.0	0.0
Frequency	5.9	4.7	3.5	2.7
Flight Strips	5.6	4.5	0.7	1.0
Equipment	3.3	4.0	4.5	4.9
Hand-off	3.6	3.1	2.9	1.8
Speed	2.6	2.8	4.4	2.8
Approval	1.0	0.9	1.0	1.8
Communication Format				
Statement	29.7	55.9	58.0	77.3
Answer	42.8	25.1	18.3	10.4
Question	12.2	16.4	22.9	11.3
Command Answer	5.8	0.3	0.0	0.0
Command	0.5	2.4	0.8	1.0
Communication Expression				
Verbal	77.1	69.3	93.9	69.0
Verbal & Nonverbal	14.7	16.8	5.0	24.7
Nonverbal	13.9	13.9	0.5	2.8
Equipment			0.0	0.1
Equipment & Verbal			0.6	3.4
Equipment & Nonverbal			0.0	0.1

We conducted a 2 x 3 (task load x automation) within-subjects ANOVA for the composite number of communication exchanges of team members and did not find any statistically significant results. A power analysis revealed that we had insufficient power, but we present the data and trends to add to the knowledge base of this poorly studied area. When we analyzed each of the three C^4T categories separately, we found secondary trends for task load on general communications about a specific aircraft and communications involving altitude changes (Figure 63 and Figure 64). More communications occurred for specific aircraft and altitude changes under high task load conditions. In addition, there was a secondary tendency for more information to be exchanged on specific aircraft, altitude, and route of flight under low task load, full automation (secondary trend; Figure 65 and Figure 66).

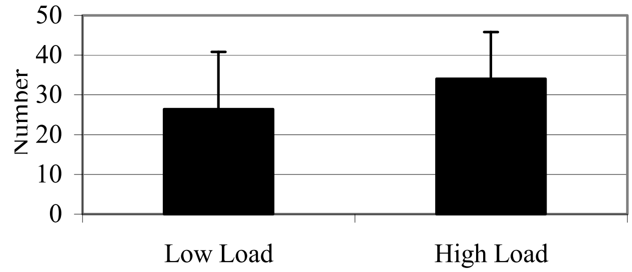

Figure 63. Number of specific aircraft communications by task load.

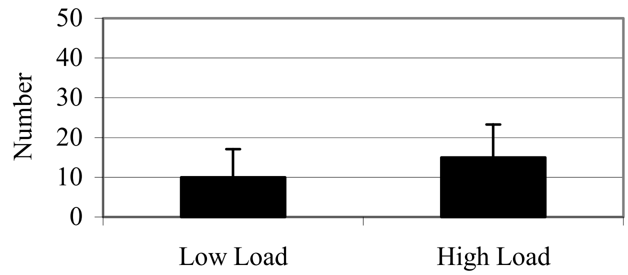

Figure 64. Number of altitude communications by task load.

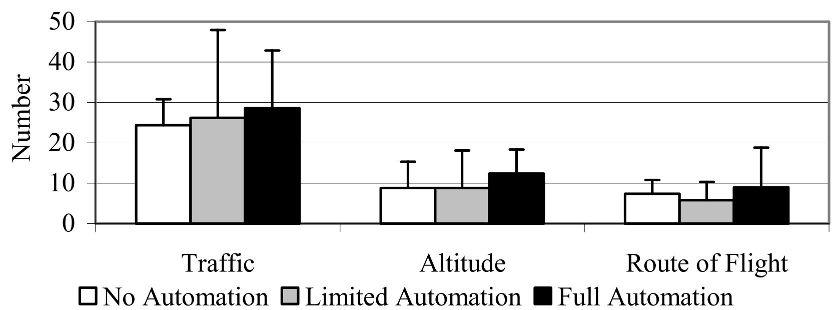

Figure 65. Number of specific aircraft, altitude, and route communications by low task load and automation .

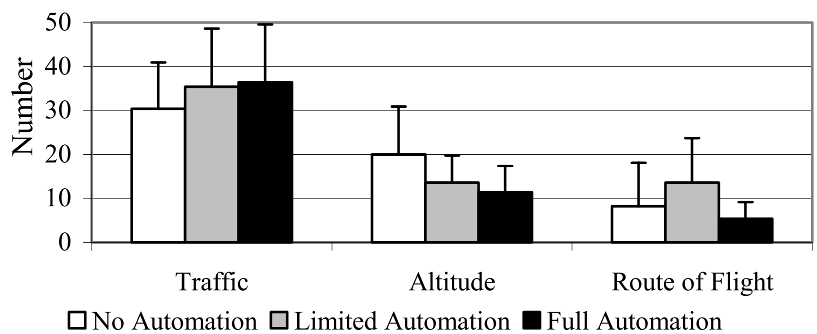

Figure 66. Number of specific aircraft, altitude, and route communications by high task load and automation.

3.4.1.2 Discussion

The C^4T profile of the laboratory experiments (collapsing across all conditions) compared favorably with data collected from the field. In both cases, the top three topics of communication were about traffic, route of flight, and aircraft altitude. Also, in both cases, voice only was the dominant mode of communications. Despite these similarities, there were notable differences in the grammatical form of communication between R-side and D-side ATCSs. In the field, but it appeared that the R-side primarily responded to statements by the D-side, in the experiment, it appeared that both R-side and D-side ATCSs were issuing statements. More nonverbal communications occurred in the field as compared to the laboratory environment. This may be because, in the field, team members are familiar with each other, but in the laboratory, they did not know each other before participating.

Although only a secondary trend, more communication exchanges occurred in the high task load conditions as compared to the low task load conditions. This was especially true concerning communications about a specific aircraft and about altitude changes. The effects of automation on communications were not as clear. There was some evidence that in the low task load conditions more communication exchanges occurred for specific aircraft, altitude, or route of flight using the full automation as compared to limited or no automation.

3.4.2 Push-to-Talk

We collected LL (ATCS to ATCS in adjoining sectors) and (R-side ATCS to simulation pilots) communications. For the purposes of analysis, these communications apply to the team of ATCSs instead of each ATCS individually.

3.4.2.1 Data Analysis and Results

We created one communications-related data set containing LL and Radar-side (R) communications. To enable measures of SA with SAGAT, scenarios had different durations. To eliminate the effect of different scenario lengths, we corrected for duration and expressed the number of communications in events per minute. We averaged the number and duration of each communication type, per scenario and team of ATCSs, and performed analyses on the eight teams. We conducted 2 x 3 (task load x automation) within-subjects MANOVAs for the LL and R communication measures, respectively. We followed any significant MANOVA results with subsequent ANOVAs and used an adjusted alpha of .025. Appendix Y contains the means, *SDs*, MANOVA, and ANOVA results.

3.4.2.1.1 Mean Number/Duration of PTT Landline Communications

The within-subjects MANOVA showed significant effects for task load, automation, and the task load x automation interaction [$\Lambda = .29$, $F(2,6) = 7.22$, $p < .05$; $\Lambda = .11$, $F(4,4) = .80$, $p < .05$; $\Lambda = .13$, $F(4,4) = 6.96$, $p < .05$, respectively, Appendix Y, Table Y-3]. Consequently, we performed univariate analyses on the number and duration of LL communications.

The ANOVA for the mean number of LL communications showed a significant main effect for task load [$F(1,7) = 16.18$, $p < .025$, Table Y-4]. ATCSs' average number of communications increased as the task load increased (Figure 67).

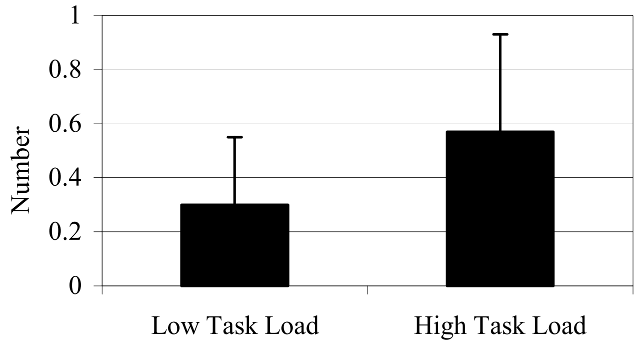

Figure 67. Mean number of LL communications per minute by task load.

The ANOVA for the mean duration of LL communications showed a significant task load x automation interaction [$F(2,14) = 4.93$, $p < .025$, Table Y-5]. The simple effect of task load within full automation was significant [$F(1,7) = 11.03$, $p < .025$, Table Y-5]. ATCSs' mean duration of communications was longest during the full automation, low task load scenario (Figure 68).

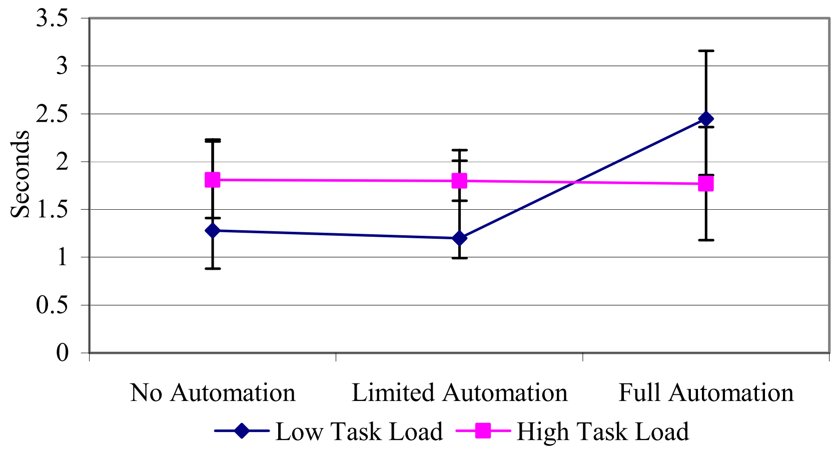

Figure 68. Mean duration of LL communications by task load and automation.

<u>3.4.2.1.2 Mean Number/Duration of PTT R-side Communications</u>

The 2 x 3 (task load x automation) within-subjects MANOVA examining R communications showed a significant effect for task load [$\Lambda = .023$, $F(2,6) = 130.21$, $p < .0001$, Table Y-6]. Because we found significant effects from the MANOVA, we performed subsequent ANOVAs.

The ANOVA for the mean number of R communications showed a significant main effect for task load [$F(1,7) = 148.07$, $p < .05$, Table Y-7]. ATCSs' mean number of communications increased as task load increased (Figure 69).

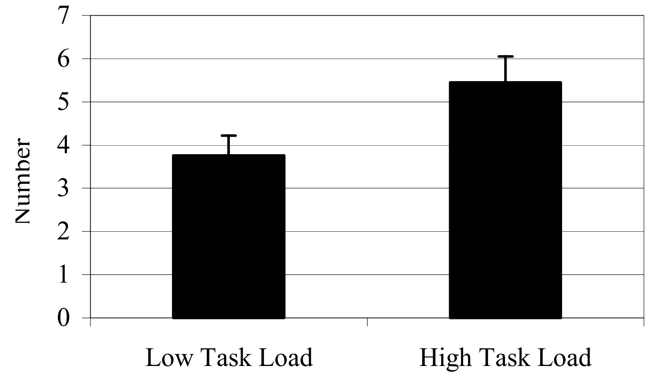

Figure 69. Mean number of R communications per minute by task load.

The ANOVA for the mean duration of R communications showed a significant main effect for task load [$F(1,7) = 11.73$, $p < .05$, Table Y-8]. ATCS's mean duration of communications increased as task load decreased (Figure 70).

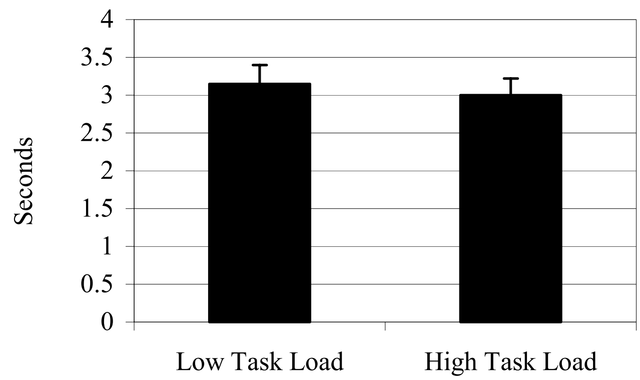

Figure 70. Mean duration of R communications by task load.

3.4.2.2 Discussion

Although not statistically significant, a trend revealed the duration of communications was longest when in the fully automated condition. Further, an interaction between task load and automation showed that task load significantly influenced the duration of communications in the full-automation condition. ATC teams' LL communications lasted longer when task load was low and full automation was present. This would seem to indicate that when task load was low and ATCSs had full use of the DST, they had more detailed ideas of how to fix potential conflicts and discussed these solutions longer with the controlling sector of the aircraft. The ATC teams' use of the LL increased with increasing task load. The higher traffic volume associated with high task load presented greater numbers of potential conflicts that ATCSs needed to resolve.

Likewise, for the R communications, the number of communications increased as task load increased. The R communications accounted for the communications between the R-side ATCSs and the simulation pilots. As the amount of traffic increased, we expected that R-side

ATCSs would increase the number of communications they initiated. The higher the volume of traffic, the more actions and clearances the R-side ATCS needs to communicate. In addition, the duration of the communications decreased when traffic task load was high. Once again, the R-side ATCS needed to issue a larger number of clearances and decreased the duration of these communications in order to get them all in. This corresponds to the Human Technology Incorporated (1991) findings that ATCSs in top-rated teams compensated for increases in workload by increasing communications while decreasing their durations. Willems and Truitt (1999) also found that the number of communications increased as task load increased; however, they did not find a change in duration nor did they examine LL communications. In the current study, even our low task load scenarios had higher volumes of traffic than what the ATCSs experienced in the Willems and Truitt study because our sector staffing included an R- and D-side, where Willems' and Truitt's study staffed the sector with a single R-side only. The more traffic, the more ATCSs need to compensate by their communications. Under increasing task loads, they need to issue more clearances but have the same amount of time to get them in, therefore shortening the amount of time they communicate.

3.5 Trust Ratings

Several researchers have proposed taxonomies to describe the dimensions of trust. Barber (1983) focused on the expectations a user brings to a trusting relationship. When experience confirms each expectation, the development of an accurate mental model of the system is more likely. Rempel, Holmes, and Zanna (1985) focused on the dynamics of trust, which evolve and change over time because of experience with a system. Zuboff (1988) and Lee and Moray (1992) synthesized the taxonomies of Barber and Rempel et al. and proposed their own taxonomies of trust. Moffa (1994) measured user trust with a questionnaire he constructed from Barber's and Rempel et al.'s trust metrics. He varied level of agreement between the user's diagnosis and that offered by a medical diagnostic aid. He varied experience with the system in the form of number of trials. Moffa found that trust decreased

- with consistent system disagreement for both levels of trials;

- with consistent system disagreement but increased with honest agreement, when pooled across trials; and,

- with more experience with the system, when pooled across level of agreement.

Moffa (1994) found no significant difference between ratings for Barber's (1983) dimensions and those for Rempel et al.'s (1985) dimensions. Whereas, Moffa studied trust and behavior of non-experts responding to discrete, case-based automation in medicine, Moffa and Stokes (1997) examined this work in light of other domains, including ATC. An ATCS works in a domain that is more continuous and process-based. Moffa and Stokes concluded that it is likely that the structure of knowledge in a domain, the workplace practices, and the professionalism of the domain affect trust, complacency, and compensatory behaviors.

We applied work done in operator trust in automation to the domain of ATC. Based on Barber (1983), Rempel et al. (1985), Zuboff (1988), and Lee and Moray (1992), we propose the following taxonomy to describe the dimensions of trust in ATC. Operator experience forms the basis of this mostly linear taxonomy.

- We assume that users of air traffic systems enter the human-automation relationship with the following expectations (from Barber):
  - The system is appropriate for the task for which it was designed, and it will remain so.
  - The system exhibits technical competence.
  - The system is responsible.
- Each user explores a new system by "trying out" its functions and features, each in turn. Exploration continues when the user applies the system to each new situation.
- With experience, the user sees how consistent (or inconsistent) the system's recurrent behaviors are (from Rempel et al.).
- With continued experience, the user shifts from evaluating specific behaviors to evaluating the dependability of the entire system (from Rempel et al.).
- Given enough experience with a predictable and dependable system that is appropriate, competent, and responsible, the user will have faith that the system will continue to be so.

3.5.1 Data Analysis and Results

Four items, rated on a 10-point scale, comprised the trust scale for the DST: the predictability of the DST's behavior, how well the DST predicted the outcome of separation strategies, the capability of the DST to predict future conflicts, and the technical knowledge incorporated in the tool. Because participants only completed these items in either the limited- or full- automation condition, we conducted a 2 x 2 x 2 (position x task load x automation) within-subjects MANOVA. Appendix Z contains the means, *SDs*, MANOVA and ANOVA tables.

At the multivariate level, a significant effect for automation occurred [$\Lambda = .38$, $F(4,12) = 4.97$, $p < .05$, Appendix Z, Table Z-2]. Because the MANOVA results indicated a significant effect for automation, we conducted ANOVAs on each of the measures and adjusted the alpha to .013 to control for error. At the univariate level, automation did not reach significance; however, we found several trends in the data using an alpha of .05. ATCSs rated the DST as having better prediction capabilities in the full automation condition than in the limited automation condition (Figure 71). ATCSs perceived the DST to behave more predictably under low task load conditions than high task load conditions, although the degree of automation affected this. For the limited automation condition, task load did not have an effect; however, in the full automation condition, ATCSs perceived the DST as more predictable when the task load was low (Figure 72). Finally, ATCSs rated the DST as more capable to predict future conflicts when the traffic task load was low (Figure 73).

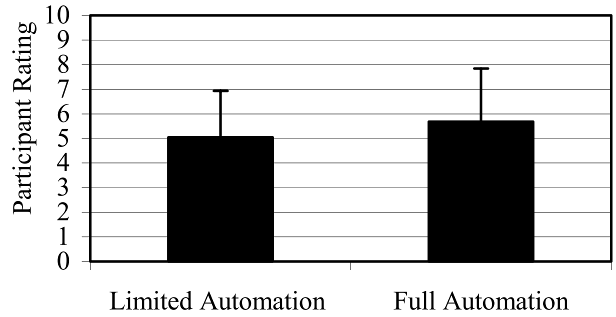

Figure 71. How well the DST predicted the outcome by automation.

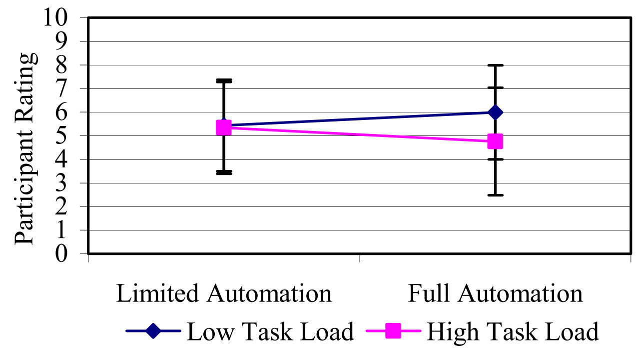

Figure 72. The predictability of the DST's behavior by task load and automation.

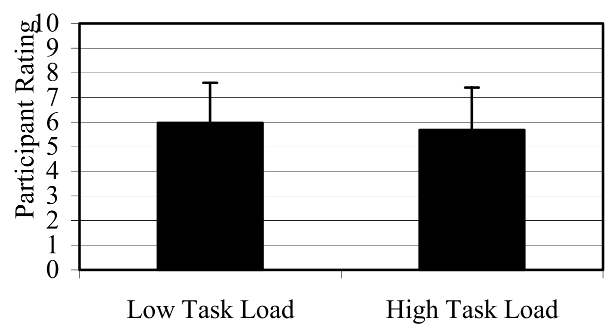

Figure 73. Capability of the DST to predict future conflicts by task load.

3.5.2 Discussion

Task load and automation affected ATCS trust in the DST. The algorithms that determined the DST's behavior obviously did not change between conditions. Therefore, ATCSs trust in the DST changed. It is quite likely that the lack of a mental model for the DST made it difficult for ATCSs to understand how the DST arrived at certain conclusions. During training, we witnessed the ATCSs' disbelief that a tool such as the DST existed and that it could predict potential conflicts between aircraft. Although training may have modified the ATCSs' expectations and beliefs regarding the DST's prediction capabilities, it is most likely that the limited training and period of use of the DST did not completely modify their expectations nor give them a chance to form a complete mental model of how the DST worked.

The graphical display and trial planning was only available under full automation. Therefore, the reduction in trust under limited automation may be related to the lack of these features. In the limited automation condition, only the aircraft list was available. In the aircraft list, potential conflicts are indicated by either red or yellow emphasis. It fails to indicate which aircraft conflict with one another. ATCSs, therefore, did not receive the feedback they needed to verify the conflicting aircraft unless they accessed the relevant information through the graphical display (see Figure 3). This, in turn, may result in a lower level of trust in the DST. As noted in the trust taxonomy, the expectations are that the system is appropriate for the task, and, in this case, limited automation is not entirely appropriate for the task. These types of issues are the focus of the human computer interface design. By providing ATCSs the ability to group aircraft when they belong to the same potential conflict, we can give them a better feel for the correctness of the information. This seems especially relevant in operations where a single R-side ATCS works a sector and often only displays the aircraft list.

Under high task load conditions (i.e., when ATCSs needed assistance most), they trusted the DST the least. During training and implementation, this may be an issue to give more emphasis. Further, as noted by Wiener (1988), users of automation may turn it off under high workloads. In the current study, ATCSs did not have the ability to turn the automation off but could disregard the information it presented. This could lead the ATCSs to believe the automation was not as predictable when task load was high and full automation was in use.

The results imply that both task load and level of automation affected ATCSs' trust in the DST in this study. ATCSs described better prediction quality for outcome of separation strategies when able to use all the features offered by the DST such as trial flight plans and the graphic plans display, relating to the appropriateness of the tool for the given task. Low task load scenarios allowed ATCSs more time to work with the DST and explore the various functions of the DST. This coincides to the trust taxonomy described previously. With increased amounts of time to "try out" the DST, the ATCSs viewed it more favorably.

3.6 Real-Time Objective Performance

To obtain the Real-Time Objective Performance variables that the Data Reduction and Analysis (DRA) system at the RDHFL previously provided, we submitted the TGF recordings to the Data Reduction and Analysis Tool (DRAT).

3.6.1 Data Reduction Analysis Tool

The automated data reduction module developed for the TGF data provided performance data for complexity and handoff efficiency, conflicts, and task load. We calculated the variables based on ATCS responsibility, not on fully active control (i.e., data block maintenance and communications active). In other words, the ATCS was responsible for any aircraft within his or her sector that they were talking to or that had already changed to the next sector's frequency but had not physically reached the sector boundary.

3.6.1.1 Data Analysis and Results

The DRAT module processed raw data files produced by TGF. DRAT provided summary, interval, and error files for each scenario. The interval and summary files comprised two data sets in spreadsheet format. The first data set contained 12 x 8 (scenarios x ATCS teams) records that included the summary variables calculated per scenario. The second data set contained 12 x 10 x 8 (scenarios x intervals x ATCS teams) records containing the summary variables calculated per 3-minute interval. For the first 10 participants, we averaged any replicated data for a team of ATCSs.

The categories of variables related to performance included Complexity and Handoff Efficiency and Conflicts. A set of ANOVAs tested the hypotheses related to performance scores on selected performance variables. These ANOVAs addressed the differences across scenarios and were of a repeated measures 2 x 3 (task load x automation) design.

3.6.1.1.1 Complexity and Handoff Efficiency Per Scenario

The number of altitude, heading, and speed changes comprised the complexity items, and the number of handoffs accepted and initiated comprised the handoff efficiency items that we derived from DRAT. We calculated the number of altitude, heading, and speed changes and the number of handoffs accepted and initiated per aircraft to adjust for the varying scenario lengths that ranged in duration from 33 to 37 minutes. We performed separate 2 x 3 (task load x automation) within-subjects ANOVAs because there were not enough degrees of freedom to conduct a MANOVA. The adjusted alpha level was .01. See Appendix AA for means, *SDs*, and ANOVA tables.

The ANOVA examining the number of speed changes showed a significant effect for task load [$F(1,7) = 19.79$, $p < .01$, Appendix AA, Table AA-8]. Simulation pilots performed more speed changes when task load was high (Figure 74). The ANOVA for the number of altitude changes did not show any significant effects, although a trend indicated that more altitude changes occurred in the low task load scenarios than in the high task load scenarios (Figure 75). Finally, the ANOVAs for the number of heading changes and the number of handoffs initiated and accepted did not show any significant effects for the experimental variables.

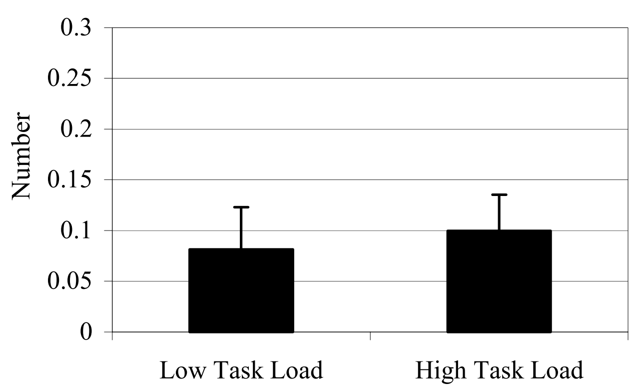

Figure 74. The number of speed changes per aircraft by task load.

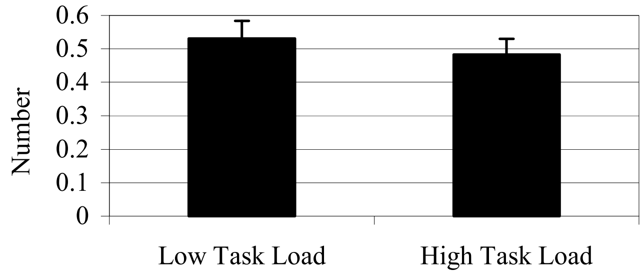

Figure 75. The number of altitude changes per aircraft by task load.

3.6.1.1.2 Complexity and Handoff Efficiency Per Interval

We conducted separate 2 x 3 x 10 (task load x automation x time interval) within-subjects ANOVAs for the number of altitude, heading, and speed changes per 3-minute interval. We conducted separate 2 x 3 x 9 (task load x automation x interval) within-subjects ANOVAs for the number of handoffs initiated and accepted per 3-minute interval. We did not include the first interval in these analyses because it included the start-up of the scenario where the system automatically accepted aircraft.

The ANOVA for the number of altitude changes per 3-minute interval showed significant effects for task load, interval, and the task load x interval interaction [$F(9,63) = 12.81$, $p < .001$, $F(1,7) = 36.36$, $p < .001$, $F(9,63) = 4.04$, $p < .001$, respectively, Figure 76, Table AA-11]. More aircraft from the high task load condition changed their altitude than aircraft in the low task load condition. Significantly more aircraft had altitude changes in the first 3-minute interval and gradually, the number of altitude changes decreased. The simple effects of task load within each interval showed significant increases in altitude changes for the third, seventh, eighth, and tenth intervals [$F(1,7) = 34.21$, 14.59, 30.64, 31.03 all *p*s < .01, respectively, Table AA-11].

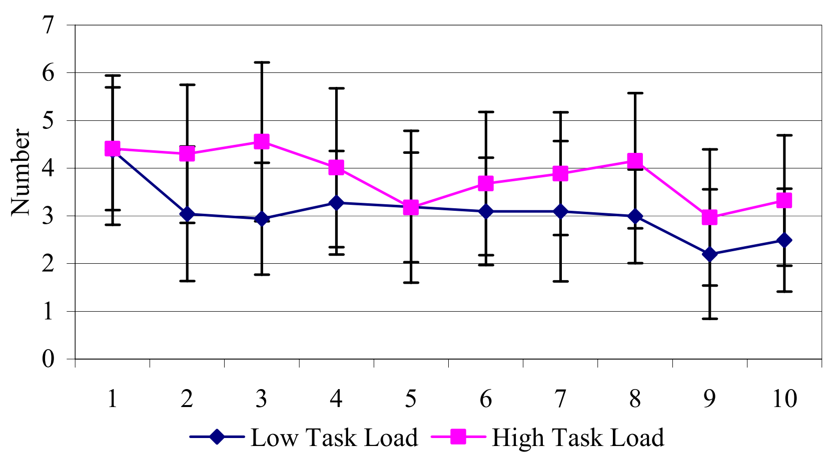

Figure 76. Number of altitude changes by task load and time interval.

The ANOVA for the number of heading changes issued showed significant effects for task load and interval [$F(1,7) = 17.32$ and $F(9,63) = 6.83$, *p*s < .01, respectively, Figure 77, Table AA-13]. More heading changes occurred in the high task load scenario. We conducted Tukey HSD post hoc tests to determine significant differences among the time interval means (Table AA-15). Several means were significantly different; therefore, we provide a general description. As the scenario progressed, the number of heading changes generally increased reaching a peak at the ninth interval and then declining at the tenth interval.

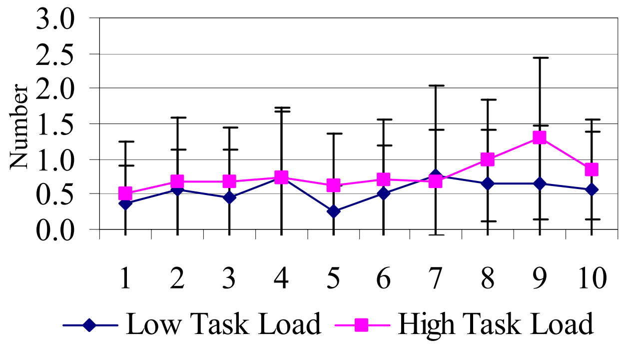

Figure 77. Number of heading changes by task load and time interval.

The ANOVA examining the number of speed changes showed a main effect for task load and time interval [$F(1,7) = 60.62$, $p < .001$, $F(9,63) = 7.54$, $p < .0001$, respectively, Figure 78, Table AA-15]. The number of speed changes increased as task load increased. Tukey HSD post hoc tests showed several significant differences among the time interval means (Table AA-16). In general, as the scenario increased, the number of speed changes increased reaching a peak at the sixth and seventh intervals and then declining.

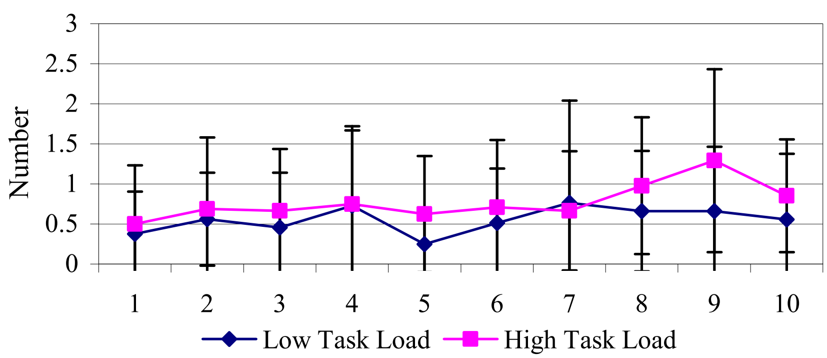

Figure 78. Number of speed changes by task load and time interval.

Task load and time interval influenced the number of handoffs accepted [$F(1,7) = 63.24$, $p < .001$ and $F(8,56) = 6.71$, $p < .001$, respectively, Figure 79, Table AA-17], and the interaction between task load and time interval was significant [$F(8,56) = 7.49$, $p < .001$, Table AA-17]. ATCSs accepted more handoffs when task load was high. Tukey HSD post hoc tests showed several significant differences among time interval means (Table AA-18). In general, ATCSs

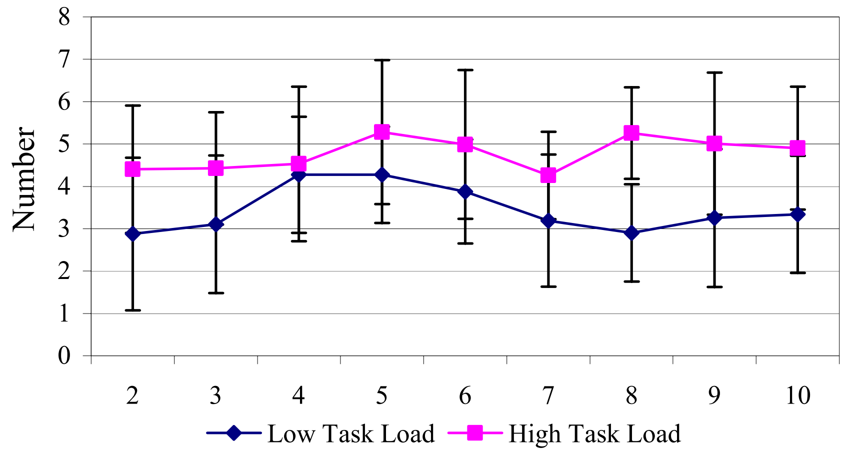

Figure 79. Number of handoffs accepted by task load and time interval.

Task load and time interval affected the number of handoffs initiated [$F(1,7) = 93.52$ and $F(8,56) = 8.91$, *p*s < .001, respectively, Figure 80, Table AA-19]. ATCSs initiated more handoffs when working traffic from the high task load scenarios. Tukey HSD post hoc tests showed significant differences among several time interval means (Table AA-20). In general, the second time interval had the lowest number of handoffs initiated.

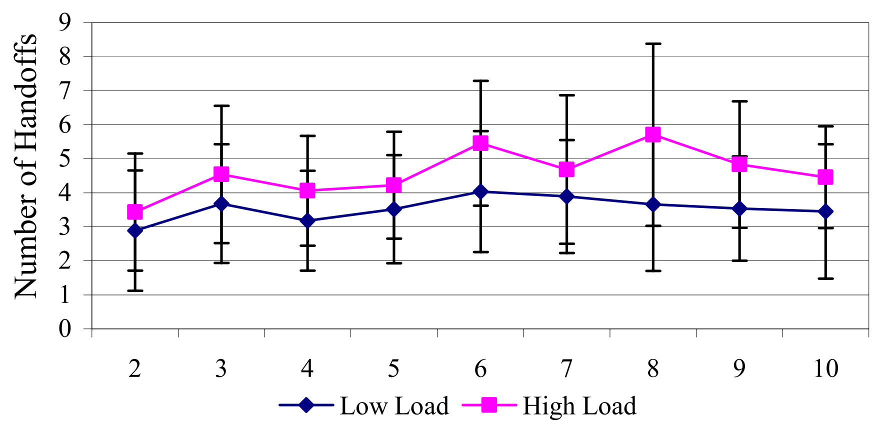

Figure 80. Number of handoffs initiated by task load and time interval.

3.6.1.1.3 Number and Duration of Conflicts

Thirty losses of separation occurred during the scenario runs. Due to the relatively low number of losses of separation, we did not conduct ANOVAs on the data. Instead, we provide descriptive statistics on the data (Appendix AA). The distribution for the mean number and duration of the losses of separation is shown in Figure 81 and Figure 82, respectively. More losses of separation occurred under higher task loads [$M = 0.45$ (0.63)] and with either limited [$M = 0.50$ (0.90)] or full automation [$M = 0.53$ (0.52)]. The mean duration of losses of separation was higher under high task load and no automation [$M = 13.3$ (31.25)].

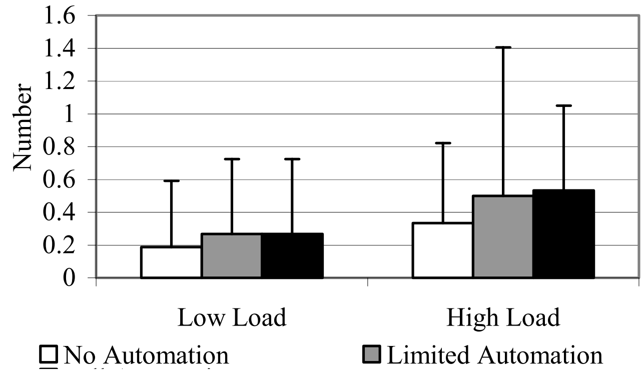

Figure 81. Number of conflicts by task load and automation.

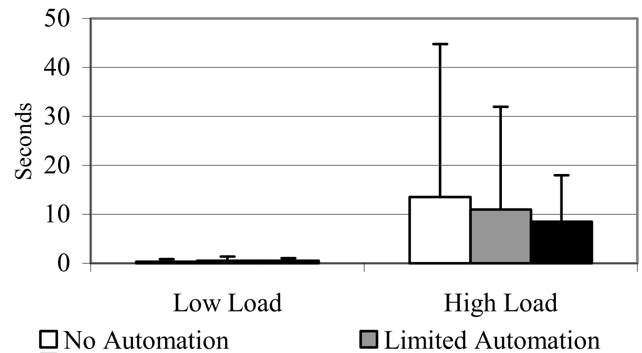

Figure 82. Duration of conflicts by task load and automation.

3.6.1.1.4 Number of aircraft handled during scenario

To prevent ATCSs from anticipating when scenarios would end, we stopped simulations at times ranging from 33 to 37 minutes. To enable comparison across experimental conditions, we accounted for the variable scenario lengths by calculating the number of aircraft handled per minute. We performed a 2 x 2 x 3 (position x task load x automation) within-subjects ANOVA. We found a main effect of task load [$F(1,15) = 270.95$, $p < .001$, Figure 83, Table AA-23]. ATCSs controlled 33% more aircraft in the high task load conditions.

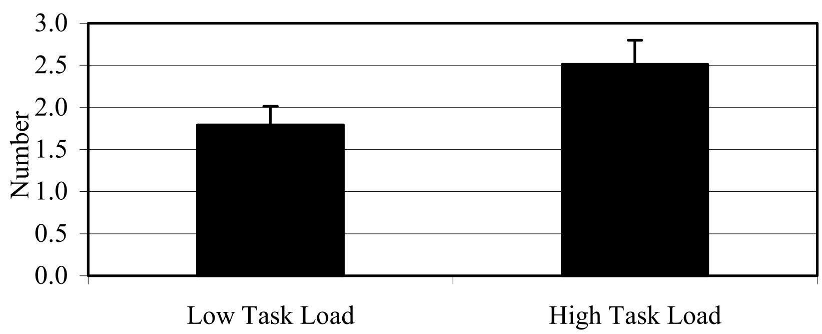

Figure 83. Number of aircraft handled per minute during scenario by task load.

3.6.1.2 Discussion

In this experiment, we observed a complex interaction of the effect of the use of automation and the assistance provided to the R-side ATCS on conflicts. Although the use of the DST can benefit the R-side ATCS strategically, it will remove some of the immediate, more tactical support. Our descriptive analysis of the instances when ATCSs lost separation showed that under full automation, these instances were more frequent, but shorter. Without automation, loss of separation occurred less frequently, but lasted longer. We explain this from the D-side ATCS activities we observed under these conditions. Without automation present, the D-side ATCS physically moved closer to the R-side ATCS and provided radar assistance. With the extra assistance, the ATCS team was able to control traffic well in the current tactical fashion. The teams resolved tactical or immediate problems easily, resulting in fewer losses of separation. Because the team now worked traffic in a tactical fashion, however, those instances that did result in a loss of separation took longer to resolve. Under full automation, the D-side ATCS had the DST available to act more strategically and may have succeeded. With more focus on the DST, however, the R-side ATCS had less assistance in the tactical environment, which may have resulted in more tactical problems that the ATCS could resolve quickly. Another possible explanation for this effect relates to earlier findings that the ATCSs' trust in the DST was low, particularly in high task load conditions. This may have translated into more losses of separation under automation because the D-side ATCS was testing the correctness of the DST; he or she waited to see if a predicted conflict actually became a conflict.

Although the number of times ATCSs did not maintain adequate separation between aircraft was high when compared to operational data, the data cannot be used for formal inferential statistical analysis. Even under high traffic conditions, only one out of two ATCSs had a loss of separation when full automation was present. Besides the fact that our data would mostly consist of observations that equal zero cases of loss of separation, our data do not indicate if a loss of separation would indeed have resulted in an OE. The loss of separation could have been the result of a pilot deviation or could have been rejected as an OE for other reasons.

Our data seem to indicate that under our experimental conditions, the high traffic task loads used in our scenarios led to an increased frequency of loss of separation. To fully understand the sequence of events that led to a loss of separation, we need to follow a different approach than we have followed in this study. We have data that we synchronized across data sets. Therefore, we can trace events that affected aircraft involved in a conflict in an attempt to explain the cause of the loss of separation. Tracing these events is labor intensive and tedious but may provide useful data that can help us understand how ATCSs loose separation between aircraft.

To allow other researchers to compare our traffic levels to those in other studies, we have expressed them in the number of aircraft handled per minute. We instructed our SMEs to create traffic levels that reflected situations where a supervisor would be about to change the staffing in the sector to only an R-side ATCS for low task load scenarios and to a three-person team (i.e., add a tracker) for high task load scenarios. Our analysis shows that our task load manipulation worked and resulted in a 33% increase in the number of aircraft handled within a fixed time interval.

To compensate for the variable scenario lengths, we calculated the number of altitude, heading, and speed changes per aircraft. This approach also tends to compensate for differences in traffic volume. Therefore, changes in the number of these aircraft maneuvers reflect a change in control strategy that the mere change in number of aircraft handled cannot explain. Without expressing aircraft maneuvers in number per aircraft, an increase in traffic task load will, of course, lead to an increase in the number of aircraft maneuvers within a scenario. This increase by itself may increase ATCS workload and alter other behavioral aspects. Our results indicate, however, that ATCSs use more speed changes per aircraft under high traffic task load conditions. In contrast, ATCSs used fewer altitude changes per aircraft when traffic task load increased. With larger traffic volumes, aircraft have less open airspace in which to maneuver, and, to compensate for this, ATCSs used speed instead of heading changes. ATCSs often indicate that an airspace with more climbing and descending aircraft is more difficult to control than airspace with mostly en route aircraft at fixed altitudes. When aircraft make a vertical transition, ATCSs have more trouble predicting their behavior and risk. Under high task load, ATCSs tried to reduce the level of uncertainty by reducing the number of vertical maneuvers per aircraft.

The complexity and handoff efficiency items, when examined by interval, showed significant effects for task load and interval. When task load was high, more altitude, heading, and speed changes occurred, and ATCSs accepted and initiated more handoffs. With higher task loads, ATCSs controlled more aircraft, thus more control actions were required, as indicated by the higher numbers for the complexity and handoff efficiency items. As time increased, the number of altitude changes gradually decreased; significantly more altitude changes occurred during the first 3-minute interval. At the beginning of the scenario, the ATCS had to work traffic that was not previously controlled. This most likely contributed to the higher number of altitude changes at the beginning of the scenario because the ATCS needed to clear the planes on either the arrival or departures routes to the proper altitudes. As the scenarios progressed, simulation pilots generally performed more heading and speed changes, whereas, ATCSs accepted and initiated more handoffs. ATCSs in our experiment sat down at the position and had very little time to take over. Early in the scenarios, ATCSs maneuvered aircraft to set the scenario up to their liking. While they were catching up, they built a bigger plan, likely foreseeing some of the future situations. This, in turn, led to less need for maneuvers that increase uncertainty such as altitude changes later in the scenario.

3.6.2 ATCSs' Interactions with the DSR

To determine how ATCSs interacted with the DSR interface, we used the SAR tapes collected during the experimental simulation runs. With the NAS Data Analysis and Reduction Tool (DART), we extracted ATCS input messages from the SAR tapes, filtered messages from the sector under study, identified whether messages came from the R- or the D-side, and stored them in text files. We then used an application in Labview (National Instruments Corporation, 2000) to parse and tally the messages and write them to data files for further analysis.

3.6.2.1 Data Analysis and Results

Due to problems with SAR tape expiration dates, we lost tapes for the last four participants and had missing tapes and data for the other 12 participants. Because of the substantial loss of the data, we did not use mean replacement but structured the dataset for between-subject analyses using an unequal N. We realize this analysis does not take advantage of eliminating between subject variance and incorporate this during the interpretation of the statistical analyses.

We divided the ATCS interactions into several groups for analysis. We performed a 2 x 2 x 3 (position x task load x automation) between-subjects ANOVA for the total number of entries into the DSR system. We grouped the number of flight plan readouts, j-rings, and route displays as information (Information Pickup) that the ATCSs gathered from the system. The number of altitude changes, interim altitudes, removal of interim altitudes, and route changes composed ATC control actions. Station-keeping activities coded from the DART included initiating and accepting handoffs and datablock movement. For these three categories, we conducted 2 x 2 x 3 (position x task load x automation) between-subjects MANOVAs. We followed any significant multivariate results with univariate analyses and adjusted the alpha level. We also assessed any results with a *p*-value of .1 to determine if anything stood out that may be of interest to examine in future studies. We did this because of the limitations due to missing data and using between-subjects analyses. Appendix BB contains the means, *SDs*, and ANOVA tables.

3.6.2.1.1 Total Number of Entries

Significant main effects for position and task load occurred for the total number of entries made into the system [$F(1,52) = 277.17$, $p < .001$ and $F(1,52) = 12.43$, $p < .001$, respectively, Appendix BB, Table BB-4, Figure 84]. When in the R-side position, ATCSs total number of entries was higher than when in the D-side position. ATCSs made more entries under high task load conditions.

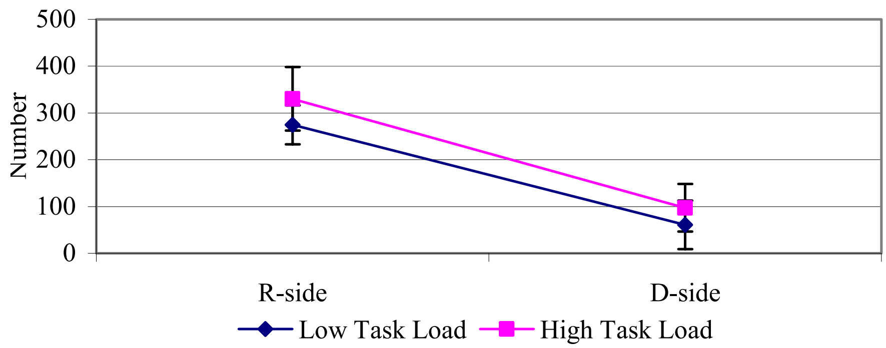

Figure 84. Total number of ATCS entries by position and task load.

3.6.2.1.2 Information Pickup

The MANOVA for the information pickup items showed a significant effect for position [$\Lambda = .48$, $F(3,50) = 17.70$, $p < .001$, Table BB-5]. The adjusted alpha was .017. The main effect of position for the number of flight plan readouts and j-rings was significant [$F(1,52) = 37.32$, $p < .001$ and $F(1,52) = 12.34$, $p < .001$, Table BB-6 and Table BB-7, respectively]. R-side ATCSs used more flight plan readouts and j-rings (Figure 85 and Figure 86, respectively). A secondary trend for position for number of route displays showed that R-side ATCSs used more route displays than D-side ATCSs (Figure 87, Table BB-8).

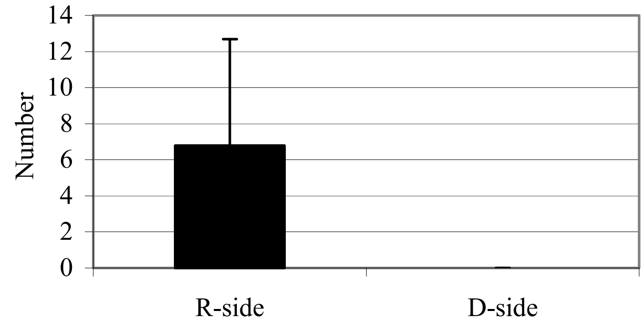

Figure 85. Number of flight plan readouts by position.

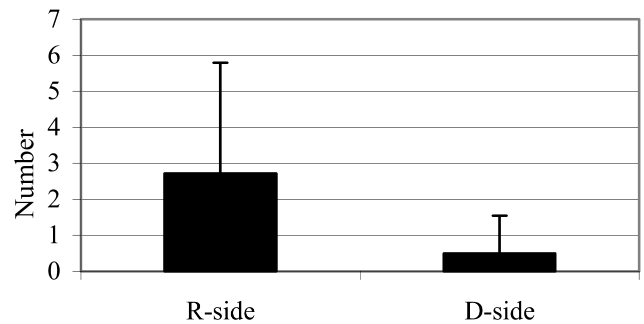

Figure 86. Number of J-rings by position.

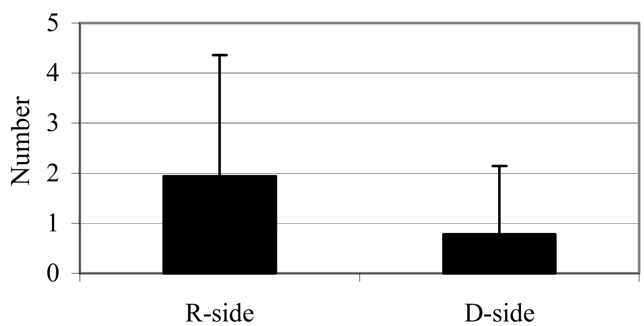

Figure 87. Number of route displays by position.

3.6.2.1.3 ATC Control Actions

The MANOVA for ATC control actions showed significant effects for position and task load [Λ = .23, $F(4,49) = 41.06$, $p < .0001$ and $\Lambda = .60$, $F(4,49) = 8.28$, $p < .001$, respectively, Table BB-9]. Because we had significant multivariate findings, we conducted subsequent univariate analyses and used an adjusted alpha of .013. We found a secondary trend for task load for the number of hard altitudes entered into the system (Figure 88, Table BB-10). ATCSs entered more hard altitudes into the system under high task load conditions.

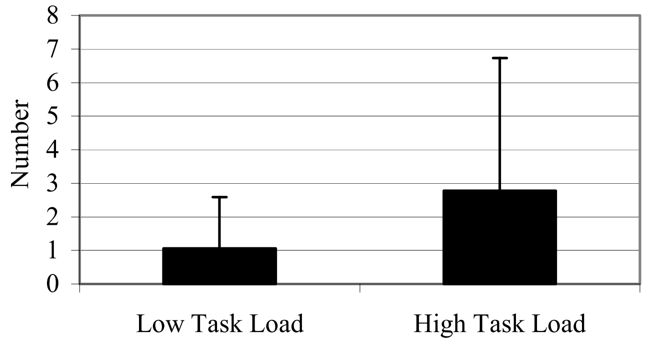

Figure 88. Number of hard altitudes entered by task load.

The ANOVA for entering interim altitudes in the system showed significant effects for position and task load [$F(1,52) = 106.81$, $p < .001$ and $F(1,52) = 14.35$, $p < .001$, respectively, Figure 89, Table BB-12]. A secondary trend in the data showed an interaction between position and task load. Under high task load conditions, R-side ATCSs entered more interim altitudes; the effect for task load diminished this for the D-side ATCS.

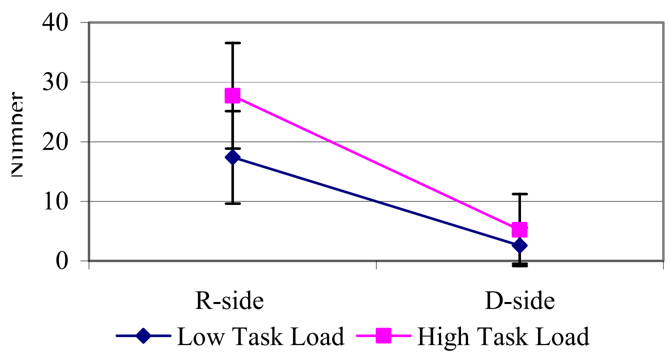

Figure 89. Number of interim altitudes by position and task load.

The effect for position in the removal of interim altitudes was significant [$F(1,52) = 64.24$, $p < .001$, Table BB-13]. A trend for automation showed more interim altitudes removed under limited automation conditions than no or full automation. In addition, the interactions between position and automation and task load and automation and the 3-way interaction were secondary trends. The limited automation condition seems to drive the interaction. Under low task loads and limited automation, R-side ATCSs removed more interim altitudes, but D-side ATCSs removed fewer interim altitudes under low task loads and limited automation (Figure 90 and Figure 91).

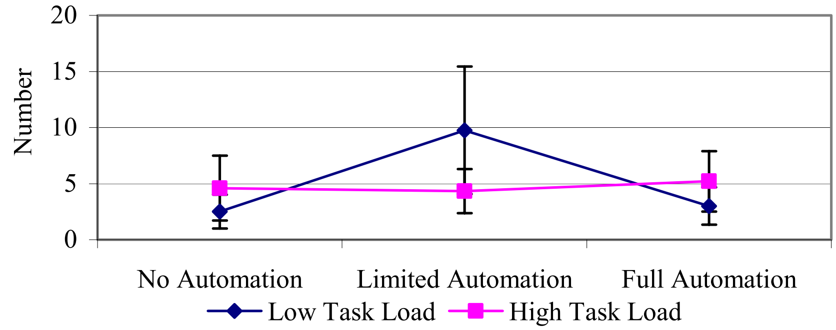

Figure 90. The interim altitudes removed by R-side position, task load, and automation.

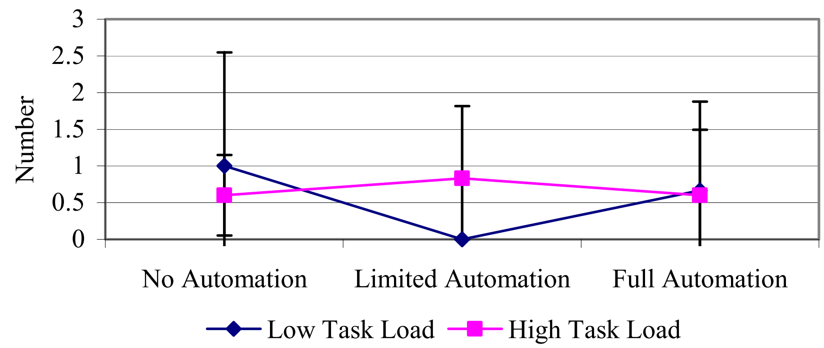

Figure 91. The number of interim altitudes removed by D-side position, task load, and automation.

3.6.2.1.4 Station-Keeping Activities

The multivariate analysis for station-keeping items showed significant effects for position and task load [$\Lambda = .19$, $F(3,50) = 70.95$, $p < .001$ and $\Lambda =.77$, $F(3,50) = 4.887$, $p < .01$, Table BB-14]. Follow-up univariate analyses showed a significant effect for position in the number of handoffs initiated [$F(1,52) = 77.17$, $p < .0001$, Table BB-15]. R-side ATCSs initiated more handoffs (Figure 92).

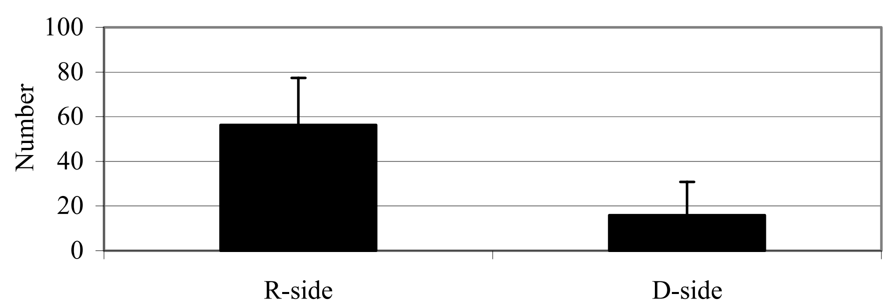

Figure 92. Number of handoffs initiated by position.

The main effects for position and task load were significant for the number of handoffs accepted [$F(1,52) = 183.73$, $p < .001$ and $F(1,52) = 14.65$, $p < .001$, respectively, Table BB-16]. R-side ATCSs accepted more handoffs and more handoffs occurred during the high task load scenarios (Figure 93). The position by task load interaction was a secondary trend. The R-side ATCSs accepted more handoffs under high task load conditions, but the effect of task load was diminished for the D-side ATCS.

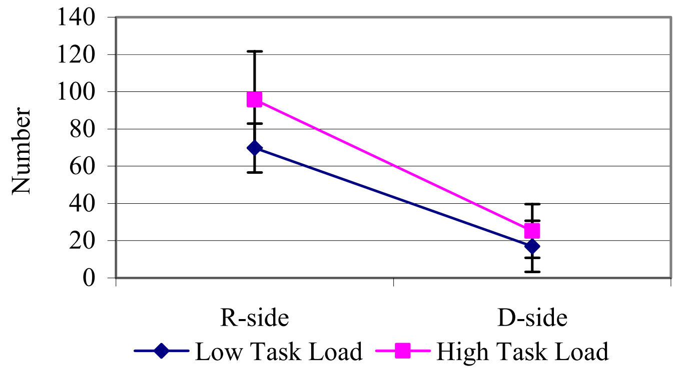

Figure 93. Number of handoffs accepted by position and task load.

The effect of position was significant for the number of datablock movements [$F(1,52) = 59.66$, $p < .001$, Table BB-17]. The R-side ATCS moved more datablocks (Figure 94). A secondary trend for task load showed that ATCSs moved more datablocks under high task load conditions (Figure 95). A secondary trend for the position x automation interaction showed that D-side ATCSs moved more datablocks under low automation conditions than limited or full automation conditions, but the number of datablocks moved by R-side ATCSs remained relatively constant.

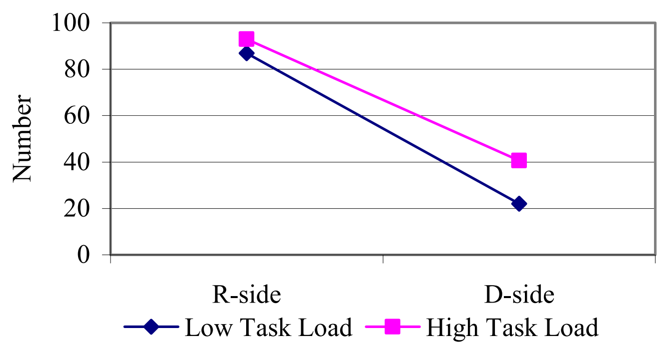

Figure 94. The number of datablock movements by position and task load.

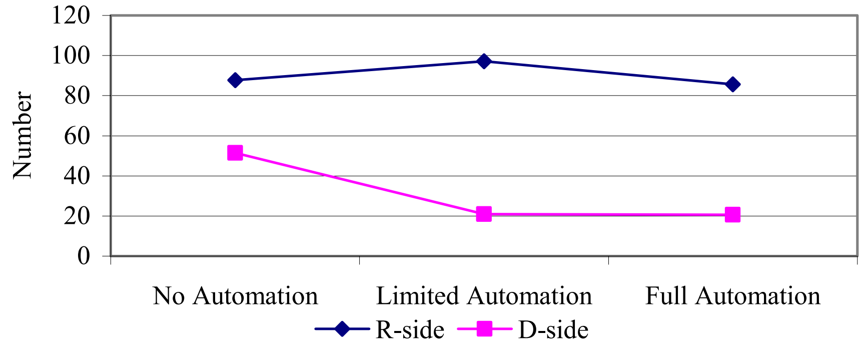

Figure 95. The number of datablock movements by position and automation.

3.6.2.2 Discussion

We used between-subjects multivariate and univariate analyses on the ATCS interactions with the DSR because of problems with substantial missing data. By conducting the between-subjects analyses, we found significant effects for position and task load on most ATCS interactions with the DSR. However, we expected to find effects of automation. The lack of statistically significant findings for automation reflects the substantial between-subject variance present in the analyses and highlights the strength of using a repeated measures design. The control strategies used by ATCSs differ greatly on an individual basis and even from ARTCC to ARTCC. Therefore, the variability between ATCS interactions with the DSR results in a larger amount of error variance that makes it more difficult to assess effects for variables. Within a repeated measures design, each participant is his or her own control and lowers this error variance. With this in mind, we presented some findings with a larger *p*-value to highlight anything that may be of interest to examine in future studies.

The effect for position was significant across the ATCS interaction categories, information pickup, ATC control actions, and station keeping. The R-side ATCSs worked directly with the DSR display and traffic and were more likely to seek additional information and perform ATC actions and station-keeping activities. Task load also had an impact across the set of items. As task load increased, more actions were necessary to control traffic. Although not statistically significant, we found an interaction between position and automation for the number of datablocks moved. Under limited or full automation conditions, D-side ATCSs moved fewer datablocks, but the number of datablocks R-side ATCSs moved remained relatively constant. This may imply that the D-side ATCSs used the automation and that this drew their attention away from assisting the R-side ATCS in the station-keeping activity of moving datablocks. Interesting, the number of datablocks moved by the R-side did not increase during the automation conditions. Perhaps the D-side ATCS was using the automation to maneuver aircraft and that less datablocks needed to be moved or that the R-side did not have enough time to move the datablocks and just ignored the task.

Also of interest are the relatively low numbers of actions D-side ATCSs performed. The D-side ATCS is supposed to assist the R-side ATCS; however, our results show that the D-side offers minimal assistance for the ATCS interaction items. The R-side ATCS performed approximately 300 control entries to roughly 80 entries made by the D-side ATCS. The D-side made substantially fewer entries for such items as flight plan readouts, route displays, interim altitudes, handoffs initiated and accepted, and datablock movements. This finding may be reflective of individual style differences and differences from ARTCC to ARTCC. Some R-side ATCSs prefer to make control entries and have the D-side support as an extra pair of eyes. In this case, the R-side ATCS requests minimal control entry assistance from the D-side ATCS. The responsibilities of the D-side ATCS also differ from facility to facility. At some ARTCCs, the D-side ATCSs are more active with control entries and initiate and accept handoffs and move datablocks. At other facilities, R-side ATCSs are responsible for control entries. The interactions of the ATCSs seem to indicate that the D-side only takes care of, at most, 25% of the observable activities. If the D-side function is to reduce the R-side workload and increase the R-side efficiency, that support function is inefficient in the configuration we tested in this experiment.

3.7 Questionnaires

We used two types of self-report questionnaires adapted from previous experiments. The questionnaires included an Entry Questionnaire (Appendix K) and a PSQ (Appendix H) (Abbott, Nataupsky, & Steinmetz, 1987; Guttman et al., 1995; Sollenberger & Stein, 1995; Stein, 1992; Willems et al., 1999). The Entry Questionnaire contained questions concerning demographic information. The PSQ contained questions about various aspects of controlling traffic during a scenario.

We administered the Entry Questionnaire and PSQ in paper and pencil format and transcribed the responses into a spreadsheet. We created three data sets. Because of problems with missing data, we divided the data sets into one with 16 participants and one with the last 6 participants for the PSQ responses. The other data set contained data from the Entry Questionnaire.

3.7.1 Entry Questionnaire

The Entry Questionnaire contained questions about participant background, demographics, and importance of provided airspace and aircraft information. We have included the information about background and demographics in section 2.1 to describe the characteristics of our participants. For most of our experiments, we asked ATCSs about importance of airspace and aircraft information because we anticipate that with the change of the ATCS working position, some of that information may change or disappear. By recording this information, the designers of future ATC systems have a baseline of what information ATCSs feel is important.

3.7.1.1 Data Analysis and Results

The analysis of the Entry Questionnaire data consisted of the calculation of means and *SDs* (Appendix CC). Table CC -1 presents the means and *SDs* for general questions from the entry questionnaire. The range of the scale was 1 to 10.

Table CC-2 presents ATCSs' ratings for the importance of various aircraft information sorted from most important to least important. ATCSs rated the current aircraft location, current altitude, and arrival airport (within sector) as the most important aircraft information.

Table CC-3 presents the ratings for the importance of radar display information. ATCSs rated sector boundaries, restricted area boundaries, and filter settings as the most important.

3.7.1.2 Discussion

ATCSs indicated that position, altitude, and destination are most important. Other information presented in the aircraft representation can be derived from this information (i.e., speed is a change of position; heading changes are related to a change of position).

3.7.2 Post-Scenario Questionnaire

We adapted the PSQ from Willems et al. (1999), and it provided information about several aspects of ATC during a particular simulation scenario (Appendix H). The PSQ contained general questions about the simulation, the perceived ATCS SA, and the NASA TLX items. Here we only discuss the items that do not provide information for other measurement constructs.

3.7.2.1 Data Analysis and Results

We conducted a MANOVA, structured as a 2 x 2 x 3 (position x task load x automation) repeated measures design, for the realism items. Because the MANOVA showed statistical significance, subsequent analyses included ANOVAs on the individual variables. We conducted separate 2 x 2 x 3 (position x task load x automation) within-subjects ANOVAs for participant performance, difficulty, ATWIT and oculometer interference, and simulation pilot performance. We conducted a 2 x 2 (position x task load) within-subject ANOVA for effectiveness of the DST in resolving conflicts when using trial flight plans, and a 2 x 2 x 2 (position x task load x automation) within-subjects ANOVA for competence using the DST. Appendix DD contains the means, *SD*s, MANOVA, and ANOVA tables.

3.7.2.1.1 Realism

Two items composed the realism category: how representative the scenario was of a typical workday and how realistic the simulation was. The MANOVA showed a significant effect for task load [Λ = .61, $F(2,14) = 4.54$, $p < .05$, Table DD-2]. An alpha level set at .025 determined significant results for subsequent ANOVAs of each measure. ATCSs indicated that the low task load scenarios were more representative of a typical workday than the high task load scenarios [$F(1,15) = 6.80$, $p < .05$, Figure 96, Table DD-4]. Using a more liberal alpha level of .05, we found several trends in the data. As automation increased, the perceived representativeness of the scenarios decreased (Figure 96). In addition, ATCSs rated the simulation less realistic when working in the full automation condition (Figure 97).

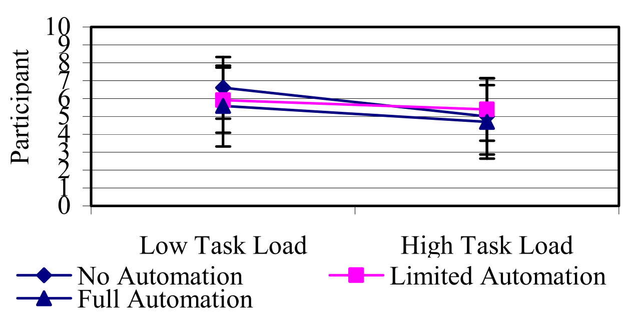

Figure 96. Representativeness by task load and automation.

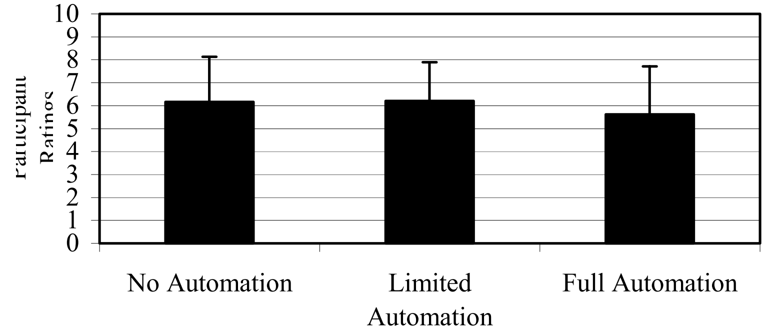

Figure 97. Realism by automation.

3.7.2.1.2 Participant Performance

Position and task load influenced perceived control quality [$F(1,15) = 5.04$, $p < .05$ and $F(1,15) = 36.08$, $p < .0001$, respectively, Table DD-5]. ATCSs perceived they had better performance when controlling traffic on the R-side and under low task load conditions (Figure 98).

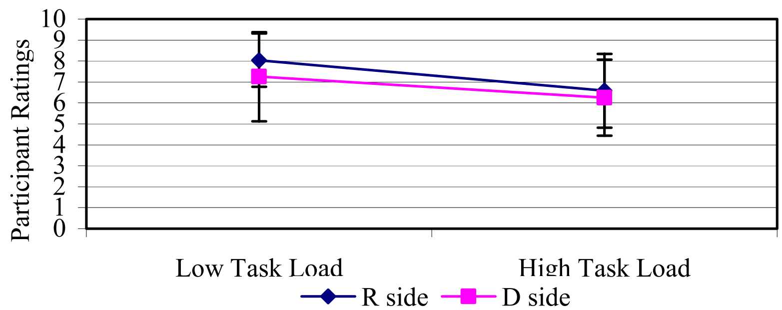

Figure 98. Performance by position and task load.

3.7.2.1.3 Difficulty

Automation had an effect on ATCSs' perceptions of scenario difficulty (Figure 99). As automation increased, the perceived difficulty of the scenario increased [$F(2,30) = 4.04$, $p < .05$,Table DD-6]. Tukey HSD post hoc tests showed ATCSs rated the scenarios more difficult

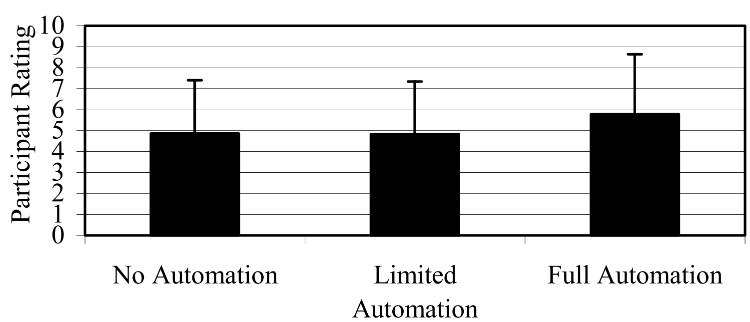

Figure 99. Difficulty by Automation.

when full automation was present than when they had only limited automation. The difficulty rating did not differ statistically between the full automation and no automation conditions or between the limited and no automation conditions. We found an additional effect for task load in the six-participant data set. When the task load was high, ATCSs rated the scenario more difficult than when they controlled low levels of traffic [$F(1,5) = 16.47$, $p < .01$, Figure 100, Table DD-7].

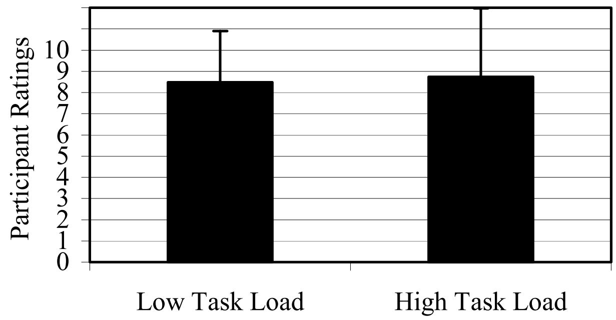

Figure 100. Difficulty by task load (N=6).

<u>3.7.2.1.4 ATWIT and Oculometer Interference</u>

Overall, the perceived interference of the ATWIT device was low (Figure 101). The device bothered the ATCSs even less when task load was low [$F(1,15) = 13.59$, $p < .01$, Table DD-8]. The task load x automation interaction was significant [$F(2,30) = 4.94$, $p < .05$, Table DD-8]. The degree of automation had an impact on the perceived level of interference when task load was high [$F(2,30) = 5.04$, $p < .05$, Table DD-9]. When full automation was present, ATCSs described the ATWIT device as more interfering than when automation was limited or not present at all. The level of automation did not have an effect in the low task load condition. The results from the data set of the last six participants differed slightly. Although overall ratings of ATWIT interference were relatively low, the experimental conditions had a strong influence on ATCSs' perceptions. A significant 3-way interaction was found [$F(2,10) = 5.60$, $p < .05$, Table DD-9]. Examination of the simple effects of task load and automation within the two positions showed a significant effect for task load and the task load x automation interaction for the R-side

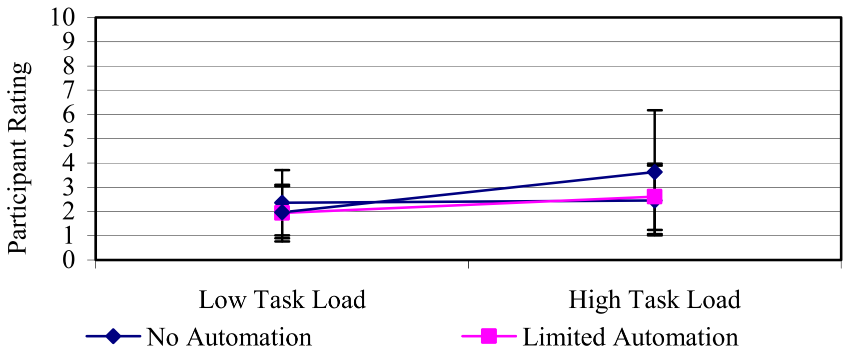

Figure 101. ATWIT interference by task load and automation.

position [$F(1,5) = 25.72$, $p < .001$ and $F(2,10) = 13.20$, $p < .001$, respectively, Figure 102, Table DD-9] but not for the D-side ATCS. When task load was high, R-side ATCSs perceived it to be more interfering than when task load was low. Further analyses of the task load x automation interaction showed no effect for automation within the low task load condition. When task load was high, full automation increased the perceived interference from the ATWIT device. However, there were no statistical differences between the effects of the limited automation and no automation conditions [$F(2,10) = 9.41$, $p < .01$, Table DD-9].

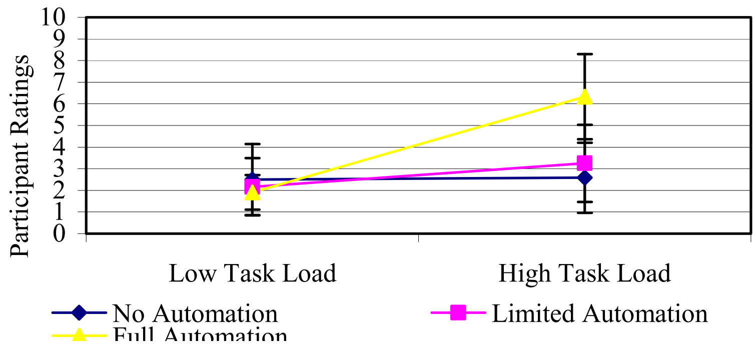

Figure 102. ATWIT interference by R-side position, task load, and automation (N=6).

Because only the ATCS in the D-side position wore the oculometer, we conducted a 2 x 3 (task load x automation) within-subjects ANOVA to examine its interference with controlling traffic. Overall, ATCSs rated the amount of interference from the oculometer as quite low, and the experimental conditions did not influence these ratings.

3.7.2.1.5 Responsiveness of the Simulation Pilots

Ratings for the responsiveness of the simulation pilots decreased when the task load was high [$F(1,15) = 14.38$, $p < .01$, Figure 103, Table DD-11]. No other effects were significant.

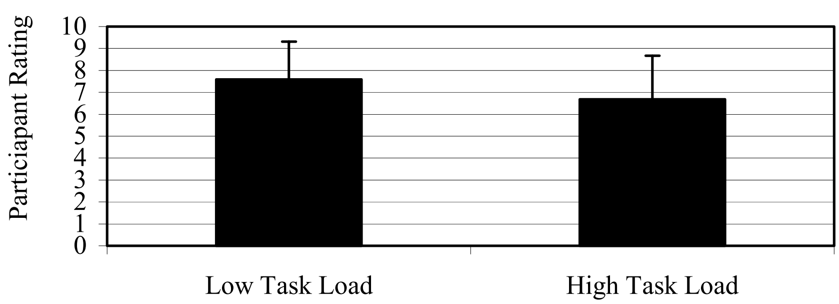

Figure 103. Responsiveness of the simulation pilots by task load.

3.7.2.1.6 Effectiveness of the DST in Resolving Conflicts When Using Trial Flight Plans

Because ATCSs only used the trial flight plan capability of the DST in the full-automation condition, we performed a 2 x 2 (position x task load) within-subjects ANOVA to assess any differences in the effectiveness of the DST in resolving conflicts when using this feature. ATCSs viewed the DST as more effective in resolving conflicts when performing D-side duties than R-side duties [$F(1,15) = 16.37$, $p < .01$, Figure 104, Table DD-12].

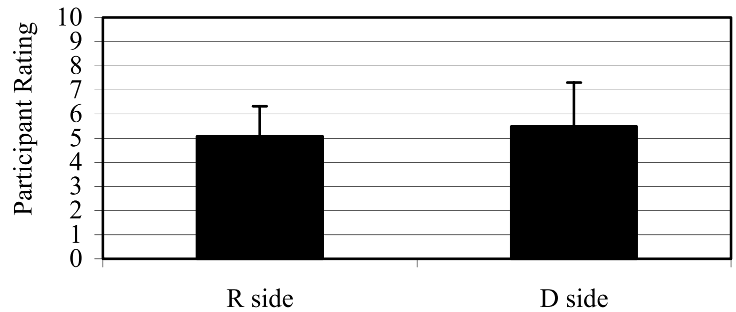

Figure 104. Effectiveness of DST by position.

3.7.2.1.7 Competence Felt When Using the DST

A 2 x 2 x 2 (position x task load x automation) within-subjects ANOVA showed that ATCSs felt more competent using the DST when the traffic task load was low than high [$F(1,5) = 9.31$, $p < .01$, Figure 105, Table DD-13].

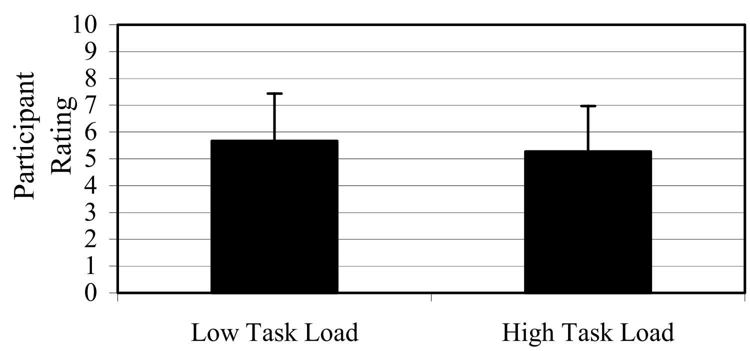

Figure 105. Competence felt using the DST by task load.

3.7.2.2 Discussion

ATCSs found full automation conditions most difficult. The ATCS participants did not have experience using the DST prior to participating in the study. They received several days of DST training and one practice day with the DST in the actual experimental environment. They did not have enough time to automate the necessary behaviors required to operate the tool, making it more difficult for them to use. In contrast, they were much more familiar working in an environment with flight strips, similar to the no-automation condition. The differences in familiarity between use of the flight strips and use of the DST would contribute to the differences in perceived difficulty. Further, the N=6 participant data set showed an effect for task load. ATCSs viewed the high task load conditions as more difficult than the low task load conditions.

With full automation present, ATCSs indicated that simulations were less realistic. The participants currently work in a setting similar to the no-automation condition. They use flight strips and do not have DSTs; therefore, it may not be surprising to find the condition with full automation and use of the DST rated as less realistic by them.

Even though ATCSs indicate that traffic levels in the field often surpass predicted levels for 2005, the low task load scenarios in this experiment were closer to normal traffic than high task load scenarios. The ATCSs rated the low task load scenarios as more representative and realistic of a typical workday. This finding seems congruent with what the ATCS actually experiences currently in the field. The ATCS participants came from level 10, 11, and 12 facilities even though we had requested only 11 and 12s. In the field environment, the volume of traffic a facility gets influences the rating level for that facility. As traffic volume increases, the rating level of the facility increases. Even at the highest-level facilities, a traffic "push" or peak in traffic volume may typically last 60 minutes and usually occurs several times a day (A. Nagy, personal communication, January 9, 2001). Even with a push in traffic, the amount of traffic fluctuates. We designed the scenarios with traffic levels that an ATCS from a level 11 or 12 facility would see; however, the traffic push did not fluctuate as much as it would in the field. The ATCS participants worked the scenario traffic for at least 30 minutes at a constant push. We designed the low traffic scenarios for 11 and 12 ATCSs, and even the low traffic was more than a level 10 ATCS would see in the field. ATCSs from level 12 facilities were capable of handling the scenario traffic levels with ease and indicated that they had expected more challenges. In contrast, ATCSs from level 10 facilities found the scenarios quite difficult and found it hard to keep up. Future traffic levels in the field will increase, and we built our traffic scenarios to reflect this. Currently, some facilities or certain sectors within facilities already experience these high volumes of traffic. Because the scenarios rotated under the various automation conditions, the effect of the task load level was solely responsible for the differences seen in the ratings of representative and realism and not from the effect of a particular scenario.

When asked about their performance, ATCSs indicated that they controlled traffic better when sitting on the R-side and when working low traffic. R-side ATCSs have direct and active control of the traffic, whereas the D-side ATCSs assist the R-side and do not directly deal with the planes. A D-side ATCS is in a more monitoring function, therefore feels less in control. This results in a perception of a lower control quality, although, as a team, the ATCSs may have controlled excellently. It is as if the D-side ATCSs feel that they can help more if they could only have more of an R-side function. In debriefings from the current and past studies, Certified Professional Controllers (CPCs) stated that they like to be in control of the traffic situation and work the R-side position relative to the D-side position. Our participant controllers are radar certified, therefore, they are used to doing R-side activities, and their mindset is focused on more radar tasks. When they move to the D-side position, they are still aware of what they could do as an R-side and have a sense of not doing as much as their capabilities. The caveat is that all of our ATCSs were radar certified. In addition, the low traffic task load scenarios allowed the ATCS time to gain the "picture" and maintain it.

ATCSs indicated that ATWIT hardly interfered. ATCSs felt that there was more interference when task load was high and even made more so with full automation. When examining the 6-participant data set, the 3-way interaction was significant and indicated that the task load x automation interaction was significant for R-side ATCSs but not for D-side ATCSs. When sitting on the R-side, the highest interference from the ATWIT device occurred when task load was high and full automation was present. R-side ATCSs may have perceived more interference from the ATWIT under these conditions because the D-side ATCS directed their attention to the use of the DST, which led to less focus on the usual D-side tasks of assisting the R-side. In the debriefings, the participants confirmed this and stated that they rarely looked at the radar display or assisted the R-side through the usual D-side responsibilities. Because the R-side ATCS did not have the usual assistance from the D-side ATCS, the R-side ATCS may have been more sensitive to the ATWIT device when task load was high. In addition, the ATWIT device and SAVANT occurred concurrently and resulted in a blind window of 15 seconds in which the ATCSs could not view the radar screen. Under high task load traffic, the effect of the "blind" window would be stronger; time away from the scope would be increasingly harder with increasing amounts of traffic.

In contrast to the ATWIT, ATCSs did not rate the oculometer as more interfering in one condition versus another. The ATCS wore the oculometer, which provided constant discomfort but did not require specific physical responses as the ATWIT did.

Even though both ATCSs used the DST, once at the R-side they put less confidence in it than on the D-side. The R-side ATCS may be unaware of the work the D-side was doing. ATCSs indicated in the debriefings that, when on the R-side, they did not pay attention to what the D-side was doing with the DST. The D-side ATCSs used the DST, and the R-side ATCS controlled traffic. The actual use of the DST most likely led to perceptions of it being more effective. The ATCSs felt more competent using the tool when the task load was low. We required that the ATCS participants needed to be novice DST users, therefore the low task load scenarios gave the participants more time and opportunity to work with the DST enabling them to feel like they had the chance to fully utilize it.

3.8 Subject Matter Expert Rating Forms

Sollenberger, Stein, and Gromelski (1997) developed and evaluated a method to assess ATCS performance. They designed a rating form to measure the effectiveness of new or enhanced ATC systems in simulation research. The rating form uses an 8-point format and a comment section for each of the questions. Sollenberger et al. showed that most of the rating scales were very reliable. The OTS ratings consisted of six categories: Maintaining Safe and Efficient Traffic Flow, Maintaining Attention and SA, Prioritizing, Providing Control Information, Technical Knowledge, and Communication related questions (Appendix I).

3.8.1 Data Analysis and Results

The SME completed one rating form for each team of ATCSs per condition. We only had one SME available for ratings, and he focused mainly on the R-side. We assumed that any changes in the team should be reflected in the R-side. We averaged the two observations per condition and conducted analyses on the eight teams. We analyzed each set of questions separately in 2 x 3 (task load x automation) within-subjects MANOVAs. We include the means, *SD*s, MANOVA, and ANOVA tables in Appendix EE. Here we only discuss the items that do not provide information for other measurement constructs.

3.8.1.1 Maintaining Safe and Efficient Traffic Flow

This category contained three items: maintaining separation and resolving potential conflicts, sequencing arrival and departure efficiently, and using control instructions effectively and efficiently. The MANOVA results showed a significant effect for task load across the set of DVs [$\Lambda = .16$, $F(3,5) = 8.72$, $p < .05$, Table EE-2, Table EE-3, Table EE-4, and Table EE-5]. We did not find any significant findings for automation. Using an adjusted alpha of .017, subsequent ANOVAs showed that ATCSs sequenced arrival and departure aircraft more efficiently in the low task load condition and maintained separation and resolved potential conflicts better under low task load conditions, although the latter finding did not reach statistical significance (i.e., it was a secondary trend; Figure 106).

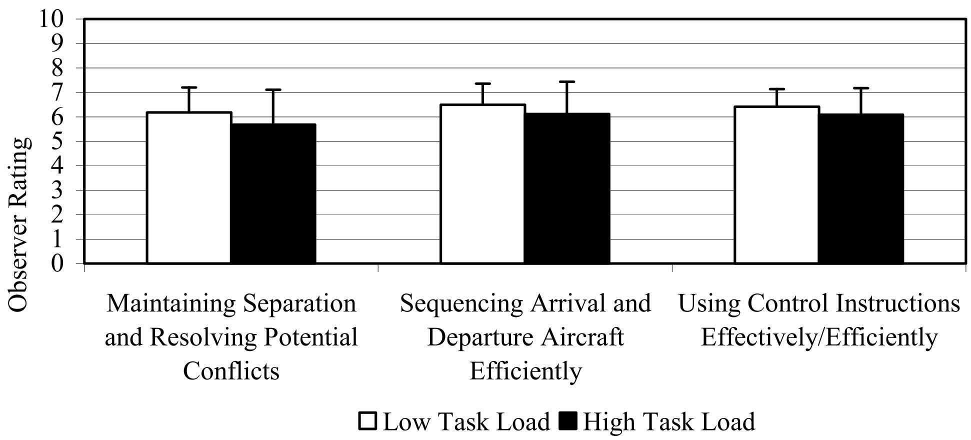

Figure 106. Maintaining safe and efficient traffic flow by task load.

3.8.1.1.1 Prioritizing

For the prioritizing items, taking actions in an appropriate order of importance, preplanning control actions, and handling control tasks for several aircraft, the MANOVA showed a significant effect of task load [$\Lambda = .17$, $F(3,5) = 7.96$, $p < .05$, Table EE-10]. The SME rated preplanning control actions higher in the low task load condition [$F(1,7) = 31.29$, $p < .001$, Figure 107, Table EE-12]. Using a more liberal alpha of .05, secondary trends indicated that ATCSs took actions in an appropriate order of importance and handled control tasks for several aircraft better under low task load. When no automation was present, the SME rated ATCSs lower for taking actions in an appropriate order of importance (Figure 108).

Because ATCSs used flight strips in only the no-automation condition, we used a one-way ANOVA to examine the effects of task load. Results indicated that task load did not affect the marking of flight strips.

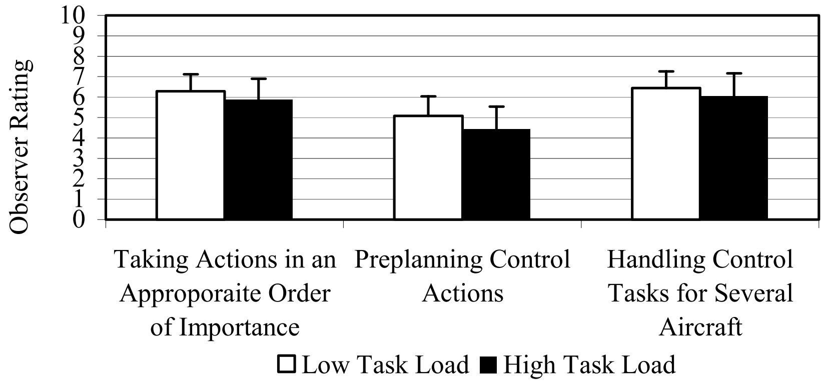

Figure 107. Prioritizing by Task Load.

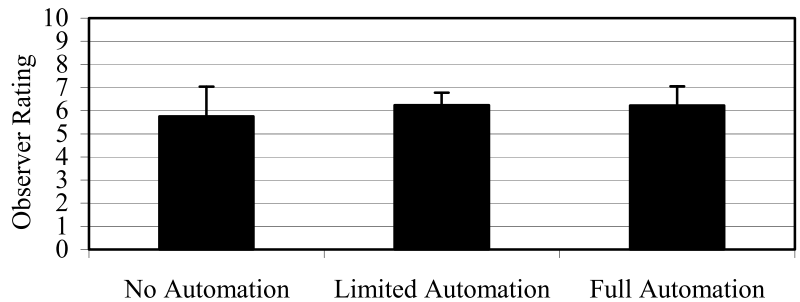

Figure 108. Taking action in an appropriate order of importance by automation.

3.8.1.1.2 Providing Control Information

The SME indicated that all items related to ATCSs' providing control information were lower under high task load conditions [Λ = .12, $F(3,5) = 11.89$, $p < .01$, Table EE-15]. ATCSs provided essential ATC information, additional ATC information, and coordination better when controlling traffic in the low task load scenarios [$F(1,7) = 49.00$, $p < .001$, $F(1,7) = 17.31$, $p < .01$, and $F(1,7) = 10.86$, $p < .05$, Table EE-16, Table EE-17, and Table EE-18, respectively, Figure 109]. A trend indicated that ATCSs received higher ratings for providing additional information when they were in the limited- or full- automation conditions relative to the no-automation condition (Figure 110).

3.8.1.1.3 Technical Knowledge

Showing knowledge of LOAs and SOPs, showing knowledge of aircraft capabilities and limitations, and showing effective use of equipment composed the technical knowledge items. The MANOVA showed a significant effect of task load [Λ = .14, $F(3,5) = 9.76$, $p < .05$, Table EE-19]. ATCSs demonstrated more effective use of the equipment when task load was low [$F(1,7) = 30.77$, $p < .001$, Figure 111, Table EE-22]. No other effects were significant.

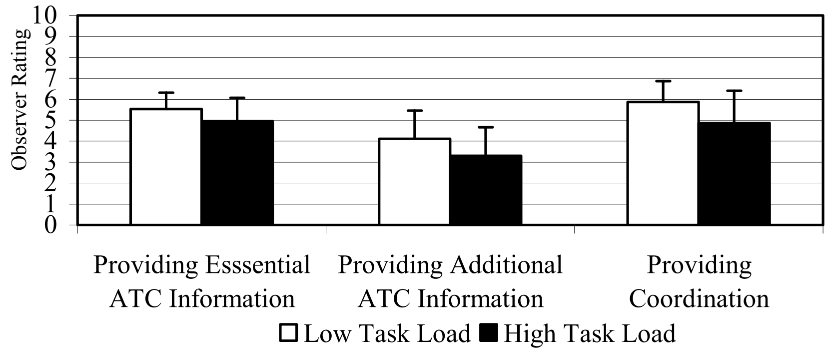

Figure 109. Providing control information by task load.

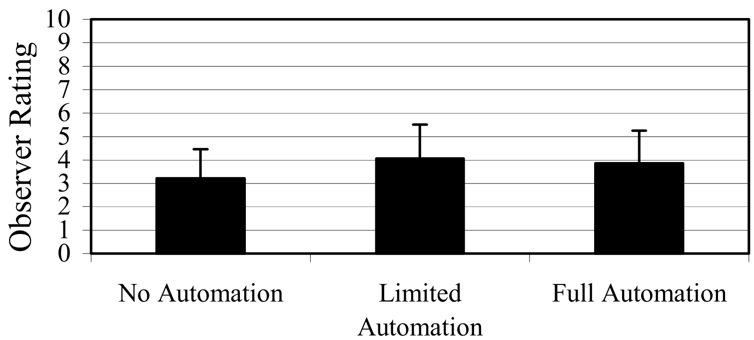

Figure 110. Providing additional ATC information by automation.

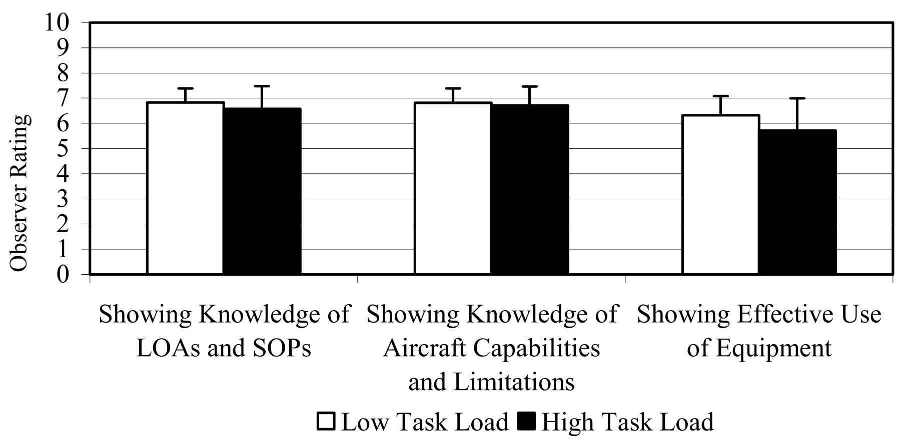

Figure 111. Technical knowledge by task load.

3.8.1.1.4 Communicating

We found no statistically significant effect for the experimental conditions on the communications items at the multivariate level. A primary trend showed that when no automation was present, an increase in task load led to a decrease in ratings for communicating clearly and efficiently. These effects diminished in the two automation conditions (Figure 112 Figure 113, respectively). Secondary trends indicated that the SME rated ATCSs higher for communicating clearly and efficiently and listening for pilot read backs and requests when they were in the full- or limited- automation conditions relative to the no-automation condition (Figure 114). Another secondary trend for using proper phraseology showed that an increase in task load led to a lower rating that was diminished in the two automation conditions (Figure 115 and Figure 116, respectively).

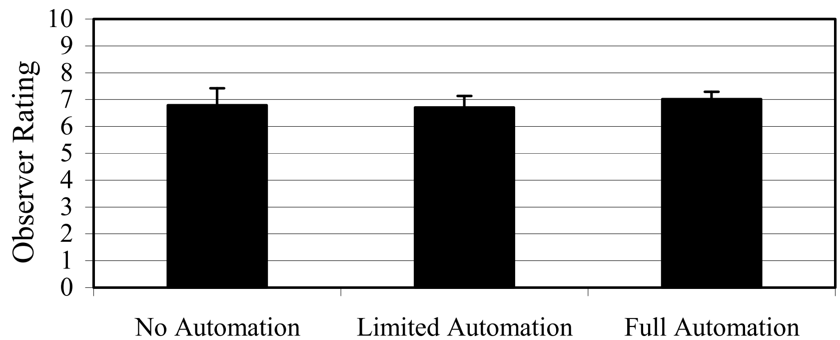

Figure 112. Communicating clearly and efficiently by low task load and automation.

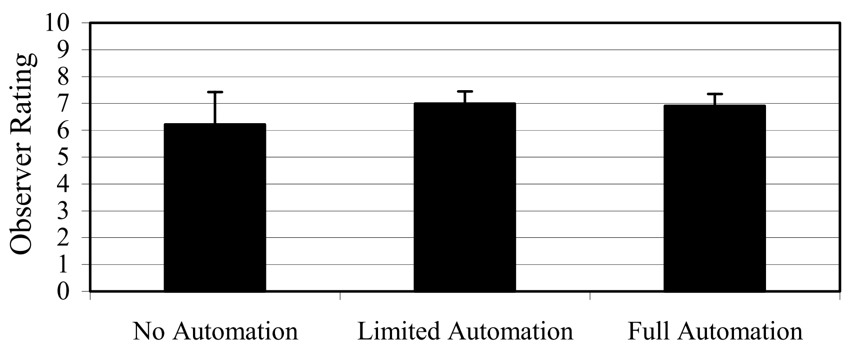

Figure 113. Communicating clearly and efficiently by high task load and automation.

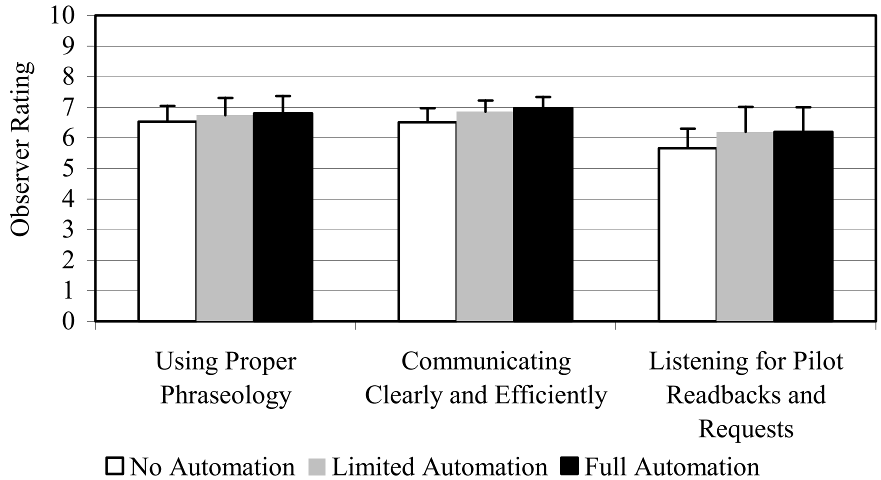

Figure 114. Communicating by automation.

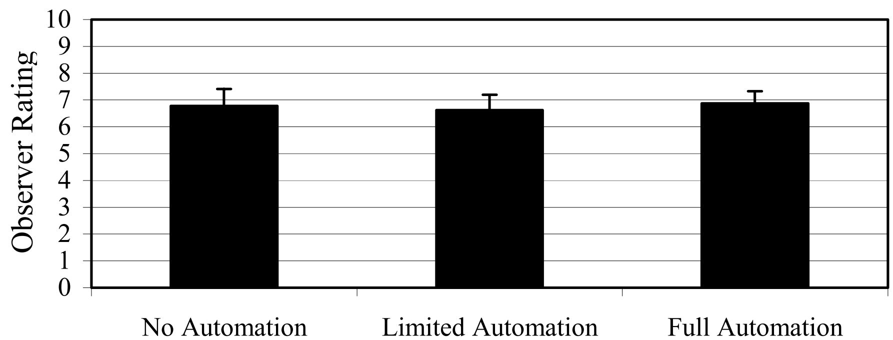

Figure 115. Using proper phraseology by low task load and automation.

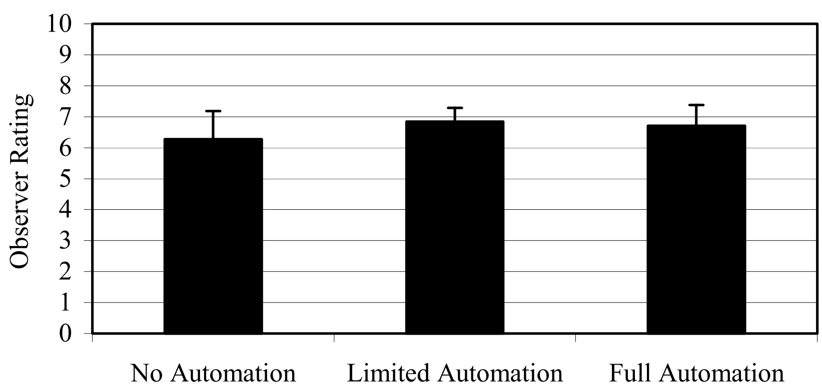

Figure 116. Using proper phraseology by high task load and automation.

3.8.1.2 Discussion

These results imply that with the assistance of automation, the D-side was able to free some cognitive resources of the R-side. This, in turn, allowed the ATCS to provide additional control information and communicate better. Other research groups have used cognitive engineering techniques to design future ATC working positions (Leroux, 1998). Although the DST used in this study does not explicitly assist in managing the ATCS's cognitive resources, it seemed to affect the distribution of the use of these resources within the ATCS team. The trends further indicated that automation led to the ATCSs gaining a better focus on the "big picture" and putting control items in the appropriate order of importance.

Although modest, the SME's ratings indicated better ATCSs performance along several control actions with full automation present. Please note that the automation was used by the D-side ATCS and may have impacted his or her performance; however, due to limited resources, we did not have the ATC SME focus primarily on the D-side for OTS ratings. The ratings presented in this section are primarily for the R-side ATCS.

The data clearly show that increasing task load decreased ATCSs' performance across the OTS rating form items. It would seem that an increase in task load may lead the ATCS to loose focus on the big picture. This leads the ATCS to implement less appropriate planning strategies and to communicate less effectively as he or she tries to gain the picture.

4. GENERAL DISCUSSION

For each of the constructs discussed in this report, we present a description, results, and a discussion of the implications of these results. In this general discussion, we place these results in a broader framework, presenting the effects of changes in automation, task load, and responsibilities (D- vs. R-side) on ATCSs. However, before we discuss the results in a general context, there are several spin-offs of this experiment that are worth mentioning here:

- Creation of a crosscutting team that involves test and evaluation, design and engineering, simulation, and human factors groups that speak a common language and work together to make things happen.
- Development of Generic Center Airspace on the Host Computer System (HCS) and the DST.
- Transfer of RDHFL human factors instruments to operational hardware and software.
- Development of a new SA assessment instrument.

The existing generic center airspace used for testing of new equipment and software is not suited for human-in-the-loop experiments that look at concepts from a human factors perspective. The airspace layout and naming convention makes it difficult if not impossible for ATCSs to learn. The ZCY airspace documentation that we were able to retrieve dated back to the early 1970s. We used the foundation of the ZCY adaptation, wiped the airspace clean, and built the Generic Center Airspace on top of it.

The creation of the airspace not only made it available to others on the HCS and the DST but now sets the standard for future studies (i.e., the Generic Center Airspace will form the foundation for future simulations). Although we can create new sectors, navigational aids, and so on, they will be additions to the Generic Center Airspace.

Over the years, we have developed a toolbox of human factors instruments at the RDHFL that we have used extensively on our in-house simulator platform. This experiment resulted in a transition of most if not all of these instruments to an environment with operational hardware and software. The successful use of these instruments in this experiment is a demonstration that studies that involve operational hardware and software can use our instruments as well. For operational test and evaluation, this means that human factors measures are available to evaluate new procedures and systems that we have tested in our laboratory settings.

The development of SAVANT created an SA assessment tool that takes the advantages of several existing instruments and ties them into a real-time environment while taking advantage of having the actual situational data available to probe participants. The implementation of SAVANT was a technical challenge as well; our programmers needed to create a remote display that was a real-time duplicate of the DSR display at the sector.

This experiment involved the operation of operational hardware and software, changes in NAS adaptations that require in-depth knowledge of the NAS architecture, programming across several platforms, and transferring data using multiple protocols, and, on top of all this, all systems needed to work in a controlled experimental fashion. The demonstration that this can work is an achievement in itself and shows the abilities of the multidisciplinary environment of the FAA William J. Hughes Technical Center.

4.1 Automation-Related Observations

We have seen in most of our data sets that with increased levels of automation, D-side ATCSs pulled away from the radar display. In the eye movement data, we have objective information on how much time the D-side ATCS spent looking at the radar display. The visual scanning information confirmed what other data sources had suggested, that is, D-side ATCSs spent less time looking at the radar display when automation levels increased. The fact that the D-side ATCS divided his or her visual attention between the radar display and the DST display suggests that they also spent a considerable amount of time in transition between these two displays. During this transition time, the D-side ATCS cannot absorb detailed information.

Recommendation: To provide the D-side ATCS with information about the current situation and to reduce the amount of time spent in transition between information displays, we should investigate if we can integrate data from automation tools into a radar display. This enhanced radar display would become the primary information display for the D-side ATCS.

In this study, we explained to our participants the purpose of the tool and emphasized the strategic nature of the tool. As a result, we observed that ATCSs took advantage of these capabilities and reached out to adjacent sectors. It is interesting to note that when we discussed this finding with ATCSs from ARTCCs that are already using a prototype DST, they did not agree with our findings. It seems that in the field, the tool is much less used in the strategic fashion for which we had trained our participants. Instead, ATCSs wait until aircraft are under control of the sector before taking action. As mentioned earlier, this may be the result of the timing of the assignment of a D-side ATCSs to a sector.

Recommendation: The balance in roles and responsibilities between ATC sector team members has evolved over many years. It may be more efficient to not disturb that balance and create a new position that is responsible for strategic planning.

When we provided limited automation, ATCS SA was lower than without or with full automation. The D-side ATCS without the graphical representation of the DST and without the familiar flight progress strips may have lost some SA under those conditions. The R-side ATCS in turn may have received normal D-side assistance during simulations without automation but less effective assistance under limited-automation conditions. This in turn led to a reduction in SA for the R-side ATCS. When full automation was available, although the R-side may have lost some of the D-side assistance on radar tasks, he or she gained assistance of the D-side on strategically resolving conflicts. The subjective ratings on SA for potential violations showed that ATCSs felt that with full automation, their SA was lower under high traffic task load conditions for which we also found an indication in the reduced score on the sector-based SAVANT questions. This finding shows that the benefit that the strategic solution provides under low traffic task load conditions no longer offsets the loss of D-side assistance on radar tasks. Within a sector team, a delicate balance between member roles and responsibilities exists.

The introduction of automation may seem to have a benefit under low traffic task load conditions but may disrupt that balance when ATCSs work under high traffic task load conditions.

Recommendation: When testing the benefits of automation tools, ATCSs need to be able to work in team configurations that reflect their normal environment. The traffic that forms the basis for these tests should include levels that one expects during a push of traffic.

Recommendation: The lower SA found for ATCSs using limited automation may indicate that a potential loss of SA may occur with the introduction of electronic flight strips. It may prove to be beneficial to examine the impact of electronic flight strips on ATCS SA within a controlled lab setting.

We found that under full automation and low task load, ATCSs had longer LL communications than under high task load conditions. This seems to reflect that the D-side ATCS took more time to communicate with adjacent sectors to try to solve potential problems strategically. Under high task load conditions, the D-side ATCS needed to revert to the traditional D-side responsibilities that do not include the use of the DST. Without the DST, less information about potential conflicts is available and communication with other sectors was less.

The participants in this study had no previous experience with the DST. During the first week, we trained the ATCS on the concepts and use of the DST. Although experts in their field, our participants were novice users of the DST. The ATCSs indicated that simulations with full automation present were less realistic. The situation in the field most resembles the simulation conditions that did not include the DST. It is therefore not surprising ATCSs felt that the full automation condition was least realistic and most difficult. The perceived difficulty may be because the use of the DST was new and therefore involved conscious effort, whereas other overlearned behaviors are automated and effortless. Besides increasing the perceived difficulty ratings for the D-side ATCS, this also may have led to an indirect increase in the perceived difficulty for the R-side ATCS; the D-side had to redirect some of the resources normally used to support the R-side, leaving the R-side ATCS with less assistance. The workload ratings and work by Bressolle et al. (2000) support the finding that the introduction of automation tools may lead to an increase in workload.

Recommendation: Additional research is needed into how long an increase in workload lasts for novice DST users and whether, with more experience, this effect diminishes. More information regarding the training needs of the ATCS for learning DSTs is needed.

In an ideal environment, the advanced automation tools will adapt to the amount of support an operator needs, returning some of the responsibilities to the operator when the task load permits. AA does not change the way the operator works but assists the operator with his or her tasks. The implementation of the DST used in the current study could not support the ATCS in an AA environment because it created tasks that were not compatible with the way ATCSs worked in the non-automated environment. In an environment without DSTs, ATCSs use flight progress strips to determine the potential for loss of separation. Therefore, AA should assist ATCSs in more efficient use of their current working methods. One way to do so may be to assist ATCSs in identifying potential conflicts. In the DST used in this study, the electronic equivalent of the flight progress strips identified the number of potential conflicts by aircraft but did not provide information on which aircraft these potential conflicts were with. ATCSs need to move to either a textual or graphical display to identify conflict pairs. Under high task load conditions, where the automation should assist ATCSs without interrupting their working habits, the flight progress equivalent did not seem to provide the information in a compatible format. In the current experiment, we saw that it pulled the D-side ATCS away from his or her role as a radar associate, leaving the R-side ATCS with less help than under conditions without the automation. The DST has all the information needed to assist the ATCSs but may be able to assist better by integrating the information into their working methods. We suggest using the same information provided by the DST and presenting it in a format that is more compatible with ATCS automated behaviors.

Recommendation: When implementing automation, designers should give special care to its compatibility with ATCS automated behaviors and working habits.

One of the observations we made during the simulations was that the D-side ATCS spent a lot of time merely separating data blocks (i.e., preventing data block information to overlap and become unreadable). Even though this may not be a very difficult task, when the introduction of automation pulled the D-side ATCS away from the normal D-side functions, the R-side ATCS now needed to separate data blocks him- or herself, leaving less time and fewer resources to focus on separating aircraft. ATCS interaction data showed that R-side ATCSs did not move more datablocks under full automation; however, the number of datablocks moved by D-side ATCSs decreased under the automation conditions. Unfortunately, ATCSs currently use leader line orientation and length as memory joggers. It seems that introducing automatic data block separation and an alternative to indication of the information now stored in the orientation and length of the leader line will greatly reduce the workload of the ATCSs.

Recommendation: Provide an automatic data block separation tool and ways to store information needed to remind ATCSs of planned or executed actions. This will free the D-side from that task and allows the D-side to concentrate on ATC-related matters without taking away assistance an R-side ATCS may need.

4.2 Task Load-Related Observations

After each of the scenarios that included automation, we asked ATCSs if they trusted the information that the tool provided. Under high task load conditions, when ATCSs needed assistance most, they indicated that they trusted that information the least. Wiener (1988) has found that operators may turn automation off under high workload. In our study, ATCSs may not have turned off the automation but may have disregarded the information and reverted to established operational procedures.

Recommendation: During training, we must make ATCSs aware that the behavior of the tool does not depend on the traffic task load. Training must make the ATCS feel so comfortable with the tool that the ATCS has realistic expectations of where the tool can assist and how much the ATCS can trust the provided information. Training must also focus on reducing ATCS workload while they become proficient using the DST. When a tool like a DST is in use, a change in formal roles and responsibilities may be necessary to prevent the D-side ATCS to revert to currently established procedures that prevent the use of the DST.

The ATCSs themselves and the OTS raters felt that an increase in task load decreased SA and performance in general. The objective SAGAT measure also showed decreases in SA, although this was not true for Level 2 SA. With more aircraft in the scenario, ATCSs may have needed more cognitive resources than available, leading to a loss of SA. Interestingly, we found that ATCSs did attempt to set up the high traffic scenarios to become more structured, potentially reducing the loss of SA. With an increase in traffic, ATCSs used more speed changes and less altitude changes. ATCSs often indicate that airspace with more climbing and descending aircraft is more difficult to control. When an aircraft makes a vertical transition, ATCSs seem to have more trouble predicting its behavior and risk. Under high traffic levels, ATCSs tried to reduce the level of uncertainty of the situation. In the current system, ATCSs do not have tools available to indicate or identify aircraft that pose a high risk. Therefore, ATCSs need to monitor all aircraft in the sector. It is likely that ATCSs have developed strategies that distribute their attention in such a way that high-risk aircraft receive most attention, but, under high traffic task load conditions, they seem to be less able to maintain the overall picture. Tools that distinguish between aircraft that have little risk of separation or sequencing problems vs. aircraft that are likely to get into situations that involve increased risk situations may be able to assist ATCSs to better focus their attention on those aircraft.

Recommendation: With the anticipated increase in traffic in the next decade, ATCSs will have increasing difficulty maintaining good SA. Developing tools that support ATCSs to focus their attention on aircraft that are at risk to get into potential conflicts or otherwise complex ATC situations may allow ATCSs to make most efficient use of their cognitive resources.

ATCSs indicted that low task load scenarios were closer to their normal traffic levels. In the field, a traffic push may typically last up to 1 hour and occurs several times per day. Even within these traffic pushes, the amount of traffic fluctuates. In this experiment, we designed the simulations to have an almost constant level of traffic because that was one of the variables we manipulated. The levels of traffic in our experiment may not reflect current conditions in the field. However, the results found for the high traffic load scenarios may give us insight in what to expect when future traffic levels further increase. Additional tools for spacing and sequencing may then create traffic pushes that are at a constant level and last for an hour or more. As one may expect, the air-ground communications increased in number and decreased in duration with an increase of task load. Even though the FAA provides ATCSs with tools to make conflict detection and resolution easier, the increase in traffic will still require ATCSs to verbally confirm the arrival of an aircraft in the sector and tell the aircraft to switch their radio to the next sector frequency. Without a solution to reduce the necessity to verbally acknowledge the entry and exit of aircraft, the capacity of the speech channel will form a major bottleneck in ATC.

Recommendation: Alternatives to verbal acknowledgement of aircraft entering and leaving the sector can alleviate the task load on the speech channel and free ATCS cognitive resources. Implementation of data functions that support these alternatives will likely increase ATCS capacity to handle more aircraft in the sector.

We found that ATCSs had very low levels of Level 3 SA (i.e., they were not able to project future situations). Although ATCSs in the current system work mostly tactically (i.e., within a timeframe of 3-5 minutes), we had expected that the use of the DST would extend this horizon, resulting in differences between the non-automated and automated conditions. One would expect that with more experience in the use of the DST, ATCSs would indeed extend their time-horizon and become more strategic, resulting in higher Level 3 SA. Although the data may be available on the use of the prototypes that are in operation in the field, there are no studies available that show this shift in ATCS SA.

Recommendation: Use existing operational data collected since the implementation of the prototypes in the field to determine if ATCSs have adopted a strategic style of control. If the field data are absent, we suggest creating an experiment that uses experienced DST users to determine if they have better Level 3 SA.

4.3 ATCS Position Related Observations

In the field, when an R- and a D-side work a sector, the D-side is responsible for coordination with adjacent sectors. In our simulations, coordination using the LL increased with an increase in task load. With full automation available, the D-side ATCS had the opportunity to review in more detail potential conflicts and coordinate control actions with adjacent sectors before aircraft entered the sector. The increase in duration of the LL communication events reflects that this happened in more detail when task load was low. Another explanation for the difference in LL communication durations is that during high task load scenarios, the R-side ATCS required full assistance of the D-side. For this experiment, we had not changed the roles and responsibilities of the D-side ATCS. Therefore, the primary duty of a D-side ATCS was to assist the R-side ATCS. This reflected established procedures in most ARTCCs. By pulling the D-side ATCS closer into the support of the tactical management of traffic in the sector, the D-side had less time to coordinate with adjacent sectors. With full automation present, the D-side ATCS still tried to resolve potential conflicts through coordination with adjacent sectors but did so more briefly.

Without providing a change in the role and responsibilities of the D-side ATCS when providing automation tools, the D-side will revert to established procedures. These established procedures will pull the D-side back into tactical control of aircraft, leaving the DST not used to its full potential when it would be able to assist ATCSs most. This is not to say that the R-side ATCS does not need the D-side assistance. Under the high task load conditions that we tested here, the R-side needed someone to separate data blocks and an extra pair of eyes.

In the field, traffic levels fluctuate and are relatively low before a push in traffic occurs. Therefore, one can bring in a D-side ATCS before an anticipated push in traffic. The D-side can then communicate longer with adjacent sectors and prepare for or resolve potential conflicts. One way of changing the role of the D-side may be to implement traffic task load indices that indicate transitions (e.g., relative low traffic task load) with a transition to a push expected in 30 minutes. During those 30 minutes, a D-side ATCS could focus on resolving potential conflicts and be more distant from tactical control. This could reduce the peak of complexity of the push. During the push itself, the D-side ATCS pulls back from a strategic into a more tactical mode and fully assists the R-side ATCS.

Recommendation: When a DST is present and one anticipates a push in traffic, bring a D-side ATCS in earlier, so that that position can make maximum use of the strategic capabilities of the DST.

This was the first study that examined eye movements in D-side ATCSs. Our results indicate that D-side ATCSs display a visual scanning behavior that is very different from R-side ATCSs. As one would expect, the D-side ATCS did not spend as much time looking at the radar display as the R-side. We also found that the D-side ATCSs focused more on FDBs than R-side ATCSs did in our previous studies. Unfortunately, we did not collect eye movement data on R- and D-side ATCSs simultaneously. Therefore, we had to make our inferences about how they may scan for information differently from results obtained from several studies.

Recommendation: When providing information relevant to D-side ATCSs, it may be most efficient to display this in or near the FDB.

Recommendation: To better understand how the scan for information differs between R- and D-side ATCSs, we should measure their eye movements simultaneously.

Although the D-side ATCS is an integral part of the sector team, our participants indicated that they felt they controlled traffic less well when working as the D-side ATCS. This likely reflects the fact that the D-side does not directly control traffic. Our participants were radar-certified, ATCSs. They felt they could have assisted better when at the radar display (i.e., more actively involved in controlling traffic). When working the D-side, our ATCS participants also indicated that high traffic task loads resulted in less of an increase in workload than when working the R-side. Our analyses of the ATCS interaction with the DSR interface clearly showed that R-side ATCSs performed the majority of the computer entries. This combination of observations reflects spare capacity within the sector team on the part of the D-side. When assisting the R-side, the D-side may be able to take on more but is not able to do so because of the division of roles and responsibilities. We confirmed this feeling of being less able to directly influence the traffic situation by objective data from SAGAT. SAGAT results indicated that for a given ATCS, their SA was lower when working on the D-side SA compared to working on the R-side.

Recommendation: More research in the roles and responsibilities and the workload of D-side ATCSs is necessary to determine if that position can be of more assistance during high traffic situations through a change in roles and responsibilities.

REFERENCES

Abbott, T. S., Nataupsky, M., & Steinmetz, G. G. (1987). *Effects of combining vertical and horizontal information into a primary flight display* (NASA Technical Paper 2783). Washington, DC: National Aeronautics and Space Administration.

Albright, C. A., & Lewandowsky, S. (1995). Momentum in a complex monitoring task. In R. S. Jensen, & L. A. Rakovan (Eds.), *Proceedings of the Eighth International Symposium on Aviation Psychology* (pp. 364-368). Columbus: The Ohio State University.

Albright, C. A., Truitt, T. R., Barile, A. L., Vortac, O. U., & Manning, C. A. (1994). Controlling traffic without flight progress strips: Compensation, workload, performance, and opinion. *Air Traffic Control Quarterly, 2,* 229-248.

Andre, A. D., Wickens, C. D., Moorman, L., & Boschelli, M. M. (1991). Display formatting techniques for improving SA in the aircraft cockpit. *The International Journal of Aviation Psychology, 1*, 205-218.

Applied Science Laboratories, Inc. (1991). 4100H Eye Tracking System with Head Mounted Optics Specifications [Computer hardware]. Bedford, MA: Applied Science Laboratories, Inc.

Bailey, L. L., Willems, B. F., & Peterson, L. M. (in review). ATC: *The effects of workload and decision support automation on En route R-side and D-side communications exchanges.*

Barber, B. (1983). *The logic and limits of trust.* New Brunswick, NJ: Rutgers University Press.

Becker, A. B., Warm, J. S., & Dember, W. N. (1991). Effects of feedback on perceived workload in vigilance performance. *In Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 1491-1494). Santa Monica, CA: Human Factors Society.

Billings, C. E. (1988). Toward human centered automation. In S. D. Norman & H. W. Orlady (Eds.), *Flight deck automation: Promises and realities* (pp. 167-190). Moffet Field, CA: NASA-Ames Research Center.

Billings, C. E. (1991). *Human-centered aircraft automation: A concept and guidelines* (NASA Technical Memorandum 103885). Moffet Field, CA: NASA Ames Research Center.

Bressolle, M. C., Benhacene, R., Boudes, N., & Pari, R. (2000, June). Advanced decision aids for Air Traffic Controllers: Understanding different working methods from a cognitive point of view. Paper presented at *3rd USA/Europe Air Traffic Management R & D Seminar*, Napoli, Italy.

Busquets, A. M., Parrish, R. V., Williams, S. P., & Nold, D. E. (1994). Comparison of pilots' acceptance and spatial awareness when using EFIS vs. pictorial display formats for complex, curved landing approaches. In R. D. Gilson, D. J. Garland, & J. M. Koonce (Eds.), *Situational awareness in complex systems* (pp. 139-167). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Carmody, D. P., Nodine, C. F., & Kundel, H. L. (1981). Finding lung nodules with and without comparative visual scanning. *Perception & Psychophysics, 29*(6), 594-598.

Carmody, M. A., & Gluckman, J. P. (1993). Task specific effects of automation and automation failure on performance, workload and situational awareness. In R. S. Jensen & D. Neumeister (Eds.), *Proceedings of the Seventh International Symposium on Aviation Psychology* (pp. 167-171). Columbus: Department of Aviation, The Ohio State University.

Crabtree, M. S., Marcelo, R. A. Q., McCoy, A. L., & Vidulich, M. A. (1993). An examination of a subjective situational awareness measure during training on a tactical operations simulator. In R. S. Jensen & D. Neumeister (Eds.), *Proceedings of the Seventh International Symposium on Aviation Psychology* (pp. 891-895). Columbus: The Ohio State University.

Danaher, J. W. (1980). Human error in Air Traffic Control system operations. *Human Factors, 22*(5), 535-545.

Dauphin Technology. (1999). Orasis [Computer hardware]. Platine, IL: Dauphin Technology.

DeGroot, A. D. (1965). *Thought and choice in chess* (2nd ed.). The Hague: Mouton.

Dittmar, M. L., Warm, J. S., Dember, W. N., & Ricks, D. F. (1993). Sex differences in vigilance performance and perceived workload. *The Journal of General Psychology, 120*(3), 309-322.

Durso, F. T., & Gronlund, S. D. (1999). Situation awareness. In F.T. Durso, R. S., Nickerson, R. W. Schvaneveldt, S. T. Dumais, D. S. Lindsay, & M. T. H Chi (Eds.), *Handbook of Applied Cognition*. John Wiley & Sons Ltd.

Durso, F. T., Gronlund, S. D., & Lewandowsky, S. (1993). *Should Hal open the pod bay doors? An argument for modular automation*. Unpublished manuscript, Oklahoma University.

Durso, F. T., Hackworth, C. A., Truitt, T. R., Crutchfield, J. M., Nikolic, D., & Manning, C. A. (1998). Situation awareness as a predictor of performance in en route air traffic controllers. *Quarterly Journal of Air Traffic Control, 6*(1)*,* 1-20.

Durso, F. T., Truitt, T. R., Hackworth, C. A., Crutchfield, J. M. (1998). En Route operational errors and situation awareness. *The International Journal of Aviation Psychology, 8*(2), 177-194.

Durso, F. T., Truitt, T. R., Hackworth, C. A., Crutchfield, J. M., Ohrt, D. D., Hamic, J. M., & Manning, C. A. (1995). Factors characterizing en route operational errors: Do they tell us anything about situation awareness? In D. J. Garland and M. R. Endsley (Eds.), *Experimental analysis and measurement of situation awareness* (pp. 189-195). Daytona Beach: Embry Riddle Aeronautical University Press.

Ellis, S. R. (1986). Statistical dependency in visual scanning. *Human Factors, 28,* 421-438.

Endsley, M. (1987). The application of human factors to the development of expert systems for advanced cockpits. In *Proceedings of the Human Factors Society 31st Annual Meeting* (pp. 1388-1392). Santa Monica, CA: Human Factors Society.

Endsley, M. R. (1988a). Design and evaluation for SA enhancement. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 97-101). Santa Monica, CA: Human Factors Society.

Endsley, M. R. (1988b). Situation Awareness Global Assessment Technique (SAGAT). In *Proceedings of the National Aerospace and Electronics Conference (NAECON)* (pp. 789-795). New York: IEEE.

Endsley, M. R. (1989a). *Final report: Situation Awareness in an advanced strategic mission* (NOR DOC 89-32). Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (1989b). *Tactical simulation test report: Addendum 1 Situation Awareness evaluations* (81203033R). Hawthorne, CA: Northrop Corporation.

Endsley, M. R. (1993). SA: A fundamental factor underlying the successful implementation of AI in the Air Traffic Control system. In D. J. Garland & J. A. Wise (Eds.), *Human Factors and Advanced Automation Technologies* (pp. 117-122). Daytona Beach, FL: Embry-Riddle Aeronautical University Press.

Endsley, M. R. (1996). Automation and SA. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 163-181). Mahwah, NJ: Lawrence Erlbaum.

Endsley, M. R. (1999). Situation awareness in aviation systems. In D. J. Garland, J. A. Wise, & V. D. Hopkin (Eds.), *Handbook of Aviation Human Factors* (pp. 257-276). Mahwah, NJ: Lawrence Erlbaum.

Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot Situation Awareness. *International Journal of Aviation Psychology, 4*(3), 241-264.

Endsley, M. R., & Kaber, D. B. (1996). *Level of automation effects on performance, SA and workload in a dynamic control task.* Lubbock, TX: Texas Tech University.

Endsley, M. R., & Kaber, D. B. (1999). Level of automation effects on performance, SA, and workload in a dynamic control task. *Ergonomics, 42*(3), 462-492.

Endsley, M. R., & Kiris, E. O. (1995). The out-of-the-loop performance problem and level of control in automation. *Human Factors, 37*(2), 381-394.

Endsley, M. R., & Rodgers, M. D. (1998). Distribution of attention, SA, and workload in a passive Air Traffic Control task: Implications for operational errors and automation. *ATC Quarterly, 6*(1), 21-44.

Endsley, M. R., & Selcon, S. J. (1997). Designing to aid decisions through SA enhancement. In *Proceedings of the 2nd Symposium on SA in Tactical Aircraft* (pp. 107-112). Patuxent River, MD: Naval Air Warfare Center.

Endsley, M. R., Selcon, S. J., Hardiman, T. D., & Croft, D. G. (1998). A comparative evaluation of SAGAT and SART for evaluations of Situation Awareness. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 82-86). Santa Monica, CA: Human Factors and Ergonomics Society.

Endsley, M. R., & Smolensky, M. W. (1998). Situation awareness in Air Traffic Control: The picture. In M. W. Smolensky & E. S. Stein (Eds.), *Human Factors in Air Traffic Control* (pp. 115-154). San Diego, CA: Academic Press.

Endsley, M. R., Sollenberger, R., Nakata, A., Hough, D., & Stein, E. (2000). *Situation awareness in Air Traffic Control: Enhanced displays for advanced operations* (DOT/FAA/CT-TN00/01). Atlantic City International Airport, NJ: William J. Hughes Technical Center.

Federal Aviation Administration. (1998). Air Traffic Control (DOT/FAA/Order 7110.65L). Washington, DC: Author.

Federal Aviation Administration. (1999). *FFP1 URET deployment* [On-line]. Available from: www.ffp1.faa.gov/uret_Report.htm

Finkelman, J. M., & Kirschner, C. (1980). An information-processing interpretation of Air Traffic Control stress. *Human Factors, 22*(5), 561-567.

Fracker, M. L. (1989). Attention allocation in Situation Awareness. In *Proceedings of the Human Factors Society 33rd Annual Meeting* (pp. 1396-1400). Santa Monica, CA: Human Factors Society.

Fracker, M. L. (1990). Attention gradients in Situation Awareness. *In Situational Awareness in Aerospace Operations (AGARD-CP-478) (Conference Proceedings #478)* (pp. 6/1-6/10). Neuilly Sur Seine, France: NATO – AGARD.

Galinsky, T. L., Rosa, R. R., Warm, J. S., & Dember, W. N. (1993). Psychophysical determinants of stress in sustained attention. *Human Factors, 35*(4), 603-614.

Garland, D. J., Stein, E. S., Blanchard, J. W., & Wise, J. A. (1992). Situational awareness in the future Air Traffic Control environment. Paper presented at the *37th Air Traffic Control Association Conference,* Atlantic City International Airport, NJ.

Gluckman, J. P., Warm, J. S., Dember, W. N., & Rosa, R. R. (1993). Demand transitions and sustained attention. *Journal of General Psychology, 120*(3), 323-337.

Gosling, G. D. (1992, June). Artificial intelligence in Air Traffic Control. Paper presented at the *NASA/FAA Workshop on Artificial Intelligence and Human Factors in Air Traffic Control and Aviation Maintenance*, Daytona Beach, FL.

Gronlund, S. D. (1992). Will changing access to flight progress strip information affect cognition performance? An empirical answer. In *Proceedings of the NASA/FAA Workshop on Artificial Intelligence and Human Factors in Air Traffic Control and Aviation Maintenance: Book of Abstracts* (pp. 11). Daytona Beach, FL: Embry-Riddle University.

Gross, J. R. (1998). *Effect of controller participation in air-to-ground and ground-to-air radio communications on radar screen data block fixation durations.* Unpublished manuscript, Embry-Riddle Aeronautical University.

Guttman, J.A., Stein, E. S., & Gromelski, S. (1995). *The influence of generic airspace on Air Traffic Controller performance.* (DOT/FAA/CT-TN95/38). Atlantic City International Airport, NJ: Federal Aviation Administration Technical Center.

Hansman, R. J., Wanke, C., Kuchar, J., Mykityshyn, M., Hahn, E., & Midkiff, A. (1992, September). Hazard alerting and situational awareness in advanced air transport cockpits. Paper presented at the *18th ICAS Congress,* Beijing, China.

Harris, W. C., Goernert, P. N., Hancock, P. A., & Arthur, E. (1994). The comparative effectiveness of adaptive automation and operator initiated automation during anticipated and unanticipated taskload increases. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: Current research and trends* (pp. 40-44). Hillsdale, NJ: LEA.

Hart, S. G., & Sheridan, T. B. (1984). Pilot workload, performance, and aircraft control automation. In *Human factors considerations in high performance aircraft* (AGARD-CP-371) (pp. 18/1-18/12). Neuilly Sur Seine, France: NATO-AGARD.

Hart, S.G., & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam: North-Holland.

Helbing, H. (1997). A preliminary cognitive model of en-route Air Traffic Control. In R. S. Jensen, & L. Rakovan (Eds.), *Proceedings of the Ninth International Symposium on Aviation Psychology* (pp. 391-396). Columbus: The Ohio State University.

Hopkin, V. D. (1991, January). Automated flight strip usage: Lessons learned from the function of paper strips. Paper presented at the *NASA/FAA Conference on Challenges in Aviation Human Factors: The National Plan*, Washington, DC.

Hopkin, V. D. (1998). The impact of automation on Air Traffic Control Specialist. In M. W. Smolensky & E. S. Stein (Eds.), *Human factors in Air Traffic Control* (pp. 391-419). San Diego, CA: Academic Press.

Human Technology, Inc. (1991). *Analysis of controller communication in en route air traffic control.* Report to the Federal Aviation Administration. McLean, VA.

Ignizio, J. P. (1991). *An introduction to expert systems.* New York: McGraw-Hill.

Kaber, D. B. (1996). *The effect of level of automation and adaptive automation on performance in dynamic control environments.* Unpublished doctoral dissertation, Texas Tech University, Lubbock.

Kaber, D. B., & Endsley, M. R. (1997). The combined effect of level of automation and adaptive automation on human performance with complex, dynamic control systems. In *Proceedings of the Human Factors and Ergonomics Society 41st Annual Meeting* (pp. 205-209). Santa Monica, CA: Human Factors and Ergonomics Society.

Kanki, B. G., & Prinzo, O. V. (Eds.) (1996). *Methods and metrics of voice communications*. Department of Transportation, Federal Aviation Administration, Office of Aviation Medicine, Washington, DC.

Kibbe, M. (1988). Information transfer from intelligent EW displays. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (pp. 107-110). Santa Monica, CA: Human Factors Society.

Kinney, G. C., Spahn, M. J., & Amato, R. A. (1977). *The human element in Air Traffic Control: Observations and analyses of the performance of controllers and supervisors in providing Air Traffic Control separation services* (MTR-7655). McLean, VA: METREK Division, MITRE Corporation.

Kirk, D. B., Heagy, W. S., & Yablonski, M. J. (2000). *Problem resolution: Support for free flight operations*. (MITRE No. 02001204-21). McLean, VA: Center for Advanced Aviation System Development.

Klein, G. (1999). Sources of power: How people make decisions. Cambridge: The MIT Press.

Lee, J., & Moray, N. (1992). Trust, control strategies and allocation of function in human-machine systems. *Ergonomics*, 35(10), 1243-1270.

Leroux, M. (1998). *In Depth Evaluation of ERATO Tools: Towards an Operational Prototype*. 2nd USA/Europe Air Traffic Management R&D Seminar.

Marshak, W. P., Kuperman, G., Ramsey, E. G., & Wilson, D. (1987). Situational awareness in map displays. In *Proceedings of the Human Factors Society 31st Annual Meeting* (pp. 533-535). Santa Monica, CA: Human Factors Society.

Moffa, A. J. (1994). *Trust in a Medical Expert System: An Exploratory Study.* (Master's Thesis). Melbourne: Florida Institute of Technology.

Moffa, A. J., & Stokes, A. F. (1997). Trust in a medical expert system: Can we generalize between domains? In M. Mouloua, & J. M. Koonce, (Eds.), *Human-Automation interaction: Research and practice.* Mahwah, NJ: Lawrence Erlbaum Associates.

Mogford, R. H., & Tansley, B. W. (1991). The importance of the Air Traffic Controller's mental model. Paper presented at the *Human Factors Society of Canada Annual Meeting*, Canada.

Moray, J. (1986). Monitoring behavior and supervisory control. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance: Volume I, Cognitive processes and performance* (pp. 43-1 to 43-32). New York: Wiley-Interscience.

Moray, N., & Rotenberg, I. (1989). Fault management in process control: Eye movements and action. Special Issue: Current methods in cognitive ergonomics. *Ergonomics, 32,* 1319-1342.

Morrison, R., & Wright, R. H. (1989). ATC control and communications problems: An overview of recent ASRS data. In R. S Jensen (Ed.), *Proceedings of the Fifth International Symposium on Aviation Psychology* (pp. 901-907). Columbus: Ohio State.

Mosier, K. L., & Chidester, T. R. (1991). Situation assessment and SA in a team setting. In Y. Queinnec & F. Daniellou (Eds.), *Designing for everyone* (pp. 798-800). London: Taylor and Francis.

Mosier, K. L., & Skitka, L. J. (1996). Human decision makers and automated decision aids: Made for each other? In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 201-220). Mahwah, NJ: Lawrence Erlbaum Associates.

Muir, B. M. (1988). Trust between humans and machines, and the design of decision aids. In E. Hollnagel, G. Mancini, & D.D. Woods, (Eds.), *Cognitive engineering in complex dynamic worlds* (pp. 71-83). London: Academic Press.

National Instrument Corporation. (2000). Labview (Version 5.1) [Visual programming], Austin, TX: National Instrument Corporation.

National Transportation Safety Board. (1990). *Aircraft accidents report: USAIR, Inc., Boeing 737-400, LaGuardia Airport, Flushing New York, September 20, 1989* (NTSB/AAR-90-03). Washington, DC: Author.

Nisbett, R. E., & Wilson, T. D. (1977). Telling more than we can know: Verbal reports on mental processes. *Psychological Review, 84*(3), 231-259.

Noldus Information Technology. (2000). Observer Video-Pro, (version 4.0) [Computer software]. Wageningen, The Netherlands: Noldus Information Technology. BV.

Norman, D. A. (1989). *The problem of automation: Inappropriate feedback and interaction not overautomation* (ICS Report 8904). La Jolla: Institute for Cognitive Science, U. C. San Diego.

Orthogon. (1999). ODS Toolbox (Version 5.1) [Computer software]. Langen, Germany: Orthogon.

Office of Air Traffic System Development. (1997). *AUA program master plan (Volume 2): Automation strategic planning*: Washington DC: Federal Aviation Administration.

Ohnemus, K., & Biers, D. (1993). Retrospective versus concurrent thinking-out-loud in usability testing. In *Proceedings of the 37th Annual Meeting of the Human Factors Society* (pp. 1127-1131). Santa Monica, CA: Human Factors and Ergonomics Society.

Parasuraman, R. (1992, June). Adaptive function allocation effects on pilot performance. Paper presented at the *NASA/FAA Workshop on Artificial Intelligence and Human Factors in Air Traffic Control and Aviation Maintenance*, Daytona Beach, FL.

Parasuraman, R., Bahri, T., Deaton, J. E., Morrison, J. G., & Barnes, M. (1992). *Theory and design of adaptive automation in aviation systems* (AWCADWAR-92033-60). Warminster, PA: Naval Air Warfare Center, Aircraft Division.

Parasuraman, R., Molloy, R., & Singh, I. L. (1993). Performance consequences of automation-induced complacency. *International Journal of Aviation Psychology, 3*(1), 1-23.

Peterson, L. M., Bailey, L. L., & Willems, B. (in press). *Controller-to-controller communication and coordination taxonomy (C^4T)*. Washington, DC: Department of Transportation, Federal Aviation Administration, Office of Aviation Medicine,.

Pritchett, A. R. (1997). *Pilot non-conformance to alerting system commands during closely spaced parallel approaches.* Unpublished Doctoral Dissertation, Massachusetts Institute of Technology, Cambridge.

Reid, G. B., & Nygren, T. E. (1988). The subjective workload assessment technique: A scaling procedure for measuring mental workload. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 185-218). Amsterdam: North-Holland.

Rempel, J. K., Holmes, J. G., & Zanna, M. P. (1985). Trust in close relationships. *Journal of Personality and Social Psychology, 49,* 95-112.

Riley, V. (1994). A theory of operator reliance on automation. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: Current research and trends* (pp. 8-14). Hillsdale, NJ: Lawrence Erlbaum Associates.

Rogers, W. (1994). The effect of an automated fault management decision aid on pilot Situation Awareness. Paper presented at the *First Automation and Human Performance Conference*, Washington, DC.

Rötting, M. (1999). *Blickbewegungsregistrierung und mentale Beanspruchung* [Online]. Available: http://wwwifa.kf.tu-berlin.de/FORSCHUNG/blickbewegungen.html.

Sarter, N. B., & Woods, D. D. (1991). Situation Awareness: A critical but ill-defined phenomenon. *The International Journal of Aviation Psychology, 1*(1), 45-57.

Scerbo, M. W. (1996). Theoretical perspectives on adaptive automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 37-63). Mahwah, NJ: Lawrence Erlbaum Associates.

Scerbo, M. W., Greenwald, C. Q., & Sawin, D. A. (1993). The effect of subject-controlled pacing and task type on sustained attention and subjective workload. *The Journal of General Psychology, 120*(3), 293-307.

Schroeder, D. J., & Nye, L. G. (1993). An examination of the workload conditions associated with operational errors/deviations at air route traffic control centers. In M. D. Rodgers (Ed.), *An analysis of the operational error database for Air Route Traffic Control Centers* (DOT/FAA/AM-93/22). Oklahoma City: Human Factors Research Laboratory, Civil Aeromedical Institute, Federal Aviation Administration.

Selcon, S. J. (1990). Decision support in the cockpit: Probably a good thing? In *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 46-50). Santa Monica, CA: Human Factors Society.

Selcon, S. J., & Taylor, R. M. (1990). Evaluation of the situational awareness rating technique (SART) as a tool for aircrew systems design. In *Situational Awareness in Aerospace Operations* (AGARD-CP-478) (pp. 5/1 –5/8). Neuilly Sur Seine, France: NATO – AGARD.

Selcon, S. J., Taylor, R. M., & Koritsas, E. (1991). Workload or situational awareness? TLX vs SART for aerospace systems design evaluation. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 62-66). Santa Monica, CA: Human Factors Society.

Smolensky, M. W. (1993). Toward the physiological measurement of Situation Awareness: The case for eye movement measurements. In *Proceedings of the Human Factors and Ergonomics Society 37th Annual Meeting* (pp. 41). Santa Monica, CA: Human Factors and Ergonomics Society.

Sollenberger, R. L., & Stein, E. S. (1995). *The effects of structured arrival and departure procedures on TRACON Air Traffic Controller memory and situational awareness* (DOT/FAA/CT-TN95/27). Atlantic City International Airport, NJ: Federal Aviation Administration Technical Center.

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

Standard Terminal Automation Replacement System Human Factors Team. (1997). *Standard terminal automation replacement system human factors review* (Internal Report to FAA Chief Scientific and Technical Advisor for Human Factors).

StatSoft, Inc. (2001). STATISTICA for windows (Computer program manual). Tulsa, OK: StatSoft, Inc.

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

Stein, E. S. (1992). *Air Traffic Control visual scanning* (DOT/FAA/CT-TN92/16). Atlantic City International Airport, NJ: Federal Aviation Administration Technical Center.

Stokes, A. F., & Kite, K. (1994). *Flight stress: Stress, fatigue, and performance in aviation.* Aldershot, UK: Ashgate.

Sullivan, C., & Blackman, H. S. (1991). Insights into pilot SA using verbal protocol analysis. In *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 57-61). Santa Monica, CA: Human Factors Society.

Taylor, R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In *Situational Awareness in aerospace operations* (AGARD-CP-478) (pp. 3/1 – 3/17). Neuilly Sur Seine, France: NATO – AGARD.

Tenney, Y. T., Adams, M. J., Pew, R. W., Huggins, A. W. F., & Rogers, W. H. (1992). *A principled approach to the measurement of SA in commercial aviation* (NASA Contractor Report 4451). Langely, VA: NASA Langely Research Center.

UFA, Inc. (1998). ATCoach [Simulation software]. Woburn, MA: UFA, Inc.

Vidulich, M. (in press). Testing the sensitivity of SA metrics in interface evaluations. In M. R. Endsley & D. J. Garland (Eds.), *SA analysis and measurement.* Mahwah, NJ: Lawrence Erlbaum.

Vidulich, M. A., & Hughes, E. R. (1991). Testing a subjective metric of Situation Awareness. *In Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 1307-1311). Santa Monica, CA: Human Factors Society.

Viviani, P. (1990). Eye movements in visual search: Cognitive perceptual and motor control aspects. In E. Kowler (Ed.), *Eye movements and their role in visual and cognitive processes,* (pp. 353-393). Amsterdam, The Netherlands: Elsevier Science Publishers BV.

Waterman, D. A. (1986). *A guide to expert systems.* Reading, MA: Addison-Wesley.

Wickens, C. D. (1992). *Engineering psychology and human performance* (2nd ed.). New York: Harper Collins.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed.). Upper Saddle River, NJ: Prentice-Hall Inc.

Wickens, C. D., & Kessel, C. (1979). The effect of participatory mode and task workload on the detection of dynamic system failures. *IEEE Transactions on Systems, Man and Cybernetics, SMC-9*(1), 24-34.

Wickens, C. D., & Kessel, C. (1980). The processing resource demands of failure detection in dynamic systems. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 564-577.

Wiener, E. (1988). Field studies in automation. In S. D. Norman & H. W. Orlady (Eds.), *Flight deck automation: Promises and realities* (pp. 37-55). Moffett Field, CA: NASA-Ames Research Center.

Wiener, E. L. (1985). Cockpit automation: In need of a philosophy. In *Proceedings of the 1985 Behavioral Engineering Conference* (pp. 369-375). Warrendale, PA: Society of Automotive Engineers.

Wiener, E. L. (1992, June). The impact of automation on aviation human factors. Paper presented at the *NASA/FAA Workshop on Artificial Intelligences and Human Factors in Air Traffic Control and Aviation Maintenance*, Daytona Beach, FL.

Wiener, E. L., & Curry, R. E. (1980). Flight deck automation: Promises and problems. *Ergonomics, 23*(10), 995-1011.

Wierwille, W., & Eggemeier, F. (1993). Recommendations for mental workload measurement in a test and evaluation environment. *Human Factors, 35,* 236-282.

Willems, B. F., Allen, R. C., & Stein, E. S. (1999). *Air Traffic Controller visual scanning II. Task load, visual noise, and intrusions into controlled airspace* (DOT/FAA/CT-TN99/23). Atlantic City International Airport, NJ: William J. Hughes Technical Center.

Willems, B. F., & Truitt, T. R. (1999). *Implications of reduced involvement in En Route Air Traffic Control* (DOT/FAA/CT-TN99/22). Atlantic City International Airport, NJ: William J. Hughes Technical Center.

Young, L. R. A. (1969). On adaptive manual control. *Ergonomics, 12*(4), 635-657.

Zuboff, S. (1988). *In the age of smart machines: The future of work and power.* New York: Basic Books.

Zweben, M. (1992, June). Artificial intelligence in aerospace systems: An overview of AI research at NASA Ames. Paper presented at the *NASA/FAA Workshop on Artificial Intelligence and Human Factors in Air Traffic Control and Aviation Maintenance*, Daytona Beach, FL.

ACRONYMS

AA	Adaptive Automation
ANOVA	Analysis of Variance
ARTCC	Air Route Traffic Control Center
ATC	Air Traffic Control
ATCS	Air Traffic Control Specialist
ATWIT	Air Traffic Workload Input Technique
C^4T	Controller-to-Controller Communication/Coordination Taxonomy
CA	Conflict Alert
CAMI	Civil Aeromedical Institute
CPC	Certified Professional Controller
CRD	Computer Readout Device
D2	Direct To
DART	Data Analysis and Reduction Tool
D-side	Data-side
DRA	Data Reduction and Analysis
DRAT	Data Reduction and Analysis Tool
DSR	Display System Replacement
DST	Decision Support Tool
DV	Dependent Variable
EDA	En Route Descent Advisor
FAA	Federal Aviation Administration
FDB	Full Data Block
FL	Flight Level
FPL	Full Performance Level
HCS	Host Computer System
HFE	Human Factors Engineer
I^2F	Integration and Interoperability Facility
IV	Independent Variable
LL	Landline
LOA	Letter of Agreement
MANOVA	Multivariate Analysis of Variance
MSAW	Minimum Safe Altitude Warning
NAS	National Airspace System
NASA	National Aeronautics and Space Administration
OE	Operational Error
OTS	Over-the-Shoulder
PARR	Problem Analysis, Resolution, and Ranking
POG	Point-of-Gaze
PSQ	Post-Scenario Questionnaire
PTT	Push-to-Talk
R	Radar-side Communication
R-side	Radar-side
RADAR	Radio Detecting and Ranging

RDHFL	Research Development and Human Factors Laboratory
RT	Response Time
SA	Situation Awareness
SAGAT	SA Global Assessment Technique
SAR	System Analysis Recording
SART	Situational Awareness Rating Technique
SAVANT	SA Verification ANalysis Tool
SD	Standard Deviation
SME	Subject Matter Expert
SOP	Standard Operating Procedure
SPAM	Situation Presence Assessment Method
SWORD	Subjective Workload Dominance
TGF	Target Generation Facility
TLX	Task Load Index
TMA	Traffic Management Advisor
URET	User Request Evaluation Tool
VOR	Very High Frequency Omnidirectional Range
VSCS	Voice Switching and Communication System

APPENDIX A

Participant Recruitment Form

Air Traffic Control Specialist
Decision Support Automation Research Study

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

In this study, we will measure air traffic control performance and behavior while using different levels of automation under two levels of traffic. This is fifth in a series of studies done at the Federal Aviation Administration (FAA) William J. Hughes Technical Center (WJHTC) Research Development and Human Factors Laboratory (RDHFL) to measure eye movements.

Improving the current air traffic system with new, automated technologies is necessary to reduce or manage current controller workloads, while accommodating continued growth in air traffic. We will investigate how task load and automation affect controller performance and behavior in the en route airspace.

The study will take place in a lab area of the Integration and Interoperability Facility (I^2F). We will use an integrated FAA Target Generation Facility (TGF), Host, and Display System Replacement (DSR). The controller environment will include a full DSR workstation with all operational functions, and an Air Traffic Workload Input Technique (ATWIT) device. We will make use of a generic en route ARTCC with IFR in effect.

The Human Factors Laboratory is looking for 16 volunteer controllers from continental US ARTCC facilities. The time requirement for this experiment is two weeks. You will travel on Monday of the first week and Friday of the second week. From Tuesday through Thursday of the first week, you will participate in training in a classroom setting to learn how to use the decision support tool. On Friday of the first week, you will engage in a series of simulations to learn the simulated airspace. On Monday of the second week, you will train on simulations on operational hardware and software. Finally, from Tuesday through Thursday of the second week you will participate in 13 experimental simulations.

The simulations will duplicate operational air traffic conditions. You will interact with simulation pilots and control simulated air traffic, as you would normally do in the field. The simulations will have two levels of difficulty; moderate and high traffic load. During the simulations, you will either control traffic as you would currently do on a DSR system, or use two levels of automation. The first level will require only limited flight strip marking and will indicate potential conflicts up to 20 minutes in advance. The second level will allow the D-side controller to create trial flight plans and send flight plan amendments directly to the Host.

During half of the simulations, the eye movements will be monitored. A small camera mounted on a headband will monitor eye movements. An invisible beam of infrared light will illuminate the eye. The device that monitors the eye movements may cause some discomfort. The skin area under the headband that supports the device may show some redness after wearing the device for the duration of a simulation. The intensity of the infrared beam that illuminates the eye is about one thirtieth of the intensity expected while walking outside on a sunny day and should not cause any discomfort or risk to your health.

During the simulations of the second week, we will record video images and audio. These recordings are for internal use only. Two researchers will be present to code the communications between the R- and D-side controllers. During these simulations, two observers will conduct over-the-shoulder ratings –one for the D- and one for the R-side controller.

Your only direct benefit is your opportunity to participate and all your expenses are paid under the Federal Travel regulations. The benefit for air traffic controllers derived from the results of this experiment may include a better understanding of why operational errors occur and a better understanding of how controllers reach out for information.

The records of this study are strictly confidential, and you will not be identifiable by name or description in any reports or publications about this study. Photographs and audio and video recordings will be made during the study. They are for use within the Research and Development Human Factors Laboratory only. Your data will be collected by code number and no permanent record of your name will be maintained.

Schedule

Two controllers will visit the laboratory each week beginning on or about February 7, 2000 until April 21, 2000. Participants will travel on Monday and report at 8:00 a.m. on Tuesday for training. The training sessions will conclude on Monday afternoon of the second week. Several 30-minute break periods will be taken each day and a one-hour lunch period will be taken around midday. The simulation will stand down during the weeks of February 21st for Federal Holidays.

Research psychologists and an SATCS, who will observe the participants in the control room, will conduct the simulation. A voice communication link to another room will allow controllers to issue commands to trained simulation pilots. Two participants will operate an R- and D-side position. During the simulations, participants will alternate between R- and D-positions. Headsets will be provided, however, participants may bring their own headset for use during the experiment.

Simulation Procedures

A generic ARTCC sector will be used. The sector is designed for easy learning in a short period of time. The sector consists of easily remembered fix names, airports, and simplified operating procedures.

Participants will complete a background questionnaire at the beginning of the experiment. A subject matter expert will provide an airspace and equipment briefing. Participants will control practice scenarios to ensure familiarity with the airspace and simulation equipment. Participants will then control experimental scenarios. After each experimental scenario, participants will complete a Post-Scenario Questionnaire. Following the last simulation run participants will complete a final questionnaire, and will be asked for detailed comments about the experiment.

An automated data collection system will record important simulation events and produce a set of system effectiveness measures, including safety, capacity, efficiency, and controller workload. The simulation will be video recorded for research purposes only. No video records will be released. An SATCS will make over-the-shoulder observations during each simulation.

Rights of Participants

Participation in this study is strictly voluntary and the privacy of participants will be protected. No individual names or identities will be recorded or released in any reports. Strict adherence to all Federal, Union, and ethical guidelines will be maintained throughout the study. The purpose of the study is to scientifically assess the previously cited concepts, not to evaluate the individual controllers.

Minimum Standards for Participants

Participants will be current level 5 Radar controllers with valid medical certificates. Single vision glasses may be used but contact lenses are not acceptable. Staff and supervisors are not acceptable.

Travel

Travel and overtime expenses, if needed, will be borne by the Technical Center. Fund codes will be provided when personnel are identified. Sending facilities or regions will write travel orders using Technical Center fund sites.

Points of Contact

Your support is important to the success of this project, and your cooperation will be greatly appreciated. If you have any additional questions, then please do not hesitate to contact me.

Earl S. Stein, Ph.D., Engineering Research Psychologist, Technical Project Lead, (609) 485-6389, ACB-220, Bldg. 28, Research Development and Human Factors Laboratory, FAA William J. Hughes Technical Center, Atlantic City International Airport, NJ 08405.

You may also contact Mr. Tony Buie, SATS, Project Controller subject matter expert, (609) 485-4869.

APPENDIX B

Informed Consent Form

I, ____________________________, understand that the Federal Aviation Administration sponsors and Ben Willems direct this study, entitled "Decision Support Automation Research (DSAR) in the En Route ATC Environment I."

Nature and Purpose

I will volunteer as a participant in the project above. The purpose is to explore active controllers' use of different levels of automation. The time requirement for this experiment is four days. I will travel on Monday and Friday. On the 2 test days of the experiment, I will participate in 6 practice and 13 experiment simulations of 45 minutes each.

Experimental Procedures

During half of the simulations, the movements of my eyes will be monitored. A small camera mounted on a headband will monitor my eye movements. An invisible beam of infrared light will illuminate my eye.

The simulations will mimic operational air traffic conditions. I will interact with simulation pilots and control simulated air traffic like I would normally do in the field.

Discomforts and Risks

The device that monitors the eye movements may cause some discomfort. The skin area under the headband that supports the device may show some redness after wearing the device for the duration of a simulation. The intensity of the infrared beam that illuminates the eye is about one thirtieth of the intensity expected while walking outside on a sunny day and should not cause any discomfort or risk to my health.

Benefits

I understand that the only direct benefit to me is to participate in research in Atlantic City, NJ.

The benefit derived from the results of this experiment for controllers may include a better understanding of why operational errors occur, which could lead to new ways to assist ATC students.

Participant's Responsibilities

During the experiment, it will be my responsibility to control the simulated air traffic as if I was controlling traffic at my home facility. I will answer any questions asked during the experiment to the best of my abilities. I will not discuss the content of the experiment with anyone until the completion of the experiment.

Participant's Assurances

I understand that my participation in this study is voluntary. Ben Willems has adequately answered any questions I have about this study, my participation, and the procedures involved. I understand that Ben Willems will be available to answer any questions concerning procedures throughout this study. I understand that if new findings develop during the course of this research that may relate to my decision to continue to participation, I will be informed.

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

I understand that records of this study are strictly confidential, and that I will not be identifiable by name or description in any reports or publications about this study. Photographs and audio and video recordings are for use within the Research and Development Human Factors Laboratory only. Any of the materials that may identify me as a participant cannot be used for purposes other than internal Research and Development Human Factors Laboratory without my written permission.

I understand I can withdraw from the study at any time without penalty or loss of benefits to which I may be entitled. I also understand that the researcher of this study may terminate my participation if he feels this to be in my best interest.

If I have questions about this study or need to report any adverse effects from the research procedures, I will contact Ben Willems at (609) 485-4191 during Monday through Friday or at (609) 404-1650 in the evening or on weekends.

I may also contact Dr. Earl Stein (609) 485-6389, the Air Traffic Human Factors Technical Lead, at any time with questions or concerns.

I have read this consent document. I understand its contents, and I freely consent to participate in this study under the conditions described. I have received a copy of this consent form.

Research Participant: ______________________Date: __________

Investigator: ____________________________Date: __________

Witness: _______________________________Date: __________

APPENDIX C

Airspace

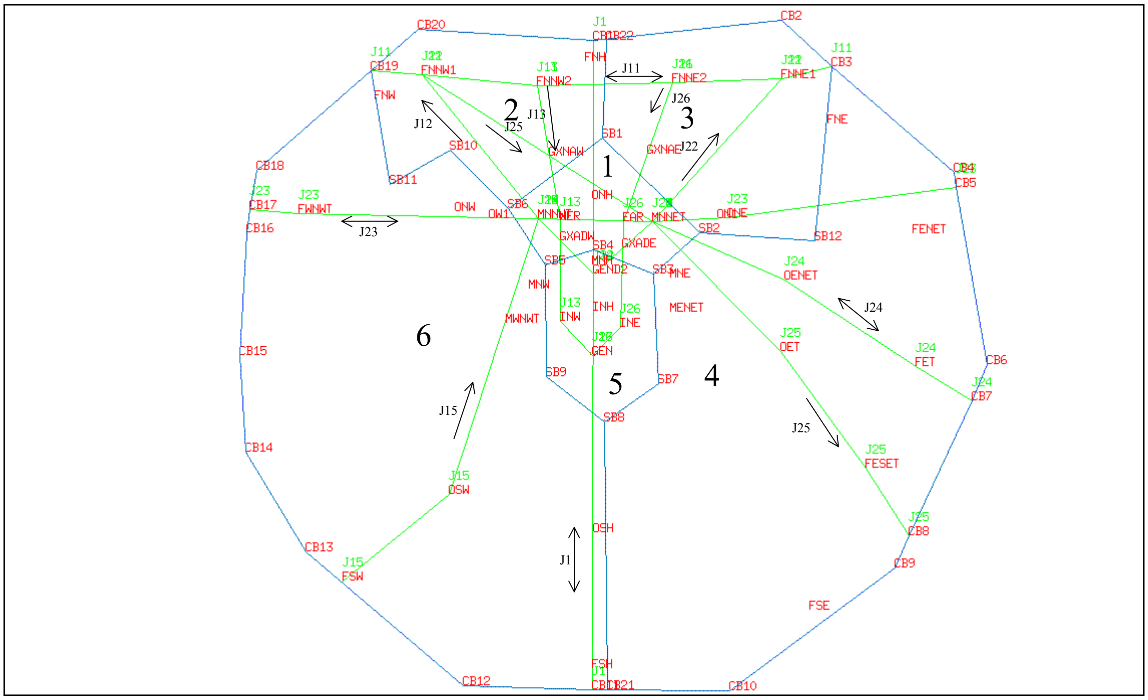

Figure C-1. Genera Center airspace showing center and sector boundaries, jetways, and fixes. Numbers 1 through 6 indicate the different sectors that make up Genera Center.

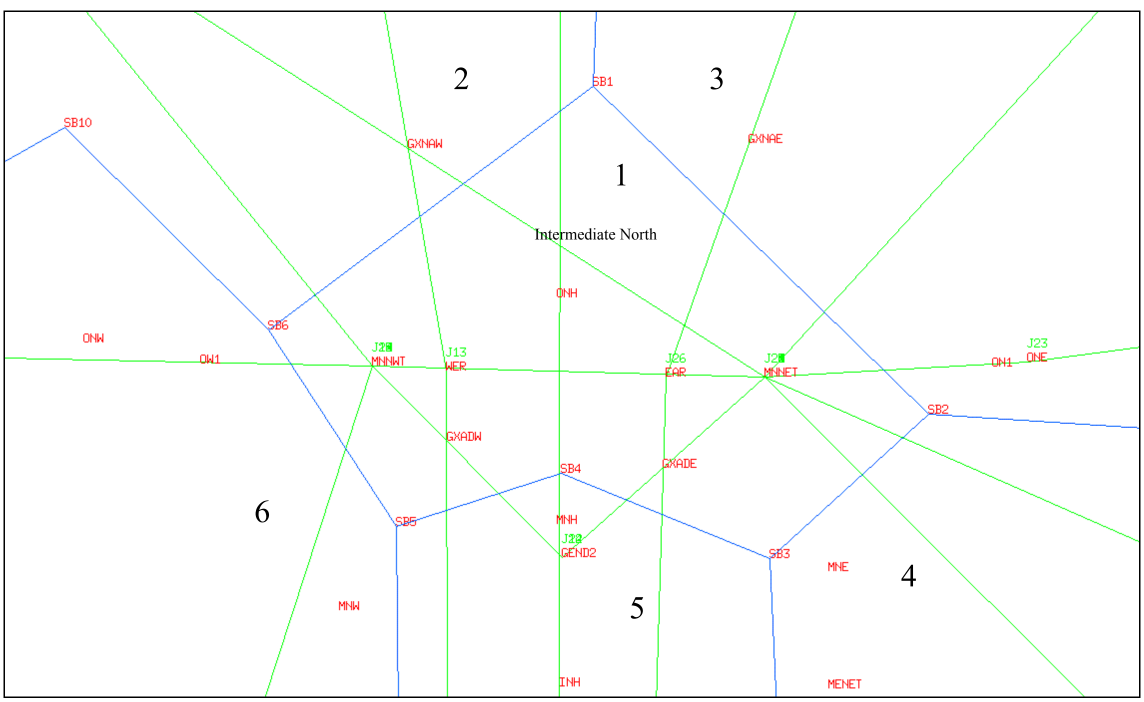

Figure C-2. Intermediate North sector (# 1) within Genera Center. This sector will be used during the study and is FL 240 and above.

APPENDIX D

DST Display Windows

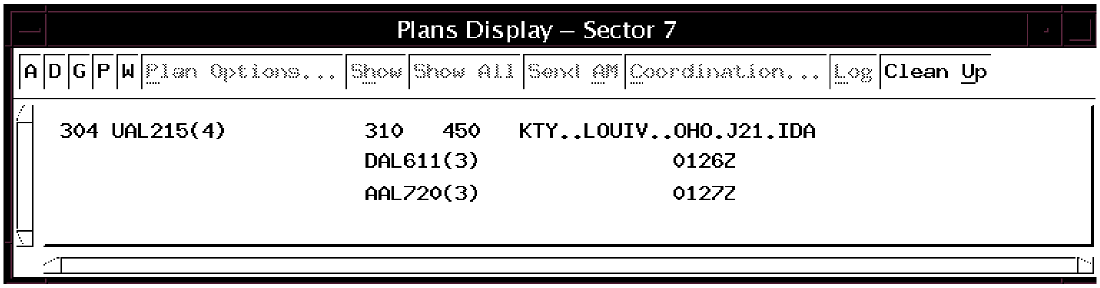

Figure D-1. Plans display window from DST.

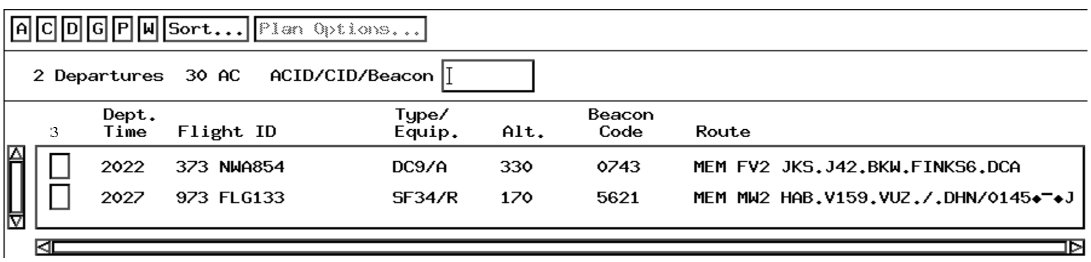

Figure D-2. Departure list display window from DST.

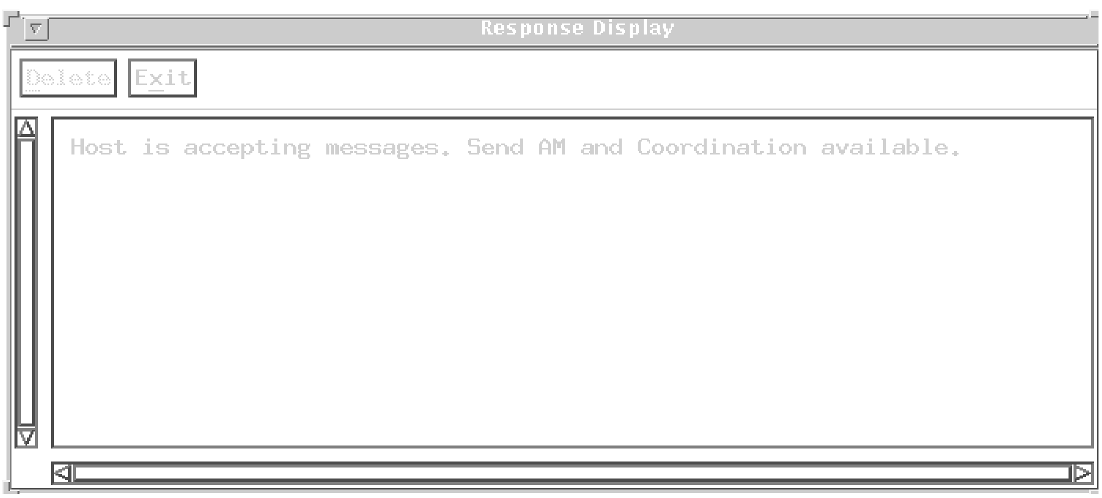

Figure D-3. Response display window from DST.

APPENDIX E

Implementation of Savant

SAVANT's display closely matched the DSR display. We presented the ATCS with several queries and recorded the controllers' responses by having them click on an aircraft. Figure E-1 presents the differences between the representations between the DSR and the SAVANT display. SAVANT recorded objects the ATCS chooses as the correct response. All ATCS entries carried a time stamp.

SAVANT ran on a separate computer, one for the D-side and one for the R-side controller. The computer display presented a replication of the airspace including targets and data blocks. SAVANT indicated targets of interest on the screen by highlighting them in a separate color and with portions of their data blocks hidden. Participants responded to queries about the status of the targets. They responded directly on the computer by clicking on the target that has the highest values for the queried variable. When SAVANT started, the computer screen was blank. Each session lasted 15 seconds (3 seconds to display the question and 9 seconds to respond).

The airspace boundaries, airways, Very High Frequency Omnidirectional Range (VORs), etc. were static and can be read in at the beginning of the simulation. The aircraft information was overlaid at the time of the query session. The aircraft messages in the NAS are the 1525 messages. We then needed to reconstruct the aircraft representations and mask some information based on the type of query we wanted to present.

SAVANT uses the ODS Toolbox (Orthogon, 1999) as the human-computer-interface. SAVANT obtained part of the data for the replication of the DSR screen from the target generation facility (TGF). SAVANT needed other data to correctly place data blocks. That data included the orientation and length of the leader line relative to the raw radar return, the history trail, and the vector line. A token ring sniffer listened in on two of the three DSR token rings and presented the DSR messages on a socket. Middleware grabbed the appropriate messages and presented it to DESIREE. Figure E-1 presents an example of both the DSR and the SAVANT representation of an aircraft pair. The implied query in this particular case would be "Which aircraft is higher?" The following sections will briefly discuss the implementation of the sniffer, middleware, and DESIREE.

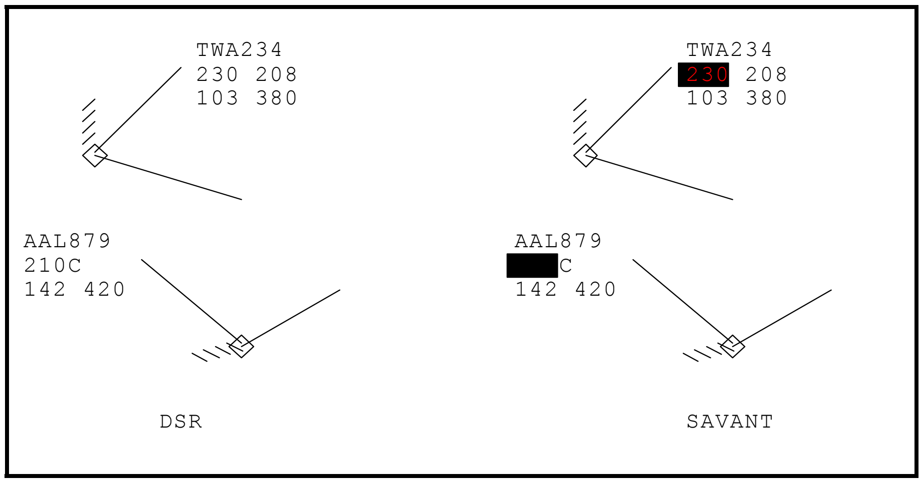

Figure E-1. Comparison of DSR and SAVANT aircraft representation for situation awareness queries.

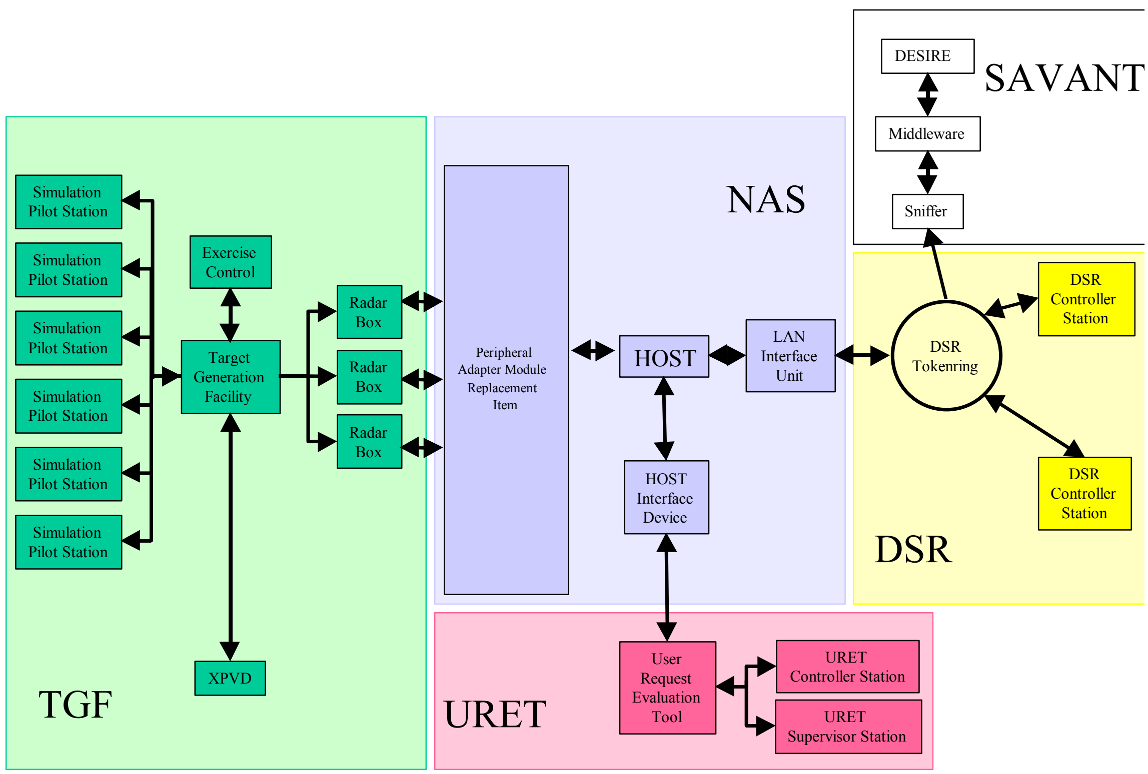

Figure E-2. Schematic layout of DSAR1 hardware and software.

Sniffer

The first component is a custom Token Ring server (TRSRV) that obtains the HCS/DSR data and forwards it to a middle-ware component that processes the data, and then passes the information to a back-end display processor. This data is identified in Interface Control Document NAS-IR-2104-4001 (FAA, 1999).

The HCS and DSR transfer data in data structures (packets) called LIU messages or System Requests (SRs). The HCS sends the data through the HCS channel attachment to the Local Communications Network (LCN) Interface Unit (LIU). This RISC 6000 system processes the SRs and forwards them across the LCN, the token ring Broadcast (B) and token ring Point-to-Point (P) rings, to the appropriate consoles based on transmission class and destination name. The data that originate at the console systems travel the reverse path to the HCS.

The TRSRV monitors the token rings in the promiscuous mode so all LAN traffic is available to the server. We then filter the SRs according to a pre-selected subset of SRs that SAVANT will use. The TRSRV encapsulates these selected SRs in UDP/IP packets and forwards them to an active client, the middle-ware component.

The TRSRV starts when a client sends an arbitrary packet to UDP port 3900. It then forwards the selected SR packets to that client. The client can terminate the TRSRV operation by exiting. This will cause the TRSRV to receive an ICMP “connection refused” status, which is the signal to terminate token ring monitoring. If the TRSRV receives a request on UDP port 3900 from another client while it is already in dialogue with a current client, all packets will be forwarded to the new client exclusively from that time.

The SAVANT tool currently uses a subset of the available SRs: 1524, 1509, 1525, and 0758. The Add Target Position (#1524) SR contains primary and beacon radar for display on the Host Situation Display at DSR R-positions. Also the Process Limited Data Block (#1509) SR contains Limited Data Blocks for display on the Host Situation Display at DSR R-positions. In addition the Process Track Data Group Block (#1525) SR contains Full Data Blocks for display on the Host Situation Display at DSR R-positions. And to complete the subset the Process Operator Host Command (#0758) SR contains an input to the Host from a DSR R, D, or A-position. This subset of SRs can convey the air traffic control situation as presented to the controller, and the controller responses as the controller responds to the situation.

Middleware

Purpose

The middleware serves two purposes. First, it provides DESIREE with the data to allow a replication of the aircraft displayed on DSR. Secondly, it records the data needed to fully replicate the aircraft displayed on DSR for posttest data reduction.

Implementation

The middleware process uses JAVA/CORBA/OrbixWeb, running on a Sun workstation under Solaris 2.7. DESIREE launches the middleware process. The middleware is an invoked process for DESIREE. The middleware then connects to the sniffer by sending an arbitrary "wake-up" packet to UDP port 3900.

Procedure

After the sniffer wakes up, it puts the filtered SR data on the UDP port. The middleware receives the filtered SR data from the UDP port. The data format on the UDP port is similar to the format specified in NAS-IR-2104-4001. However, the Full Data Block (FDB) message 1525 changes to message 1672 on the HCS portion of the DSR token ring. The difference between the 1672 and the 1525 message is that the 1672 message carries FDB data for one sector only, while the 1525 message carries FDB data for multiple sectors.

To provide DESIREE with data, the middleware picks up data packets from the UDP port, populates CORBA objects with the data received from the sniffer, and DESIREE picks up the objects.

To record for posttest data reduction, the middleware writes the data to disk in XML format. The development team chose the XML format, because XML parsers are readily available. The file names are date and time stamped to the actual local time. The recorded data allows the eye movement data reduction software to correlate point of gaze data with the position of aircraft radar returns and data blocks. A separate software package uses the XML files to transform the data into a format that is compatible with the existing eye movement data reduction and replay software packages.

DESIREE

DESIREE is the RDHFL front end to our in-house ATC simulator. Although DESIREE uses the ODS Toolbox (Orthogon, 1999), it has adopted a CORBA environment. In the implementation of SAVANT, DESIREE incorporated the middleware component, a file reader component, and two display components.

The **file reader component** enabled researchers to create text files that contained events with the corresponding event-time. These events included the swapping of displays, presentation of a question, highlighting of objects, and hiding of object elements. DESIREE swapped [need a picture here to show the physical layout of the DSR/Host/SAVANT/URET setup] the DSR and URET displays for SAVANT displays by controlling a wide bandwidth matrix switch through a serial port connection.

DESIREE received the SR data by populating aircraft objects through its **middleware component**. It displayed the aircraft objects in real time through one of the **display components** on the same Sony high resolution (2000 x 2000 pixel) display. The controllers used these objects to respond to our queries. The other **display component** recorded the current displayed objects to a file on disk.

APPENDIX F

ATWIT Instructions

ATWIT instructions given before calibration of the oculometer.

One purpose of this research is to obtain an accurate evaluation of controller workload. By workload, we mean all the physical and mental effort that you must exert to do your job. This includes maintaining the "picture," planning, coordinating, decision making, communicating, and whatever else is required to maintain a safe and expeditious traffic flow. Every five minutes the ATWIT device, located to the side of the radar display, will emit a brief tone and ten buttons will appear. The buttons will remain visible for only a limited amount of time. Tell us how hard you are working by pushing the buttons numbered from 1 to 10 on the ATWIT.

I will review what these buttons mean in terms of your workload. At the low end of the scale (1 or 2), your workload is low - you can accomplish everything easily. As the numbers increase, your workload is getting higher. Numbers 3, 4, and 5 represent increasing levels of moderate workload where the chance of error is still low but steadily increasing. Numbers 6, 7, and 8 reflect relatively high workload where there is some chance of making errors. At the high end of the scale are numbers 9 and 10, which represent a very high workload, where it is likely that you will have to leave some tasks unfinished.

All controllers, no matter how proficient and experienced, will be exposed at one time or another to all levels of workload. It does not detract from a controller's professionalism when he indicates that he is working very hard or that he is hardly working. Feel free to use the entire scale and tell us honestly how hard you are working. Do not sacrifice the safe and expeditious flow of traffic in order to respond to the ATWIT device. Remember, your workload rating should not reflect how much you are working during the course of the scenario. Instead, your rating should reflect how much workload you are experiencing during the instant when you are prompted to make the rating.

Do you have any questions about using the ATWIT device?

APPENDIX G

ATCS Roles and Responsibilities

Radar	Radar Associate (RA)	Flight Data (D)	Non-Radar
Ensure separation	Ensure separation	Operate interphones	Ensure separation
Initiate control instructions	Initiate control instructions	Assist the RA-position in managing flight progress strips	Initiate control instructions
Monitor and operate radios	Operate interphones	Receive/process and distribute flight progress strips	Monitor and operate radios
Accept and initiate automated handoffs	Accept and initiate automated handoffs, and ensure radar position is made aware of the actions	Ensure flight data processing equipment is operational	Accept and initiate transfer of control, communications, and flight data
Assist the RA position with non-automated handoff actions when needed	Assist the R-side position by accepting or initiating automated handoffs which are necessary for the continued smooth operation of the sector, and ensure that the R-side is made immediately aware of any action taken	Request/receive and disseminate weather, NOTAM's, NAS status, traffic management, and Special Use Airspace status messages	Ensure computer entries are completed on instructions or clearances issued or received
Assist the RA position in coordination when needed	Coordinate including point outs	Manually prepare flight progress strips when automation systems are not available	Ensure strip marking is completed on instructions or clearances issued or received
Scan radar display. Correlate with flight progress strip information	Monitor radios when not performing higher priority duties	Enter flight data into computer	Facilities utilizing nonradar positions may modify the standards contained in the radar associate
Ensure computer entries are completed on instructions or clearances you issue or receive	Scan Flight Progress Strips. Correlate with radar data.	Forward flight data via computer	
Ensure strip marking is completed on instructions or clearances you issue or receive	Manage Flight Progress Strips.	Assist facility/sector in meeting situation objectives	
Adjust equipment at R-side to be usable by all members of the team	Ensure computer entries are completed on instructions issued or received by the R-side when aware of those instructions.		
The R-side shall not be responsible for G/G communications when precluded by VSCS split functionality	Ensure strip marking is completed on instruction issued or received by the R-side when aware of them.		
	Adjust equipment at RA-position to be usable by all members of the team		

APPENDIX H

Post-Scenario Questionnaire

Please complete the following:

I. NASA-TLX

1. Circle the number that best describes the **mental demand** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high
2. Circle the number that best describes the **physical demand** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high
3. Circle the number that best describes the **temporal demand** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high
4. Circle the number that best describes your **performance** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high
5. Circle the number that best describes your **effort** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high
6. Circle the number that best describes your level of **frustration** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	extremely high

II. Scenario

7. Circle the number below that best describes **how hard you were working** during this scenario.	extremely hard	1 2 3 4 5 6 7 8 9 10	Not hard
8. Circle the number that best describes **how well you controlled traffic** during this scenario.	extremely poorly	1 2 3 4 5 6 7 8 9 10	Extremely well
9. Circle the number that best describes your **overall situational awareness** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	Extremely high
10. Circle the number that best describes your **situational awareness for current aircraft locations** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	Extremely high
11. Circle the number that best describes your **situational awareness for projected aircraft locations** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	Extremely high
12. Circle the number that best describes your **situational awareness for potential violations** during this scenario.	extremely low	1 2 3 4 5 6 7 8 9 10	Extremely high
13. Circle the number that best describes **how difficult** this scenario was.	extremely difficult	1 2 3 4 5 6 7 8 9 10	extremely easy
14. Circle the number that best describes **how competent** you felt using the decision support tool(s).	extremely incompetent	1 2 3 4 5 6 7 8 9 10	extremely competent
15. Circle the number that best describes **how predictable** the behavior of the decision support tool(s) was.	extremely unpredictable	1 2 3 4 5 6 7 8 9 10	extremely predictable
16. Circle the number that best describes **how well** the decision support tool(s) predicted the outcome of separation strategies.	Extremely poorly	1 2 3 4 5 6 7 8 9 10	extremely well
17. Circle the number that best describes how **effective** the decision support tool(s) was in resolving conflicts **when using trial flight plans**.	extremely ineffective	1 2 3 4 5 6 7 8 9 10	extremely effective
18. Circle the number that best describes your view of the capability of the tool(s) to **predict** future conflicts.	Extremely unable	1 2 3 4 5 6 7 8 9 10	extremely able
19. Circle the number that best describes the **technical knowledge** incorporated in the tool.	extremely ignorant	1 2 3 4 5 6 7 8 9 10	extremely knowledgeable
20. Circle the number that best describes your **overall confidence** in the decision support tool(s).	No confidence	1 2 3 4 5 6 7 8 9 10	complete confidence

III. Simulation

21. Circle the number that best describes **how realistic** the simulation was.	extremely unrealistic	1	2	3	4	5	6	7	8	9	10	extremely realistic
22. Circle the number that best describes **how representative** the scenario was of a typical workday.	Not representative	1	2	3	4	5	6	7	8	9	10	extremely representative
23. Circle the number that best describes if the **ATWIT device interfered** with controlling traffic.	no interference	1	2	3	4	5	6	7	8	9	10	extreme interference
24. Circle the number that best describes if the **oculometer interfered** with controlling traffic.	no interference	1	2	3	4	5	6	7	8	9	10	extreme interference
25. Circle the number that best describes **how well the simulation-pilots responded** to clearances and callbacks.	extremely poorly	1	2	3	4	5	6	7	8	9	10	Extremely well
26. Do you have any other **comments about your experiences during the simulation**?												

APPENDIX I

OTS Rating Forms

INSTRUCTIONS FOR QUESTIONS 1-20

This form was designed to be used by instructor certified ATC specialist to evaluate the effectiveness of controllers working in simulation environments. Observers will rate the effectiveness of controllers in several different performance areas using the scale show below. When making your ratings, please try to use the entire scale range as much as possible. You are encouraged to write down observations and you may make preliminary ratings during the course of the scenario. However, we recommend that you wait until the scenario is finished before making your final ratings. The observations you make do not need to be restricted to the performance areas covered in this form ands may include other areas that you think are important. Also, please write down any comments that may improve this evaluation form. Your identity will remain anonymous; so do not write your name on the form.

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

I. MAINTAINING SAFE AND EFFICIENT TRAFFIC FLOW

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

II. MAINTAINING ATTENTION AND SA

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

III. PRIORITIZING

8. Taking Actions in an Appropriate Order of Importance - resolving situations that need immediate attention before handling low priority tasks - issuing control instructions in a prioritized, structured, and timely manner	1	2	3	4	5	6	7	8
9. Preplanning Control Actions - scanning adjacent sectors to plan for future and conflicting traffic - studying pending flight strips in bay	1	2	3	4	5	6	7	8
10. Handling Control Tasks for Several Aircraft - shifting control tasks between several aircraft when necessary - avoiding delays in communications while thinking or planning control actions	1	2	3	4	5	6	7	8
11. Marking Flight Strips while Performing Other Tasks - marking flight strips accurately while talking or performing other tasks - keeping flight strips current	1	2	3	4	5	6	7	8

IV. PROVIDING CONTROL INFORMATION

12. Providing Essential ATC Information - providing mandatory services and advisories to pilots in a timely manner - exchanging essential information	1	2	3	4	5	6	7	8
13. Providing Additional ATC Information - providing additional services when workload is not a factor - exchanging additional information	1	2	3	4	5	6	7	8
14. Providing Coordination - providing effective and timely coordination - using proper point-out procedures	1	2	3	4	5	6	7	8

V. TECHNICAL KNOWLEDGE

15. Showing Knowledge of LOAs and SOPs - controlling traffic as depicted in current LOAs and SOPs - performing hand-off procedures correctly	1	2	3	4	5	6	7	8
16. Showing Knowledge of Aircraft Capabilities and Limitations - avoiding clearances that are beyond aircraft performance parameters - using appropriate speed, vectoring, and/or altitude assignments to separate aircraft with varied flight capabilities	1	2	3	4	5	6	7	8
17. Showing Effective Use of Equipment - updating data blocks - using equipment capabilities	1	2	3	4	5	6	7	8

VI. COMMUNICATING

18. Using Proper Phraseology - using words and phrases specified in ATP 7110.65 - using ATP phraseology that is appropriate for the situation - avoiding the use of excessive verbiage	1	2	3	4	5	6	7	8
19. Communicating Clearly and Efficiently - speaking confidently and at the proper volume and rate for pilots to understand - speaking fluently while scanning or performing other tasks - clearance delivery is complete, correct and timely	1	2	3	4	5	6	7	8
20. Listening for Pilot Readbacks and Requests - correcting pilot readback errors - processing requests correctly in a timely manner - acknowledging pilot or other controller requests promptly	1	2	3	4	5	6	7	8

APPENDIX J

SA Global Assessment Technique Questions

1. Enter the location of all aircraft by indicating:
   - the aircraft in track control
   - other aircraft in the sector
   - aircraft that will be in track control in the next 2 minutes
2. Enter the aircraft callsign (for the aircraft indicated on the sector map in question 1).
3. Enter the aircraft altitude (for the aircraft indicated on the sector map in question 1).
4. Enter the aircraft ground speed (for the aircraft indicated on the sector map in question 1).
5. Enter the aircraft heading (for the aircraft indicated on the sector map in question 1).
6. Enter the aircraft's next sector (for the aircraft indicated on the sector map in question 1). (insert list of sector names here as choices followed by "landing in sector")
7. Enter the aircraft's current direction of change in each column (for the aircraft indicated on the sector map in question 1).

Altitude Change	Turn
climbing	right turn
descending	left turn
level	straight

8. Enter the aircraft type (for the aircraft indicated on the sector map in question 1).
9. Enter the aircraft's activity in the sector (for the aircraft indicated on the sector map in question 1).
   En route
   inbound to airport
   outbound from airport

10. Which pairs of aircraft have lost or will lose separation if they stay on their current assigned courses?
11. Which aircraft have been issued clearances that have not been completed?
12. Did the aircraft receive its clearance correctly (for each entered in question 11)?
13. Which aircraft are currently conforming to their clearances (for each entered in question 11)?
14. Which aircraft will be handed off to another sector/facility in the next 2 minutes?
15. Enter the aircraft that are not in communication with you.
16. Enter the aircraft that will violate special airspace separation standards if they stay on their current assigned paths.
17. For which aircraft is weather currently impacting or will be impacting on in the next 5 minutes?

APPENDIX K

Entry Questionnaire

Note: We provided space for comments after each question and at the end of the questionnaire.

Please complete the following:

1. What is your age in years? ____________ years
2. Are you wearing corrective lenses during this experiment? ☐ Yes ☐ No
3. How many years have you actively controlled traffic? ____________ years
4. How many years have you controlled traffic at your current facility? ____________ years
5. How many months in the past year have you actively controlled traffic? ____________ months
6. What is your current position as a controller? ☐ Developmental ☐ Full Performance Level ☐ Other: ____________
7. Please list other facilities where you have worked: ____
____

8. Do you **search the DSR display in one special way for information**? ☐ Yes ☐ No
   If it depends on certain factors, what are they?

9. Please circle the number that best describes your **current skill as a controller**.	not skilled	1	2	3	4	5	6	7	8	9	10	extremely skilled
10. Please circle the number that best describes your **motivation** to participate in this study.	not motivated	1	2	3	4	5	6	7	8	9	10	extremely motivated
11. How often do you use **vertical separation?**	never	1	2	3	4	5	6	7	8	9	10	always
12. How often do you use **vectoring** for separation?	never	1	2	3	4	5	6	7	8	9	10	always
13. How often do you use **speed control** for separation?	never	1	2	3	4	5	6	7	8	9	10	always
14. Please circle the number that best describes your **experience with video games**.	Not experienced	1	2	3	4	5	6	7	8	9	10	extremely experienced

Please circle the number that best describes the **importance** of the following **aircraft** information.

15. Aircraft CallSign	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
16. Aircraft Type	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
17. Aircraft Beacon Code	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
18. Controller Ownership	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
19. Entry Altitude	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
20. Entry Airspeed	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
21. Entry Fix	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
22. Exit Altitude	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high

23. Exit Airspeed	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
24. Exit Fix	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
25. Arrival Airport (within sector)	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
26. Departure Airport (within sector)	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
27. Current Altitude	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
28. Current Airspeed	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
29. Current Heading	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
30. Current Aircraft Location	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
31. Most Recently Assigned Altitude	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
32. Most Recently Assigned Airspeed	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
33. Most Recently Assigned Heading	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
34. Aircraft Holding/Spinning	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
35. Aircraft Waiting for Hand-off/Release	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
36. Aircraft Near Exit Fix/Arrival Airport	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
37. Density of Aircraft on Radar Display	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high

Please circle the number that best describes the **importance** of the following **radar display** information.

38. System Clock	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
39. VORs	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
40. Fixes	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
41. Airports	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
42. Restricted Area Boundaries	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
43. ILS Approaches	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
44. Holding Patterns	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
45. Obstructions	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
46. Sector Boundaries	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
47. Filter Settings	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
48. Future Aircraft List	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high
49. Collision Alert	extremely low	1	2	3	4	5	6	7	8	9	10	extremely high

APPENDIX L

Exit Questionnaire

Please complete the following:

1. Circle the number that best describes the **Genera Center Airspace hands-on training**.	extremely poor 1 2 3 4 5 6 7 8 9 10 extremely good
2. Circle the number that best describes the **Decision Support Tool hands-on training**.	extremely poor 1 2 3 4 5 6 7 8 9 10 extremely good
3. Was there anything that you found particularly unique in the simulation that you would not see at your home facility?	
4. Were you **constantly** aware of wearing the oculometer or did you tune it out?	
5. How do you decide whether or not to suppress data?	
6. Is there anything about the study that we should have asked or that you would like to comment about?	

Thank you for your participation.

APPENDIX M

Biweekly Schedule

Monday	Travel to the FAA WJH Technical Center
Tuesday	Introduction to the experiment and first day of DST training: Classroom
Wednesday	Second day of DST Training: Classroom
Thursday	Third day of DST Training: Hands-on
Friday	Airspace training
Monday	Training on TGF/DSR/URET/SAVANT/SAGAT
Tuesday	Experimental Simulations
Wednesday	Experimental Simulations
Thursday	Experimental Simulations
Friday	Travel home

APPENDIX N

Statistical Methods

This section provides background information on the statistical methods used in this report. These methods are powerful tools that allow researchers to determine the most probable outcomes of an experiment based on limited sample sizes. The following paragraphs explain general concepts in statistics, the methods used in this study and some important considerations in order to use them effectively.

The purpose of any statistical experiment is to determine the effect of certain factors on one or more outcome variables (or **dependent variables** [**DV**s]). An example of a DV is the number of altitude changes an ATCS makes. This DV could be affected by the type of airspace (terminal, en route, or oceanic), the number of aircraft flying through or to the sector, or many other factors. The manipulated factors of an experiment are the **independent variables** (or **IV**s). Each manipulation of an IV (e.g. 25 planes or 45 planes) forms a separate experimental **condition**. Each trial under a particular condition is termed an **observation**.

Experiments can include one or more IVs. When an experiment includes more than one IV, multiple IVs can affect the outcome differently, called an **interaction**. It would be impossible to study the effects of type of airspace and number of aircraft independently. When such interactions between IVs occur, the researcher will study the effect by holding one variable constant while varying the others, called testing for **simple effects.** In this way, the researcher obtains a picture of how the variables interact by examining the outcome of each manipulation. When researchers study the effect of each IV separately (no interactions), it is termed an analysis of **main effects**. We will study main effects only in the absence of interactions.

The number of values for the IVs included in an experiment depends on several practical considerations. For example, if a researcher is studying the decision-making patterns of controllers as a function of type of airspace, the values of the IV, type of airspace (tower, TRACON, en route), are clear. In other cases, the answer would depend upon what type of outcome the researchers needed from the results of the experiment as well as some practical considerations. Different values of IVs, termed **levels**, can increase the number of experimental conditions and thus increase the resources needed to complete the experiment. One can imagine the complexity and length of an experiment in which researchers studied controllers with experience ranging from 1 to 50 years, thus creating 50 incremental levels. It would be far simpler and easier to study the effect of controller experience by using only three categories: Developmental, ATCS, and Supervisor.

What is the number of observations required for each test condition? Increasing the number of observations increases the statistical power of the experiment. Increased statistical power means that an increased probability exists that the outcome of an experiment will likely be true for the entire population. However, increasing the number of observations comes at the expense of greater numbers of participants, more time, or both. An efficient experimental design should include enough observations for reasonable statistical power without including unnecessary observations that could dramatically increase demands for resources unless there was an increased need for power.

With insight into statistical terminology as well as some background into considerations involved in experimental design, it is now useful to look into several different categories of experiments and statistical methods used to determine significant outcomes. For simplicity, each of the following categories involves only a single IV (the experience level of controllers). In increasing level of complexity, we will examine three categories of experiments:

1. Observations on a single DV under two conditions (*T*-test)

2. Observations on a single DV under multiple conditions (ANOVA)

3. Observations on multiple DVs under multiple conditions (MANOVA)

We will discuss each of these categories below.

Observations on a Single Variable under Two Conditions

When a researcher wants to compare two conditions, he or she takes the average of multiple observations on a single variable under two conditions and performs a ***T*-test**. However, an average value can often be misleading. Within a group of such observations, some differences will exist in the individual observations that contributed to the average. Some Developmental controllers may be faster learners than others and will use less altitude changes in order to control traffic. The average number of altitude changes for all developmental controllers can include a wide range of values. The differences between the individual times and the mean number of times represent the **variability** of the data. As the variability in the data increases, the mean value is less useful to the researcher because many of the individual values are far from the mean. Figure N-1 illustrates the variability of data.

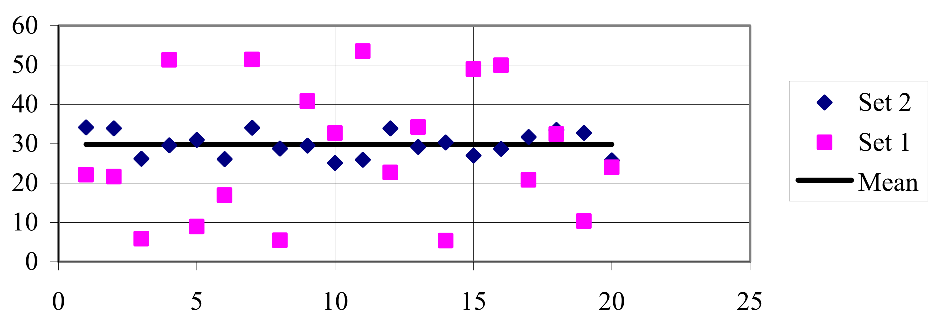

Figure N-1. Two sets of observations with the same means, but very different values.

If a researcher wants to compare two samples, the comparison not only involves comparing the averages but also the variability within the observations. For this reason, the true mean (the mean a researcher would calculate if he or she sampled the number of altitude changes for all Developmental controllers in the world) differs from the sample mean. A researcher must ask if the difference in the means of these two sets of observations is a true difference, or caused by chance. This is where **probability** theory aids the researcher. Statistics can help the researcher determine the probability that the two means for the entire population (all controllers) are different from the sample (limited number of controllers). The statistical test used in this case is the ***t*-test**. The t-test compares two averages and checks if the two averages are different due to chance alone. It is important to recognize that the t-test never gives the researcher 100% assurance that the two means actually differ. It is common practice to accept a 95% assurance (or in other words a 5% risk) as sufficient guarantee.

<u>SUMMARY OF A *T*-TEST</u>. An experiment includes multiple observations on a single variable under two conditions. The average values (means) of the two conditions take variability into consideration. The analyses determine the probability that the means differ due to chance alone.

Example: When one compares the number of altitude changes between Developmental and ATCS controllers at a local center, the comparison involves multiple observations. The multiple observations consist of the number of altitude changes of each individual within each experience level. The variable is the number of altitude changes. The conditions include the two levels of experience. Figure N-2 is a graphical display of this example. Although it shows a difference in number of altitude changes between the two groups, some individual observations overlap. The researcher uses a t-test to see if chance caused this difference.

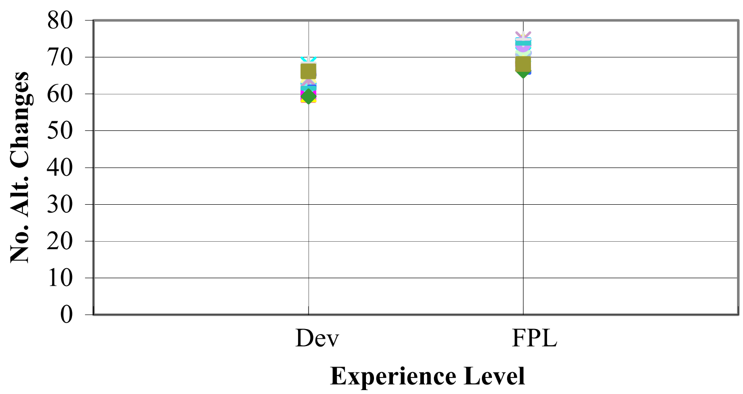

Figure N-2. Multiple observations of altitude changes as a function of experience level.

Observations on a Single Variable under Multiple Conditions

Where the t-test compared the averages between two conditions, the **analysis of variance** (ANOVA) compares averages of a single variable between multiple conditions (i.e. the number of altitude changes including Developmental, ATCS, and Supervisors). An ANOVA tests if these averages are different due to chance alone. The basic test results in an **F value** for a single DV (the number of altitude changes). The value of F ranges from 0 to infinity (∞). A large F value may indicate that the IV (experience level) has a powerful effect on the DV (number of altitude changes) with less likelihood that differences between means occurred by chance. The strength of association (e.g. η) or percent of variance explained is an indication of the difference in the strength of effects between conditions. A difference between means is **significant** if there is a very high probability that the means are actually different (usually greater than 95%).

Sometimes, there is a significant difference where the F value is relatively low. This indicates that the IV does not have a very strong effect.

An ANOVA can show that there is a difference in means not caused by chance alone. If the ANOVA indicated that the number of altitude changes varies with experience level, is the mean number of altitude changes for Developmental controllers different than ATCS controllers? The mean for Developmental controllers differs significantly from those of ATCS but not significantly from those for Supervisors. Therefore, the researcher runs another test to compliment the ANOVA, called a **post-hoc comparison**. Researchers will use post-hoc comparisons to determine which of the pairs of means differ significantly.

SUMMARY OF AN ANOVA. The **analysis of variance** (ANOVA) compares averages of a single DV between multiple conditions and tests if these averages are different due to chance alone. The test results in an F value. A large F value indicates less likelihood and a small value indicates increased likelihood that differences between means occurred by chance. A difference between means is **significant** if there is a very high probability that the means are actually different. A post-hoc comparison determines which means differ.

Example: When a researcher compares the number of altitude changes between Developmental, ATCS and Supervisors at a local center, the comparison involves multiple observations. The multiple observations are the number of altitude changes of each individual within each group. The variable is the number of altitude changes. The conditions are the three experience levels. Figure N-3 displays the data related to this example. Some differences in number of altitude changes exist between experience levels, but there is overlap between observations in each experience level. An ANOVA would determine if these differences were due to chance alone. If the ANOVA indicated that there is some difference in experience levels regarding number of altitude changes, post-hoc comparisons would indicate which means associated with which experience levels differ.

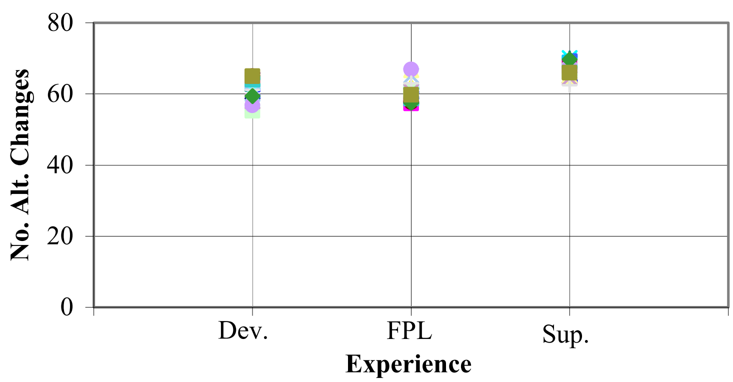

Figure N-3. Multiple observations of altitude changes as a function of experience level.

Observations on a Multiple Variable under Multiple Conditions

Where an ANOVA compares averages between multiple conditions for a single variable (a **univariate** test), the **multivariate ANOVA** (MANOVA) compares averages for several DVs simultaneously and tests if these averages are different due to chance alone. Suppose that researchers wanted to examine number of altitude changes **and** the number of heading changes. Also, suppose that the researchers wanted to test both of these measures as a function of experience level. This presents a case of multiple DVs (number of altitude changes and number of heading changes) and multiple conditions (Developmental, ATCS and Supervisor). The experimenters focus on how experience level affects the **set** of variables (number of altitude changes **and** number of heading changes). A researcher would not do two ANOVAs for each of the DVs (number of altitude changes **and** number of heading changes) because misleading outcomes result from multiple ANOVAs. A MANOVA is more appropriate.

The basic MANOVA test results in a value called **Wilk's Lambda** (Λ) that includes the effects of more than one DV (both number of altitude changes and number of heading changes). The value of Wilk's Lambda ranges from zero to one. The lower the value of Λ, the more powerful the effect of the IV (experience level) on the set of DVs and the less likely it is that the differences between means occurred by chance. Sometimes, there is a significant difference where Wilk's Lambda is relatively high. This indicates that the effect is not that strong.

After a significant result of a MANOVA test, researchers then conduct ANOVA tests (one for number of altitude changes and one for number of heading changes). Figure N-4 depicts an example of the steps taken during a MANOVA. The example shown in Figure N-4 includes two DVs.

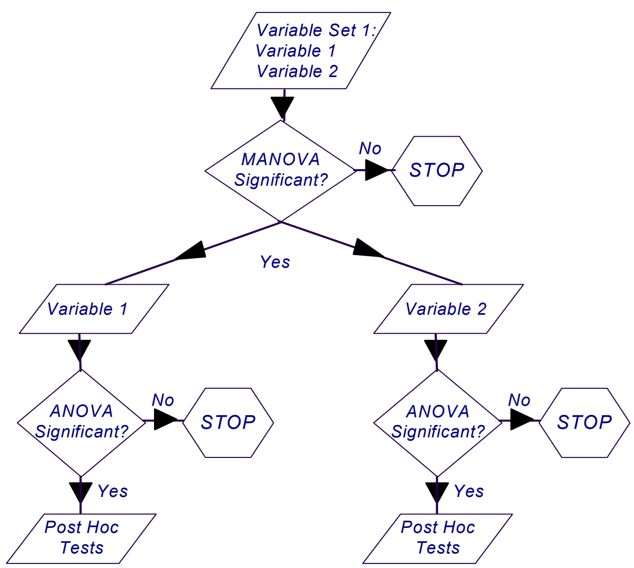

Figure N-4. Example of the steps in a Multivariate Analysis of Variance (MANOVA).

<u>SUMMARY OF A MANOVA.</u> The **multivariate ANOVA** (MANOVA) compares averages for several variables simultaneously and tests if these averages are different due to chance alone. The basic MANOVA results in a value called **Wilk's Lambda** (Λ) that includes the effects of more than one DV. The lower the value of Λ, the more powerful the effect of the IV on the set of DVs and the less likely it is that the differences between means occurred by chance. After a significant result of a MANOVA that indicates that at least two means are statistically different for the system, researchers then conduct ANOVAs.

Example: When one compares the number of altitude changes and number of heading changes between Developmental, ATCS and Supervisor at a local center, the comparison involves multiple observations of two variables. The multiple observations are the number of altitude changes and number of heading changes of each individual within each experience level. The DVs are the number of altitude changes and the number of heading changes. The three experience levels form the conditions. Figure N-5 displays the data for this example. Without looking at the individual variables, one can see that the three experience levels differ. A MANOVA would determine if chance alone caused these differences. If the differences are beyond chance (or are significant in statistical terms), the researcher then runs ANOVAs on the individual variables.

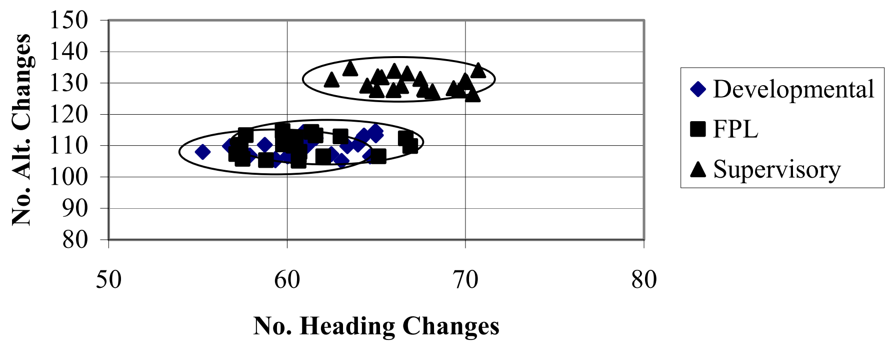

Figure N-5. Multiple observations of number of altitude changes and number of heading changes as a function of experience level.

Summary

The preceding paragraphs give some insight into the statistical methods used by researchers. Statistical methods are very powerful tools for the researcher. They tell the researcher if the experimental conditions affect the dependent measures tested. The type of statistical test that the researcher uses varies with the type of experiment. A good researcher will design experiments so they can use these techniques fully.

APPENDIX O

SAVANT SA Results

Table O-1. Current and Future SA (Correct Responses Only): Means and Standard Deviations (seconds).

	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto1	Current RT	4.88	1.16	4.22	1.22	4.55	1.22	4.74	1.26	5.27	1.06	5.00	1.17	4.81	1.19	4.74	1.24	4.78	1.21
	Future RT	5.19	1.04	5.63	1.02	5.41	1.04	5.18	0.82	5.74	0.98	5.55	0.93	5.18	0.92	5.68	0.99	5.43	0.98
Auto2	Current RT	4.87	1.17	4.91	1.07	4.89	1.10	4.76	0.73	5.20	1.69	4.98	1.30	4.81	0.96	5.06	1.40	4.94	1.20
	Future RT	4.95	0.98	5.40	0.88	5.24	0.93	5.47	1.26	5.56	0.78	5.51	1.04	5.21	1.15	5.46	0.84	5.35	0.98
Auto3	Current RT	4.72	0.93	4.49	0.71	4.61	0.82	4.74	0.66	4.84	1.27	4.79	0.99	4.73	0.79	4.66	1.02	4.70	0.91
	Future RT	4.89	1.15	7.61	2.07	5.32	1.62	5.01	1.51	5.14	1.03	5.07	1.27	4.95	1.32	5.53	1.49	5.17	1.40
Auto Coll.	Current RT	4.82	1.07	4.54	1.04	4.68	1.06	4.75	0.90	5.10	1.35	4.92	1.15	4.78	0.99	4.82	1.23	4.80	1.11
	Future RT	5.01	1.04	5.61	1.12	5.31	1.12	5.22	1.22	5.48	0.95	5.35	1.10	5.11	1.13	5.55	1.04	5.33	1.11

Table O-2. Current and Future SA (All Responses): Means and Standard Deviations (seconds).

	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	Current RT	4.97	0.81	5.04	1.55	5.00	0.98	4.97	1.35	5.42	0.71	5.19	1.09	4.97	1.10	5.27	0.96	5.10	1.03
	Current Correct	58.40	30.20	69.20	32.20	63.80	31.20	54.36	37.40	58.60	19.20	56.40	29.30	56.30	33.50	63.90	26.67	60.10	30.30
	Future RT	5.27	1.04	5.79	1.07	5.53	1.07	5.21	0.71	5.78	0.63	5.49	0.72	5.24	0.88	5.78	0.86	5.51	0.90
	Future Correct	81.10	19.70	75.20	18.60	78.10	19.10	79.30	13.40	67.80	24.80	73.60	20.40	80.20	16.60	71.53	21.90	75.90	19.80
Auto 2	Current RT	5.14	1.33	5.29	1.14	5.21	1.22	5.35	1.49	6.00	1.30	5.68	1.41	5.25	1.39	5.65	1.25	5.45	1.33
	Current Correct	69.80	20.90	66.10	31.40	68.00	26.30	54.99	37.64	63.70	36.70	59.30	36.80	62.40	30.90	64.90	33.60	63.60	32.00
	Future RT	5.02	0.62	5.49	1.15	5.25	0.94	5.37	1.32	5.49	0.85	5.43	1.10	5.19	1.03	5.49	1.00	5.34	1.01
	Future Correct	82.20	21.20	77.30	15.20	79.70	18.30	78.00	23.80	79.00	18.10	78.40	20.80	80.10	22.30	78.15	16.40	79.10	19.40
Auto 3	Current RT	4.93	0.87	4.91	0.55	4.92	0.18	4.99	0.68	4.99	1.40	4.99	1.08	4.96	0.77	4.95	1.05	4.95	0.91
	Current Correct	52.50	33.60	76.20	22.80	64.40	30.70	48.80	37.30	66.40	37.80	57.60	38.00	50.70	35.00	71.30	31.10	61.00	34.40
	Future RT	4.97	1.02	5.63	1.06	5.30	1.08	5.26	0.92	1.00	0.76	5.18	0.95	5.11	0.97	5.36	1.05	5.24	1.02
	Future Correct	77.60	28.10	77.30	19.50	77.40	23.80	76.90	24.40	76.20	20.50	76.50	22.20	77.30	25.90	76.70	19.70	77.30	22.80
Auto Coll.	Current RT	5.01	1.01	5.08	0.98	5.05	0.99	5.10	1.21	5.47	1.22	5.29	1.22	5.06	1.11	5.27	1.12	5.17	1.12
	Current Correct	60.30	29.00	70.50	28.80	65.40	29.20	52.70	36.70	62.90	31.90	57.80	34.60	56.50	33.10	66.70	30.50	61.60	32.20
	Future RT	5.08	0.90	5.64	1.08	5.36	1.03	5.28	1.00	5.45	0.87	5.37	0.94	5.18	0.95	5.54	0.98	5.36	0.98
	Future Correct	80.30	22.90	76.60	17.50	78.40	20.38	78.11	20.78	74.30	21.40	76.20	21.00	79.20	21.70	75.40	19.50	77.30	20.70

Table O-3. SAGAT type SA: Means and Standard Deviations (percent) (N=16).

N=16	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto1	Percent Answered	105.40	23.50	105.30	35.00	105.40	23.00	106.60	21.80	87.80	31.10	97.20	28.10	106.00	22.30	96.60	33.80	101.30	28.80
	Percent Correct	34.70	17.90	37.50	18.80	36.10	18.10	22.50	16.70	27.80	16.50	25.20	16.50	28.60	18.10	32.60	18.10	30.60	18.10
Auto2	Percent Answered	102.50	29.00	88.90	25.30	95.70	27.70	102.70	32.30	111.90	36.30	107.30	34.10	102.60	30.20	100.40	32.90	101.50	31.40
	Percent Correct	33.90	13.30	32.30	10.20	33.10	11.70	28.10	15.00	25.50	15.60	26.80	15.10	31.00	14.20	28.90	13.50	30.00	13.80
Auto3	Percent Answered	105.70	22.90	89.50	45.70	97.60	36.50	93.90	51.40	92.60	53.40	93.20	51.60	99.80	39.60	91.00	48.00	95.40	44.40
	Percent Correct	28.10	14.90	24.60	17.40	26.30	16.00	16.40	14.20	27.00	18.30	21.70	17.00	22.20	15.50	25.80	17.60	24.00	16.50
Auto Coll.	Percent Answered	104.50	24.80	94.60	36.30	99.60	31.30	101.10	36.80	97.40	41.90	99.20	39.30	102.80	31.30	96.00	39.00	99.40	35.40
	Percent Correct	32.20	15.50	31.50	16.50	31.90	15.90	22.40	15.70	26.80	16.50	24.60	16.20	27.30	16.30	29.10	16.60	28.20	16.40

Table O-4. SAGAT type SA: Means and Standard Deviations (percent) (N=6).

N=6	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto1	Percent Answered	120.90	23.40	127.20	40.70	124.00	31.80	123.60	20.50	103.30	21.00	113.40	22.40	122.20	21.00	115.20	33.30	118.70	27.50
	Percent Correct	41.70	16.50	53.90	20.30	47.80	18.70	26.10	14.60	27.70	19.50	26.90	16.40	33.90	17.00	40.80	23.40	37.30	20.30
Auto2	Percent Answered	112.90	34.60	100.90	32.60	106.90	32.60	99.50	35.40	122.80	42.60	111.20	39.30	106.20	34.10	111.80	37.90	109.00	35.40
	Percent Correct	37.70	7.40	27.60	5.10	32.70	8.10	25.00	17.80	35.90	11.10	30.50	15.30	31.40	14.60	31.70	9.30	3.16	12.00
Auto3	Percent Answered	119.00	29.30	108.40	21.70	113.70	25.20	101.30	75.00	117.40	24.20	109.40	53.80	110.10	55.10	112.90	22.40	111.50	41.10
	Percent Correct	34.90	18.90	33.70	13.10	34.30	15.50	14.90	16.80	33.50	8.70	24.20	16.00	24.90	20.00	33.60	10.60	29.20	16.30
Auto Coll.	Percent Answered	117.60	27.90	112.20	32.70	114.90	30.00	108.10	47.70	114.50	30.10	111.30	39.40	112.90	38.80	113.30	31.00	113.10	34.90
	Percent Correct	38.10	14.50	38.40	17.70	38.20	15.90	22.10	16.30	32.40	13.50	27.20	15.70	30.00	17.30	35.40	15.80	32.70	16.60

Table O-5. Current SA (Correct Responses Only): ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	3.206	0.547	5.864	1	15	.029
Load	0.020	2.165	0.009	1	15	.925
Automation	1.014	1.080	0.939	2	30	.402
Position X Load	4.964	1.547	3.209	1	15	.093
Position X Automation	0.618	0.637	0.969	2	30	.391
Load X Automation	0.600	1.104	0.543	2	30	.587
Position X Load X Automation	0.760	1.038	0.732	2	30	.489

Table O-6. Future SA (Correct Responses Only): ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.070	0.943	0.074	1	15	.790
Load	8.924	0.826	10.807	1	15	.005
Automation	1.032	1.096	0.942	2	30	.401
Position X Load	1.403	0.897	1.563	1	15	.230
Position X Automation	1.077	1.164	0.925	2	30	.408
Load X Automation	0.144	1.021	0.141	2	30	.869
Position X Load X Automation	0.689	1.353	0.509	2	30	.606

Table O-7. Current SA (All Responses): MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Position	.610	0.390	0.639	0.639	4.476	2	14	.031
Load	.628	0.372	0.593	0.593	4.154	2	14	.038
Automation	.639	0.361	0.564	0.564	1.693	4	12	.216
Position X Load	.956	0.044	0.046	0.046	0.324	2	14	.728
Position X Automation	.889	0.111	0.125	0.125	0.376	4	12	.822
Load X Automation	.639	0.361	0.566	0.566	1.697	4	12	.215
Position X Load X Automation	.841	0.159	0.189	0.189	0.568	4	12	.691

Table O-8. Response Time for Current SA (All Responses): ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	2.800	0.507	5.520	1	15	.033
Load	2.248	1.328	1.693	1	15	.213
Automation	4.087	1.028	3.977	2	30	.029
Position X Load	1.093	2.055	0.532	2	30	.477
Position X Automation	0.635	0.959	0.662	2	30	.523
Load X Automation	0.685	1.431	0.479	2	30	.624
Position X Load X Automation	0.239	0.980	0.243	2	30	.785

Table O-9. Percentage Correct for Current SA (All Responses): ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.277	0.076	3.622	1	15	.076
Load	0.502	0.121	4.145	1	15	.060
Automation	0.022	0.131	0.165	2	30	.849
Position X Load	0.000	0.160	0.000	2	30	.998
Position X Automation	0.001	0.074	0.020	2	30	.980
Load X Automation	0.140	0.112	1.256	2	30	.299
Position X Load X Automation	0.047	0.092	0.513	2	30	.604

Table O-10. Future SA (All Responses): MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Position	.938	0.062	0.066	0.066	0.464	2	14	.638
Load	.620	0.380	0.612	0.612	4.285	2	14	.035
Automation	.778	0.222	0.285	0.285	0.856	4	12	.517
Position X Load	.845	0.155	0.183	0.183	1.282	2	14	.308
Position X Automation	.899	0.101	0.113	0.113	0.338	4	12	.847
Load X Automation	.862	0.138	0.160	0.160	0.479	4	12	.751
Position X Load X Automation	.839	0.161	0.192	0.192	0.479	4	12	.751

Table O-11. Response Time for Future SA (All Responses): ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.000	0.589	0.000	1	15	.983
Load	6.078	0.803	7.570	1	15	.015
Automation	1.104	0.841	1.313	2	30	.284
Position X Load	1.600	0.660	2.427	2	30	.140
Position X Automation	0.440	0.632	0.696	2	30	.506
Load X Automation	0.456	0.707	0.645	2	30	.532
Position X Load X Automation	0.690	1.107	0.623	2	30	.543

Table O-12. Percentage Future for Current SA (All Responses): ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.043	0.043	0.989	1	15	.336
Load	0.097	0.036	2.714	1	15	.120
Automation	0.018	0.052	0.341	2	30	.713
Position X Load	0.003	0.061	0.045	2	30	.834
Position X Automation	0.005	0.052	0.087	2	30	.917
Load X Automation	0.022	0.032	0.674	2	30	.517
Position X Load X Automation	0.015	0.029	0.532	2	30	.593

Table O-13. SAGAT type Queries: MANOVA Results (N=16).

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Position	.446	0.554	1.242	1.242	8.696	2	14	.004
Load	.694	0.306	0.440	0.440	3.080	2	14	.078
Automation	.637	0.363	0.569	0.569	1.708	4	12	.213
Position X Load	.943	0.057	0.060	0.060	0.421	2	14	.664
Position X Automation	.741	0.259	0.350	0.350	1.050	4	12	.422
Load X Automation	.760	0.240	0.316	0.316	0.948	4	12	.470
Position X Load X Automation	.503	0.497	0.989	0.989	2.968	4	12	.064

Table O-14. Percent Questions Answered: ANOVA Results (N=16).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.000	0.132	0.004	1	15	.953
Load	0.223	0.061	3.631	1	15	.076
Automation	0.076	0.091	0.830	2	30	.446
Position X Load	0.048	0.142	0.335	1	15	.571
Position X Automation	0.176	0.096	1.835	2	30	.177
Load X Automation	0.025	0.085	0.298	2	30	.745
Position X Load X Automation	0.194	0.117	1.667	2	30	.206

Table O-15. Percent Correct: ANOVA Results (N=16).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.255	0.019	13.405	1	15	.002
Load	0.016	0.010	1.545	1	15	.233
Automation	0.086	0.018	4.699	2	30	.017
Position X Load	0.032	0.036	0.901	1	15	.358
Position X Automation	0.017	0.021	0.811	2	30	.454
Load X Automation	0.019	0.022	0.837	2	30	.443
Position X Load X Automation	0.025	0.025	0.993	2	30	.382

Table O-16. SAGAT type Queries: MANOVA Results (N=6).

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Position	.201	0.799	3.969	3.969	7.939	2	4	.040
Load	.523	0.477	0.912	0.912	1.823	2	4	.274
Automation	.201	0.799	3.987	3.987	1.993	4	2	.361
Position X Load	.847	0.153	0.181	0.181	0.362	2	4	.717
Position X Automation	.205	0.795	3.879	3.879	1.939	4	2	.368
Load X Automation	.102	0.898	8.812	8.812	4.406	4	2	.193
Position X Load X Automation	.022	0.978	43.673	43.673	21.836	4	2	.044

Table O-17. Percent Questions Answered: ANOVA Results (N=6).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.023	0.222	0.103	1	5	.762
Load	0.000	0.087	0.005	1	5	.948
Automation	0.061	0.113	0.541	2	10	.599
Position X Load	0.062	0.127	0.485	1	5	.517
Position X Automation	0.033	0.150	0.222	2	10	.804
Load X Automation	0.026	0.060	0.434	2	10	.660
Position X Load X Automation	0.168	0.058	2.883	2	10	.103

Table O-18. Percent Correct: ANOVA Results (N=6).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.220	0.013	16.970	1	5	.009
No Automation	0.525	0.006	91.380	1	5	.000
Limited Automation	0.006	0.024	0.242	1	5	.644
Full Automation	0.122	0.030	4.046	1	5	.100
Load	0.051	0.013	4.078	1	5	.099
Automation	0.042	0.011	3.778	2	10	.060
D-Side	0.024	0.014	1.698	2	10	.232
R-side	0.165	0.025	6.604	2	10	.015
Position x Load	0.046	0.051	0.893	1	5	.388
Position x Automation	0.053	0.008	6.230	2	10	.017
Load x Automation	0.011	0.024	0.485	2	10	.630
Position x Load x Automation	0.048	0.026	1.861	2	10	.206

APPENDIX P

SAGAT SA Results

Table P-1. Level 1 SA Items: Means and Standard Deviations (percent).

Level 1	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Automation 1	Aircraft Position	25.68	9.56	27.23	8.85	26.46	9.10	23.20	13.07	27.60	8.95	25.47	11.25	24.46	11.34	27.45	8.76	25.96	10.17
	Callsign Letter	30.95	39.76	11.81	24.81	21.48	33.22	21.40	33.76	2.00	2.76	11.81	24.41	26.25	36.43	6.92	17.65	16.67	29.08
	Callsign Number	0.20	0.28	0.20	0.28	0.20	0.27	0.20	0.28	0.20	0.28	0.20	0.27	0.20	0.27	0.20	0.27	0.20	0.27
	Altitude	27.77	33.03	18.22	24.94	23.02	28.97	35.40	38.43	26.10	33.60	30.84	35.76	31.64	35.39	22.19	29.18	26.95	32.39
	Speed	0.79	0.77	0.79	0.77	0.79	0.75	0.40	0.72	0.70	0.77	0.63	0.73	0.63	0.73	0.79	0.75	0.71	0.74
	Heading	64.42	46.54	56.61	45.29	60.59	45.28	78.10	46.30	56.60	45.29	68.10	46.05	71.61	46.10	56.61	44.56	64.42	45.41
	Vertical Change	83.92	36.69	82.93	41.24	83.43	38.40	87.10	39.33	87.10	43.14	87.13	40.65	85.57	37.61	85.10	41.59	85.34	39.36
	Turning	95.51	32.44	98.51	24.75	97.21	28.68	97.10	24.87	97.10	33.59	97.11	29.20	96.35	28.80	97.87	29.11	97.16	28.73
	Aircraft Type	15.78	24.50	15.78	24.50	15.78	24.10	12.70	24.83	14.70	24.62	13.79	24.33	14.29	24.28	15.28	24.16	14.79	24.03
	Level of Control	0.60	1.13	3.96	4.53	2.28	3.67	10.40	24.88	3.90	4.53	7.18	17.66	5.51	17.62	3.96	4.46	4.73	12.75
Automation 2	Aircraft Position	30.79	15.18	26.54	7.30	28.68	11.88	23.50	9.83	19.60	8.31	21.63	9.16	27.20	13.05	23.13	8.44	25.17	11.08
	Callsign Letter	11.81	24.81	13.14	24.64	12.47	24.33	11.80	24.81	3.30	3.28	7.59	17.62	11.81	24.41	8.26	17.58	10.04	21.13
	Callsign Number	0.20	0.28	0.34	0.33	0.27	0.30	0.20	0.28	0.30	0.33	0.27	0.30	0.20	0.27	0.34	0.32	0.27	0.30
	Altitude	29.40	32.68	13.56	9.02	21.55	24.16	37.00	38.43	38.60	37.94	37.84	37.57	33.25	35.19	26.32	28.40	29.80	31.81
	Speed	0.79	0.77	1.41	0.82	1.10	0.81	0.90	0.77	0.90	0.80	0.94	0.77	0.86	0.75	1.18	0.81	1.02	0.79
	Heading	56.61	45.29	46.13	38.46	51.47	41.52	39.40	39.66	54.60	41.20	47.20	40.17	48.26	42.36	50.43	39.33	49.35	40.56
	Vertical Change	74.44	41.11	68.85	37.16	71.70	38.63	79.50	44.10	77.00	42.71	78.31	42.72	77.05	42.02	73.08	39.57	75.10	40.53
	Turning	98.24	24.60	97.64	24.17	97.95	24.01	98.70	24.87	93.20	38.62	96.55	32.67	98.51	24.35	95.73	32.08	97.30	28.51
	Aircraft Type	6.03	5.12	26.35	32.78	16.28	24.03	5.00	4.90	51.50	45.85	29.14	35.79	5.52	4.94	39.30	40.20	22.76	30.53
	Level of Control	1.12	2.24	3.46	4.39	2.29	3.62	1.10	2.24	3.40	4.39	2.29	3.62	1.12	2.20	3.46	4.32	2.29	3.59
Automation 3	Aircraft Position	25.31	10.22	26.25	7.70	25.78	8.92	20.10	8.63	22.10	10.15	21.15	9.31	22.75	9.66	24.19	9.11	23.47	9.34
	Callsign Letter	3.36	3.28	2.69	3.06	3.03	3.13	13.10	24.64	2.60	3.06	7.93	17.60	8.26	17.58	2.69	3.01	5.48	12.64
	Callsign Number	10.14	24.95	0.27	0.31	5.21	17.65	0.30	0.33	0.20	0.31	0.30	0.31	5.24	17.64	0.27	0.30	2.76	12.48
	Altitude	31.02	31.86	27.77	33.03	29.40	31.97	8.40	8.36	27.70	33.03	18.22	24.54	19.89	24.27	27.77	32.49	23.85	28.55
	Speed	10.58	24.88	10.74	24.85	10.66	24.46	0.70	0.77	1.00	0.82	0.94	0.79	5.69	17.60	5.93	17.58	5.81	17.45
	Heading	27.93	28.16	67.21	45.06	48.82	39.77	37.20	34.32	74.10	45.05	57.14	42.04	32.61	31.05	70.76	44.44	53.04	40.72
	Vertical Change	84.89	31.44	88.01	38.65	86.49	34.71	82.90	41.24	56.70	37.49	71.07	40.61	83.92	36.09	74.44	40.44	79.42	38.36
	Turning	97.31	23.90	93.26	38.62	95.51	31.92	98.50	24.75	97.90	24.41	98.24	24.20	97.95	24.01	95.94	32.24	97.03	28.33
	Aircraft Type	16.77	24.35	13.79	24.74	15.28	24.16	6.00	5.12	5.00	4.90	5.52	4.94	11.42	17.66	9.42	17.77	10.42	17.58
	Level of Control	1.52	2.27	3.66	4.33	2.59	3.57	1.50	2.27	3.60	4.33	2.59	3.57	1.52	2.23	3.66	4.26	2.59	3.54
Automation Collapsed	Aircraft Position	27.27	11.91	26.68	7.82	26.97	10.02	22.33	10.55	23.17	9.58	22.75	10.03	24.81	11.45	24.93	8.88	24.87	10.22
	Callsign Letter	15.49	27.60	9.22	20.06	12.36	24.08	15.49	27.60	2.69	3.00	9.11	19.96	15.49	27.46	5.96	14.42	10.74	22.08
	Callsign Number	3.52	14.41	0.27	0.30	1.89	10.19	0.25	0.29	0.27	0.30	0.26	0.29	1.88	10.19	0.27	0.30	1.08	7.21
	Altitude	29.40	31.86	19.89	24.24	24.67	28.36	27.22	32.32	30.90	34.42	29.07	33.22	28.31	31.93	25.43	29.81	26.88	30.83
	Speed	4.06	14.38	4.32	14.36	4.19	14.29	0.73	0.73	0.94	0.78	0.84	0.76	2.40	10.18	2.63	10.17	2.51	10.15
	Heading	50.41	41.59	56.97	42.67	53.73	41.99	53.21	42.37	62.22	43.59	57.80	42.90	51.81	41.78	59.62	42.95	55.78	42.37
	Vertical Change	81.33	36.20	80.62	39.14	80.98	37.50	83.34	41.05	74.92	42.05	79.32	41.64	82.35	38.53	77.85	40.51	80.16	39.53
	Turning	97.13	26.84	96.81	29.79	96.97	28.21	98.19	24.30	96.35	32.36	97.35	28.68	97.69	25.58	96.59	30.95	97.16	28.38
	Aircraft Type	12.87	19.97	18.67	27.26	15.78	23.84	7.95	14.76	24.41	32.53	16.24	25.69	10.42	17.54	21.55	29.91	16.01	24.72
	Level of Control	1.08	1.94	3.70	4.33	2.39	3.59	4.35	14.42	3.70	4.33	4.02	10.59	2.72	10.29	3.70	4.31	3.21	7.90

Table P-2. Level 2 SA: Means and Standard Deviations (percent).

Level 2	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Automation 1	Next Fix	52.90	45.74	52.90	45.74	52.90	45.00	25.90	33.71	52.90	45.74	39.88	40.66	39.88	40.66	52.90	45.00	46.52	42.80
	Flight Profile	86.09	39.61	83.56	38.68	84.85	38.54	71.20	44.27	47.50	40.36	60.07	42.15	79.26	42.79	67.62	42.01	73.71	42.25
	Separation	84.10	41.19	71.29	37.83	78.11	39.47	84.10	41.19	71.90	43.51	78.39	42.16	84.10	40.52	71.61	40.11	78.25	40.51
	Open Clearance	45.38	36.13	27.24	23.24	36.49	30.18	16.40	13.88	27.80	20.88	22.20	18.52	31.27	28.74	27.57	21.74	29.42	25.29
	Clearance Received	70.74	39.11	91.30	21.01	82.38	33.09	51.20	28.66	83.10	36.79	68.90	35.24	61.46	34.54	87.54	29.98	76.05	34.51
	Clearance Conformance	61.36	38.21	89.66	22.34	77.47	34.16	48.20	27.42	75.10	36.97	62.64	34.93	54.97	33.02	83.12	30.51	70.44	34.61
	In Communication	17.19	24.57	20.86	23.91	19.03	23.88	36.30	38.17	23.70	25.17	30.12	32.03	26.93	32.28	22.28	24.20	24.61	28.33
	Flight Plan Conformance	99.22	24.75	99.04	33.58	99.13	29.21	99.20	24.82	97.20	33.46	98.42	29.21	99.26	24.39	98.24	32.98	98.80	28.98
Automation 2	Next Fix	16.40	25.52	20.76	25.45	18.58	25.12	35.30	39.22	20.70	25.45	28.13	32.90	25.99	33.17	20.76	25.04	23.38	29.20
	Flight Profile	47.58	40.36	83.56	38.68	67.62	42.01	80.50	45.51	77.90	35.90	79.26	40.35	65.65	44.66	80.84	36.84	73.71	41.23
	Separation	91.30	26.47	77.24	30.45	85.03	29.37	85.00	24.40	85.50	36.55	85.27	30.62	88.36	25.15	81.58	33.41	85.15	29.78
	Open Clearance	37.56	32.73	39.88	27.90	38.72	30.06	39.00	31.87	24.00	19.28	31.66	26.26	38.29	31.79	32.10	24.43	35.21	28.13
	Clearance Received	74.59	36.75	85.43	37.90	80.34	37.18	80.70	35.43	83.90	31.62	82.38	33.09	77.77	35.65	84.69	34.34	81.37	34.93
	Clearance Conformance	72.50	36.62	72.06	40.91	72.28	38.19	78.90	35.55	82.00	31.95	80.48	33.29	75.79	35.64	77.27	36.49	76.54	35.79
	In Communication	18.92	25.70	13.18	12.71	16.06	20.04	22.50	26.56	20.60	23.91	21.57	24.87	20.70	25.72	16.94	18.96	18.82	22.46
	Flight Plan Conformance	99.30	24.82	99.92	2.73	99.68	17.57	99.90	2.45	99.80	2.80	99.93	2.64	99.73	17.61	99.90	2.73	99.83	12.54
Automation 3	Next Fix	39.45	38.20	18.58	25.53	29.19	32.74	56.60	44.02	18.50	25.53	38.42	37.86	48.26	41.04	18.58	25.12	33.85	35.25
	Flight Profile	80.99	32.68	71.24	40.83	76.33	36.69	60.40	32.83	60.40	32.83	60.48	32.29	71.51	33.60	66.03	36.75	68.82	35.01
	Separation	94.86	18.68	80.02	38.71	88.57	31.65	93.60	18.23	79.40	35.43	87.51	30.02	94.27	18.19	79.74	36.57	88.05	30.65
	Open Clearance	31.83	25.90	26.84	25.44	29.34	25.44	33.40	32.34	36.50	32.29	35.00	31.84	32.63	28.83	31.74	28.89	32.18	28.64
	Clearance Received	81.58	30.13	83.94	31.62	82.78	30.41	63.40	36.90	77.30	38.51	70.74	37.37	73.17	33.84	80.76	34.86	77.10	34.20
	Clearance Conformance	84.88	27.42	82.01	31.95	83.47	29.34	61.30	36.53	71.10	36.62	66.39	36.11	74.28	33.27	76.85	34.18	75.58	33.47
	In Communication	23.16	25.61	12.14	8.29	17.68	19.11	9.90	7.20	17.90	24.38	13.96	17.89	16.58	19.20	15.06	17.97	15.82	18.46
	Flight Plan Conformance	99.94	2.62	99.89	2.80	99.92	2.69	99.10	2.73	97.20	2.82	98.30	2.76	99.66	2.64	98.99	2.77	99.37	2.71
Automation Collapsed	Next Fix	36.71	38.07	31.25	34.67	33.99	36.27	39.73	39.45	31.25	34.67	35.53	37.05	38.23	38.58	31.25	34.48	34.76	36.57
	Flight Profile	73.50	39.88	79.79	39.02	76.74	39.37	71.24	41.21	62.85	37.21	67.16	39.14	72.38	40.39	71.86	38.72	72.12	39.46
	Separation	90.56	30.33	76.30	35.26	84.16	33.80	88.01	29.49	79.29	38.32	83.92	34.62	89.32	29.78	77.82	36.64	84.04	34.12
	Open Clearance	38.33	31.34	31.39	25.76	34.88	28.57	29.77	27.71	29.56	24.44	29.66	26.02	34.08	29.61	30.47	24.99	32.28	27.32
	Clearance Received	75.82	35.07	87.09	30.62	81.85	33.34	66.08	34.64	81.58	35.14	74.33	35.33	71.12	34.86	84.44	32.95	78.23	34.44
	Clearance Conformance	73.69	34.98	81.88	32.98	77.95	34.07	63.73	34.25	76.29	34.80	70.29	34.90	68.87	34.63	79.17	33.77	74.24	34.50
	In Communication	19.77	24.86	15.41	16.13	17.59	20.90	23.08	27.57	20.79	24.06	21.93	25.74	21.43	26.13	18.10	20.44	19.77	23.42
	Flight Plan Conformance	99.57	20.05	99.71	19.94	99.65	19.89	99.57	14.42	98.49	19.87	99.11	17.34	99.57	17.40	99.22	19.80	99.41	18.61

Table P-3. Level 3 SA: Means and Standard Deviations (percent).

Level 3	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	Handoff Needed	20.21	23.11	8.52	13.43	14.39	19.43	13.30	19.50	17.10	16.43	15.26	17.86	16.80	21.31	12.86	15.39	14.83	18.52
	Clearance Needed	10.11	15.23	14.81	15.44	12.47	15.17	41.50	38.43	28.70	32.87	35.22	35.53	26.18	30.61	21.82	25.56	24.01	28.03
Auto 2	Handoff Needed	8.78	13.67	10.63	9.79	9.70	11.73	19.00	26.86	10.80	10.43	14.95	20.20	13.93	21.24	10.74	9.95	12.33	16.48
	Clearance Needed	27.29	34.34	27.21	27.29	27.25	30.54	29.30	34.11	25.30	24.98	27.36	29.42	28.31	33.69	26.29	25.75	27.30	29.74
Auto 3	Handoff Needed	20.75	24.59	21.46	29.93	21.10	26.95	17.30	24.96	20.00	25.54	18.68	24.86	19.05	24.44	20.74	27.39	19.89	25.75
	Clearance Needed	26.07	27.29	19.62	25.75	22.86	26.28	32.30	33.13	16.30	16.62	24.42	26.19	29.22	29.90	17.99	21.32	23.64	26.04
Auto Coll.	Handoff Needed	16.60	21.11	13.56	19.72	15.09	20.38	16.58	23.49	16.02	18.36	16.30	20.97	16.59	22.22	14.79	18.99	15.69	20.62
	Clearance Needed	21.22	26.66	20.58	23.22	20.90	24.87	34.46	34.77	23.51	25.30	29.03	30.43	27.91	31.12	22.05	24.18	24.99	27.85

Table P-4. Aircraft Position: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.091	0.009	9.917	1	15	.007
Load	0.000	0.016	0.004	1	15	.948
Automation	0.011	0.011	1.033	2	30	.368
Position X Load	0.003	0.009	0.286	1	15	.601
Position X Automation	0.016	0.012	1.287	2	30	.291
Load X Automation	0.024	0.011	2.136	2	30	.136
Position X Load X Automation	0.001	0.007	0.103	2	30	.903

Table P-5. Callsign Letter: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.051	0.102	0.503	1	15	.489
Load	0.442	0.083	5.340	1	15	.035
Automation	0.205	0.110	1.873	2	30	.171
Position X Load	0.051	0.104	0.495	1	15	.492
Position X Automation	0.090	0.061	1.484	2	30	.243
Load X Automation	0.123	0.141	0.873	2	30	.428
Position X Load X Automation	0.013	0.144	0.089	2	30	.915

Table P-6. Callsign Number: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.013	0.012	1.029	1	15	.327
Load	0.013	0.013	0.971	1	15	.340
Automation	0.014	0.013	1.052	2	30	.362
Position X Load	0.013	0.013	0.998	1	15	.334
Position X Automation	0.013	0.013	1.007	2	30	.377
Load X Automation	0.014	0.013	1.034	2	30	.368
Position X Load X Automation	0.013	0.013	1.020	2	30	.373

Table P-7. Altitude: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.100	0.201	0.498	1	15	.491
Load	0.043	0.329	0.130	1	15	.723
Automation	0.061	0.175	0.349	2	30	.708
Position X Load	0.224	0.492	0.455	1	15	.510
Position X Automation	0.342	0.246	1.394	2	30	.264
Load X Automation	0.151	0.159	0.948	2	30	.399
Position X Load X Automation	0.059	0.170	0.345	2	30	.711

Table P-8. Speed: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.054	0.024	2.213	1	15	.158
Load	0.000	0.027	0.010	1	15	.922
Automation	0.052	0.024	2.221	2	30	.126
Position X Load	0.000	0.027	0.000	1	15	.991
Position X Automation	0.049	0.023	2.093	2	30	.141
Load X Automation	0.000	0.027	0.000	2	30	1.000
Position X Load X Automation	0.000	0.028	0.004	2	30	.996

Table P-9. Heading: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.116	0.526	0.220	1	15	.646
Load	0.425	0.364	1.167	1	15	.297
Automation 1	0.308	0.237	1.304	1	15	.271
Automation 2	0.005	0.243	0.020	1	15	.888
Automation 3	1.648	0.172	9.562	1	15	.007
Automation	0.592	0.404	1.466	2	30	.247
Low Load	0.889	0.147	6.037	2	30	.006
High Load	0.282	0.289	0.974	2	30	.389
Position x Load	0.013	0.380	0.034	1	15	.857
Position x Automation	0.116	0.425	0.272	2	30	.764
Load x Automation	1.749	0.469	3.727	2	30	.036
Position x Load x Automation	0.244	0.428	0.570	2	30	.571

Table P-10. Vertical Change: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.037	0.535	0.069	1	15	.796
Load	0.271	0.374	0.725	1	15	.408
Automation	0.485	0.565	0.858	2	30	.434
Position X Load	0.191	0.541	0.353	1	15	.561
Position X Automation	0.620	0.398	1.556	2	30	.227
Load X Automation	0.089	0.286	0.313	2	30	.734
Position X Load X Automation	0.290	0.276	1.051	2	30	.362

Table P-11. Awareness of Aircraft Turning: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.012	0.300	0.041	1	15	.843
Load	0.104	0.153	0.682	1	15	.422
Automation	0.002	0.227	0.009	2	30	.991
Position X Load	0.055	0.211	0.259	1	15	.618
Position X Automation	0.126	0.253	0.498	2	30	.613
Load X Automation	0.152	0.176	0.868	2	30	.430
Position X Load X Automation	0.094	0.162	0.579	2	30	.566

Table P-12. Aircraft Type: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.001	0.076	0.014	1	15	.909
Load	0.611	0.194	3.145	1	15	.096
Automation 1	0.001	0.150	0.005	1	15	.942
Automation 2	0.972	0.072	13.450	1	15	.002
Automation 3	0.003	0.049	0.065	1	15	.802
Automation	0.258	0.206	1.251	2	30	.301
Low Load	0.032	0.068	0.473	2	30	.627
High Load	0.432	0.122	3.538	2	30	.042
Position x Load	0.141	0.068	2.065	1	15	.171
Position x Automation	0.220	0.123	1.792	2	30	.184
Load x Automation	0.671	0.174	3.849	2	30	.033
Position x Load x Automation	0.093	0.116	0.802	2	30	.458

Table P-13. Level of Control: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.013	0.012	1.043	1	15	.323
Load	0.005	0.013	0.351	1	15	.563
Automation	0.011	0.012	0.927	2	30	.407
Position X Load	0.013	0.016	0.815	1	15	.381
Position X Automation	0.013	0.016	0.818	2	30	.451
Load X Automation	0.008	0.015	0.511	2	30	.605
Position X Load X Automation	0.013	0.013	1.008	2	30	.377

Table P-14. Next Fix: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.013	0.479	0.027	1	15	.872
Load	0.266	0.432	0.615	1	15	.445
Automation	0.987	0.239	4.123	2	30	.026
Position X Load	0.013	0.195	0.066	1	15	.801
Position X Automation	0.321	0.290	1.108	2	30	.343
Load X Automation	0.866	0.300	2.889	2	30	.071
Position X Load X Automation	0.321	0.327	0.982	2	30	.386

Table P-15. Flight Profile: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.920	0.173	5.310	1	15	.036
Load	0.003	0.400	0.007	1	15	.935
Automation	0.104	0.424	0.246	2	30	.784
Position X Load	0.537	0.309	1.735	1	15	.208
Position X Automation	1.250	0.414	3.017	2	30	.064
Load X Automation	0.688	0.342	2.012	2	30	.151
Position X Load X Automation	0.477	0.407	1.170	2	30	.324

Table P-16. Awareness of Aircraft Separation: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.001	0.276	0.003	1	15	.954
Load	2.172	0.425	5.112	1	15	.039
Automation	0.529	0.221	2.390	2	30	.109
Position X Load	0.131	0.189	0.693	1	15	.418
Position X Automation	0.004	0.165	0.024	2	30	.977
Load X Automation	0.129	0.211	0.610	2	30	.550
Position X Load X Automation	0.090	0.443	0.203	2	30	.817

Table P-17. Open Clearances: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.146	0.170	0.857	1	15	.369
Load	0.070	0.115	0.607	1	15	.448
Automation	0.060	0.111	0.539	2	30	.589
Position X Load	0.062	0.105	0.590	1	15	.454
Position X Automation	0.180	0.224	0.806	2	30	.456
Load X Automation	0.013	0.136	0.095	2	30	.910
Position X Load X Automation	0.246	0.193	1.277	2	30	.294

Table P-18. Clearances Received: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.700	0.457	1.533	1	15	.235
Load	2.204	0.301	7.322	1	15	.016
Automation	0.135	0.315	0.429	2	30	.655
Position X Load	0.015	0.215	0.072	1	15	.793
Position X Automation	0.293	0.225	1.302	2	30	.287
Load X Automation	0.434	0.203	2.137	2	30	.136
Position X Load X Automation	0.084	0.257	0.328	2	30	.723

Table P-19. Clearance Conformance: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.629	0.463	1.359	1	15	.262
Load	1.136	0.395	2.879	1	15	.110
Automation 1	1.275	0.138	9.269	1	15	.008
Automation 2	0.004	0.129	0.033	1	15	.859
Automation 3	0.088	0.157	0.559	1	15	.466
Automation	0.149	0.268	0.557	2	30	.579
Low Load	0.381	0.113	3.383	2	30	.047
High Load	0.055	0.114	0.484	2	30	.621
Position x Load	0.025	0.214	0.118	1	15	.736
Position x Automation	0.714	0.253	2.824	2	30	.075
Load x Automation	0.724	0.186	3.900	2	30	.031
Position x Load x Automation	0.083	0.329	0.253	2	30	.778

Table P-20. In Communication: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.094	0.121	0.777	1	15	.392
Load	0.055	0.091	0.604	1	15	.449
Automation	0.133	0.156	0.854	2	30	.436
Position X Load	0.005	0.072	0.072	1	15	.793
Position X Automation	0.094	0.143	0.661	2	30	.524
Load X Automation	0.004	0.085	0.053	2	30	.949
Position X Load X Automation	0.133	0.175	0.761	2	30	.476

Table P-21. Flight Plan Conformance: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.116	0.145	0.800	1	15	.385
Load	0.050	0.149	0.333	1	15	.572
Automation	0.150	0.098	1.523	2	30	.234
Position X Load	0.116	0.131	0.883	1	15	.362
Position X Automation	0.139	0.099	1.407	2	30	.261
Load X Automation	0.045	0.087	0.514	2	30	.603
Position X Load X Automation	0.000	0.116	0.002	2	30	.998

Table P-22. Level 3 SA: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Position	.922	0.078	0.084	0.084	0.591	2	14	.567
Load	.918	0.082	0.089	0.089	0.625	2	14	.550
Automation	.872	0.128	0.147	0.147	0.440	4	12	.777
Position X Load	.906	0.094	0.104	0.104	0.728	2	14	.500
Position X Automation	.673	0.327	0.486	0.486	1.458	4	12	.275
Load X Automation	.953	0.047	0.050	0.050	0.149	4	12	.960
Position X Load X Automation	.663	0.337	0.509	0.509	1.526	4	12	.256

Table P-23. Handoff Needed: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.007	0.050	0.145	1	15	.709
Load	0.016	0.053	0.299	1	15	.592
Automation	0.098	0.092	1.067	2	30	.357
Position X Load	0.008	0.157	0.048	1	15	.829
Position X Automation	0.024	0.058	0.416	2	30	.663
Load X Automation	0.015	0.057	0.271	2	30	.764
Position X Load X Automation	0.067	0.113	0.588	2	30	.562

Table P-24. Clearances Needed: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.339	0.275	1.231	1	15	.285
Load	0.176	0.133	1.319	1	15	.269
Automation	0.028	0.135	0.206	2	30	.815
Position X Load	0.140	0.095	1.477	1	15	.243
Position X Automation	0.274	0.161	1.698	2	30	.200
Load X Automation	0.039	0.242	0.160	2	30	.853
Position X Load X Automation	0.021	0.151	0.140	2	30	.870

APPENDIX Q

PSQ SA Results

Table Q-1. PSQ SA: Means and Standard Deviations.

SA	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	10) Current Aircraft Locations	7.37	1.62	6.62	1.62	7.00	1.58	6.50	2.15	6.12	2.09	6.31	2.09	6.94	1.93	6.37	1.85	6.65	1.90
	11) Projected Aircraft Locations	7.08	1.75	6.52	1.66	6.80	1.71	6.95	2.10	6.33	2.24	6.64	2.16	7.02	1.91	6.42	1.94	6.72	1.93
	12) Potential Violations	7.48	1.56	6.48	2.00	6.98	1.83	6.79	2.37	6.29	2.05	6.54	2.20	7.14	2.00	6.39	1.99	6.76	2.02
Auto 2	10) Current Aircraft Locations	7.09	1.75	6.62	1.23	6.86	1.51	6.47	2.24	5.75	1.90	6.11	2.07	6.78	2.00	6.18	1.63	6.48	1.84
	11) Projected Aircraft Locations	7.43	1.40	6.72	1.37	7.07	1.41	6.43	2.41	6.22	1.96	6.32	2.17	6.93	2.01	6.47	1.69	6.70	1.85
	12) Potential Violations	7.53	1.87	6.90	1.29	7.21	1.61	6.78	2.31	6.71	1.84	6.75	2.06	7.16	2.11	6.80	1.57	6.98	1.85
Auto 3	10) Current Aircraft Locations	7.47	1.42	6.00	2.02	6.73	1.88	6.56	2.09	5.06	2.04	5.81	2.01	7.01	1.82	5.53	2.05	6.27	2.06
	11) Projected Aircraft Locations	7.68	1.24	5.89	1.85	6.78	1.80	7.14	1.53	5.45	2.13	6.30	2.01	7.41	1.40	5.67	1.98	6.54	1.91
	12) Potential Violations	7.72	1.36	5.98	2.06	6.85	1.93	7.92	1.67	5.54	2.30	6.73	2.31	7.82	1.50	5.76	2.16	6.79	2.11
Auto Coll.	10) Current Aircraft Locations	7.31	1.57	6.41	1.65	6.86	1.67	6.51	2.12	5.64	2.01	6.08	2.10	6.91	1.90	6.03	1.87	6.47	1.93
	11) Projected Aircraft Locations	7.40	1.47	6.37	1.65	6.88	1.64	6.84	2.03	6.00	2.10	6.42	2.10	7.12	1.78	6.19	1.89	6.65	1.89
	12) Potential Violations	7.58	1.58	6.45	1.81	7.01	1.78	7.16	2.16	6.18	2.08	6.67	2.17	7.37	1.90	6.32	1.95	6.84	1.99

Table Q-2. Situational Awareness: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Position	.523	0.477	0.912	0.912	3.953	3	13	.033
Load	.318	0.682	2.146	2.146	9.300	3	13	.002
Automation	.495	0.505	1.021	1.021	1.701	6	10	.218
Position X Load	.968	0.032	0.033	0.033	0.143	3	13	.932
Position X Automation	.478	0.522	1.091	1.091	1.819	6	10	.192
Load X Automation	.475	0.525	1.106	1.106	1.844	6	10	.187
Position X Load X Automation	.691	0.309	0.447	0.447	0.745	6	10	.627

Table Q-3. Situational Awareness for Current Aircraft: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	29.673	2.620	11.326	1	15	.004
Load	37.206	3.123	11.913	1	15	.004
Automation	2.355	2.400	0.981	2	30	.386
Position X Load	0.011	3.093	0.004	1	15	.952
Position X Automation	0.234	1.513	0.155	2	30	.857
Load X Automation	4.372	2.352	1.859	2	30	.173
Position X Load X Automation	0.403	4.149	0.097	2	30	.908

Table Q-4. Situational Awareness for Projected Aircraft: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	1.370	2.379	4.359	1	15	.054
Load	41.607	3.259	12.766	1	15	.003
Automation	0.607	2.253	0.270	2	30	.766
Position X Load	0.387	2.537	0.153	1	15	.702
Position X Automation	1.417	1.553	0.912	2	30	.412
Load X Automation	7.888	3.049	2.587	2	30	.092
Position X Load X Automation	0.335	2.810	0.119	2	30	.888

Table Q-5. Situational Awareness for Potential Violations: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	5.616	3.617	1.553	1	15	.232
Load	53.303	2.400	22.209	1	15	.000
Automation	0.907	2.495	0.363	2	30	.698
Position X Load	0.243	3.470	0.070	1	15	.795
Position X Automation	0.596	1.227	0.486	2	30	.620
Load X Automation	12.714	3.087	4.119	2	30	.026
Position X Load X Automation	1.818	3.625	0.502	2	30	.611

APPENDIX R

SA from OTS

Table R-1. SA from OTS: Means and Standard Deviations.

Attention and SA	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	4) Awareness of Aircraft Positions	5.77	0.91	5.02	1.46	5.40	1.25	5.77	0.91	5.02	1.46	5.40	1.25	5.77	0.89	5.02	1.43	5.40	1.24
	5) Ensuring Positive Control	6.38	0.96	5.38	1.54	5.88	1.36	6.38	0.96	5.38	1.54	5.88	1.36	6.38	0.94	5.38	1.52	5.88	1.35
	6) Pilot Deviations from Control Instructions	5.79	1.11	5.35	1.25	5.57	1.18	5.79	1.11	5.35	1.25	5.57	1.18	5.79	1.09	5.35	1.23	5.57	1.17
	7) Correcting Errors in a Timely Manner	6.07	1.06	5.44	1.37	5.75	1.25	6.07	1.06	5.44	1.37	5.75	1.25	6.07	1.05	5.44	1.34	5.75	1.24
Auto 2	4) Awareness of Aircraft Positions	6.19	0.62	5.36	1.18	5.77	1.02	6.19	0.62	5.36	1.18	5.77	1.02	6.19	0.61	5.36	1.16	5.77	1.01
	5) Ensuring Positive Control	6.19	0.65	6.00	0.63	6.10	0.64	6.19	0.65	6.00	0.63	6.10	0.64	6.19	0.64	6.00	0.62	6.10	0.64
	6) Pilot Deviations from Control Instructions	5.97	0.37	5.71	0.99	5.84	0.75	5.97	0.37	5.71	0.99	5.84	0.75	5.97	0.36	5.71	0.98	5.84	0.74
	7) Correcting Errors in a Timely Manner	6.13	0.62	5.94	0.77	6.04	0.69	6.13	0.62	5.94	0.77	6.04	0.69	6.13	0.61	5.94	0.76	6.04	0.69
Auto 3	4) Awareness of Aircraft Positions	6.50	0.87	5.21	1.16	5.86	1.20	6.50	0.87	5.21	1.16	5.86	1.20	6.50	0.85	5.21	1.14	5.86	1.19
	5) Ensuring Positive Control	6.44	0.51	5.69	0.95	6.06	0.84	6.44	0.51	5.69	0.95	6.06	0.84	6.44	0.50	5.69	0.93	6.06	0.83
	6) Pilot Deviations from Control Instructions	6.28	0.62	5.92	1.44	6.10	1.10	6.28	0.62	5.92	1.44	6.10	1.10	6.28	0.61	5.92	1.41	6.10	1.10
	7) Correcting Errors in a Timely Manner	6.69	0.47	5.88	0.96	6.29	0.85	6.69	0.47	5.88	0.96	6.29	0.85	6.69	0.46	5.88	0.94	6.29	0.84
Auto Coll.	4) Awareness of Aircraft Positions	6.15	0.85	5.20	1.25	5.68	1.17	6.15	0.85	5.20	1.25	5.68	1.17	6.15	0.84	5.20	1.25	5.68	1.16
	5) Ensuring Positive Control	6.34	0.72	5.69	1.11	6.01	0.99	6.34	0.72	5.69	1.11	6.01	0.99	6.34	0.72	5.69	1.11	6.01	0.99
	6) Pilot Deviations from Control Instructions	6.01	0.77	5.66	1.23	5.84	1.04	6.01	0.77	5.66	1.23	5.84	1.04	6.01	0.77	5.66	1.23	5.84	1.04
	7) Correcting Errors in a Timely Manner	6.30	0.80	5.75	1.06	6.03	0.97	6.30	0.80	5.75	1.06	6.03	0.97	6.30	0.79	5.75	1.06	6.03	0.97

Table R-2. Maintaining Attention and SA: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Load	.074	0.926	12.517	12.517	12.517	4	4	.016

Table R-3. Maintaining Awareness of Aircraft Positions: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p*-level
Load	11.021	0.271	4.692	1	7	.000
Automation	0.956	0.569	1.680	2	14	.222
Load X Automation	0.341	1.090	0.313	2	14	.736

Table R-4. Ensuring Positive Control: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p*-level
Load	5.005	.089	56.529	1	7	.000
Automation	0.226	0.772	0.292	2	14	.751
Load X Automation	0.694	0.562	1.236	2	14	.320

Table R-5. Detecting Pilot Deviations From Control Instructions: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p*-level
Load	1.505	0.350	4.295	1	7	.077
Automation	1.107	0.539	2.054	2	14	.165
Load X Automation	0.032	0.459	0.069	2	14	.933

Table R-6. Correcting Own Errors in a Timely Manner: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p*-level
Load	3.521	0.176	2.051	1	7	.003
Automation	1.134	0.459	2.473	2	14	.120
Load X Automation	0.415	0.371	1.119	2	14	.354

APPENDIX S

ATWIT Results

Table S-1. ATWIT Ratings: Means and Standard Deviations.

Ratings	Interval	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Automation 1	1	3.01	1.64	4.82	2.00	3.92	2.02	2.64	1.22	4.70	1.62	3.67	1.75	2.82	1.43	4.76	1.79	3.79	1.88
	2	3.27	1.70	5.02	1.96	4.15	2.01	2.52	1.29	4.65	1.66	3.58	1.82	2.90	1.54	4.83	1.80	3.86	1.92
	3	3.46	1.94	5.16	2.01	4.31	2.13	2.65	1.43	4.41	1.20	3.53	1.58	3.05	1.73	4.78	1.67	3.92	1.90
	4	3.60	1.91	5.23	1.95	4.42	2.07	2.79	1.23	4.44	1.44	3.61	1.56	3.20	1.63	4.83	1.74	4.01	1.86
	5	3.59	1.76	5.46	2.32	4.53	2.24	2.84	1.36	4.29	1.39	3.57	1.54	3.21	1.59	4.88	1.97	4.05	1.97
	6	3.21	1.65	5.27	1.47	4.24	1.86	2.72	1.43	4.78	1.42	3.75	1.75	2.96	1.54	5.02	1.44	3.99	1.81
	7	3.13	1.86	5.50	2.36	4.32	2.42	2.57	1.51	4.44	1.63	3.50	1.81	2.85	1.69	4.97	2.07	3.91	2.16
	8	3.24	2.14	5.18	2.29	4.21	2.39	2.49	1.45	4.12	1.41	3.31	1.63	2.87	1.84	4.65	1.95	3.76	2.08
	9	3.19	1.90	5.25	2.24	4.22	2.30	2.69	1.54	4.31	1.74	3.50	1.81	2.94	1.72	4.78	2.03	3.86	2.08
	10	3.04	1.83	5.67	2.20	4.36	2.40	2.48	1.62	4.22	1.53	3.35	1.78	2.76	1.72	4.95	2.00	3.85	2.16
	Interval Collapsed	3.27	1.80	5.26	2.05	4.27	2.17	2.64	1.38	4.43	1.48	3.54	1.69	2.96	1.63	4.85	1.83	3.90	1.97
Automation 2	1	2.95	1.20	5.62	1.75	4.24	1.99	2.83	1.29	4.75	1.27	3.76	1.59	2.89	1.23	5.18	1.56	4.00	1.81
	2	2.66	1.35	5.78	1.86	4.17	2.25	2.59	1.45	5.50	1.01	4.00	1.92	2.63	1.38	5.64	1.48	4.08	2.08
	3	2.78	1.51	5.50	1.97	4.10	2.21	2.78	1.56	5.03	1.49	3.87	1.89	2.78	1.51	5.27	1.73	3.98	2.04
	4	3.06	1.37	5.73	2.18	4.35	2.23	3.19	1.89	5.13	1.49	4.13	1.95	3.12	1.63	5.43	1.86	4.24	2.08
	5	2.73	1.35	5.71	1.99	4.17	2.25	2.73	1.25	5.11	1.35	3.88	1.76	2.73	1.28	5.41	1.70	4.03	2.01
	6	2.84	1.36	5.56	1.57	4.16	1.99	2.72	1.48	5.10	1.41	3.87	1.87	2.78	1.40	5.33	1.49	4.01	1.92
	7	2.44	1.37	5.80	1.61	4.07	2.25	2.63	1.41	5.07	1.38	3.81	1.85	2.54	1.37	5.44	1.52	3.94	2.05
	8	2.30	1.53	5.66	1.51	3.93	2.26	2.30	1.44	5.19	1.27	3.70	1.99	2.30	1.46	5.42	1.39	3.81	2.12
	9	2.19	1.38	5.67	1.72	3.87	2.33	2.37	1.50	5.33	1.68	3.81	2.17	2.28	1.42	5.50	1.68	3.84	2.23
	10	2.16	1.43	5.37	1.39	3.71	2.14	2.35	1.37	5.44	1.83	3.84	2.23	2.25	1.39	5.40	1.60	3.78	2.17
	Interval Collapsed	2.61	1.38	5.54	1.71	4.08	2.13	2.65	1.45	5.10	1.39	3.87	1.88	2.63	1.42	5.32	1.57	3.98	2.01
Automation 3	1	3.74	1.98	5.37	2.22	4.61	2.24	3.48	2.07	4.96	1.59	4.27	1.95	3.61	2.00	5.16	1.91	4.44	2.09
	2	3.49	2.00	5.50	2.25	4.56	2.34	3.09	1.30	5.23	1.29	4.22	1.67	3.29	1.67	5.36	1.81	4.39	2.02
	3	3.29	1.35	5.79	2.20	4.62	2.22	2.89	1.21	5.38	1.47	4.21	1.84	3.09	1.28	5.59	1.85	4.42	2.03
	4	3.71	1.60	5.76	2.13	4.80	1.49	3.18	1.49	5.41	1.49	4.36	1.85	3.44	1.54	5.59	1.82	4.58	2.00
	5	3.63	1.73	5.63	2.29	4.69	2.26	3.16	1.62	5.57	1.81	4.44	2.09	3.40	1.66	5.60	2.03	4.57	2.16
	6	3.55	2.14	5.97	2.46	4.84	2.59	2.69	1.65	5.33	2.00	4.09	2.26	3.12	1.93	5.65	2.23	4.46	2.44
	7	3.54	2.20	5.71	2.47	4.69	2.56	2.74	1.58	4.89	1.69	3.88	1.95	3.14	1.93	5.30	2.12	4.29	2.29
	8	2.99	1.68	5.64	2.48	4.40	2.50	2.93	1.66	5.05	1.72	4.05	1.98	2.96	1.65	5.34	2.12	4.23	2.25
	9	3.13	2.33	6.00	2.48	4.66	2.78	2.53	1.64	5.47	2.21	4.09	2.44	2.83	2.00	5.73	2.33	4.37	2.61
	10	2.65	2.11	5.91	2.53	4.38	2.84	2.45	1.67	5.20	1.78	3.91	2.20	2.55	1.87	5.56	2.19	4.15	2.53
	Interval Collapsed	3.42	1.85	5.83	2.33	4.62	2.43	2.99	1.56	5.32	1.72	4.15	2.01	3.20	1.72	5.58	2.06	4.39	2.24
Automation Collapsed	1	3.23	1.63	5.26	2.00	4.32	2.04	2.97	1.57	4.81	1.48	3.85	1.78	3.10	1.60	5.04	1.76	4.08	1.94
	2	3.13	1.70	5.43	2.02	4.39	2.14	2.73	1.34	5.12	1.37	3.95	1.79	2.93	1.54	5.27	1.72	4.11	2.01
	3	3.17	1.62	5.49	2.04	4.44	2.14	2.77	1.38	4.95	1.42	3.89	1.75	2.97	1.51	5.22	1.77	4.11	1.99
	4	3.45	1.63	5.57	2.06	4.62	2.09	3.05	1.54	5.00	1.50	4.02	1.76	3.25	1.59	5.29	1.82	4.28	1.99
	5	3.31	1.64	5.60	2.17	4.50	2.20	2.90	1.39	5.00	1.60	3.94	1.80	3.11	1.53	5.30	1.92	4.21	2.05
	6	3.19	1.72	5.61	1.89	4.45	2.14	2.71	1.49	5.07	1.63	3.86	1.93	2.95	1.62	5.34	1.78	4.16	2.08
	7	3.02	1.85	5.67	2.15	4.38	2.34	2.64	1.47	4.80	1.57	3.66	1.86	2.83	1.67	5.23	1.92	4.05	2.16
	8	2.84	1.81	5.49	2.12	4.23	2.33	2.57	1.51	4.78	1.53	3.68	1.84	2.71	1.66	5.14	1.87	3.93	2.15
	9	2.83	1.91	5.64	2.16	4.24	2.42	2.53	1.53	5.04	1.94	3.75	2.11	2.68	1.73	5.34	2.06	4.03	2.32
	10	2.62	1.81	5.66	2.09	4.11	2.42	2.42	1.52	4.95	1.76	3.59	2.08	2.52	1.66	5.31	1.95	3.93	2.29
	Interval Collapsed	3.10	1.72	5.54	2.05	4.32	2.26	2.76	1.47	4.95	1.58	3.85	1.88	2.93	1.61	5.25	1.85	4.09	2.09

Table S-2. ATWIT Response Times: Means and Standard Deviations (seconds).

R Times	Interval	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Automation 1	1	1.92	0.74	2.19	1.00	2.06	0.88	2.32	0.94	2.75	1.19	2.53	1.08	2.12	0.86	2.47	1.12	2.30	1.00
	2	1.75	1.06	1.65	1.30	1.70	1.17	2.35	0.70	2.63	0.87	2.49	0.79	2.05	0.94	2.14	1.20	2.10	1.07
	3	2.38	1.11	1.93	1.05	2.15	1.09	2.21	0.71	2.57	0.84	2.39	0.79	2.30	0.92	2.25	0.99	2.27	0.95
	4	1.91	1.17	2.27	1.10	2.08	1.13	2.21	0.78	2.53	1.44	2.37	1.15	2.05	0.99	2.40	1.27	2.22	1.14
	5	2.08	1.32	1.84	1.00	1.96	1.16	2.88	1.96	2.69	1.02	2.78	1.54	2.48	1.69	2.27	1.08	2.37	1.41
	6	1.88	0.77	2.57	1.64	2.22	1.86	2.21	0.72	3.20	3.35	2.70	2.43	2.04	0.75	2.88	2.61	2.46	1.95
	7	2.06	1.11	2.46	2.85	2.26	2.13	3.48	4.53	2.40	0.77	2.94	3.24	2.77	3.32	2.43	2.05	2.60	2.74
	8	2.07	0.75	2.35	1.52	2.21	1.19	2.65	2.60	2.20	0.84	2.42	1.92	2.36	1.91	2.27	1.21	2.32	1.58
	9	2.30	1.44	2.37	1.38	2.33	1.39	3.87	3.91	2.54	1.27	3.21	2.94	3.08	3.01	2.46	1.31	2.77	2.32
	10	1.59	0.68	2.47	0.94	2.03	0.92	2.61	1.18	2.42	1.13	2.52	1.14	2.10	1.08	2.45	1.02	2.27	1.06
	Interval Collapsed	1.99	1.04	2.21	1.47	2.10	1.27	2.68	2.26	2.59	1.44	2.64	1.89	2.34	1.79	2.40	1.46	2.37	1.63
Automation 2	1	1.66	0.80	2.32	1.18	1.98	1.04	2.93	1.54	2.70	0.67	2.82	1.19	2.30	1.37	2.51	0.96	2.40	1.18
	2	2.02	0.64	2.34	1.68	2.17	1.24	2.28	0.57	2.49	1.04	2.38	0.83	2.15	0.61	2.41	1.37	2.28	1.05
	3	2.08	1.92	2.05	0.48	2.07	1.40	2.47	1.25	2.36	1.06	2.42	1.15	2.28	1.61	2.20	0.82	2.24	1.28
	4	1.64	0.89	1.86	0.83	1.75	0.85	2.34	1.24	2.03	0.65	2.19	0.99	1.99	1.12	1.94	0.74	1.97	0.94
	5	1.62	0.72	2.10	0.91	1.85	0.84	2.34	1.30	2.49	1.14	2.41	1.21	1.98	1.10	2.29	1.03	2.13	1.07
	6	2.69	1.33	2.39	1.05	2.54	1.19	2.79	1.38	2.85	1.40	2.82	1.37	2.74	1.33	2.62	1.24	2.68	1.28
	7	2.07	0.92	2.25	0.97	2.16	0.93	2.57	1.37	2.31	0.84	2.44	1.14	2.32	1.18	2.28	0.90	2.30	1.04
	8	1.77	1.04	2.49	1.12	2.12	1.12	2.51	1.38	2.24	0.59	2.38	1.07	2.14	1.26	2.37	0.89	2.25	1.09
	9	2.98	4.02	2.12	1.03	2.57	2.96	2.37	0.71	3.00	1.86	2.67	1.40	2.68	2.86	2.56	1.54	2.62	2.30
	10	1.90	0.93	2.62	1.15	2.24	1.09	2.74	1.80	2.81	1.42	2.77	1.60	2.32	1.47	2.71	1.27	2.51	1.38
	Interval Collapsed	2.04	1.65	2.27	1.03	2.15	1.38	2.53	1.28	2.52	1.10	2.53	1.19	2.29	1.49	2.39	1.08	2.34	1.30
Automation 3	1	1.91	0.77	2.14	0.74	2.03	0.75	3.65	2.77	2.87	1.14	3.23	2.07	2.78	2.18	2.50	1.01	2.63	1.66
	2	1.96	1.09	2.70	2.94	2.35	2.27	4.09	3.46	2.44	0.75	3.22	2.53	3.02	2.74	2.57	2.12	2.78	2.42
	3	2.20	1.14	1.64	0.54	1.90	0.90	3.13	1.81	3.43	3.35	3.29	2.70	2.67	1.56	2.53	2.53	2.60	2.12
	4	2.07	1.95	2.10	0.98	2.09	1.49	3.70	1.57	2.47	1.12	3.04	1.47	2.88	1.93	2.28	1.05	2.56	1.54
	5	2.05	1.31	2.44	1.49	2.25	1.40	2.68	0.67	2.14	0.61	2.39	0.69	2.36	1.08	2.29	1.13	2.32	1.10
	6	2.44	1.37	2.28	1.16	2.36	1.25	2.73	1.15	2.27	0.58	2.49	0.91	2.59	1.25	2.28	0.90	2.42	1.08
	7	2.48	1.52	2.29	1.47	2.38	1.47	2.80	0.94	2.39	0.77	2.58	0.86	2.64	1.25	2.34	1.16	2.48	1.20
	8	2.36	2.05	1.90	0.93	2.11	1.55	3.30	1.75	3.05	1.56	3.17	1.63	2.83	1.93	2.47	1.40	2.64	1.67
	9	2.55	1.47	2.89	1.56	2.73	1.51	2.96	1.15	2.27	0.74	2.59	1.00	2.76	1.31	2.58	1.24	2.66	1.27
	10	1.52	0.72	3.46	4.43	2.55	3.37	2.77	0.73	2.78	1.42	2.77	1.13	2.14	0.95	3.12	3.26	2.66	2.49
	Interval Collapsed	2.17	1.35	2.38	2.05	2.28	1.74	3.13	1.77	2.62	1.49	2.88	1.66	2.65	1.64	2.50	1.80	2.58	1.72
Automation Collapsed	1	1.83	0.76	2.21	0.97	2.04	0.91	2.95	1.92	2.77	1.01	2.82	1.49	2.39	1.56	2.49	1.02	2.44	1.31
	2	1.91	0.94	2.23	2.13	2.16	1.66	2.88	2.14	2.52	0.87	2.71	1.63	2.40	1.72	2.38	1.62	2.39	1.67
	3	2.22	1.42	1.86	0.74	2.06	1.17	2.60	1.35	2.81	2.15	2.73	1.77	2.41	1.39	2.34	1.67	2.37	1.53
	4	1.87	1.38	2.08	0.97	1.97	1.18	2.73	1.38	2.35	1.12	2.50	1.26	2.30	1.44	2.21	1.05	2.26	1.26
	5	1.91	1.14	2.13	1.18	1.98	1.14	2.63	1.41	2.43	0.95	2.50	1.20	2.27	1.33	2.28	1.08	2.28	1.20
	6	2.33	1.21	2.41	1.29	2.35	1.24	2.57	1.12	2.76	2.10	2.62	1.65	2.45	1.17	2.59	1.74	2.52	1.48
	7	2.20	1.19	2.33	1.90	2.27	1.60	2.95	2.78	2.37	0.78	2.63	1.99	2.58	2.16	2.35	1.44	2.46	1.83
	8	2.06	1.37	2.23	1.21	2.14	1.27	2.81	1.97	2.51	1.15	2.62	1.58	2.44	1.73	2.37	1.18	2.40	1.48
	9	2.61	2.59	2.48	1.36	2.49	1.98	3.07	2.44	2.59	1.35	2.85	1.95	2.84	2.51	2.53	1.35	2.68	2.01
	10	1.67	0.79	2.87	2.75	2.26	2.03	2.70	1.30	2.67	1.31	2.70	1.30	2.19	1.19	2.77	2.15	2.48	1.76
	Interval Collapsed	2.07	1.37	2.28	1.57	2.18	1.48	2.78	1.83	2.58	1.36	2.68	1.61	2.43	1.65	2.43	1.47	2.43	1.57

Table S-3. Mean ATWIT Ratings and Response Times: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Position	.359	0.641	1.784	1.784	12.488	2	14	.001
Load	.113	0.887	7.831	7.831	54.820	2	14	.000
Automation	.433	0.567	1.307	1.307	3.920	4	12	.029
Position X Load	.615	0.385	0.626	0.626	4.382	2	14	.033
Position X Automation	.704	0.296	0.420	0.420	1.259	4	12	.339
Load X Automation	.793	0.207	0.260	0.260	0.781	4	12	.558
Position X Load X Automation	.841	0.159	0.189	0.189	0.567	4	12	.691

Table S-4. Mean ATWIT Ratings: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	1.502	2.715	3.869	1	15	.068
Load	257.917	2.340	11.205	1	15	.000
Automation	4.419	1.229	3.596	2	30	.040
Position X Load	0.743	1.192	0.624	1	15	.442
Position X Automation	1.105	1.130	0.978	2	30	.388
Load X Automation	2.604	1.427	1.825	2	30	.179
Position X Load X Automation	0.175	2.177	0.080	2	30	.923

Table S-5. Mean ATWIT Response Times: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	12.165	0.490	24.804	1	15	.000
Low Load	4.066	0.131	3.941	1	15	.000
High Load	0.691	0.132	5.235	1	15	.037
Load	0.002	0.526	0.004	1	15	.953
R - Side	0.372	0.142	2.625	1	15	.126
D - Side	0.331	0.133	2.481	1	15	.136
Automation	1.072	0.420	2.551	2	30	.095
Position x Load	2.107	0.300	7.028	1	15	.018
Position x Automation	0.223	0.256	0.872	2	30	.428
Load x Automation	0.301	0.438	0.687	2	30	.511
Position x Load x Automation	0.275	0.270	1.019	2	30	.373

Table S-6. Mean ATWIT Ratings and Response Times: MANOVA Results (N=6).

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	p-level
Position	.494	0.506	1.023	1.023	2.045	2	4	.244
Load	.043	0.957	22.233	22.233	44.465	2	4	.002
Automation	.417	0.583	1.400	1.400	0.700	4	2	.660
Position X Load	.139	0.861	6.177	6.177	12.354	2	4	.019
Position X Automation	.690	0.310	0.450	0.450	0.225	4	2	.904
Load X Automation	.379	0.621	1.640	1.640	0.820	4	2	.614
Position X Load X Automation	.924	0.076	0.082	0.082	0.041	4	2	.994

Table S-7. Mean ATWIT Ratings: ANOVA Results (N=6).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	13.258	4.037	3.284	1	5	.130
Low Load	0.403	0.647	0.623	1	5	.466
High Load	5.468	0.792	6.905	1	5	.047
Load	10.750	1.436	7.163	1	5	.000
R - Side	24.499	0.232	105.708	1	5	.000
D - Side	1.536	0.340	31.000	1	5	.003
Automation	1.221	0.924	1.321	2	10	.310
Position x Load	4.354	0.279	15.610	1	5	.011
Position x Automation	0.748	1.462	0.512	2	10	.614
Load x Automation	0.002	0.700	0.003	2	10	.997
Position x Load x Automation	0.106	1.071	0.099	2	10	.907

Table S-8. Mean ATWIT Response Times: ANOVA Results (N=6).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	3.466	0.680	5.096	1	5	.074
Load	0.301	0.838	0.360	1	5	.575
Automation	0.632	0.281	2.250	2	10	.156
Position X Load	1.304	0.387	3.372	1	5	.126
Position X Automation	0.159	0.157	1.010	2	10	.399
Load X Automation	0.411	0.315	1.304	2	10	.314
Position X Load X Automation	0.021	0.390	0.054	2	10	.948

Table S-9. Maximum ATWIT Ratings: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	8.689	3.201	2.715	1	15	.120
Load	317.354	2.186	145.178	1	15	.000
Automation	6.993	2.032	3.441	2	30	.045
Position X Load	0.901	2.028	0.445	1	15	.515
Position X Automation	1.757	1.409	1.247	2	30	.302
Load X Automation	2.609	2.022	1.290	2	30	.290
Position X Load X Automation	0.483	3.773	0.128	2	30	.880

Table S-10. Maximum ATWIT Ratings: ANOVA Results (N=6).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	16.056	2.922	5.494	1	5	.066
Load	117.556	1.556	75.571	1	5	.000
Automation	1.542	0.958	1.609	2	10	.248
Position X Load	5.556	2.022	2.747	1	5	.158
Position X Automation	1.431	1.747	0.819	2	10	.468
Load X Automation	0.514	0.964	0.533	2	10	.603
Position X Load X Automation	0.014	3.631	0.004	2	10	.996

Table S-11. Mean ATWIT Ratings by Time: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	104.605	26.742	3.912	1	15	.067
Interval	2.567	1.414	1.815	9	135	.071
Low Load	0.808	0.101	8.037	9	135	.000
High Load	0.157	0.237	0.660	9	135	.744
Load	2617.031	22.459	116.527	1	15	.000
Interval 1	29.416	0.542	54.301	1	15	.000
Interval 2	43.008	0.746	57.680	1	15	.000
Interval 3	41.443	0.471	88.016	1	15	.000
Interval 4	32.214	0.443	72.757	1	15	.000
Interval 5	39.915	0.466	85.705	1	15	.000
Interval 6	44.459	0.340	13.941	1	15	.000
Interval 7	47.119	0.402	117.230	1	15	.000
Interval 8	45.519	0.480	94.852	1	15	.000
Interval 9	56.478	0.584	96.747	1	15	.000
Interval 10	61.636	0.474	13.036	1	15	.000
Automation	45.734	1.201	4.483	2	30	.020
Position x Interval	0.224	1.028	0.218	9	135	.991
Position x Load	5.545	11.564	0.479	1	15	.499
Interval x Load	3.232	0.754	4.286	9	135	.000
Position x Automation	17.138	11.410	1.502	2	30	.239
Interval x Automation	0.312	0.589	0.529	18	270	.943
Load x Automation	22.701	13.763	1.649	2	30	.209
Position x Interval x Load	0.456	0.601	0.758	9	135	.655
Position x Interval x Automation	0.885	0.777	1.139	18	270	.314
Position x Load x Automation	3.600	21.355	0.169	2	30	.846
Interval x Load x Automation	1.083	0.703	1.541	18	270	.076
Position x Interval x Load x Automation	0.299	0.802	0.373	18	270	.992

Table S-12. Mean ATWIT Response Times by Time: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	125.673	4.948	25.397	1	15	.000
Low Load	4.572	0.121	37.696	1	15	.000
High Load	0.547	0.120	4.555	1	15	.050
Interval	3.088	2.214	1.395	9	135	.196
Load	0.391	4.188	.093	1	15	.764
R - Side	0.393	0.136	2.896	1	15	.109
D - Side	0.596	0.075	7.930	1	15	.013
Automation	8.151	4.010	2.033	2	30	.149
Position x Interval	1.369	2.013	0.680	9	135	.726
Position x Load	28.655	2.322	12.340	1	15	.003
Interval x Load	2.555	2.173	1.176	9	135	.315
Position x Automation	0.854	2.362	0.362	2	30	.700
Interval x Automation	1.715	2.129	0.806	18	270	.693
Load x Automation	4.107	4.019	1.022	2	30	.372
Position x Interval x Load	2.228	2.053	1.085	9	135	.378
Position x Interval x Automation	2.185	1.872	1.167	18	270	.289
Position x Load x Automation	3.830	2.752	1.392	2	30	.264
Interval x Load x Automation	1.720	1.843	0.933	18	270	.539
Position x Interval x Load x Automation	2.697	2.006	1.344	18	270	.160

Table S-13. Mean ATWIT Ratings by Time: ANOVA Results (N=6).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	132.582	4.371	3.284	1	5	.130
Low Load	0.403	0.647	0.623	1	5	.466
High Load	5.468	0.792	6.905	1	5	.047
Interval	2.675	1.309	2.044	9	45	.056
Load	1007.501	14.359	7.163	1	5	.000
R - Side	24.499	0.232	105.708	1	5	.000
D - Side	1.536	0.340	31.000	1	5	.003
Automation	12.206	9.243	1.321	2	10	.310
Position x Interval	0.217	1.110	0.195	9	45	.994
Position x Load	43.541	2.789	15.610	1	5	.011
Interval x Load	0.954	0.719	1.327	9	45	.250
Position x Automation	7.480	14.623	0.512	2	10	.614
Interval x Automation	0.434	0.765	0.567	18	90	.914
Load x Automation	0.018	7.003	0.003	2	10	.997
Position x Interval x Load	0.321	0.780	0.411	9	45	.922
Position x Interval x Automation	0.732	1.017	0.720	18	90	.782
Position x Load x Automation	1.055	1.714	0.099	2	10	.907
Interval x Load x Automation	1.396	.914	1.527	18	90	.099
Position x Interval x Load x Automation	0.411	0.946	0.434	18	90	.976

Table S-14. Mean ATWIT Response Times by Time: ANOVA Results (N=6).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	34.665	6.803	5.096	1	5	.074
Interval	2.718	1.802	1.508	9	45	.174
Load	3.013	8.379	0.360	1	5	.575
Automation	6.323	2.811	2.250	2	10	.156
Position X Interval	2.363	2.540	0.931	9	45	.508
Position X Load	13.036	3.866	3.372	1	5	.126
Interval X Load	0.634	2.254	0.281	9	45	.977
Position x Automation	1.587	1.571	1.010	2	10	.399
Interval x Automation	1.371	1.849	0.741	18	90	.760
Load x Automation	4.106	3.149	1.304	2	10	.314
Position x Interval x Load	1.917	2.484	0.772	9	45	.643
Position x Interval x Automation	1.621	1.913	0.847	18	90	.641
Position x Load x Automation	0.209	3.905	0.054	2	10	.948
Interval x Load x Automation	2.064	1.676	1.231	18	90	.255
Position x Interval x Load x Automation	2.252	2.455	0.917	18	90	.560

APPENDIX T

NASA TLX Results

Table T-1. NASA TLX: Means and Standard Deviations.

NASA TLX	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Automation 1	1) Mental Demand	4.23	1.55	6.67	2.03	5.45	2.16	3.73	1.96	6.23	1.61	4.98	2.17	3.98	1.75	6.45	1.81	5.22	2.16
	2) Physical Demand	2.48	1.39	4.29	1.96	3.39	1.91	2.54	1.69	4.17	2.40	3.35	2.20	2.51	1.52	4.23	2.16	3.37	2.04
	3) Temporal Demand	4.08	1.74	6.39	2.30	5.23	2.32	3.39	1.60	6.08	1.94	4.73	2.22	3.73	1.68	6.23	2.10	4.98	2.27
	4) Performance	7.75	1.00	6.88	1.86	7.31	1.53	7.56	1.63	6.94	1.39	7.25	1.52	7.66	1.33	6.91	1.61	7.28	1.52
	5) Effort	6.20	2.07	7.45	1.49	6.82	1.88	5.64	2.65	7.26	1.22	6.45	2.19	5.92	2.36	7.36	1.34	6.64	2.03
	6) Frustration	2.20	1.61	3.89	2.36	3.05	2.16	2.02	1.33	3.64	2.09	2.83	1.91	2.11	1.46	3.77	2.19	2.94	2.03
Automation 2	1) Mental Demand	3.44	1.78	6.78	1.77	5.11	2.43	3.38	1.78	6.72	1.40	5.05	2.31	3.41	1.75	6.75	1.57	5.08	2.36
	2) Physical Demand	2.06	0.90	4.08	2.44	3.07	2.08	2.38	1.57	3.77	1.93	3.07	1.87	2.22	1.27	3.93	2.17	3.07	1.96
	3) Temporal Demand	3.33	1.60	6.15	2.05	4.74	2.31	3.21	1.41	6.34	1.68	4.77	2.20	3.27	1.49	6.25	1.84	4.76	2.24
	4) Performance	7.94	1.34	7.00	1.51	7.47	1.48	6.44	2.28	7.19	1.38	6.81	1.89	7.19	1.99	7.10	1.42	7.14	1.72
	5) Effort	5.67	2.29	7.34	1.50	6.50	2.08	5.36	2.37	7.27	1.35	6.32	2.13	5.52	2.30	7.30	1.41	6.41	2.09
	6) Frustration	1.84	0.94	3.66	1.70	2.75	1.64	2.52	1.68	4.34	2.31	3.43	2.19	2.18	1.39	4.00	2.03	3.09	1.95
Automation 3	1) Mental Demand	4.32	2.32	7.48	2.23	5.90	2.76	3.92	1.85	7.11	1.64	5.51	2.36	4.12	2.08	7.30	1.93	5.71	2.55
	2) Physical Demand	2.56	1.99	4.67	3.18	3.61	2.82	2.35	1.17	4.54	2.71	3.45	2.34	2.46	1.61	4.60	2.91	3.53	2.57
	3) Temporal Demand	4.21	2.05	7.20	2.44	5.71	2.69	3.76	1.85	6.89	1.88	5.33	2.43	3.99	1.94	7.05	2.15	5.52	2.55
	4) Performance	7.88	1.09	6.25	2.18	7.06	1.88	6.94	1.44	5.44	1.93	6.19	1.84	7.41	1.34	5.84	2.07	6.63	1.90
	5) Effort	5.61	2.49	8.07	1.55	6.84	2.39	6.26	1.69	7.07	1.42	6.67	1.59	5.94	2.11	7.57	1.55	6.76	2.01
	6) Frustration	2.02	1.23	4.52	2.84	3.27	2.50	2.08	1.45	5.27	2.39	3.67	2.53	2.05	1.32	4.89	2.61	3.47	2.50
Automation Collapsed	1) Mental Demand	4.00	1.91	6.98	2.01	5.49	2.46	3.68	1.84	6.69	1.56	5.18	2.27	3.84	1.87	6.83	1.80	5.33	2.37
	2) Physical Demand	2.37	1.48	4.35	2.53	3.36	2.29	2.42	1.46	4.16	2.34	3.29	2.13	2.40	1.46	4.25	2.43	3.32	2.21
	3) Temporal Demand	3.87	1.81	6.58	2.26	5.23	2.45	3.45	1.61	6.44	1.83	4.94	2.28	3.66	1.72	6.51	2.05	5.09	2.36
	4) Performance	7.86	1.13	6.71	1.86	7.28	1.63	6.98	1.84	6.52	1.74	6.75	1.79	7.42	1.58	6.62	1.79	7.02	1.73
	5) Effort	5.83	2.25	7.62	1.52	6.72	2.11	5.75	2.26	7.20	1.31	6.48	1.97	5.79	2.24	7.41	1.42	6.60	2.04
	6) Frustration	2.02	1.27	4.02	2.33	3.02	2.12	2.21	1.48	4.42	2.32	3.31	2.23	2.11	1.38	4.22	2.32	3.17	2.17

Table T-2. NASA TLX: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Position	.389	0.611	1.574	1.574	2.623	6	10	.086
Low Load	.013	0.987	74.668	74.668	124.446	6	10	.000
High Load	.115	0.885	7.663	7.663	12.771	6	10	.000
Load	.038	0.962	25.283	25.283	42.138	6	10	.000
D Side	.093	0.907	9.801	9.801	21.561	5	11	.000
R Side	.016	0.984	59.699	59.699	99.498	6	10	.000
Automation	.102	0.898	8.820	8.820	2.940	12	10	.154
Position x Load	.144	0.856	5.921	5.921	9.869	6	4	.001
Position x Automation	.408	0.592	1.450	1.450	0.483	12	4	.852
Load x Automation	.384	0.616	1.604	1.604	0.535	12	4	.819
Position x Load x Automation	.250	0.750	3.001	3.001	1.000	12	4	.555

Table T-3. Mental Demand: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	4.490	3.561	1.261	1	15	.279
Load	43.087	2.890	148.815	1	15	.000
Automation	6.971	2.514	2.773	2	30	.079
Position X Load	0.010	2.359	0.004	1	15	.950
Position X Automation	0.738	1.919	0.384	2	30	.684
Load X Automation	3.403	1.801	1.890	2	30	.169
Position X Load X Automation	0.004	4.286	0.001	2	30	.999

Table T-4. Physical Demand: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.209	4.757	0.044	1	15	.837
Load	165.639	5.484	3.205	1	15	.000
Automation	3.459	1.503	2.302	2	30	.117
Position X Load	0.709	0.810	0.875	1	15	.364
Position X Automation	0.126	1.683	0.075	2	30	.928
Load X Automation	0.997	1.836	0.543	2	30	.587
Position X Load X Automation	0.511	3.207	0.159	2	30	.853

Table T-5. Temporal Demand: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	3.829	2.312	1.656	1	15	.218
Load	388.493	3.610	107.615	1	15	.000
Automation	9.674	2.313	4.183	2	30	.025
Position X Load	0.896	2.418	0.370	1	15	.552
Position X Automation	1.240	2.540	0.488	2	30	.619
Load X Automation	1.455	1.728	0.842	2	30	.441
Position X Load X Automation	0.064	4.549	0.014	2	30	.986

Table T-6. Performance: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	13.547	2.882	4.700	1	15	.047
Load	3.880	1.436	21.507	1	15	.000
Automation	7.643	0.769	9.943	2	30	.000
Position X Load	5.672	3.073	1.846	1	15	.194
Position X Automation	2.829	1.578	1.793	2	30	.184
Load X Automation	8.666	2.265	3.826	2	30	.033
Position X Load X Automation	3.013	2.464	1.223	2	30	.309

Table T-7. Effort: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	2.899	1.809	1.603	1	15	.225
Load	126.090	3.721	33.890	1	15	.000
Automation	1.958	1.943	1.007	2	30	.377
Position X Load	1.402	1.418	0.988	1	15	.336
Position X Automation	0.201	1.456	0.138	2	30	.872
Load X Automation	0.492	1.729	0.285	2	30	.754
Position X Load X Automation	5.153	3.083	1.672	2	30	.205

Table T-8. Frustration: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	4.047	4.447	0.910	1	15	.355
Load	212.782	4.289	49.606	1	15	.000
Automation	4.828	2.762	1.748	2	30	.191
Position X Load	0.534	1.633	0.327	1	15	.576
Position X Automation	3.436	2.698	1.273	2	30	.295
Load X Automation	6.580	2.683	2.452	2	30	.103
Position X Load X Automation	0.708	3.617	0.196	2	30	.823

APPENDIX U

PSQ Workload Results

Table U-1. PSQ Workload: Means and Standard Deviations.

Workload	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Automation 1	7) How Hard You Were Working	7.32	2.05	6.00	2.50	6.66	2.35	6.25	3.15	6.57	1.41	6.41	2.41	6.79	2.67	6.29	2.02	6.54	2.36
Automation 2	7) How Hard You Were Working	6.57	2.96	5.20	2.26	5.88	2.69	7.13	2.15	5.57	1.90	6.35	2.15	6.85	2.56	5.38	2.06	6.12	2.42
Automation 3	7) How Hard You Were Working	6.57	2.50	4.07	2.60	5.32	2.81	6.25	2.74	5.38	2.36	5.82	2.56	6.41	2.59	4.72	2.53	5.57	2.68
Auto. Coll.	7) How Hard You Were Working	6.82	2.50	5.09	2.54	5.95	2.65	6.55	2.69	5.84	1.96	6.19	2.37	6.68	2.59	5.46	2.29	6.07	2.51

Table U-1. Working Hard: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	2.747	4.772	0.576	1	15	.460
Load	71.256	28.484	2.502	1	15	.135
Automation	15.090	4.302	3.508	2	30	.043
Position X Load	12.523	3.193	3.922	1	15	.066
Position X Automation	2.876	2.428	1.185	2	30	.320
Load X Automation	6.390	2.852	2.241	2	30	.124
Position X Load X Automation	4.385	4.869	0.901	2	30	.417

Table U-2. Working Hard: ANOVA Results (N=6).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	11.681	7.114	1.642	1	5	.256
Load	19.125	6.358	29.902	1	5	.003
Automation	7.097	4.147	1.711	2	10	.230
Position X Load	1.125	3.892	2.602	1	5	.168
Position X Automation	0.014	2.397	0.006	2	10	.994
Load X Automation	5.542	2.625	2.111	2	10	.172
Position X Load X Automation	0.292	4.308	0.068	2	10	.935

APPENDIX V

Visual Scanning Data After First Data Reduction and Analysis Pass

V.1 Visual Scanning Variables

The oculometer recorded eye movements during both practice scenarios and experimental scenarios. O-1 provides a summary of the eye movement measures.

Table V–1. Visual Scanning Variables

1. Conditional information – Aircraft
2. Conditional information – Location
3. Conditional information – Range
4. Conditional information - Tightness
5. Eye motion workload
6. Pupil motion workload
7. Visual efficiency
8. Mean number of fixations
9. Mean duration of fixations
10. Mean fixation area
11. Mean distance of saccades
12. Mean duration of saccades
13. Mean number of dwells
14. Mean dwell area
15. Mean duration of dwells
16. Number of fixations on target
17. Mean duration of fixations on target
18. Number of fixations off target
19. Mean duration of fixations off target
20. Number of fixations on radar returns
21. Mean duration of fixations on radar returns
22. Number of fixations on data blocks
23. Mean duration of fixations on data blocks
24. Number of fixations on other static objects
25. Mean duration of fixations on other static objects
26. Number of fixations on PVD
27. Mean duration of fixations on PVD
28. Number of fixations on SCRD
29. Mean duration of fixations on SCRD
30. Number of fixations on map
31. Mean duration of fixations on map
32. Number of fixations on flight strips
33. Mean duration of fixations on flight strips
34. Number of fixations on keyboard
35. Mean duration of fixations on keyboard
36. Number of fixations on trackball
37. Mean duration of fixations on trackball
38. Number of fixations on ATWIT
39. Mean duration of fixations on ATWIT

V.1.1 Fixations

A fixation is a sequence of at least 6 oculometer samples with an intersample distance of less than 1 degree of visual angle. At 1-meter distance, this corresponds to a circle with an 8.73 mm radius. The distance between two samples is the norm of the vectorial difference of the sample coordinates. If 2 fixations are not separated by either a blink or a saccade (see definitions below), these fixations should be combined within one fixation. In summary:

Fixation if:

$D = \sqrt{((x_i - x_{i+1})^2 + (y_i - y_{i+1})^2)} > 8.73$ mm

with D the distance between to subsequent samples x and y the horizontal and vertical point of gaze coordinates in mm respectively

and:

$n > 6$ with n the number of samples in a sequence

and

separated by a blink or a saccade

Related to a fixation the following variables need to be calculated: Fixation Duration and Fixation Area. Fixation Area is an approximation of the area covered by the POG due to eye movements within a fixation.

Fixation Duration:

$FIXDUR = t_{sample} * \Sigma samples$

with t_{sample} where the duration of a sample ($^1/_{60}$ second) and Σsample is the total number of samples within a fixation

Fixation Area:

$FIXAREA = (\max(x_{fix}) - \min(x_{fix})) * (\max(y_{fix}) - \min(y_{fix}))$

with x_{fix} and y_{fix} the sequences of horizontal and vertical POG coordinates within a fixation respectively

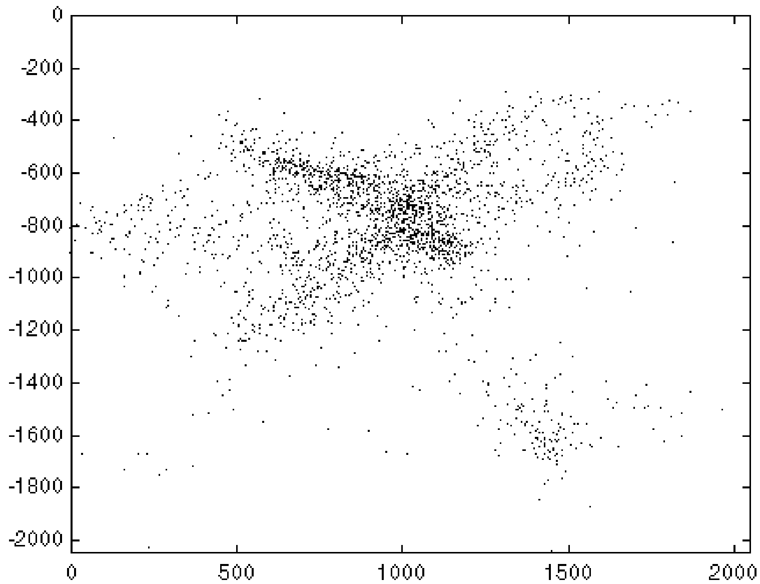

Figure V-1. Example of Fixation Distribution. The units for horizontal and vertical coordinates are in pixels. The top left corner corresponds with the top left corner of the radar display.

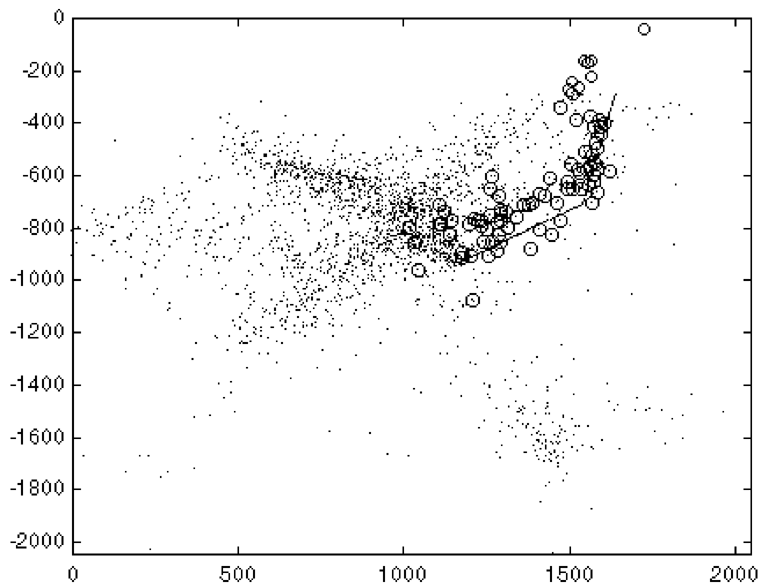

Figure V-2. Example of Fixation Distribution. The flight path of a departure, BTA3721 is superimposed. The circles represent fixations that were identified as fixation on flight BTA3721. The units for horizontal and vertical coordinates are in pixels

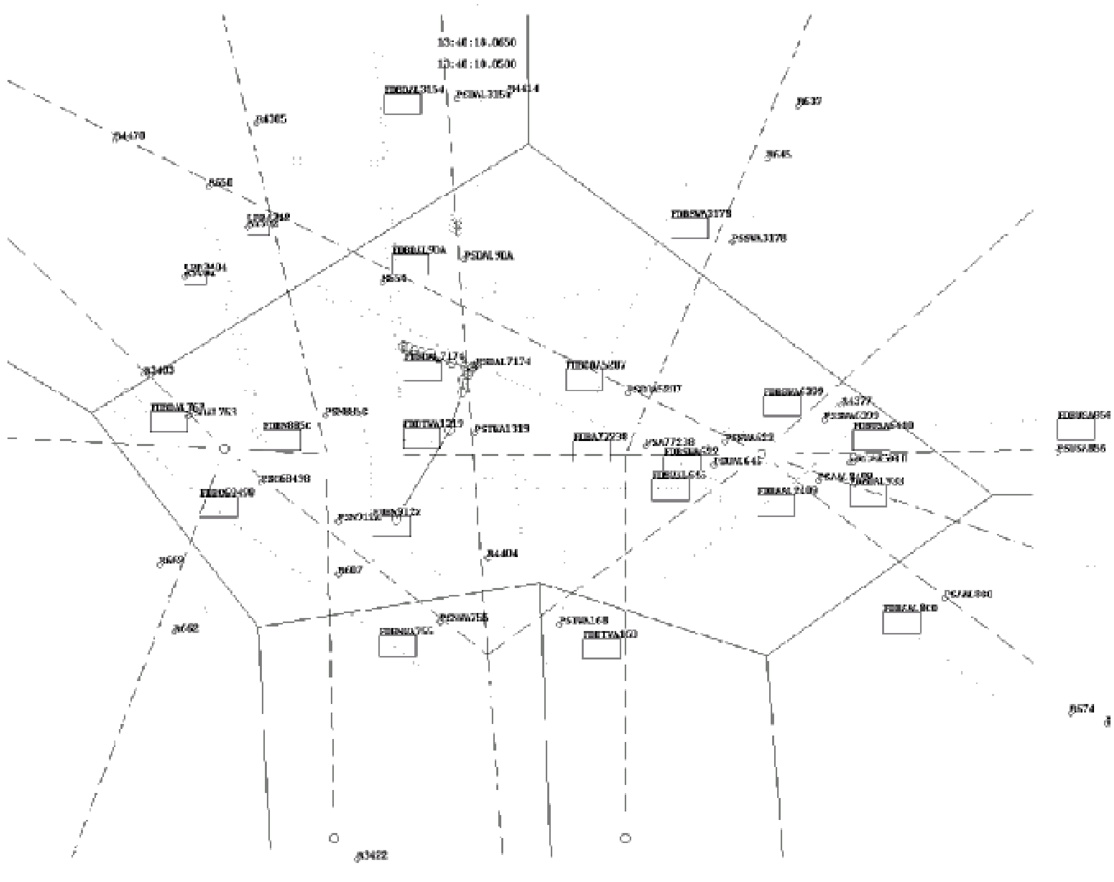

Figure V-3. Example of point of one second of point of gaze data integrated with simulator data on position of radar returns and data blocks.

V.1.2 Blink

A blink is the complete or partial closure of the eye. The oculometer will suggest that the velocity at the start and end of a blink was greater than 700 degrees per second, which corresponds with 6.108 $^{m}/_{s}$. This is physically impossible, but it does give us a way to determine start and end of a blink. A blink starts after the last sample of the previous fixation and stops before the first sample of the next fixation. In summary:

Blink if:

$$VEL = \sqrt{((x_i - x_{i+1})^2 + (y_i - y_{i+1})^2)} / t_{sample} \quad > 6.108\ ^{m}/_{s}$$

with VEL being the a crude estimate of the tangential velocity and x and y the horizontal and vertical point of gaze coordinates in mm respectively. The index denotes the current sample i and next sample i+1 respectively

and:

n >12 with n the number of samples in a sequence

Related to a blink the following variables need to be calculated: Fixation Duration and Blink Distance. Blink Distance is the distance covered by the POG due to eye movements during a blink.

Blink Duration:

$$BLNKDUR = t_{sample} * \Sigma samples$$

with t_{sample} where the duration of a sample ($^{1}/_{60}$ second) and Σsample is the total number of samples within a blink

Blink Distance:

$$BLNKDST = (x_n - x_p) * (y_n - y_p)$$

with x and y the horizontal and vertical point of gaze coordinates in mm respectively. The index denotes the last sample of the previous fixation p and first sample of the next fixation n respectively

V.1.3 Saccade

A saccade is the ballistic movement of the eye from one fixation to the next. A saccade is characterized by fast eye movements of up to 700 degrees per second. The cut-off for a saccade is a difference in distance between two subsequent saccades that is greater or equal to 8.73 mm, lasts at least 3 samples (or a velocity of 0.524 $^{m}/_{s}$), and the velocity is less or equal to 700 degrees per second (6.108 $^{m}/_{s}$). The saccade will start at the end of the last sample of the previous fixation and will end at the beginning of the first sample of the next fixation. In summary:

$$0.524 > VEL > 6.108\ ^{m}/_{s}$$

and:

$$n > 2$$

Related to saccades a number of variables need to be calculated: Saccade Duration, Saccade Distance, and Saccade Velocity. The saccade distance is the angular distance traveled during a saccade in degrees. The saccade velocity is the average velocity within a saccade in degrees per second.

Saccade Duration:

$$SACDUR = t_{sample} * \Sigma samples$$

with t_{sample} where the duration of a sample ($^{1}/_{60}$ second) and Σsample is the total number of samples within a saccade

Saccade Distance:

$$SACDST = (x_n - x_p)*(y_n - y_p)$$

with x and y the horizontal and vertical point of gaze coordinates in mm respectively. The index denotes the last sample of the previous fixation p and first sample of the next fixation n respectively

Saccade Velocity:

$$SACVEL = \Sigma (\sqrt{((x_i - x_{i+1})^2 + (y_i - y_{i+1})^2)}) / t_{sample} * n_{saccade}$$

with t_{sample} where the duration of a sample ($^1/_{60}$ second) and $n_{saccade}$ is the number of samples within the saccade

V.1.4 Dwell

A dwell is defined as a sequence of fixations that return to a location within 1 degree of visual from a target location or within 1 degree of visual angle if the POG does not rest on a target. This way included in a dwell are also moving targets.

Related to dwells a number of variables need to be calculated: Dwell Duration and Dwell Area. Dwell Duration is the duration between the start of the first sample of the first fixation and the end of the last sample of the last fixation within a dwell sequence. Dwell Area is an approximation of the area covered by the POG within a dwell.

Dwell Duration:

$$DDUR = t_{n,fix\ m} - t_{1,fix\ 1}$$

with $t_{1,fix\ 1}$ is the start of the first sample of the first fixation and $t_{n,fix\ m}$ is the end (sample n) of the last fixation (fixation m).

Dwell Area:

$$DAREA = (\max(x_{fix}) - \min(x_{fix}))*(\max(y_{fix}) - \min(y_{fix}))$$

with x_{fix} and y_{fix} the sequences of horizontal and vertical POG coordinates within a dwell respectively

V.1.5 Visual Efficiency

Visual efficiency is defined as the proportion of the total scanning time that is spent fixating.

Visual Efficiency:

$$VISEFF = (\text{mean}(FIXDUR) * N_{fix}) / (\text{mean}(FIXDUR) * N_{fix} + \text{mean}(SACDUR) * N_{sac})$$

In fact, this is nothing more than the portion of the time that the eye is fixed once the blinks are removed:

Visual Efficiency:

$$VISEFF = \Sigma FIXDUR / (\Sigma FIXDUR + \Sigma SACDUR)$$

with ΣFIXDUR the sum of the duration of the fixations, ΣSACDUR the sum of the duration of the saccades and TIME the total time in seconds.

V.1.6 Eye Motion Workload

Eye Motion Workload is defined as the average saccade motion in degrees by the number of saccades, or:

Eye Motion Workload:

$$EYEMWL = \text{mean}(SACDST) * N_{sac} / TIME$$

with N_{sac} the number of saccades within the interval under study and TIME the total time in seconds.

In fact, this is nothing more than the total distance traveled divided by the total the time:

Eye Motion Workload:

$$EYEMWL = \Sigma SACDST / TIME$$

with ΣSACDST the sum of the distance of the saccades in degrees and TIME the total time in seconds.

V.1.7 Pupil Motion Workload

Pupil Motion Workload is defined as the sum of the average pupil diameter within a fixation divided by the total time within the interval under consideration.

Pupil Motion Workload

$$PUPMWL = \Sigma\|mean(PUPDIAM)_{fix\ i} - mean(PUPDIAM)_{fix\ i+1})\| / TIME$$

with PUPDIAM the pupil diameter in mm based on a conversion from ASL arbitrary units to mm of 0.044 mm per ASL unit. The index fix i and fix i+1 denote the i-th and the i+1th fixation respectively

It seems if the author of the article that this measure was based on was after the "distance" traveled during an interval. I is of course possible to separate the oculometer samples that do not include blinks and then to calculate the cumulative sum of the pupil diameter differences. This may be a more accurate estimate of pupil workload:

Pupil Average Work:

for fixations or saccades:

$$PUPAW = \Sigma\|PUPDIAM_i - PUPDIAM_{i+1}\|$$

with i and i+1 oculometer sample i an i+1 respectively. In this case the oculometer samples that occur during blinks are removed from the time series of data.

V.1.8 Conditional Information

The conditional information is defined by Brillouin (1962) as described in Ellis (1986). The formula will here be given without getting too much into the details:

$$CONINF = \Sigma\ p_i * [\Sigma\ p_{i,j} * \log_2 (p_{i,j})] \quad \text{with } i \neq j$$

with p_i is simple probability of viewing target i, and $p_{i,j}$ is the probability of a transition from target i to target j. Simple probability was defined by Ellis (1986) as the percentage of time spent on each particular target or jumping between each target. Here we will calculate it not as a percentage of time, but the ratio of the number of times on a target and the total number of fixations and the number of transitions and the total number of saccades for p_i and $p_{i,j}$ respectively.

APPENDIX W

Visual Scanning

Table W-1. Fixations: Means and Standard Deviations.

Fixations	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	Number	2976.08	309.37	3046.82	208.26	3011.45	261.90
	Mean Duration	435.58	63.87	428.99	62.61	432.28	62.31
	Stdev Duration	237.98	126.30	210.71	70.43	224.34	101.54
	Mean Area	0.96	0.35	0.99	0.29	0.98	0.32
	Stdev Area	2.45	1.71	2.86	1.61	2.66	1.65
	Efficiency	0.76	0.05	0.75	0.08	0.76	0.06
Automation 2	Number	2873.96	240.67	2968.88	297.09	2921.42	270.30
	Mean Duration	490.29	51.44	453.95	74.80	472.12	65.79
	Stdev Duration	326.01	88.36	288.16	152.84	307.09	124.30
	Mean Area	1.12	0.33	0.96	0.25	1.04	0.30
	Stdev Area	3.27	2.16	2.50	1.35	2.89	1.81
	Efficiency	0.81	0.04	0.75	0.08	0.78	0.07
Automation 3	Number	2851.32	336.75	2963.58	255.52	2907.45	299.52
	Mean Duration	498.55	61.68	479.46	43.34	489.00	53.33
	Stdev Duration	328.52	105.91	286.50	65.96	307.51	89.38
	Mean Area	0.97	0.24	0.92	0.25	0.95	0.24
	Stdev Area	2.46	1.17	2.13	1.03	2.29	1.10
	Efficiency	0.81	0.04	0.79	0.04	0.80	0.04
Automation Collapsed	Number	2900.45	297.04	2993.09	253.63	2946.77	278.65
	Mean Duration	474.81	64.48	454.13	63.79	464.47	64.64
	Stdev Duration	297.50	113.90	261.79	108.44	279.65	112.06
	Mean Area	1.02	0.31	0.96	0.26	0.99	0.29
	Stdev Area	2.73	1.74	2.50	1.35	2.61	1.55
	Efficiency	0.79	0.05	0.77	0.07	0.78	0.06

Table W-2. Saccades: Means and Standard Deviations.

Saccades	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	Number	2637.26	275.24	2711.22	342.83	2674.24	308.12
	Mean Duration	148.27	24.25	154.00	37.57	151.13	31.24
	Stdev Duration	72.19	44.90	64.65	32.26	68.42	38.65
	Mean Distance	4.71	0.76	4.56	0.44	4.63	0.61
	Stdev Distance	55.39	23.33	52.85	16.17	54.12	19.79
	Eye	5.91	0.83	5.86	0.90	5.89	0.85
Automation 2	Number	2550.01	288.01	2701.53	379.21	2625.77	340.06
	Mean Duration	132.38	22.07	161.71	54.91	147.04	43.78
	Stdev Duration	54.88	44.23	127.54	288.11	91.21	206.09
	Mean Distance	4.45	0.50	4.50	0.56	4.48	0.52
	Stdev Distance	41.57	15.45	42.84	12.60	42.20	13.88
	Eye	5.34	0.65	5.63	0.67	5.49	0.67
Automation 3	Number	2600.28	502.89	2785.95	322.79	2693.11	426.24
	Mean Duration	130.87	17.42	136.86	26.90	133.87	22.50
	Stdev Duration	53.56	30.52	45.01	23.03	49.28	26.95
	Mean Distance	4.22	0.41	4.29	0.44	4.26	0.42
	Stdev Distance	52.58	75.17	37.37	12.71	44.98	53.59
	Eye	5.16	0.86	5.62	0.67	5.39	0.79
Automation Collapsed	Number	2595.85	364.24	2732.90	343.67	2664.37	358.91
	Mean Duration	137.17	22.43	150.86	41.88	144.01	34.12
	Stdev Duration	60.21	40.48	79.07	168.10	69.64	121.99
	Mean Distance	4.46	0.60	4.45	0.49	4.45	0.54
	Stdev Distance	49.85	45.71	44.35	15.09	47.10	33.97
	Eye	5.47	0.84	5.70	0.75	5.59	0.80

Table W-3. Blinks: Means and Standard Deviations.

Blinks	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	Number	306.31	144.19	273.65	170.83	289.98	156.38
	Mean Duration	471.57	536.47	543.22	601.34	507.39	561.74
	Stdev Duration	292.26	549.73	375.23	611.32	333.75	573.44
	Mean Distance	14.85	3.44	15.40	5.16	15.12	4.32
	Stdev Distance	192.36	179.56	129.34	86.58	160.85	142.31
Automation 2	Number	305.94	190.65	269.96	135.31	287.95	163.65
	Mean Duration	503.67	523.39	547.06	597.49	525.37	552.97
	Stdev Duration	391.57	544.47	412.75	608.88	402.16	568.29
	Mean Distance	15.77	3.41	16.10	3.55	15.93	3.42
	Stdev Distance	186.53	148.52	125.33	48.21	155.93	112.98
Automation 3	Number	239.02	171.25	268.13	167.68	253.57	167.38
	Mean Duration	569.41	588.39	493.50	554.11	531.46	563.53
	Stdev Duration	412.30	604.91	338.17	557.17	375.24	573.31
	Mean Distance	14.59	4.82	14.36	4.90	14.48	4.78
	Stdev Distance	378.64	868.51	136.63	67.21	257.63	618.30
Automation Collapsed	Number	283.76	169.17	270.58	155.36	277.17	161.69
	Mean Duration	514.88	539.87	527.93	572.66	521.41	553.61
	Stdev Duration	365.38	557.35	375.38	581.05	370.38	566.34
	Mean Distance	15.07	3.89	15.29	4.55	15.18	4.21
	Stdev Distance	252.51	515.94	130.43	67.81	191.47	371.13

Table W-4. Scene Based Fixations: Means and Standard Deviations.

SBF	Window	DST Display						Radar Display						Display Collapsed				Display/Load Collapsed	
		Low Load		High Load		Load Collapsed		Low Load		High Load		Load Collapsed		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	Number	381.90	365.00	470.93	406.53	426.42	382.72	1525.65	458.01	1501.75	473.55	1513.70	458.43	953.78	709.62	986.34	680.21	970.06	689.73
	Percent	12.90	12.49	15.90	14.08	14.40	13.18	51.66	15.68	49.42	15.46	50.54	15.36	32.28	24.12	32.66	22.39	32.47	23.09
	Mean Duration	384.08	76.53	394.58	92.82	389.33	83.85	505.81	67.65	500.26	75.43	503.04	70.54	444.95	94.19	447.42	99.02	446.18	95.87
Auto 2	Number	1011.97	379.44	1069.25	391.99	1040.61	380.61	1212.40	258.27	1258.43	482.67	1235.42	381.51	1112.18	335.12	1163.84	443.07	1138.01	390.56
	Percent	34.99	12.19	36.78	15.91	35.88	13.97	42.38	8.66	41.83	14.72	42.10	11.88	38.68	11.06	39.30	15.29	38.99	13.24
	Mean Duration	533.86	86.69	499.55	101.03	516.71	94.23	515.14	58.41	473.19	89.33	494.16	77.24	524.50	73.33	486.37	94.76	505.44	86.22
Auto 3	Number	1538.31	526.88	1444.03	395.33	1491.17	460.70	875.43	334.22	999.72	386.93	937.57	361.22	1206.87	549.34	1221.87	446.10	1214.37	496.46
	Percent	54.04	16.60	49.35	15.02	51.70	15.76	30.69	11.25	33.29	11.50	31.99	11.26	42.37	18.31	41.32	15.48	41.84	16.83
	Mean Duration	534.91	74.25	523.20	58.83	529.06	66.17	500.77	51.56	495.51	50.73	498.14	50.39	517.84	65.23	509.35	55.84	513.60	60.39
Auto Coll.	Number	977.39	636.63	994.74	561.85	986.06	597.30	1204.50	442.59	1253.30	486.43	1228.90	463.22	1090.94	557.18	1124.02	538.63	1107.48	546.80
	Percent	33.98	21.77	34.01	20.26	33.99	20.92	41.58	14.76	41.51	15.23	41.54	14.92	37.78	18.89	37.76	18.22	37.77	18.51
	Mean Duration	484.29	105.62	472.44	101.50	478.36	103.21	507.24	58.60	489.65	72.98	498.45	66.43	495.76	85.74	481.05	88.36	488.41	87.14

Table W-5. Scene Based Saccades: Means and Standard Deviations.

SBS	Window	DST Display						Radar Display						Display Collapsed				Display/Load Collapsed	
		Low Load		High Load		Load Collapsed		Low Load		High Load		Load Collapsed		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	Number	343.21	315.67	413.49	365.97	378.35	338.08	1270.77	416.69	1233.61	452.67	1252.19	428.40	806.99	595.20	823.55	580.98	815.27	583.51
	Percent	13.01	11.88	15.54	13.93	14.28	12.80	48.60	16.43	45.67	15.55	47.14	15.80	30.80	22.93	30.61	21.10	30.71	21.86
	Mean Distance	4.25	0.98	3.44	0.57	3.85	0.89	3.11	0.33	2.87	0.36	2.99	0.36	3.68	0.93	3.16	0.55	3.42	0.80
Auto 2	Number	952.52	337.77	991.92	392.72	972.22	360.88	971.90	266.54	1023.11	448.82	997.50	364.04	962.21	299.46	1007.52	415.15	984.86	359.80
	Percent	37.23	12.08	37.15	15.52	37.19	13.68	38.20	9.23	37.53	15.23	37.87	12.39	37.72	10.59	37.34	15.13	37.53	12.95
	Mean Distance	3.35	0.37	3.56	0.48	3.46	0.43	3.18	0.38	2.97	0.31	3.07	0.36	3.26	0.38	3.26	0.50	3.26	0.44
Auto 3	Number	1487.80	611.09	1409.40	409.86	1448.60	513.39	689.42	303.85	822.15	349.58	755.78	329.17	1088.61	624.39	1115.77	478.97	1102.19	552.19
	Percent	56.82	16.99	50.94	14.81	53.88	15.96	26.92	11.09	29.21	11.15	28.06	11.00	41.87	20.73	40.07	16.98	40.97	18.82
	Mean Distance	3.53	0.40	3.49	0.25	3.51	0.33	3.13	0.25	2.82	0.34	2.98	0.33	3.33	0.39	3.15	0.45	3.24	0.43
Auto Coll.	Number	927.84	640.86	938.27	562.03	933.05	599.57	977.36	406.32	1026.29	444.42	1001.83	424.27	952.60	534.31	982.28	505.91	967.44	519.15
	Percent	35.69	22.62	34.55	20.63	35.12	21.54	37.91	15.25	37.47	15.39	37.69	15.24	36.80	19.22	36.01	18.16	36.40	18.66
	Mean Distance	3.71	0.75	3.50	0.44	3.60	0.62	3.14	0.32	2.89	0.34	3.01	0.35	3.43	0.64	3.19	0.50	3.31	0.58

Table W-6. Object Based: Means and Standard Deviations.

Object Based	Window	Full Data Block						Aircraft Position Symbol						Information Collapsed				Info/Load Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	Number	892.16	240.29	922.46	229.14	907.31	231.47	417.67	131.79	468.60	185.26	443.13	160.25	654.91	307.32	695.53	308.50	675.22	306.14
	Mean Duration	541.18	73.97	528.02	77.72	534.60	74.93	669.52	97.69	676.15	75.41	672.84	85.91	605.35	107.31	602.08	106.48	603.72	106.06
Auto 2	Number	725.97	243.93	776.02	312.49	751.00	276.92	322.85	66.84	384.85	161.32	353.85	125.48	524.41	269.98	580.44	315.16	552.42	292.47
	Mean Duration	554.28	59.99	500.38	97.96	527.33	84.46	673.80	78.91	688.10	75.26	680.95	76.20	614.04	91.87	594.24	128.36	604.14	111.18
Auto 3	Number	500.96	224.96	597.47	217.24	549.22	223.00	260.04	103.99	329.35	140.23	294.70	126.45	380.50	211.42	463.41	225.62	421.96	220.88
	Mean Duration	531.70	65.61	519.42	58.99	525.56	61.69	703.75	81.87	701.29	62.37	702.52	71.60	617.72	113.87	610.36	110.01	614.04	111.13
Auto Coll.	Number	706.36	282.51	765.32	284.64	735.84	283.63	333.52	121.28	394.27	169.94	363.89	149.99	519.94	286.15	579.79	298.58	549.87	293.21
	Mean Duration	542.39	66.00	515.94	78.97	529.16	73.60	682.36	86.08	688.52	70.51	685.44	78.33	612.37	103.78	602.23	114.32	607.30	109.01

Table W-7. Structure: Means and Standard Deviations.

Structure	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Auto 1	Object-Based	0.13	0.05	0.07	0.02	0.10	0.04
	Range-Based	1.00	0.01	1.00	0.01	1.00	0.01
	Box-Based	4.24	0.17	4.14	0.13	4.19	0.15
	Ring-Based	0.72	0.04	0.71	0.03	0.72	0.03
Auto 2	Object-Based	0.12	0.04	0.08	0.02	0.10	0.04
	Range-Based	1.00	0.01	1.00	0.02	1.00	0.02
	Box-Based	4.21	0.14	4.10	0.23	4.16	0.19
	Ring-Based	0.71	0.06	0.70	0.06	0.71	0.06
Auto 3	Object-Based	0.12	0.04	0.07	0.02	0.10	0.04
	Range-Based	1.00	0.02	1.00	0.02	1.00	0.02
	Box-Based	4.09	0.24	4.00	0.25	4.04	0.24
	Ring-Based	0.69	0.06	0.70	0.03	0.69	0.05
Auto Collapsed	Object-Based	0.12	0.04	0.07	0.02	0.10	0.04
	Range-Based	1.00	0.01	1.00	0.02	1.00	0.01
	Box-Based	4.18	0.19	4.08	0.21	4.13	0.21
	Ring-Based	0.71	0.06	0.70	0.04	0.71	0.05

Table W-8. Limited Data Block: Means and Standard Deviations.

LDB	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Auto 1	Number	22.09	27.58	31.61	27.11	26.85	27.33
	Percent	0.71	0.82	1.02	0.85	0.87	0.84
	Mean Duration	461.95	233.05	441.90	189.13	451.92	209.03
	Stdev Duration	214.17	281.43	200.93	287.45	207.55	279.91
Auto 2	Number	16.27	12.73	26.86	22.16	21.57	18.57
	Percent	0.56	0.44	0.88	0.72	0.72	0.61
	Mean Duration	349.57	173.60	366.81	169.53	358.19	169.02
	Stdev Duration	91.75	80.31	112.43	89.84	102.09	84.47
Auto 3	Number	20.11	24.51	23.77	23.92	21.94	23.90
	Percent	0.67	0.78	0.78	0.79	0.72	0.77
	Mean Duration	381.72	201.14	361.95	215.79	371.83	205.45
	Stdev Duration	110.84	89.78	161.33	354.50	136.09	255.67
Auto Collapsed	Number	19.49	22.19	27.42	24.17	23.45	23.42
	Percent	0.65	0.69	0.90	0.78	0.77	0.74
	Mean Duration	397.75	205.29	390.22	191.88	393.98	197.69
	Stdev Duration	138.92	181.28	158.23	265.32	148.58	226.22

Table W-9. Fixation Number: Means and Standard Deviations.

Number	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	D-side Comm. Panel	2.63	2.89	3.96	6.25	3.29	4.84
	R-side Comm. Panel	3.22	4.81	2.42	3.28	2.82	4.07
	D-side Map	8.61	8.66	7.79	5.89	8.20	7.30
	R-side Map	7.28	10.14	6.89	9.05	7.08	9.45
	Keyboard Area	216.11	133.17	330.49	162.00	273.30	157.03
	CRD	14.41	16.65	20.88	21.77	17.65	19.35
	ATWIT	19.21	11.63	24.38	13.50	21.80	12.67
	Flight Strip Bay 1	33.54	44.19	39.85	58.11	36.70	50.88
	Flight Strip Bay 2	763.52	493.87	637.47	401.74	700.50	447.46
Automation 2	D-side Comm. Panel	4.69	6.75	4.46	4.33	4.58	5.58
	R-side Comm. Panel	2.60	5.33	0.61	0.97	1.60	3.90
	D-side Map	28.05	47.37	9.86	9.02	18.95	34.79
	R-side Map	18.46	38.88	7.08	9.60	12.77	28.45
	Keyboard Area	189.17	115.96	227.93	138.22	208.55	127.04
	CRD	17.29	10.42	14.63	14.02	15.96	12.22
	ATWIT	20.90	10.57	23.38	8.78	22.14	9.64
	Flight Strip Bay 1	10.04	8.93	25.60	69.21	17.82	49.18
	Flight Strip Bay 2	358.40	157.00	327.66	148.80	343.03	151.28
Automation 3	D-side Comm. Panel	10.77	17.94	11.57	23.28	11.17	20.45
	R-side Comm. Panel	3.23	5.51	1.04	1.81	2.13	4.18
	D-side Map	19.98	20.00	9.24	9.05	14.61	16.21
	R-side Map	10.20	26.36	2.90	3.56	6.55	18.87
	Keyboard Area	114.99	76.47	206.67	124.24	160.83	111.66
	CRD	11.51	10.64	8.48	6.63	9.99	8.85
	ATWIT	21.63	6.85	21.21	11.67	21.42	9.42
	Flight Strip Bay 1	10.35	10.14	11.98	15.47	11.16	12.90
	Flight Strip Bay 2	234.91	143.51	246.77	127.18	240.84	133.52
Automation Collapsed	D-side Comm. Panel	6.03	11.49	6.66	14.27	6.35	12.89
	R-side Comm. Panel	3.02	5.12	1.36	2.32	2.19	4.04
	D-side Map	18.88	30.54	8.96	8.00	13.92	22.76
	R-side Map	11.98	27.57	5.62	7.96	8.80	20.43
	Keyboard Area	173.42	116.98	255.03	149.62	214.23	139.74
	CRD	14.40	12.84	14.66	15.94	14.53	14.40
	ATWIT	20.58	9.74	22.99	11.31	21.79	10.57
	Flight Strip Bay 1	17.98	28.37	25.81	53.06	21.89	42.50
	Flight Strip Bay 2	452.28	379.93	403.97	304.46	428.12	343.31

Table W-10. Fixation Percent: Means and Standard Deviations.

Percent	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	D-side Comm. Panel	0.09	0.10	0.13	0.21	0.11	0.16
	R-side Comm. Panel	0.11	0.18	0.08	0.10	0.09	0.14
	D-side Map	0.29	0.29	0.26	0.20	0.27	0.24
	R-side Map	0.25	0.38	0.23	0.30	0.24	0.34
	Keyboard Area	7.24	4.56	10.63	4.99	8.93	5.01
	CRD	0.47	0.49	0.70	0.77	0.58	0.64
	ATWIT	0.65	0.40	0.81	0.45	0.73	0.43
	Flight Strip Bay 1	1.07	1.36	1.26	1.83	1.16	1.59
	Flight Strip Bay 2	25.28	15.43	20.59	12.47	22.93	14.00
Automation 2	D-side Comm. Panel	0.16	0.22	0.16	0.17	0.16	0.19
	R-side Comm. Panel	0.09	0.19	0.02	0.03	0.06	0.14
	D-side Map	0.94	1.53	0.35	0.37	0.65	1.13
	R-side Map	0.66	1.40	0.23	0.32	0.45	1.02
	Keyboard Area	6.55	4.03	7.50	4.22	7.02	4.09
	CRD	0.60	0.36	0.49	0.46	0.55	0.41
	ATWIT	0.72	0.37	0.80	0.32	0.76	0.34
	Flight Strip Bay 1	0.35	0.32	0.81	2.16	0.58	1.53
	Flight Strip Bay 2	12.55	5.68	11.03	5.21	11.79	5.42
Automation 3	D-side Comm. Panel	0.38	0.63	0.37	0.71	0.37	0.66
	R-side Comm. Panel	0.11	0.17	0.04	0.06	0.07	0.13
	D-side Map	0.69	0.64	0.32	0.32	0.51	0.53
	R-side Map	0.33	0.80	0.10	0.12	0.21	0.58
	Keyboard Area	3.99	2.50	6.82	3.91	5.41	3.53
	CRD	0.39	0.36	0.28	0.22	0.34	0.30
	ATWIT	0.78	0.30	0.74	0.47	0.76	0.39
	Flight Strip Bay 1	0.36	0.35	0.43	0.62	0.40	0.50
	Flight Strip Bay 2	8.24	4.83	8.26	4.24	8.25	4.47
Automation Collapsed	D-side Comm. Panel	0.21	0.40	0.22	0.44	0.21	0.42
	R-side Comm. Panel	0.10	0.18	0.04	0.08	0.07	0.14
	D-side Map	0.64	0.99	0.31	0.30	0.47	0.74
	R-side Map	0.41	0.95	0.19	0.26	0.30	0.71
	Keyboard Area	5.93	3.97	8.31	4.62	7.12	4.45
	CRD	0.49	0.41	0.49	0.55	0.49	0.48
	ATWIT	0.72	0.36	0.78	0.41	0.75	0.38
	Flight Strip Bay 1	0.59	0.88	0.83	1.67	0.71	1.34
	Flight Strip Bay 2	15.36	12.13	13.29	9.62	14.32	10.94

Table W-11. Fixation Duration: Means and Standard Deviations.

Duration	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	D-side Comm. Panel	116.75	126.13	74.23	93.06	95.49	111.15
	R-side Comm. Panel	99.82	113.75	81.62	77.96	90.72	96.37
	D-side Map	195.35	123.15	161.91	92.36	178.63	108.42
	R-side Map	177.53	140.52	257.13	459.17	217.33	336.46
	Keyboard Area	259.32	68.79	276.88	68.13	268.10	67.93
	CRD	306.92	192.47	309.33	131.87	308.12	162.29
	ATWIT	327.88	230.50	293.47	108.43	310.67	178.05
	Flight Strip Bay 1	199.13	73.08	181.87	105.83	190.50	89.89
	Flight Strip Bay 2	321.40	42.40	300.10	47.46	310.75	45.57
Automation 2	D-side Comm. Panel	176.08	219.76	153.99	121.70	165.04	175.10
	R-side Comm. Panel	86.61	88.00	29.42	51.90	58.02	76.77
	D-side Map	199.73	95.84	221.16	126.46	210.45	110.91
	R-side Map	193.21	178.01	149.45	121.14	171.33	151.42
	Keyboard Area	265.34	64.86	255.86	35.64	260.60	51.70
	CRD	394.46	185.54	301.16	178.77	347.81	185.38
	ATWIT	290.61	95.85	348.51	121.74	319.56	111.73
	Flight Strip Bay 1	171.59	120.73	152.29	110.02	161.94	114.04
	Flight Strip Bay 2	338.77	57.78	302.70	31.43	320.74	49.28
Automation 3	D-side Comm. Panel	143.42	82.94	129.49	128.02	136.45	106.35
	R-side Comm. Panel	93.63	98.47	40.04	81.04	66.83	92.79
	D-side Map	224.37	101.74	165.13	111.11	194.75	109.03
	R-side Map	124.38	129.53	114.70	117.72	119.54	121.85
	Keyboard Area	233.62	66.51	260.95	32.75	247.29	53.41
	CRD	300.04	126.48	263.67	135.00	281.86	130.00
	ATWIT	306.42	64.75	346.31	272.43	326.36	195.83
	Flight Strip Bay 1	198.38	129.16	176.22	160.10	187.30	143.53
	Flight Strip Bay 2	313.42	48.09	289.38	46.56	301.40	48.14
Automation Collapsed	D-side Comm. Panel	145.42	152.60	119.24	117.72	132.33	136.20
	R-side Comm. Panel	93.35	98.61	50.36	73.58	71.86	89.20
	D-side Map	206.48	106.03	182.73	111.90	194.61	109.08
	R-side Map	165.04	150.52	173.76	283.11	169.40	225.57
	Keyboard Area	252.76	66.76	264.56	48.07	258.66	58.17
	CRD	333.81	172.63	291.39	148.22	312.60	161.45
	ATWIT	308.30	146.51	329.43	181.19	318.87	164.24
	Flight Strip Bay 1	189.70	108.85	170.13	125.64	179.91	117.34
	Flight Strip Bay 2	324.53	49.92	297.40	41.95	310.96	47.85

Table W-12. Saccade Number: Means and Standard Deviations.

Number	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	D-side Comm. Panel	1.37	1.92	2.37	3.30	1.87	2.71
	R-side Comm. Panel	1.04	1.72	0.70	1.11	0.87	1.44
	D-side Map	5.76	5.81	5.85	4.58	5.81	5.15
	R-side Map	5.33	7.82	4.94	7.70	5.14	7.64
	Keyboard Area	178.26	128.45	283.54	180.82	230.90	163.29
	CRD	7.98	11.38	13.16	14.73	10.57	13.21
	ATWIT	11.70	9.16	14.70	9.48	13.20	9.29
	Flight Strip Bay 1	22.90	32.87	27.28	45.88	25.09	39.32
	Flight Strip Bay 2	539.73	432.07	438.29	346.43	489.01	388.66
Automation 2	D-side Comm. Panel	2.12	2.96	2.68	3.72	2.40	3.32
	R-side Comm. Panel	0.73	1.33	0.13	0.24	0.43	0.99
	D-side Map	21.14	39.35	7.48	9.03	14.31	28.93
	R-side Map	9.58	16.43	5.76	9.28	7.67	13.27
	Keyboard Area	166.33	122.45	220.16	174.56	193.24	150.82
	CRD	10.67	7.83	9.10	9.22	9.88	8.45
	ATWIT	13.07	8.31	16.26	7.92	14.67	8.15
	Flight Strip Bay 1	3.84	4.93	9.66	21.98	6.75	15.95
	Flight Strip Bay 2	199.80	119.86	169.98	94.06	184.89	107.06
Automation 3	D-side Comm. Panel	6.56	12.05	8.81	18.01	7.68	15.12
	R-side Comm. Panel	0.38	0.43	0.23	0.52	0.30	0.48
	D-side Map	14.73	16.21	7.45	8.98	11.09	13.41
	R-side Map	7.57	20.31	1.46	2.28	4.51	14.55
	Keyboard Area	93.16	72.41	188.95	134.34	141.06	116.78
	CRD	6.91	8.39	5.48	5.11	6.19	6.87
	ATWIT	15.08	6.92	13.51	8.09	14.29	7.45
	Flight Strip Bay 1	4.78	5.29	4.27	7.03	4.53	6.12
	Flight Strip Bay 2	122.98	93.68	136.67	87.87	129.82	89.61
Automation Collapsed	D-side Comm. Panel	3.35	7.46	4.62	10.97	3.99	9.35
	R-side Comm. Panel	0.72	1.28	0.35	0.75	0.53	1.06
	D-side Map	13.88	25.09	6.93	7.69	10.40	18.78
	R-side Map	7.49	15.50	4.05	7.19	5.77	12.14
	Keyboard Area	145.92	114.76	230.88	165.83	188.40	148.14
	CRD	8.52	9.27	9.24	10.71	8.88	9.97
	ATWIT	13.28	8.13	14.82	8.42	14.05	8.27
	Flight Strip Bay 1	10.50	20.98	13.74	30.66	12.12	26.18
	Flight Strip Bay 2	287.50	316.95	248.31	249.42	267.91	284.37

Table W-13. Saccade Percent: Means and Standard Deviations.

Percent	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	D-side Comm. Panel	0.05	0.07	0.08	0.11	0.06	0.09
	R-side Comm. Panel	0.04	0.07	0.02	0.03	0.03	0.05
	D-side Map	0.22	0.21	0.22	0.17	0.22	0.19
	R-side Map	0.20	0.33	0.17	0.26	0.19	0.29
	Keyboard Area	6.63	4.73	10.08	5.74	8.35	5.46
	CRD	0.29	0.37	0.49	0.55	0.39	0.47
	ATWIT	0.45	0.37	0.54	0.34	0.49	0.35
	Flight Strip Bay 1	0.85	1.23	1.00	1.80	0.92	1.52
	Flight Strip Bay 2	20.16	15.49	15.99	11.99	18.08	13.79
Automation 2	D-side Comm. Panel	0.08	0.10	0.10	0.13	0.09	0.12
	R-side Comm. Panel	0.03	0.05	0.01	0.01	0.02	0.04
	D-side Map	0.79	1.37	0.29	0.38	0.54	1.02
	R-side Map	0.39	0.67	0.20	0.32	0.29	0.52
	Keyboard Area	6.40	4.74	7.66	5.41	7.03	5.05
	CRD	0.41	0.29	0.32	0.32	0.37	0.31
	ATWIT	0.52	0.35	0.60	0.29	0.56	0.32
	Flight Strip Bay 1	0.15	0.18	0.36	0.80	0.25	0.58
	Flight Strip Bay 2	7.89	4.84	6.47	3.71	7.18	4.31
Automation 3	D-side Comm. Panel	0.24	0.45	0.28	0.55	0.26	0.50
	R-side Comm. Panel	0.02	0.02	0.01	0.02	0.01	0.02
	D-side Map	0.55	0.54	0.27	0.31	0.41	0.45
	R-side Map	0.26	0.65	0.05	0.08	0.15	0.46
	Keyboard Area	3.52	2.57	6.57	4.28	5.04	3.80
	CRD	0.25	0.25	0.19	0.17	0.22	0.21
	ATWIT	0.60	0.31	0.49	0.31	0.54	0.31
	Flight Strip Bay 1	0.18	0.19	0.17	0.29	0.18	0.24
	Flight Strip Bay 2	4.86	3.52	4.96	3.34	4.91	3.38
Automation Collapsed	D-side Comm. Panel	0.12	0.28	0.15	0.34	0.14	0.31
	R-side Comm. Panel	0.03	0.05	0.01	0.02	0.02	0.04
	D-side Map	0.52	0.87	0.26	0.29	0.39	0.66
	R-side Map	0.28	0.56	0.14	0.24	0.21	0.44
	Keyboard Area	5.52	4.30	8.11	5.28	6.81	4.96
	CRD	0.31	0.31	0.33	0.39	0.32	0.35
	ATWIT	0.52	0.34	0.54	0.31	0.53	0.32
	Flight Strip Bay 1	0.39	0.78	0.51	1.18	0.45	1.00
	Flight Strip Bay 2	10.97	11.52	9.14	8.84	10.06	10.26

Table W-14. Saccade Duration: Means and Standard Deviations.

Duration	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	D-side Comm. Panel	707.25	1319.33	911.18	1272.79	809.21	1279.39
	R-side Comm. Panel	634.75	894.46	344.52	535.89	489.64	740.15
	D-side Map	2703.05	2529.92	2689.57	2209.82	2696.31	2336.66
	R-side Map	2680.03	2911.01	2520.49	3152.75	2600.26	2986.05
	Keyboard Area	2818.91	955.97	2703.07	631.98	2760.99	799.33
	CRD	1505.60	1682.51	1635.63	869.56	1570.61	1319.09
	ATWIT	1233.51	666.79	1175.88	534.71	1204.70	595.26
	Flight Strip Bay 1	1670.76	1573.31	1501.31	871.76	1586.04	1254.14
	Flight Strip Bay 2	2853.83	576.93	2916.60	952.17	2885.21	775.09
Automation 2	D-side Comm. Panel	1216.61	1713.92	718.49	725.34	967.55	1319.09
	R-side Comm. Panel	589.89	1184.04	85.32	152.62	337.60	869.10
	D-side Map	3125.15	1628.48	2906.64	2537.37	3015.90	2100.19
	R-side Map	2626.53	2453.71	3228.24	3831.18	2927.39	3179.45
	Keyboard Area	3059.28	667.58	3114.74	718.45	3087.01	682.79
	CRD	1604.61	927.95	1637.97	1340.51	1621.29	1134.22
	ATWIT	1468.61	747.89	1246.68	442.38	1357.64	614.86
	Flight Strip Bay 1	738.13	769.84	1010.95	864.71	874.54	817.18
	Flight Strip Bay 2	3053.89	940.83	2847.68	607.04	2950.78	785.86
Automation 3	D-side Comm. Panel	798.51	885.04	957.43	1351.06	877.97	1126.40
	R-side Comm. Panel	255.43	318.57	137.74	309.65	196.59	314.76
	D-side Map	2831.45	2198.90	1943.80	1744.94	2387.62	2004.05
	R-side Map	2471.99	3163.32	1569.41	2183.80	2020.70	2712.88
	Keyboard Area	2627.37	1057.79	2978.97	752.20	2803.17	920.37
	CRD	1593.64	1146.25	2100.72	1181.98	1847.18	1173.93
	ATWIT	1419.48	553.03	1406.18	848.49	1412.83	704.55
	Flight Strip Bay 1	2919.02	5923.20	818.28	948.61	1868.65	4307.03
	Flight Strip Bay 2	3194.61	1381.53	3587.12	1922.13	3390.86	1658.61
Automation Collapsed	D-side Comm. Panel	907.46	1339.12	862.36	1130.68	884.91	1232.96
	R-side Comm. Panel	493.36	874.30	189.19	377.46	341.27	687.05
	D-side Map	2886.55	2112.84	2513.33	2181.44	2699.94	2144.30
	R-side Map	2592.85	2797.77	2439.38	3138.55	2516.12	2958.36
	Keyboard Area	2835.19	907.13	2932.26	709.04	2883.72	811.30
	CRD	1567.95	1264.76	1791.44	1144.35	1679.70	1204.94
	ATWIT	1373.87	654.59	1276.25	626.85	1325.06	639.37
	Flight Strip Bay 1	1775.97	3604.37	1110.18	923.47	1443.08	2638.42
	Flight Strip Bay 2	3034.11	1008.87	3117.13	1303.72	3075.62	1160.26

Table W-15. DST Display Number: Means and Standard Deviations.

Number	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	Clock	0.82	1.76	1.09	1.95	0.96	1.83
	Response Display	0.35	0.75	0.46	0.83	0.41	0.78
	Aircraft List	105.42	226.64	140.56	251.44	122.99	236.15
	Graphic Plans Display	33.58	72.19	44.77	80.09	39.17	75.22
	Trial Plans	0.47	1.00	0.62	1.11	0.54	1.04
	Plans Display	3.40	7.31	4.54	8.11	3.97	7.62
	Other Displays	208.31	61.70	239.48	149.62	223.89	113.69
Automation 2	Clock	3.07	5.21	2.40	2.64	2.74	4.08
	Response Display	1.97	3.43	0.59	0.90	1.28	2.56
	Aircraft List	804.23	339.45	765.68	345.06	784.96	337.27
	Graphic Plans Display	34.20	71.90	46.71	79.05	40.45	74.60
	Trial Plans	1.22	2.32	2.75	5.54	1.98	4.25
	Plans Display	24.71	48.29	32.72	67.10	28.72	57.65
	Other Displays	113.00	70.98	178.98	123.86	145.99	104.81
Automation 3	Clock	8.47	20.11	4.38	10.46	6.42	15.90
	Response Display	0.90	1.82	0.72	1.15	0.81	1.50
	Aircraft List	829.43	314.49	756.98	251.82	793.21	282.66
	Graphic Plans Display	482.02	484.57	438.14	415.54	460.08	444.60
	Trial Plans	2.00	2.77	1.78	2.32	1.89	2.52
	Plans Display	22.91	18.11	15.46	16.28	19.19	17.36
	Other Displays	153.16	70.24	197.00	279.81	175.08	201.91
Automation Collapsed	Clock	4.12	12.21	2.63	6.34	3.37	9.71
	Response Display	1.07	2.33	0.59	0.96	0.83	1.79
	Aircraft List	579.70	446.88	554.41	407.29	567.05	425.48
	Graphic Plans Display	183.27	351.89	176.54	306.75	179.90	328.37
	Trial Plans	1.23	2.21	1.72	3.56	1.47	2.96
	Plans Display	17.01	31.00	17.57	40.99	17.29	36.15
	Other Displays	158.16	77.18	205.15	194.13	181.65	148.83

Table W-16. DST Display Percent: Means and Standard Deviations.

Percent	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	Clock	24.69	53.08	32.92	58.89	28.81	55.31
	Response Display	24.67	53.05	32.90	58.85	28.79	55.27
	Aircraft List	28.36	60.96	37.81	67.63	33.08	63.52
	Graphic Plans Display	25.79	55.44	34.39	61.51	30.09	57.77
	Trial Plans	24.68	53.06	32.91	58.86	28.79	55.28
	Plans Display	24.78	53.28	33.04	59.11	28.91	55.52
	Other Displays	6.99	1.99	8.01	5.43	7.50	4.06
Automation 2	Clock	24.77	53.05	32.97	58.86	28.87	55.28
	Response Display	24.73	53.02	32.90	58.85	28.82	55.26
	Aircraft List	52.59	50.00	59.41	56.37	56.00	52.53
	Graphic Plans Display	25.81	55.43	34.45	61.47	30.13	57.75
	Trial Plans	24.70	53.05	32.99	58.81	28.85	55.25
	Plans Display	25.51	52.95	34.07	58.55	29.79	55.08
	Other Displays	3.93	2.44	6.06	4.01	5.00	3.44
Automation 3	Clock	33.19	58.74	24.81	53.03	29.00	55.21
	Response Display	32.92	58.84	24.69	53.04	28.80	55.26
	Aircraft List	62.86	53.85	50.80	50.76	56.83	51.84
	Graphic Plans Display	49.01	54.56	39.39	50.40	44.20	51.90
	Trial Plans	32.95	58.83	24.73	53.04	28.84	55.26
	Plans Display	33.70	58.72	25.20	53.08	29.45	55.23
	Other Displays	5.49	2.63	6.79	9.84	6.14	7.11
Automation Collapsed	Clock	27.55	53.99	30.23	55.90	28.89	54.68
	Response Display	27.44	54.00	30.16	55.89	28.80	54.68
	Aircraft List	47.93	55.88	49.34	58.11	48.64	56.71
	Graphic Plans Display	33.54	55.08	36.08	56.83	34.81	55.68
	Trial Plans	27.45	54.01	30.21	55.88	28.83	54.68
	Plans Display	28.00	54.02	30.77	55.89	29.38	54.69
	Other Displays	5.47	2.64	6.95	6.79	6.21	5.18

Table W-17. DST Display Duration: Means and Standard Deviations.

Duration	Window	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	Clock	286.43	615.81	381.91	683.18	334.17	641.63
	Response Display	253.77	545.59	338.36	605.28	296.06	568.46
	Aircraft List	315.48	678.27	420.65	752.48	368.07	706.71
	Graphic Plans Display	282.65	607.69	376.87	674.17	329.76	633.16
	Trial Plans	277.72	597.09	370.30	662.41	324.01	622.12
	Plans Display	291.31	626.29	388.41	694.81	339.86	652.55
	Other Displays	365.94	57.80	370.39	75.14	368.17	65.98
Automation 2	Clock	510.29	613.75	634.51	645.35	572.40	622.71
	Response Display	337.97	526.23	360.23	598.52	349.10	554.49
	Aircraft List	771.91	459.49	818.90	524.88	795.41	485.84
	Graphic Plans Display	342.77	600.24	452.33	652.19	397.55	619.06
	Trial Plans	465.51	734.06	444.49	639.47	455.00	677.28
	Plans Display	504.73	607.29	554.31	641.61	529.52	615.05
	Other Displays	351.81	113.02	389.66	104.15	370.74	108.62
Automation 3	Clock	658.93	667.75	539.19	693.86	599.06	672.61
	Response Display	353.83	599.18	311.06	543.36	332.45	563.07
	Aircraft List	856.05	497.92	767.68	458.24	811.87	472.85
	Graphic Plans Display	769.66	450.08	664.85	441.04	717.25	441.56
	Trial Plans	619.85	702.71	551.87	580.51	585.86	634.97
	Plans Display	835.91	522.02	594.31	526.17	715.11	529.98
	Other Displays	429.17	95.60	443.03	123.55	436.10	108.89
Automation Collapsed	Clock	485.22	638.35	518.54	668.28	501.88	650.25
	Response Display	315.19	547.66	336.55	570.87	325.87	556.53
	Aircraft List	647.81	592.40	669.08	606.31	658.45	596.33
	Graphic Plans Display	465.03	587.79	498.02	598.39	481.52	590.21
	Trial Plans	454.36	680.69	455.56	619.49	454.96	647.37
	Plans Display	543.98	617.35	512.34	618.01	528.16	614.63
	Other Displays	382.31	95.99	401.03	105.35	391.67	100.69

Table W-18. Fixations: MANOVA Results.

	Wilks	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Load	.506	0.494	0.975	0.975	1.625	6	10	.237
Automation	.129	0.871	6.727	6.727	2.242	12	4	.227
Load X Automation	.404	0.596	1.475	1.475	0.492	12	4	.847

Table W-19. Fixation Number: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	205977.011	47952.372	4.295	1	15	.056
Automation	101956.010	70402.277	1.448	2	30	.251
Load X Automation	3480.456	61352.030	0.057	2	30	.945

Table W-20. Fixation Mean Duration: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.010	0.003	3.829	1	15	.069
Automation	0.027	0.004	6.853	2	30	.004
Load X Automation	0.002	0.004	0.509	2	30	.606

Table W-21. Fixation SD Duration: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.031	0.012	2.553	1	15	.131
Automation	0.073	0.013	5.714	2	30	.008
Load X Automation	0.000	0.008	0.056	2	30	.945

Table W-22. Fixation Mean Area: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.074	0.030	2.411	1	15	.141
Automation	0.077	0.044	1.753	2	30	.191
Load X Automation	0.074	0.064	1.160	2	30	.327

Table W-23. Fixation SD Area: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	1.266	1.341	0.944	1	15	.347
Automation	2.861	1.195	2.394	2	30	.108
Load X Automation	2.897	1.583	1.830	2	30	.178

Table W-24. Visual Efficiency: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.017	0.003	6.483	1	15	.022
Automation	0.011	0.003	3.632	2	30	.039
Load X Automation	0.004	0.003	1.097	2	30	.347

Table W-25. Saccades: MANOVA Results.

	Wilks	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Load	.452	0.548	1.214	1.214	2.023	6	10	.155
Automation	.073	0.927	12.709	12.709	4.236	12	4	.087
Load X Automation	.360	0.640	1.778	1.778	0.593	12	4	.783

Table W-26. Saccades Number: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	450781.205	69993.299	6.440	1	15	.023
Automation	38616.885	131725.035	0.293	2	30	.748
Load X Automation	26212.230	101018.952	0.259	2	30	.773

Table W-27. Saccades Mean Duration: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.004	0.001	4.493	1	15	.051
Automation	0.003	0.001	2.360	2	30	.112
Load X Automation	0.001	0.001	1.464	2	30	.247

Table W-28. Saccades SD Duration: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.009	0.016	0.523	1	15	.481
Automation	0.014	0.016	0.880	2	30	.425
Load X Automation	0.017	0.014	1.237	2	30	.305

Table W-29. Saccades Mean Distance: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.004	0.236	0.019	1	15	.893
Automation	1.158	0.217	5.340	2	30	.010
Load X Automation	0.107	0.235	0.455	2	30	.639

Table W-30. Saccades SD Distance: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	725.268	821.772	0.883	1	15	.362
Automation	1244.119	1077.287	1.155	2	30	.329
Load X Automation	595.648	1224.598	0.486	2	30	.620

Table W-31. Eye Motion Workload: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	1.317	0.463	2.843	1	15	.112
Automation	2.251	0.596	3.780	2	30	.034
Load X Automation	0.551	0.610	0.903	2	30	.416

Table W-32. Blinks: MANOVA Results.

	Wilks	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Load	.723	0.277	0.384	0.384	0.844	5	11	.546
Automation	.424	0.576	1.361	1.361	0.817	10	6	.630
Load X Automation	.767	0.233	0.304	0.304	0.182	10	6	.991

Table W-33. Blinks Number: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	4169.333	12275.937	0.340	1	15	.569
Automation	13392.969	20692.823	0.647	2	30	.531
Load X Automation	10749.561	24963.446	0.431	2	30	.654

Table W-34. Blinks Mean Duration: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.004	0.232	0.018	1	15	.896
Automation	0.005	0.134	0.037	2	30	.963
Load X Automation	0.049	0.473	0.104	2	30	.902

Table W-35. Blinks SD Duration: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.002	0.268	0.009	1	15	.926
Automation	0.038	0.165	0.230	2	30	.796
Load X Automation	0.050	0.476	0.105	2	30	.900

Table W-36. Blinks Mean Distance: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	1.120	19.292	0.058	1	15	.813
Automation	17.008	8.198	2.075	2	30	.143
Load X Automation	1.308	11.728	0.111	2	30	.895

Table W-37. Blinks SD Distance: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	357674.224	112335.839	3.184	1	15	.095
Automation	105250.398	128555.372	0.819	2	30	.451
Load X Automation	86300.386	146171.334	0.590	2	30	.560

Table W-38. Scene Based Fixations: MANOVA Results.

	Wilks	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Load	.780	0.220	0.281	0.281	1.220	3	13	.342
Automation	.140	0.860	6.168	6.168	10.280	6	10	.001
Display	.392	0.608	1.551	1.551	6.719	3	13	.006
Load X Automation	.668	0.332	0.498	0.498	0.830	6	10	.573
Load X Display	.896	0.104	0.116	0.116	0.504	3	13	.686
Automation X Display	.074	0.926	12.484	12.484	20.807	6	10	.000
Load X Auto X Display	.775	0.225	0.291	0.291	0.485	6	10	.806

Table W-39. Scene Based Fixations Number: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load	52501.932	23137.518	2.269	1	15	.153
Automation	999761.109	40745.657	24.537	2	30	.000
Radar Display	1328190.318	42815.702	31.021	2	30	.000
DST Display	4570467.771	85387.260	53.526	2	30	.000
Display	2830465.333	482126.300	5.871	1	15	.029
No Automation	9457444.133	127484.516	74.185	1	15	.000
Limited Automation	303615.281	201150.131	1.509	1	15	.238
Full Automation	2451728.320	128088.770	19.141	1	15	.001
Load x Automation	5377.126	58004.627	0.093	2	30	.912
Load x Display	11875.521	213391.265	0.056	1	15	.817
Automation x Display	10797555.068	215660.268	50.067	2	30	.000
Load x Auto x Display	115364.255	268737.666	0.429	2	30	.655

Table W-40. Scene Based Fixations Percent: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load	0.000	0.002	0.001	1	15	.982
Automation	0.148	0.005	29.094	2	30	.000
Radar Display	0.138	0.005	26.252	2	30	.000
DST Display	0.561	0.009	62.150	2	30	.000
Display	0.274	0.064	4.268	1	15	.057
No Automation	1.045	0.015	70.339	1	15	.000
Limited Automation	0.031	0.025	1.261	1	15	.279
Full Automation	1.113	0.008	143.952	1	15	.000
Load x Automation	0.001	0.006	0.232	2	30	.795
Load x Display	0.000	0.028	0.000	1	15	.984
Automation x Display	1.249	0.023	53.233	2	30	.000
Load x Auto x Display	0.017	0.030	0.574	2	30	.570

Table W-41. Scene Based Fixations Mean Duration: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load	0.010	0.007	1.540	1	15	.234
Automation	0.087	0.008	11.338	2	30	.000
Radar Display	0.000	0.002	0.183	2	30	.834
DST Display	0.096	0.004	24.540	2	30	.000
Display	0.019	0.010	1.870	1	15	.192
No Automation	0.103	0.004	27.022	1	15	.000
Limited Automation	0.004	0.003	1.237	1	15	.284
Full Automation	0.008	0.002	4.545	1	15	.050
Load x Automation	0.007	0.006	1.137	2	30	.334
Load x Display	0.000	0.003	0.125	1	15	.729
Automation x Display	0.105	0.004	29.123	2	30	.000
Load x Auto x Display	0.001	0.003	0.185	2	30	.832

Table W-42. Scene Based Saccades: MANOVA Results.

	Wilks	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Load	.258	0.742	2.871	2.871	12.441	3	13	.000
Automation	.128	0.872	6.788	6.788	11.313	6	10	.001
Display	.143	0.857	5.970	5.970	25.869	3	13	.000
Load X Automation	.288	0.712	2.472	2.472	4.120	6	10	.024
Load X Display	.957	0.043	0.044	0.044	0.193	3	13	.900
Automation X Display	.075	0.925	12.399	12.399	20.665	6	10	.000
Load X Auto X Display	.390	0.610	1.562	1.562	2.603	6	10	.087

Table W-43. Scene Based Saccades Number: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load	42271.444	26148.328	1.617	1	15	.223
Automation	1331755.641	61728.259	21.574	2	30	.000
Radar Display	985900.911	37370.679	26.382	2	30	.000
DST Display	4600148.583	93418.945	49.242	2	30	.000
Display	227012.521	372171.310	0.610	1	15	.447
No Automation	6108823.195	98975.962	61.720	1	15	.000
Limited Automation	5113.133	177532.999	0.029	1	15	.868
Full Automation	3839913.281	109427.681	35.091	1	15	.000
Load x Automation	3382.203	72021.773	0.047	2	30	.954
Load x Display	17787.000	205936.367	0.086	1	15	.773
Automation x Display	9840343.349	199850.988	49.238	2	30	.000
Load x Auto x Display	103618.734	231039.184	0.448	2	30	.643

Table W-44. Scene Based Saccades Percent: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load	0.003	0.003	0.880	1	15	.363
Automation	0.175	0.006	30.650	2	30	.000
Radar Display	0.146	0.006	24.649	2	30	.000
DST Display	0.633	0.009	71.364	2	30	.000
Display	0.032	0.059	0.539	1	15	.474
No Automation	0.864	0.014	60.837	1	15	.000
Limited Automation	0.000	0.026	0.014	1	15	.907
Full Automation	0.533	0.013	39.678	1	15	.000
Load x Automation	0.001	0.007	0.168	2	30	.846
Load x Display	0.001	0.027	0.022	1	15	.885
Automation x Display	1.382	0.024	57.957	2	30	.000
Load x Auto x Display	0.019	0.030	0.638	2	30	.535

Table W-45. Scene Based Saccades Mean Distance: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load	2.606	0.110	23.643	1	15	.000
No Automation	2.189	0.135	16.215	1	15	.001
Limited Automation	0.000	0.054	0.000	1	15	.998
Full Automation	0.248	0.044	5.593	1	15	.032
Automation	0.586	0.167	3.510	2	30	.043
Low Load	0.796	0.116	6.889	2	30	.003
High Load	0.063	0.057	1.116	2	30	.341
Radar Display	0.045	0.028	1.592	2	30	.220
DST Display	0.717	0.169	4.254	2	30	.024
Display	16.755	0.212	79.058	1	15	.000
No Automation	5.866	0.256	22.878	1	15	.000
Limited Automation	1.166	0.037	31.828	1	15	.000
Full Automation	2.285	0.040	57.332	1	15	.000
Load x Automation	1.134	0.178	6.371	2	30	.005
Load x Display	0.015	0.120	0.127	1	15	.726
Automation x Display	0.939	0.227	4.139	2	30	.026
Load x Auto x Display	1.139	0.214	5.317	2	30	.011

Table W-46. Scene Based Saccades Mean Distance- Simple Effects of Load x Automation x Display: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load X Auto X Display	1.139	0.214	5.317	2	30	.011
Low Load						
Automation	1.593	0.231	6.889	2	30	.003
Display	7.879	0.217	36.338	1	15	.000
Automation X Display	2.059	0.295	6.968	2	30	.003
High Load						
Automation	0.127	0.114	1.116	2	30	.341
Display	8.891	0.115	77.268	1	15	.000
Automation x Display	0.020	0.146	0.137	2	30	.873
No Automation						
Load	4.378	0.270	16.215	1	15	.001
Display	11.732	0.513	22.878	1	15	.000
Load x Display	1.316	0.274	4.813	1	15	.044
Limited Automation						
Load	0.000	0.107	0.000	1	15	.998
Display	2.331	0.073	31.828	1	15	.000
Load x Display	0.705	0.181	3.901	1	15	.067
Full Automation						
Load	0.497	0.089	5.593	1	15	.032
Display	4.570	0.080	57.332	1	15	.000
Load x Display	0.272	0.094	2.889	1	15	.110
Radar Display						
Load	2506319.448	11842.404	211.639	1	15	.000
Automation	2503368.923	11837.607	211.476	2	30	.000
Load x Automation	2502110.968	11843.250	211.269	2	30	.000
DST Display						
Load	1.111	0.162	6.844	1	15	.019
Automation	1.435	0.337	4.254	2	30	.024
Load x Automation	2.253	0.272	8.272	2	30	.001

Table W-47. Object Based: MANOVA Results.

	Wilks	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Load	.607	0.393	0.646	0.646	4.523	2	14	.031
Automation	.216	0.784	3.620	3.620	10.861	4	12	.001
Information	.024	0.976	40.780	40.780	285.458	2	14	.000
Load X Automation	.964	0.036	0.038	0.038	0.114	4	12	.975
Load X Information	.651	0.349	0.536	0.536	3.755	2	14	.049
Automation X Information	.213	0.787	3.705	3.705	11.116	4	12	.001
Load X Auto X Information	.687	0.313	0.455	0.455	1.366	4	12	.303

Table W-48. Object Based Number: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load	171945.060	49607.606	3.466	1	15	.082
Automation	1026609.536	41133.251	24.958	2	30	.000
FDB	515681.078	21181.183	24.346	2	30	.000
Position Symbol	89344.786	5274.668	16.938	2	30	.000
Information	6640527.497	31684.559	209.582	1	15	.000
No Automation	1723673.634	8238.504	209.222	1	15	.000
Limited Automation	1261790.561	16008.223	78.821	1	15	.000
Full Automation	518241.746	3374.005	153.598	1	15	.000
Load x Automation	7332.341	50546.272	0.145	2	30	.866
Load x Information	38.569	5994.906	0.006	1	15	.937
Automation x Information	183442.193	11778.452	15.574	2	30	.000
Load x Auto x Information	2595.622	12241.572	0.212	2	30	.810

Table W-49. Object Based Duration: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load	0.005	0.003	1.914	1	15	.187
FDB	0.006	0.001	5.372	1	15	.035
Position Symbol	0.000	0.001	0.530	1	15	.478
Automation	0.002	0.004	0.555	2	30	.580
Information	1.172	0.006	200.854	1	15	.000
Low Load	0.157	0.001	203.558	1	15	.000
High Load	0.238	0.002	123.505	1	15	.000
Load x Automation	0.001	0.007	0.164	2	30	.850
Load x Information	0.013	0.002	5.641	1	15	.031
Automation x Information	0.006	0.005	1.302	2	30	.287
Load x Auto x Information	0.004	0.002	2.393	2	30	.109

Table W-50. Structure Conditional Index: MANOVA Results.

	Wilks	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Load	.102	0.898	8.840	8.840	26.520	4	12	.000
Automation	.247	0.753	3.049	3.049	3.049	8	8	.068
Load X Automation	.845	0.155	0.183	0.183	0.183	8	8	.986

Table W-51. Object Conditional Index: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.056	0.001	66.707	1	15	.000
Automation	0.000	0.001	0.186	2	30	.831
Load X Automation	0.000	0.001	0.292	2	30	.749

Table W-52. Range Conditional Index: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.000	0.000	0.000	1	15	.991
Automation	0.000	0.000	0.245	2	30	.784
Load X Automation	0.000	0.000	0.120	2	30	.888

Table W-53. Box Conditional Index: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.224	0.042	5.335	1	15	.036
Automation	0.185	0.022	8.403	2	30	.001
Load X Automation	0.001	0.032	0.044	2	30	.957

Table W-54. Ring Conditional Index: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.000	0.002	0.197	1	15	.664
Automation	0.005	0.003	1.778	2	30	.186
Load X Automation	0.001	0.002	0.443	2	30	.646

APPENDIX X

Controller-to-Controller Coordination Communications Taxonomy

Table X-1. Controller-to-Controller Coordination Communications Taxonomy (C^4T): Coordination-Communication Topic.

ATC Coordination-Communication Topic	Definitions and Examples
Approval	Communications about intersector control/approval requests (“Get me control for descent on that aircraft.” “APREQ N1234 climbing to FL330.”)
Handoff	Communications relating to the transfer of radar identification of a particular aircraft (“Handoff N1234.” “Did you handoff N1234.”)
Point Out	Communications relating to the transfer of radar identification of a particular aircraft when radio communications will be retained (“Point out N1234 to 22.”)
Traffic	Communications about a traffic situation involving a specific aircraft. Includes conflict, spacing, other protected air space or terrain and the resolution of that situation. (“Are you watching that aircraft.”)
Altitude	Communications about altitude not in relation to traffic (“N1234 is requesting flight level 220.”)
Route	Communications regarding headings and/or amendments to route, not in relation to traffic situations (“N1234 is on a 330 heading.” “Next sector, 27, wants N1234 over WEVER.”)
Speed	Communications about speed not in relation to traffic situations (“These three aircraft are slowed to 250 knots.”)
Weather	Communications about weather display or weather updates (Often communicated nonverbally by passing written information. “Sector 22 says continuous moderate turbulence above FL290.”)
Frequency	Communications about an aircraft’s radio communications transfer or frequency assignment (“Have you switched N1234 yet?” “Tell them to switch to N1234.”)
Flow Messages	Communications about traffic flow restrictions not referring to a specific aircraft (“The next sector is requesting 25 miles in trail.”) (due to radar outage)
Flight Strips	Communications about flight progress strips (“Where is that strip?”) Often communicated nonverbally
Equipment	Communications about any ATC hardware (The radar is out of service.”)
Aircraft ID Identification of Aircraft	Communications involving identifying a specific aircraft (Who was that who called?” “That was N1234 who called.”)

Table X-2. Controller-To-Controller Coordination Communications Taxonomy (C^4T): Coordination-Communication Format.

ATC Coordination-Communication Format	Definitions
Question	A direct inquiry about the state or status of sector events.
Answer	A response to a direct or implied question
Statement	Providing information, without being asked, about the state or status of sector events.
Command	A direct order to perform a specific act

Table X-3. Controller-To-Controller Coordination Communications Taxonomy (C^4T): Coordination-Communication Expression.

ATC Coordination-Communication Expression	Definitions
Verbal	Use of voice only communication.
Nonverbal	Use of only body movement communication.
Mixed	Communication that contains both a verbal and non verbal component.
Electronic	Not used. Communication that is electronically transferred.

Table X-4. Frequency for Topic of Communication: Means and Standard Deviations (number).

	Topic of Communication	Load				Load Collapsed	
		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs
Automation 1	Total Communications	63.20	23.27	69.00	31.63	66.10	26.35
	Approval	1.00	0.00	2.00	0.00	1.50	0.71
	Handoff	3.33	1.53	1.50	0.71	2.60	1.52
	Traffic	24.40	6.43	30.40	10.55	27.40	8.82
	Altitude	8.80	6.53	20.00	10.86	14.40	10.31
	Route of Flight	7.40	3.36	8.20	9.88	7.80	6.97
	Speed	3.20	2.68	3.67	1.53	3.38	2.20
	Frequency	3.40	1.94	3.75	3.59	3.56	2.60
	Flight Strips	4.33	4.16	4.00	1.41	4.20	3.03
	Equipment	6.00	8.66	5.50	6.36	5.80	6.91
Automation 2	Total Communications	62.60	27.63	72.60	5.43	67.60	19.49
	Approval	1.00	0.00	3.00	1.41	2.33	1.53
	Handoff	5.00	5.66	2.33	1.52	3.40	3.36
	Traffic	26.20	21.71	35.40	13.18	30.80	17.61
	Altitude	8.80	9.26	13.60	6.19	11.20	7.84
	Route of Flight	5.80	4.55	13.60	10.11	9.70	8.46
	Speed	4.80	5.93	3.50	1.00	4.22	4.29
	Frequency	2.67	0.58	1.75	0.50	2.14	0.69
	Flight Strips	N/A	N/A	N/A	N/A	N/A	N/A
	Equipment	5.00	3.87	3.50	2.08	4.33	3.12
Automation 3	Total Communications	61.00	27.07	60.60	23.07	60.80	23.71
	Approval	1.33	0.58	6.00	0.00	2.50	2.38
	Handoff	5.00	4.58	1.50	0.71	3.60	3.78
	Traffic	28.60	14.28	36.40	13.16	32.50	13.58
	Altitude	12.40	5.94	11.40	5.94	11.90	5.63
	Route of Flight	9.00	9.80	5.40	3.78	7.20	7.25
	Speed	2.50	0.71	4.75	5.56	4.00	4.47
	Frequency	2.75	0.96	3.25	1.26	3.00	1.07
	Flight Strips	N/A	N/A	N/A	N/A	N/A	N/A
	Equipment	5.00	2.45	6.33	3.51	5.57	2.76
Automation Collapsed	Total Communications	62.27	24.15	67.40	21.75	64.83	22.73
	Approval	1.20	0.45	3.50	1.91	2.22	1.72
	Handoff	4.38	3.46	1.86	1.07	3.20	2.86
	Traffic	26.40	14.42	34.07	11.76	30.23	13.50
	Altitude	10.00	7.06	15.00	8.31	12.50	7.99
	Route of Flight	7.40	6.20	9.07	8.58	8.23	7.40
	Speed	3.75	4.05	4.00	3.22	3.87	3.60
	Frequency	3.00	1.35	2.92	2.19	2.96	1.78
	Flight Strips	2.67	3.20	2.00	1.67	2.33	2.46
	Equipment	5.25	4.58	4.89	3.41	5.10	4.02

Table X-5. Frequency for Expression of Communication: Means and Standard Deviations (number).

	Expression of Communication	Load				Load Collapsed	
		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs
Automation 1	Question	9.00	4.36	8.40	2.97	8.70	3.53
	Answer	9.60	8.99	7.60	6.43	8.60	7.44
	Statement	36.20	11.12	51.80	23.99	44.00	19.45
Automation 2	Question	10.20	3.56	13.60	2.88	11.90	3.54
	Answer	8.80	4.18	11.00	4.95	9.50	4.60
	Statement	42.40	19.63	45.80	6.38	44.10	13.88
Automation 3	Question	12.20	7.16	13.40	5.03	12.80	5.87
	Answer	8.80	5.22	9.40	6.91	9.10	5.78
	Statement	39.60	16.85	42.20	13.03	40.90	14.26
Automation Collapsed	Question	10.47	5.05	11.80	4.28	11.13	4.65
	Answer	8.80	6.03	9.33	5.88	9.07	5.85
	Statement	39.40	15.28	46.60	15.54	43.00	15.58

Table X-6. Frequency for Communication Expression: Means and Standard Deviations (number).

	Communication Expression	Load				Load Collapsed	
		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs
Automation 1	Verbal	46.40	18.30	52.80	24.10	49.60	20.45
	Mixed Verbal/Nonverbal	6.80	7.05	13.80	8.14	10.30	8.07
Automation 2	Verbal	53.60	24.09	55.20	8.76	54.40	17.11
	Mixed Verbal/Nonverbal	6.60	3.51	14.60	7.06	10.60	6.74
Automation 3	Verbal	51.60	23.55	56.40	17.90	54.00	19.88
	Mixed Verbal/Nonverbal	7.80	4.66	8.40	8.14	8.10	6.26
Automation Collapsed	Verbal	50.53	20.73	54.80	16.79	52.67	18.66
	Mixed Verbal/Nonverbal	7.07	4.92	12.27	7.76	9.67	6.91

Table X-7. Total Number of Communications: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	197.633	188.717	1.047	1	4	.364
Automation	127.633	326.175	0.391	2	8	.688
Load X Automation	68.433	406.892	0.168	2	8	.848

Table X-8. Traffic: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	440.833	39.167	11.255	1	4	.028
Automation	67.433	119.892	0.562	2	8	.591
Load X Automation	6.433	182.892	0.035	2	8	.966

Table X-9. Altitude: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	187.500	17.583	10.664	1	4	.031
Automation	28.300	35.300	0.802	2	8	.482
Load X Automation	93.100	61.683	1.509	2	8	.278

Table X-10. Route of Flight: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	20.833	12.417	1.678	1	4	.265
Automation	17.033	22.117	0.770	2	8	.494
Load X Automation	82.633	60.467	1.367	2	8	.309

Table X-11. Speed: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	1.333	0.333	4.000	1	1	.295
Automation	16.000	5.333	3.000	2	2	.250
Load X Automation	34.333	30.333	1.132	2	2	.469

Table X-12. Frequency: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.333	5.333	0.063	1	1	.844
Automation	10.083	4.083	2.469	2	2	.288
Load X Automation	1.083	10.083	0.107	2	2	.903

Table X-13. Question: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	13.333	10.333	1.290	1	4	.319
Automation	46.433	17.183	2.702	2	8	.127
Load X Automation	10.033	17.533	0.572	2	8	.586

Table X-14. Answer: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	2.133	8.467	0.252	1	4	.642
Automation	2.033	13.242	0.154	2	8	.860
Load X Automation	15.633	32.342	0.483	2	8	.634

Table X-15. Statement: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	388.800	57.883	6.717	1	4	.061
Automation	33.100	74.558	0.444	2	8	.656
Load X Automation	132.700	210.408	0.631	2	8	.557

Table X-16. Verbal: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	136.533	176.783	0.772	1	4	.429
Automation	70.933	132.142	0.537	2	8	.604
Load X Automation	14.933	275.308	0.054	2	8	.948

Table X-17. Mixed Verbal/Nonverbal: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	202.800	62.217	3.260	1	4	.145
Automation	18.633	21.217	0.878	2	8	.452
Load X Automation	40.300	22.967	1.755	2	8	.233

APPENDIX Y

Push-to-Talk

Table Y-1. Push-to-Talk: Means and Standard Deviations for Number and Duration of Landline Communications (number; seconds).

Table LL		LL Communication					
		Low Load		High Load		Load Coll.	
		Means	SDs	Means	SDs	Means	SDs
Auto 1	N	0.26	0.29	0.43	0.19	0.35	0.25
	D	1.28	0.95	1.81	0.40	1.56	0.73
Auto 2	N	0.19	0.18	0.60	0.37	0.40	0.35
	D	1.20	0.92	1.80	0.21	1.50	0.71
Auto 3	N	0.44	0.22	0.68	0.48	0.56	0.38
	D	2.45	0.71	1.77	0.59	2.11	0.72
Auto Coll.	N	0.30	0.25	0.57	0.36	0.44	0.34
	D	1.68	1.01	1.79	0.42	1.74	0.76

Table Y-2. Push-to-Talk: Means and Standard Deviations for Number and Duration of R-side Communications (number; seconds).

Table R		R Communication					
		Low Load		High Load		Load Coll.	
		Means	SDs	Means	SDs	Means	SDs
Auto 1	N	3.66	0.57	5.38	0.76	4.58	1.11
	D	3.18	0.24	3.09	0.18	3.13	0.21
Auto 2	N	3.80	0.36	5.40	0.62	4.60	0.97
	D	3.18	0.31	2.87	0.27	3.03	0.32
Auto 3	N	3.81	0.48	5.54	0.46	4.68	1.00
	D	3.10	0.22	3.01	0.19	3.05	0.20
Auto Coll.	N	3.76	0.46	5.45	0.60	4.62	1.00
	D	3.15	0.25	3.00	0.22	3.07	0.24

Table Y-3. Mean Number/Duration PTT Landline Communications: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Load	.293	0.707	2.408	2.408	7.225	2	6	.025
Automation	.114	0.886	7.800	7.800	7.800	4	4	.036
Load X Automation	.126	0.874	6.959	6.959	6.959	4	4	.043

Table Y-4. Mean Number PTT Landline Communications: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.752	0.046	16.180	1	7	.005
Automation	0.183	0.051	3.606	2	14	.055
Load X Automation	0.045	0.098	0.457	2	14	.642

Table Y-5. Mean Duration PTT Landline Communications: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.130	0.088	1.472	1	7	.264
No Automation	0.869	0.331	2.627	1	7	.149
Limited Automation	1.103	0.338	3.265	1	7	.114
Full Automation	1.842	0.167	11.031	1	7	.013
Automation	1.680	0.571	2.942	2	14	.086
Low Task Load	3.519	0.791	4.447	2	14	.032
High Task Load	0.002	0.153	0.015	2	14	.985
Load x Automation	1.842	0.374	4.931	2	14	.024

Table Y-6. Mean Number/Duration PTT Radar-side (R) Communications: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Load	.023	0.977	43.402	43.402	13.206	2	6	.000
Automation	.650	0.350	0.538	0.538	0.538	4	4	.719
Load X Automation	.484	0.516	1.067	1.067	1.067	4	4	.476

Table Y-7. Mean Number PTT Radar-side (R) Communications: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	29.900	0.202	148.068	1	7	.000
Automation	0.040	0.256	0.156	2	14	.857
Load X Automation	0.109	0.320	0.341	2	14	.717

Table Y-8. Mean Duration PTT Radar-Side (R) Communications: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.248	0.021	11.735	1	7	.011
Automation	0.041	0.024	1.680	2	14	.222
Load X Automation	0.047	0.015	3.202	2	14	.072

APPENDIX Z

DST Trust Results

Table Z-1. DST Trust Means and Standard Deviations.

	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 2	15) Predictable Behavior	5.25	1.52	4.85	1.98	5.05	1.75	5.61	2.32	5.82	1.84	5.72	2.06	5.43	1.94	5.34	1.95	5.38	1.93
	16) Predicting Outcome of Sep. Strategies	5.14	1.19	4.83	2.13	4.99	1.71	5.02	2.41	5.19	1.76	5.11	2.08	5.08	1.87	5.01	1.93	5.05	1.89
	18) Predicting Future Conflicts	5.77	1.54	5.41	2.20	5.59	1.88	5.77	2.31	5.95	1.97	5.86	2.11	5.77	1.93	5.68	2.07	5.72	1.99
	19) Technical Knowledge	5.80	1.51	5.87	1.63	5.84	1.55	6.00	2.39	6.00	1.67	6.00	2.03	5.90	1.97	5.93	1.62	5.92	1.79
Auto 3	15) Predictable Behavior	5.48	1.86	4.72	2.14	5.10	2.01	6.51	2.05	4.80	2.47	5.65	2.40	5.99	1.99	4.76	2.27	5.37	2.21
	16) Predicting Outcome of Sep. Strategies	5.84	1.86	5.03	1.60	5.44	1.76	6.70	2.23	5.17	2.58	5.94	2.49	6.27	2.06	5.10	2.11	5.69	2.15
	18) Predicting Future Conflicts	5.85	2.05	5.41	1.91	5.63	1.96	6.72	2.06	5.61	2.47	6.16	2.31	6.28	2.07	5.51	2.17	5.90	2.14
	19) Technical Knowledge	5.93	1.65	5.62	1.58	5.77	1.60	6.87	1.79	5.87	2.09	6.37	1.98	6.40	1.76	5.75	1.83	6.07	1.81
Auto Collapsed	15) Predictable Behavior	5.36	1.36	4.89	1.77	5.12	1.58	5.83	1.82	5.33	1.79	5.58	1.81	5.59	1.61	5.11	1.78	5.35	1.71
	16) Predicting Outcome of Sep. Strategies	5.46	1.28	5.08	1.52	5.27	1.41	5.70	1.99	5.25	1.77	5.48	1.89	5.58	1.67	5.17	1.64	5.37	1.67
	18) Predicting Future Conflicts	5.82	1.45	5.56	1.66	5.69	1.55	6.11	1.80	5.80	1.79	5.96	1.79	5.97	1.63	5.68	1.72	5.82	1.68
	19) Technical Knowledge	5.91	1.27	5.82	1.29	5.87	1.27	6.29	1.74	5.95	1.52	6.12	1.63	6.10	1.53	5.89	1.40	5.99	1.47

Table Z-2. DST Trust Items: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Position	.837	0.163	0.195	0.195	0.586	4	12	.679
Load	.650	0.350	0.537	0.537	1.612	4	12	.235
Automation	.376	0.624	1.656	1.656	4.969	4	12	.013
Position X Load	.733	0.267	0.365	0.365	1.094	4	12	.403
Position X Automation	.783	0.217	0.277	0.277	0.832	4	12	.530
Load X Automation	.569	0.431	0.756	0.756	2.268	4	12	.122
Position X Load X Automation	.777	0.223	0.286	0.286	0.859	4	12	.516

Table Z-3. The Predictability of the DST's Behavior: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	11.925	6.821	1.748	1	15	.206
Load	14.073	1.951	7.212	1	15	.017
Limited Automation	0.070	0.767	0.092	1	15	.766
Full Automation	12.154	0.736	16.517	1	15	.001
Automation	0.003	2.113	0.001	1	15	.971
Low Load	2.508	1.025	2.448	1	15	.139
High Load	2.681	0.559	4.796	1	15	.045
Position x Load	0.241	3.117	0.077	1	15	.785
Position x Automation	0.108	2.211	0.049	1	15	.828
Load x Automation	1.375	1.054	9.844	1	15	.007
Position x Load x Automation	4.883	1.824	2.676	1	15	.123

Table Z-4. The Predictability of the DST's Behavior: ANOVA Results (N=6).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	3.201	9.017	0.355	1	5	.577
Load	14.083	2.779	5.068	1	5	.074
Automation	0.941	2.691	0.350	1	5	.580
Position X Load	0.464	4.282	0.108	1	5	.755
Position X Automation	3.000	1.947	1.541	1	5	.270
Load X Automation	7.302	1.958	3.728	1	5	.111
Position X Load X Automation	8.333	0.813	1.254	1	5	.024
R side						
Load	6.554	0.493	13.297	1	5	.015
Automation	2.317	1.405	1.649	2	10	.240
Load x Automation	1.647	0.602	2.735	2	10	.113
D side						
Load	3.144	4.214	0.746	1	5	.427
Automation	2.020	3.421	0.590	2	10	.572
Load x Automation	8.595	1.960	4.385	2	10	.043

Table Z-5. How Well the DST Predicted the Outcome of the Separation Strategies: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	3.082	5.910	0.522	1	15	.481
Load	12.259	3.691	3.321	1	15	.088
Automation	13.109	1.904	6.883	1	15	.019
Position X Load	.102	2.819	0.036	1	15	.852
Position X Automation	1.137	2.801	0.406	1	15	.534
Load X Automation	9.659	2.616	3.693	1	15	.074
Position X Load X Automation	2.869	1.225	2.342	1	15	.147

Table Z-6. Capability of the Tool to Predict Future Conflicts: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	5.261	8.245	0.638	1	15	.437
Load	5.934	1.171	5.068	1	15	.040
Automation	0.950	1.594	0.596	1	15	.452
Position X Load	0.036	1.101	0.033	1	15	.859
Position X Automation	0.556	1.500	0.371	1	15	.552
Load X Automation	3.832	1.724	2.223	1	15	.157
Position X Load X Automation	2.989	2.843	1.051	1	15	.321

Table Z-7. The Technical Knowledge Incorporated in the Tool: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	4.619	5.105	0.905	1	15	.357
Load	3.121	1.277	2.445	1	15	.139
Automation	0.794	0.798	0.995	1	15	.334
Position X Load	1.132	1.410	0.802	1	15	.385
Position X Automation	1.529	1.213	1.260	1	15	.279
Load X Automation	3.777	1.202	3.142	1	15	.097
Position X Load X Automation	0.772	1.333	0.579	1	15	.458

APPENDIX AA

DRAT Results

Table AA-1. Interval Altitude: Means and Standard Deviations.

ALT	Interval	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	1	4.68	1.53	4.77	1.70	4.73	1.59
	2	3.24	1.72	4.34	1.71	3.79	1.78
	3	2.93	1.37	4.46	1.38	3.70	1.56
	4	3.31	1.24	4.34	2.03	3.82	1.74
	5	3.74	1.56	3.52	1.44	3.63	1.48
	6	2.99	1.35	3.90	1.73	3.45	1.59
	7	3.37	1.74	3.71	1.38	3.54	1.55
	8	3.12	1.13	4.15	1.52	3.63	1.42
	9	2.12	1.65	3.15	1.40	2.63	1.59
	10	2.62	1.13	3.34	1.61	2.98	1.41
	Interval Collapsed	3.21	1.56	3.97	1.64	3.59	1.64
Automation 2	1	3.99	1.81	4.14	1.10	4.07	1.48
	2	3.24	1.21	4.08	1.25	3.66	1.28
	3	2.93	1.02	4.77	1.69	3.85	1.66
	4	3.36	0.92	3.64	1.60	3.50	1.29
	5	2.74	0.96	2.83	0.88	2.78	0.91
	6	3.18	1.14	3.64	1.56	3.41	1.36
	7	3.18	1.58	4.08	1.30	3.63	1.49
	8	3.24	0.81	4.21	1.48	3.72	1.27
	9	2.11	1.10	2.89	1.12	2.50	1.16
	10	2.36	1.11	3.33	1.10	2.85	1.19
	Interval Collapsed	3.03	1.27	3.76	1.42	3.40	1.39
Automation 3	1	4.46	1.33	4.30	0.92	4.38	1.13
	2	2.65	1.26	4.49	1.40	3.57	1.61
	3	2.96	1.17	4.43	1.95	3.69	1.75
	4	3.15	1.13	4.05	1.32	3.60	1.29
	5	3.09	2.01	3.18	1.01	3.13	1.56
	6	3.11	0.92	3.49	1.24	3.30	1.09
	7	2.74	1.02	3.86	1.24	3.30	1.25
	8	2.61	0.91	4.11	1.34	3.36	1.36
	9	2.36	1.33	2.86	1.76	2.61	1.55
	10	2.49	1.05	3.30	1.42	2.90	1.30
	Interval Collapsed	2.96	1.34	3.81	1.46	3.39	1.46
Automation Collapsed	1	4.38	1.56	4.41	1.29	4.39	1.42
	2	3.04	1.41	4.30	1.45	3.67	1.56
	3	2.94	1.17	4.55	1.66	3.75	1.64
	4	3.27	1.08	4.01	1.67	3.64	1.45
	5	3.19	1.59	3.18	1.15	3.18	1.38
	6	3.10	1.13	3.68	1.50	3.39	1.35
	7	3.10	1.47	3.89	1.29	3.49	1.43
	8	2.99	0.98	4.16	1.42	3.57	1.35
	9	2.20	1.36	2.97	1.43	2.58	1.44
	10	2.49	1.08	3.32	1.36	2.91	1.29
	Interval Collapsed	3.07	1.40	3.85	1.51	3.46	1.50

Table AA-2. Interval Heading: Means and Standard Deviations.

HDG	Interval	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	1	0.33	0.58	0.69	0.85	0.51	0.74
	2	0.71	0.44	0.62	0.69	0.67	0.57
	3	0.40	0.79	0.56	0.79	0.48	0.78
	4	0.71	0.77	0.56	0.79	0.63	0.77
	5	0.27	0.42	0.69	0.99	0.48	0.78
	6	0.65	0.86	0.62	0.69	0.63	0.77
	7	0.71	1.52	0.87	1.00	0.79	1.27
	8	0.71	0.68	1.19	1.03	0.95	0.89
	9	0.83	1.03	1.44	1.35	1.13	1.22
	10	0.90	1.13	0.56	0.59	0.73	0.90
	Interval Collapsed	0.62	0.88	0.78	0.92	0.70	0.90
Automation 2	1	0.29	0.40	0.39	0.76	0.34	0.60
	2	0.35	0.43	0.64	0.91	0.50	0.72
	3	0.42	0.45	0.58	0.67	0.50	0.57
	4	0.54	0.69	0.83	1.24	0.69	1.00
	5	0.23	0.36	0.71	0.53	0.47	0.51
	6	0.35	0.43	0.77	1.02	0.56	0.80
	7	0.42	0.77	0.64	0.55	0.53	0.67
	8	0.42	0.58	0.89	0.83	0.65	0.75
	9	0.60	0.57	1.08	0.79	0.84	0.72
	10	0.48	0.58	0.96	0.71	0.72	0.68
	Interval Collapsed	0.41	0.54	0.75	0.83	0.58	0.72
Automation 3	1	0.50	0.60	0.42	0.58	0.46	0.58
	2	0.62	0.78	0.79	1.09	0.71	0.93
	3	0.56	0.79	0.85	0.86	0.71	0.82
	4	0.94	1.27	0.85	0.86	0.89	1.07
	5	0.25	0.39	0.48	0.58	0.36	0.50
	6	0.54	0.69	0.73	0.83	0.63	0.75
	7	1.17	1.37	0.48	0.58	0.82	1.09
	8	0.85	0.93	0.85	0.69	0.85	0.81
	9	0.54	0.78	1.35	1.25	0.95	1.10
	10	0.29	0.54	1.04	0.75	0.67	0.75
	Interval Collapsed	0.63	0.88	0.78	0.85	0.71	0.87
Automation Collapsed	1	0.37	0.53	0.50	0.73	0.44	0.64
	2	0.56	0.58	0.69	0.90	0.62	0.75
	3	0.46	0.68	0.67	0.77	0.56	0.73
	4	0.73	0.94	0.75	0.97	0.74	0.95
	5	0.25	0.38	0.62	0.72	0.44	0.61
	6	0.51	0.68	0.71	0.84	0.61	0.77
	7	0.76	1.28	0.67	0.74	0.71	1.04
	8	0.66	0.75	0.98	0.85	0.82	0.82
	9	0.66	0.81	1.29	1.14	0.97	1.03
	10	0.55	0.82	0.85	0.70	0.70	0.78
	Interval Collapsed	0.55	0.79	0.77	0.87	0.66	0.84

Table AA-3. Interval Speed: Means and Standard Deviations.

SPD	Interval	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	1	0.08	0.21	0.24	0.53	0.16	0.40
	2	0.26	0.56	0.36	0.45	0.31	0.50
	3	0.08	0.21	0.68	1.06	0.38	0.81
	4	0.51	0.71	0.80	1.15	0.66	0.95
	5	0.83	1.15	0.43	0.46	0.63	0.89
	6	0.58	0.87	1.24	1.05	0.91	1.01
	7	0.95	1.22	1.30	0.86	1.13	1.05
	8	0.76	1.22	0.86	1.01	0.81	1.10
	9	0.26	0.56	0.61	0.86	0.44	0.73
	10	0.58	1.08	0.93	0.98	0.75	1.03
	Interval Collapsed	0.49	0.89	0.75	0.92	0.62	0.91
Automation 2	1	0.22	0.34	0.44	0.67	0.33	0.54
	2	0.22	0.34	0.63	1.17	0.42	0.87
	3	0.40	0.57	0.75	0.72	0.58	0.66
	4	0.34	0.42	0.50	0.57	0.42	0.50
	5	0.40	0.57	0.32	0.39	0.36	0.48
	6	0.90	1.47	1.00	1.23	0.95	1.34
	7	0.72	0.83	1.38	1.56	1.05	1.27
	8	0.34	0.56	0.82	1.01	0.58	0.84
	9	0.47	0.86	0.57	0.67	0.52	0.76
	10	0.47	0.68	0.69	0.64	0.58	0.66
	Interval Collapsed	0.45	0.74	0.71	0.95	0.58	0.86
Automation 3	1	0.11	0.25	0.28	0.39	0.20	0.33
	2	0.18	0.33	0.59	0.57	0.38	0.50
	3	0.11	0.25	0.78	1.03	0.45	0.81
	4	1.05	1.28	0.90	0.75	0.98	1.03
	5	0.61	1.00	0.84	1.01	0.73	0.99
	6	0.78	0.89	1.09	1.02	0.93	0.95
	7	0.72	1.16	0.97	0.80	0.84	0.99
	8	0.53	0.86	0.78	0.62	0.65	0.75
	9	0.40	0.77	0.78	0.96	0.59	0.88
	10	0.22	0.34	1.03	1.33	0.62	1.04
	Interval Collapsed	0.47	0.84	0.80	0.89	0.64	0.88
Automation Collapsed	1	0.14	0.27	0.32	0.54	0.23	0.43
	2	0.22	0.41	0.53	0.79	0.37	0.64
	3	0.20	0.40	0.74	0.93	0.47	0.76
	4	0.64	0.91	0.74	0.86	0.69	0.88
	5	0.61	0.94	0.53	0.70	0.57	0.82
	6	0.75	1.10	1.11	1.08	0.93	1.10
	7	0.79	1.07	1.22	1.12	1.00	1.11
	8	0.54	0.92	0.82	0.88	0.68	0.90
	9	0.38	0.73	0.65	0.83	0.51	0.79
	10	0.42	0.76	0.88	1.01	0.65	0.92
	Interval Collapsed	0.47	0.82	0.75	0.92	0.61	0.88

Table AA-4. Interval Handoffs Accepted: Means and Standard Deviations.

HOs ACC	Interval	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	1	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX
	2	3.20	2.05	4.71	1.61	3.96	1.97
	3	3.01	1.60	4.71	1.61	3.86	1.80
	4	4.58	1.67	4.52	2.12	4.55	1.88
	5	4.08	1.06	5.40	1.69	4.74	1.54
	6	3.95	1.57	5.27	2.04	4.61	1.91
	7	3.39	2.00	4.08	0.68	3.74	1.51
	8	2.64	1.05	5.52	1.12	4.08	1.81
	9	3.26	2.06	4.83	1.59	4.05	1.98
	10	3.58	1.83	5.71	1.42	4.64	1.94
	Interval Collapsed	3.58	1.66	4.89	1.57	4.23	1.74
Automation 2	1	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX
	2	2.90	1.61	4.35	1.53	3.63	1.71
	3	3.22	1.40	4.04	1.03	3.63	1.28
	4	4.15	1.09	4.54	1.78	4.35	1.46
	5	4.34	0.78	5.16	1.60	4.75	1.31
	6	3.65	1.27	4.54	1.08	4.10	1.24
	7	3.03	1.06	4.54	0.87	3.78	1.23
	8	3.28	1.15	4.97	1.09	4.13	1.40
	9	3.34	1.32	4.47	1.49	3.91	1.50
	10	2.84	1.25	4.91	1.23	3.88	1.61
	Interval Collapsed	3.49	1.25	4.56	1.28	4.03	1.37
Automation 3	1	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX
	2	2.52	1.77	4.15	1.41	3.34	1.78
	3	3.08	1.93	4.53	1.25	3.81	1.76
	4	4.08	1.34	4.53	1.66	4.31	1.50
	5	4.40	1.49	5.28	1.89	4.84	1.74
	6	4.03	0.73	5.15	2.01	4.59	1.60
	7	3.15	1.56	4.15	1.41	3.65	1.55
	8	2.78	1.22	5.28	1.03	4.03	1.69
	9	3.15	1.52	5.72	1.79	4.44	2.09
	10	3.59	0.83	4.09	1.29	3.84	1.10
	Interval Collapsed	3.49	1.45	4.70	1.55	4.10	1.61
Automation Collapsed	1	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX
	2	2.88	1.80	4.40	1.50	3.64	1.82
	3	3.11	1.62	4.43	1.32	3.77	1.62
	4	4.27	1.37	4.53	1.83	4.40	1.61
	5	4.27	1.14	5.28	1.70	4.78	1.52
	6	3.88	1.22	4.99	1.76	4.43	1.61
	7	3.19	1.56	4.26	1.03	3.73	1.42
	8	2.90	1.15	5.26	1.08	4.08	1.62
	9	3.25	1.63	5.01	1.68	4.13	1.87
	10	3.34	1.38	4.90	1.45	4.12	1.62
	Interval Collapsed	3.52	1.46	4.72	1.47	4.12	1.58

We omitted the first interval of data due to automatic system handoff acceptance.

Table AA-5. Interval Handoffs Initiated: Means and Standard Deviations.

HOs INT	Interval	Low Load		High Load		Load Collapsed	
		Means	SDs	Means	SDs	Means	SDs
Automation 1	1	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX
	2	-2.57	1.94	-3.20	1.84	-2.89	1.89
	3	-4.13	1.86	-4.83	2.00	-4.48	1.93
	4	-2.82	1.43	-3.83	1.47	-3.32	1.52
	5	-3.13	1.37	-4.33	1.74	-3.73	1.65
	6	-4.13	2.03	-5.95	1.86	-5.04	2.13
	7	-4.32	1.78	-4.20	2.04	-4.26	1.88
	8	-3.88	1.96	-6.08	1.90	-4.98	2.20
	9	-3.82	1.64	-4.45	1.26	-4.14	1.47
	10	-2.82	1.28	-4.33	1.58	-3.57	1.61
	Interval Collapsed	-3.57	1.70	-4.52	1.83	-4.05	1.83
Automation 2	1	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX
	2	-3.08	2.02	-3.96	1.61	-3.52	1.85
	3	-3.39	1.42	-4.52	2.36	-3.96	2.00
	4	-3.64	1.41	-4.27	1.43	-3.96	1.44
	5	-3.95	1.57	-4.02	1.10	-3.99	1.33
	6	-3.52	1.60	-5.08	1.97	-4.30	1.94
	7	-3.89	1.82	-4.83	2.63	-4.36	2.28
	8	-3.33	1.36	-4.96	2.31	-4.14	2.04
	9	-3.64	1.32	-5.15	1.88	-4.39	1.77
	10	-4.27	2.77	-4.52	1.58	-4.39	2.22
	Interval Collapsed	-3.68	1.65	-4.54	1.83	-4.11	1.79
Automation 3	1	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX
	2	-3.01	1.33	-3.14	1.68	-3.08	1.49
	3	-3.51	1.94	-4.27	1.73	-3.89	1.85
	4	-3.08	1.53	-4.08	1.95	-3.58	1.80
	5	-3.45	1.79	-4.33	1.85	-3.89	1.85
	6	-4.45	1.67	-5.33	1.65	-4.89	1.69
	7	-3.45	1.32	-5.02	1.85	-4.24	1.77
	8	-3.77	2.47	-6.08	3.54	-4.92	3.22
	9	-3.14	1.64	-4.89	2.33	-4.02	2.17
	10	-3.27	1.30	-4.52	1.41	-3.89	1.48
	Interval Collapsed	-3.52	1.63	-4.57	2.07	-4.05	1.93
Automation Collapsed	1	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX	XXXXX
	2	-2.89	1.77	-3.43	1.72	-3.16	1.75
	3	-3.68	1.74	-4.54	2.02	-4.11	1.92
	4	-3.18	1.47	-4.06	1.61	-3.62	1.60
	5	-3.51	1.59	-4.23	1.57	-3.87	1.61
	6	-4.03	1.78	-5.45	1.83	-4.74	1.93
	7	-3.89	1.66	-4.68	2.18	-4.29	1.97
	8	-3.66	1.95	-5.70	2.67	-4.68	2.55
	9	-3.53	1.53	-4.83	1.86	-4.18	1.81
	10	-3.45	1.97	-4.45	1.50	-3.95	1.81
	Interval Collapsed	-3.59	1.66	-4.54	1.91	-4.07	1.85

We omitted the first interval of data due to automatic system handoff acceptance.

SCENARIO RESULTS

Table AA-6. Altitude: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.027	0.004	7.347	1	7	.030
Automation	0.002	0.002	1.017	2	14	.387
Load X Automation	0.003	0.002	1.909	2	14	.185

Table AA-7. Heading: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.001	0.001	1.396	1	7	.276
Automation	0.002	0.003	0.545	2	14	.592
Load X Automation	0.003	0.004	0.660	2	14	.532

Table AA-8. Speed: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.004	0.000	19.786	1	7	.003
Automation	0.001	0.001	0.562	2	14	.583
Load X Automation	0.002	0.002	1.402	2	14	.279

Table AA-9. Handoffs Accepted: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.001	0.002	0.610	1	7	.460
Automation	0.002	0.001	1.096	2	14	.361
Load X Automation	0.001	0.004	0.204	2	14	.817

Table AA-10. Handoffs Initiated: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.016	0.008	2.018	1	7	.198
Automation	0.002	0.003	0.503	2	14	.615
Load X Automation	0.005	0.004	1.297	2	14	.304

INTERVAL RESULTS

Table AA-11. Altitude: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	72.247	1.987	36.365	1	7	.001
Interval 1	0.003	0.128	0.026	1	7	.877
Interval 2	6.328	0.549	11.527	1	7	.012
Interval 3	1.393	0.304	34.206	1	7	.001
Interval 4	2.172	0.641	3.391	1	7	.108
Interval 5	0.001	0.344	0.002	1	7	.966
Interval 6	1.353	0.224	6.030	1	7	.044
Interval 7	2.496	0.171	14.589	1	7	.007
Interval 8	5.428	0.177	3.645	1	7	.001
Interval 9	2.366	0.321	7.361	1	7	.030
Interval 10	2.766	0.089	31.026	1	7	.001
Automation	2.114	1.450	1.458	2	14	.266
Interval	11.719	0.915	12.812	9	63	.000
Low Load	2.563	0.248	1.345	9	63	.000
High Load	2.368	0.311	7.611	9	63	.000
Load x Automation	0.152	1.204	0.126	2	14	.883
Load x Interval	3.075	0.762	4.035	9	63	.000
Automation x Interval	0.506	0.989	3.622	18	126	.949
Load x Automation x Interval	0.404	0.952	4.145	18	126	.980

Table AA-12. Altitude: Tukey HSD Post Hoc Test Results.

	Interval 1	Interval 2	Interval 3	Interval 4	Interval 5	Interval 6	Interval 7	Interval 8	Interval 9	Interval 10
	4.392	3.673	3.746	3.642	3.183	3.386	3.490	3.574	2.584	2.907
Interval 1		.016	.046	.010	.000	.000	.001	.003	.000	.000
Interval 2	.016		1.000	1.000	.285	.899	.995	1.000	.000	.008
Interval 3	.046	1.000		1.000	.133	.706	.948	.997	.000	.002
Interval 4	.010	1.000	1.000		.374	.948	.999	1.000	.000	.013
Interval 5	.000	.285	.133	.374		.989	.856	.603	.085	.918
Interval 6	.000	.899	.706	.948	.989		1.000	.994	.004	.313
Interval 7	.001	.995	.948	.999	.856	1.000		1.000	.001	.104
Interval 8	.003	1.000	.997	1.000	.603	.994	1.000		.000	.035
Interval 9	.000	.000	.000	.000	.085	.004	.001	.000		.816
Interval 10	.000	.008	.002	.013	.918	.313	.104	.035	.816	

Table AA-13. Heading: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load	5.802	0.335	17.323	1	7	.004
Automation	0.805	0.660	1.221	2	14	.325
Interval	1.325	0.194	6.831	9	63	.000
Load X Automation	0.443	1.398	0.317	2	14	.734
Load X Interval	0.492	0.418	1.176	9	63	.326
Automation X Interval	0.144	0.287	0.502	18	126	.953
Load X Automation X Interval	0.349	0.282	1.236	18	126	.243

Table AA-14. Heading: Tukey HSD Post Hoc Test Results.

	Interval 1	Interval 2	Interval 3	Interval 4	Interval 5	Interval 6	Interval 7	Interval 8	Interval 9	Interval 10
	.436	.624	.561	.739	.436	.610	.714	.818	.975	.704
Interval 1		.544	.926	.040	1.000	.649	.081	.003	.000	.108
Interval 2	.544		.999	.956	.544	1.000	.991	.493	.008	.996
Interval 3	.926	.999		.623	.926	1.000	.792	.140	.001	.851
Interval 4	.040	.956	.623		.040	.913	1.000	.996	.228	1.000
Interval 5	1.000	.544	.926	.040		.649	.081	.003	.000	.108
Interval 6	.649	1.000	1.000	.913	.649		.976	.393	.005	.988
Interval 7	.081	.991	.792	1.000	.081	.976		.976	.128	1.000
Interval 8	.003	.493	.140	.996	.003	.393	.976		.770	.956
Interval 9	.000	.008	.001	.228	.000	.005	.128	.770		.098
Interval 10	.108	.996	.851	1.000	.108	.988	1.000	.956	.098	

Table AA-15. Speed: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load	9.766	0.161	6.616	1	7	.000
Automation	0.153	0.879	0.174	2	14	.842
Interval	2.666	.354	7.539	9	63	.000
Load X Automation	0.077	0.673	0.114	2	14	.893
Load X Interval	0.400	0.254	1.577	9	63	.142
Automation X Interval	0.313	0.343	0.914	18	126	.564
Load X Automation X Interval	0.210	0.323	0.648	18	126	.855

Table AA-16. Speed: Tukey HSD Post Hoc Test Results.

	Interval 1	Interval 2	Interval 3	Interval 4	Interval 5	Interval 6	Interval 7	Interval 8	Interval 9	Interval 10
	.234	.373	.467	.686	.571	.932	1.005	.682	.515	.650
Interval 1		.978	.658	.015	.168	.000	.000	.016	.396	.034
Interval 2	.978		.999	.252	.828	.001	.000	.268	.975	.414
Interval 3	.658	.999		.731	.997	.010	.002	.752	1.000	.883
Interval 4	.015	.252	.731		.994	.585	.227	1.000	.921	1.000
Interval 5	.168	.828	.997	.994		.108	.022	.996	1.000	1.000
Interval 6	.000	.001	.010	.585	.108		1.000	.562	.033	.393
Interval 7	.000	.000	.002	.227	.022	1.000		.213	.006	.122
Interval 8	.016	.268	.752	1.000	.996	.562	.213		.931	1.000
Interval 9	.396	.975	1.000	.921	1.000	.033	.006	.931		.981
Interval 10	.034	.414	.883	1.000	1.000	.393	.122	1.000	.981	

Table AA-17. Handoffs Accepted: ANOVA Results.

	Ms effect	MS Error	*F*	df 1	df 2	*p*-level
Load	191.086	3.021	63.243	1	7	.000
Interval 2	9.343	0.251	37.185	1	7	.000
Interval 3	6.969	0.303	23.008	1	7	.002
Interval 4	0.265	0.119	2.234	1	7	.179
Interval 5	4.060	0.207	19.659	1	7	.003
Interval 6	4.921	0.234	21.005	1	7	.003
Interval 7	4.558	0.170	26.829	1	7	.001
interval 8	22.262	0.135	165.370	1	7	.000
interval 9	12.320	0.545	22.620	1	7	.002
interval 10	9.828	0.490	2.054	1	7	.003
automation	2.000	3.104	0.644	2	14	.540
interval	6.724	1.002	6.711	8	56	.000
low load	2.404	0.250	9.633	8	56	.000
high load	1.191	0.265	4.492	8	56	.000
Load x Automation	0.599	3.834	0.156	2	14	.857
Load x Interval	4.061	0.542	7.492	8	56	.000
Automation x Interval	0.755	0.758	0.996	16	112	.467
Load x Automation x Interval	1.094	0.950	1.152	16	112	.318

Table AA-18. Handoffs Accepted: Tukey HSD Post Hoc Test Results.

	Interval 2	Interval 3	Interval 4	Interval 5	Interval 6	Interval 7	Interval 8	Interval 9	Interval 10
	3.640	3.766	4.401	4.776	4.434	3.725	4.079	4.132	4.121
Interval 2[a]		1.000	.013	.000	.008	1.000	.453	.303	.331
Interval 3	1.000		.067	.000	.045	1.000	.834	.688	.720
Interval 4	.013	.067		.659	1.000	.040	.814	.921	.904
Interval 5	.000	.000	.659		.758	.000	.031	.060	.053
Interval 6	.008	.045	1.000	.758		.026	.724	.861	.837
Interval 7	1.000	1.000	.040	.000	.026		.724	.558	.592
Interval 8	.453	.834	.814	.031	.724	.724		1.000	1.000
Interval 9	.303	.688	.921	.060	.861	.558	1.000		1.000
Interval 10	.331	.720	.904	.053	.837	.592	1.000	1.000	

[a] We omitted the first interval of data due to automatic system handoff acceptance.

Table AA-19. Handoffs Initiated: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Load	121.610	1.300	93.526	1	7	.000
Automation	0.223	1.033	0.216	2	14	.809
Interval	11.850	1.329	8.914	8	56	.000
Load X Automation	0.406	0.726	0.560	2	14	.583
Load X Interval	2.530	1.528	1.656	8	56	.130
Automation X Interval	1.798	1.676	1.073	16	112	.389
Load X Automation X Interval	1.181	1.796	0.658	16	112	.829

Table AA-20. Handoffs Initiated: Tukey HSD Post Hoc Test Results.

	Interval 2	Interval 3	Interval 4	Interval 5	Interval 6	Interval 7	Interval 8	Interval 9	Interval 10
	-3.161	-4.108	-3.619	-3.870	-4.745	-4.286	-4.682	-4.182	-3.953
Interval 2[a]		.005	.585	.085	.000	.001	.000	.002	.035
Interval 3	.005		.497	.983	.170	.998	.285	1.000	.999
Interval 4	.585	.497		.977	.001	.128	.001	.308	.886
Interval 5	.085	.983	.977		.013	.701	.027	.919	1.000
Interval 6	.000	.170	.001	.013		.585	1.000	.310	.035
Interval 7	.001	.998	.128	.701	.585		.755	1.000	.887
Interval 8	.000	.285	.001	.027	1.000	.755		.468	.069
Interval 9	.002	1.000	.308	.919	.310	1.000	.468		.987
Interval 10	.035	.999	.886	1.000	.035	.887	.069	.987	

[a] We omitted the first interval of data due to automatic system handoff acceptance.

Table AA-21. Number and Duration of Conflicts: Means and Standard Deviations (number; seconds).

Conflicts		Load				Load Collapsed	
		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs
Auto 1	Number	0.19	0.40	0.33	0.49	0.26	0.44
	Duration	4.75	11.52	13.53	31.25	9.00	23.28
Auto 2	Number	0.27	0.46	0.50	0.90	0.37	0.69
	Duration	16.40	49.42	11.00	20.95	14.00	38.84
Auto 3	Number	0.27	0.46	0.53	0.52	0.40	0.50
	Duration	21.27	69.55	8.47	9.52	14.87	49.21
Auto Coll.	Number	0.24	0.43	0.45	0.63	0.34	0.54
	Duration	13.93	48.57	11.00	22.06	12.53	38.10

Table AA-22. Number of Aircraft Handled Per Minute: Means and Standard Deviations.

Aircraft Handled		R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	Number	1.79	0.22	2.57	0.29	2.18	0.47	1.79	0.22	2.57	0.29	2.18	0.47	1.79	0.21	2.57	0.29	2.18	0.47
Auto 2	Number	1.80	0.23	2.47	0.28	2.13	0.43	1.80	0.23	2.47	0.28	2.13	0.43	1.80	0.23	2.47	0.28	2.13	0.42
Auto 3	Number	1.79	0.24	2.49	0.29	2.14	0.44	1.79	0.24	2.49	0.29	2.14	0.44	1.79	0.24	2.49	0.29	2.14	0.44
Auto Coll.	Number	1.79	0.22	2.51	0.29	2.15	0.44	1.79	0.22	2.51	0.29	2.15	0.44	1.79	0.22	2.51	0.29	2.15	0.44

Table AA-23. Number of Aircraft Handled Per Minute: ANOVA Results.

	MS Effect	MS Error	F	df 1	df 2	p-level
Position	0.000	0.059	0.000	1	15	1.000
Load	24.898	0.092	270.952	1	15	.000
Automation	0.039	0.117	0.336	2	30	.717
Position X Load	0.000	0.040	0.000	1	15	1.000
Position X Automation	0.000	0.085	0.000	2	30	1.000
Load X Automation	0.057	0.050	1.131	2	30	.336
Position X Load X Automation	0.000	0.055	0.000	2	30	1.000

APPENDIX BB

Controller Interactions with DSR

Table BB-1. Information Pickup: Means and Standard Deviations (number).

	Information Pickup	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	Total Number of Entries	265.17	42.55	331.40	46.39	295.27	54.43	100.67	61.60	124.80	63.77	111.64	60.68	182.92	99.64	228.10	120.15	203.45	109.53
	Flight Plan Readouts	7.50	6.38	5.60	3.97	6.64	5.26	0.00	0.00	0.00	0.00	0.00	0.00	3.75	5.82	2.80	3.97	3.32	4.97
	J-rings	2.67	2.34	2.40	1.67	2.55	1.97	0.67	1.21	0.40	0.89	0.55	1.04	1.67	2.06	1.40	1.65	4.97	1.55
	Route Displays	2.33	2.42	0.60	0.55	1.55	1.97	1.17	2.04	0.40	0.55	0.82	1.54	1.75	2.22	0.50	0.53	1.18	1.76
Auto 2	Total Number of Entries	279.50	12.82	342.67	72.14	317.40	63.32	24.50	22.29	92.50	32.92	65.30	44.73	152.00	137.34	217.58	141.16	191.35	139.90
	Flight Plan Readouts	6.25	4.35	7.33	7.65	6.90	6.26	0.00	0.00	0.00	0.00	0.00	0.00	3.13	4.39	3.67	6.43	3.45	5.58
	J-rings	1.75	3.50	2.67	2.66	2.30	2.87	0.25	0.50	0.67	1.21	0.50	0.97	1.00	2.45	1.67	2.23	1.40	2.28
	Route Displays	4.00	3.56	1.50	1.76	2.50	2.76	1.25	1.89	1.17	0.47	1.20	1.55	2.63	3.02	1.33	1.56	1.85	2.28
Auto 3	Total Number of Entries	280.17	55.79	314.20	90.00	295.64	71.50	44.67	26.45	75.20	52.57	58.54	41.34	162.42	129.83	194.70	143.86	177.09	134.05
	Flight Plan Readouts	9.67	7.92	3.40	2.07	6.82	6.69	0.00	0.00	0.00	0.00	0.00	0.00	4.83	7.35	1.70	2.50	3.41	5.79
	J-rings	3.00	4.38	3.60	4.39	3.27	4.17	0.00	0.00	1.00	1.73	0.45	1.21	1.50	3.34	2.30	3.43	5.79	1.86
	Route Displays	1.67	2.66	2.00	2.92	1.82	2.64	0.17	0.41	0.60	1.34	0.36	0.92	0.92	1.98	1.30	2.26	1.09	2.07
Auto Coll.	Total Number of Entries	274.38	41.57	330.25	67.94	302.31	62.25	60.63	51.85	97.19	51.05	78.91	53.92	167.50	118.02	213.72	132.33	190.61	126.54
	Flight Plan Readouts	8.00	6.35	5.56	5.33	6.78	5.90	0.00	0.00	0.00	0.00	0.00	0.00	4.00	6.00	2.78	4.66	3.39	5.37
	J-rings	2.56	3.31	2.88	2.92	2.72	3.07	0.31	0.79	0.69	1.25	0.50	1.05	1.44	2.63	1.78	2.47	1.61	2.54
	Route Displays	2.50	2.78	1.38	1.93	1.94	2.42	0.81	1.56	0.75	1.18	0.78	1.36	1.66	2.38	1.06	1.61	1.36	2.03

Table BB-2. ATC Control Actions: Means and Standard Deviations (number).

	ATC Control Actions	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	Hard Altitude	0.67	1.21	5.40	8.29	2.82	5.86	1.17	1.33	2.20	1.64	1.64	1.50	0.92	1.24	3.80	5.88	2.23	4.22
	Interim Altitude	18.17	7.60	26.40	8.20	21.91	8.62	4.17	4.22	7.80	9.18	5.82	6.79	11.17	9.37	17.10	12.78	13.86	11.19
	Removal of Interim Altitude	2.50	1.52	4.60	2.88	3.45	2.38	1.00	1.55	0.60	0.55	0.82	1.17	1.75	1.66	2.60	2.88	2.14	2.27
	Route Change	2.50	2.07	2.20	2.77	2.36	2.29	3.17	2.93	2.20	2.95	2.73	2.83	2.83	2.44	2.20	2.70	2.55	2.52
Auto 2	Hard Altitude	1.25	2.50	1.50	1.52	1.40	1.84	0.25	0.50	2.17	2.79	1.40	2.32	0.75	1.75	1.83	2.17	1.40	2.04
	Interim Altitude	14.75	10.81	28.33	7.31	22.90	10.86	1.25	1.89	2.17	3.37	1.80	2.78	8.00	10.18	15.25	14.70	12.35	13.29
	Removal of Interim Altitude	9.75	5.68	4.33	1.97	6.50	4.55	0.00	0.00	0.83	0.98	0.50	0.85	4.88	6.40	2.58	2.35	3.50	4.43
	Route Change	5.25	3.86	3.00	3.52	3.90	3.63	1.75	2.87	2.50	3.89	2.20	3.36	3.50	3.66	2.75	3.55	3.05	3.52
Auto 3	Hard Altitude	2.33	1.86	3.80	4.66	3.00	3.32	0.50	0.84	2.00	1.00	1.18	1.17	1.42	1.68	2.90	3.31	2.09	2.60
	Interim Altitude	18.33	6.74	28.20	12.51	22.82	10.58	1.83	1.60	6.20	3.90	3.82	3.54	10.08	9.80	17.20	14.52	13.32	12.40
	Removal of Interim Altitude	3.00	1.67	5.20	2.68	4.00	2.37	0.67	1.21	0.60	0.89	0.64	1.03	1.83	1.85	2.90	3.07	2.32	2.48
	Route Change	3.83	3.97	1.00	1.41	2.55	3.30	2.67	3.08	0.60	1.34	1.73	2.57	3.25	3.44	0.80	1.32	2.14	2.92
Auto Coll.	Hard Altitude	1.44	1.86	3.44	5.27	2.44	4.02	0.69	1.01	2.13	1.89	1.41	1.66	1.06	1.52	2.78	3.95	1.92	3.09
	Interim Altitude	17.38	7.76	27.69	8.85	22.53	9.72	2.56	3.03	5.19	6.05	3.88	4.89	9.97	9.50	16.44	13.65	13.20	12.11
	Removal of Interim Altitude	4.50	4.24	4.69	2.36	4.59	3.38	0.63	1.20	0.69	0.79	0.66	1.00	2.56	3.65	2.69	2.67	2.63	3.17
	Route Change	3.69	3.30	2.13	2.73	2.91	3.08	2.63	2.83	1.81	2.93	2.22	2.86	3.16	3.07	1.97	2.79	2.56	2.97

Table BB-3. Station Keeping Activities: Means and Standard Deviations (number).

	House Keeping Activities	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	Handoffs Initiated	44.50	21.54	61.40	14.98	52.18	19.99	21.33	15.47	17.80	19.80	19.73	16.73	32.92	21.59	39.60	28.32	35.95	24.49
	Handoffs Accepted	68.33	15.93	103.80	30.53	84.45	29.03	25.17	16.52	29.20	12.17	27.00	14.15	46.75	27.34	66.50	45.01	55.73	36.90
	Datablock Movements	91.33	28.26	83.40	23.09	87.73	25.10	41.67	32.36	63.20	54.50	51.45	42.88	66.50	38.88	73.30	40.87	69.59	38.99
Auto 2	Handoffs Initiated	62.00	16.12	57.00	24.02	56.00	20.35	4.00	2.45	24.00	18.75	16.00	17.44	33.00	32.79	40.50	26.82	37.50	28.75
	Handoffs Accepted	69.75	11.00	94.17	24.88	84.40	23.31	9.75	10.72	26.83	15.38	20.00	15.73	39.75	33.61	60.50	40.32	52.20	38.29
	Datablock Movements	80.75	24.10	108.00	35.35	97.10	32.95	5.75	5.91	31.17	15.05	21.00	17.60	43.25	43.26	69.58	47.76	59.06	46.75
Auto 3	Handoffs Initiated	56.00	8.97	60.00	36.69	57.82	24.15	11.67	7.12	12.80	13.61	12.18	9.99	33.83	24.41	36.40	36.05	35.00	29.51
	Handoffs Accepted	71.17	13.51	89.60	26.18	79.55	21.40	13.50	9.29	19.40	16.52	16.18	12.72	42.33	32.08	54.50	42.36	47.86	36.70
	Datablock Movements	86.67	32.36	84.40	38.31	85.64	33.35	13.17	10.13	29.40	24.40	20.55	19.01	49.92	44.68	56.90	41.92	53.09	42.56
Auto Coll.	Handoffs Initiated	53.19	16.96	59.31	24.80	56.25	21.13	13.38	12.16	18.56	17.15	15.97	14.86	33.28	24.89	38.94	29.47	36.11	27.21
	Handoffs Accepted	69.75	13.09	95.75	25.94	82.75	24.14	16.94	13.72	25.25	14.45	21.09	14.49	43.34	29.90	60.50	41.34	51.92	36.82
	Datablock Movements	86.94	27.38	92.94	33.10	89.94	30.03	22.00	25.43	40.63	35.69	31.31	31.92	54.47	42.00	66.78	43.04	60.63	42.64

Table BB-4. Total Number of Entries: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	793291.563	2862.092	277.172	1	52	.000
Load	35588.352	2862.092	12.434	1	52	.001
Automation	4311.262	2862.092	1.506	2	52	.231
Position X Load	722.574	2862.092	0.252	1	52	.617
Position X Automation	6471.628	2862.092	2.261	2	52	.114
Load X Automation	1428.460	2862.092	0.499	2	52	.610
Position X Load X Automation	826.108	2862.092	0.289	2	52	.750

Table BB-5. Information Pickup: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Position	.485	0.515	1.062	1.062	17.705	3	50	.000
Load	.945	0.055	0.058	0.058	0.973	3	50	.413
Automation	.949	0.052	0.054	0.049	0.445	6	100	.847
Position X Load	.958	0.042	0.044	0.044	0.737	3	50	.535
Position X Automation	.976	0.025	0.025	0.018	0.207	6	100	.974
Load X Automation	.891	0.111	0.121	0.109	0.993	6	100	.434
Position X Load X Automation	.943	0.057	0.060	0.060	0.496	6	100	.810

Table BB-6. Flight Plan Readout: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	686.984	18.410	37.316	1	52	.000
Load	21.815	18.410	1.185	1	52	.281
Automation	0.105	18.410	0.006	2	52	.994
Position X Load	21.815	18.410	1.185	1	52	.281
Position X Automation	0.105	18.410	0.006	2	52	.994
Load X Automation	17.623	18.410	0.957	2	52	.391
Position X Load X Automation	17.623	18.410	0.957	2	52	.391

Table BB-7. J-Rings: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	74.613	6.047	12.339	1	52	.001
Load	2.504	6.047	0.414	1	52	.523
Automation	1.713	6.047	0.283	2	52	.754
Position X Load	0.004	6.047	0.001	1	52	.979
Position X Automation	1.577	6.047	0.261	2	52	.771
Load X Automation	1.818	6.047	0.301	2	52	.742
Position X Load X Automation	0.259	6.047	0.043	2	52	.958

Table BB-8. Route Display: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	23.488	3.943	5.957	1	52	.018
Load	8.102	3.943	2.055	1	52	.158
Automation	4.961	3.943	1.258	2	52	.293
Position X Load	5.275	3.943	1.338	1	52	.253
Position X Automation	1.178	3.943	0.299	2	52	.743
Load X Automation	4.866	3.943	1.234	2	52	.299
Position X Load X Automation	1.734	3.943	0.440	2	52	.647

Table BB-9. ATC Control Actions: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Position	.230	0.770	3.351	3.351	41.056	4	49	.000
Load	.597	0.403	0.676	0.676	8.284	4	49	.000
Automation	.779	0.222	0.281	0.273	1.626	8	98	.127
Position X Load	.878	0.122	0.139	0.139	1.700	4	49	.165
Position X Automation	.790	0.212	0.262	0.248	1.530	8	98	.157
Load X Automation	.834	0.170	0.195	0.170	1.164	8	98	.329
Position X Load X Automation	.756	0.249	0.317	0.296	1.839	8	98	.079

Table BB-10. Hard Altitudes: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	19.324	9.301	2.078	1	52	.155
Load	51.657	9.301	5.554	1	52	.022
Automation	6.448	9.301	0.693	2	52	.505
Position X Load	1.739	9.301	0.187	1	52	.667
Position X Automation	3.642	9.301	0.392	2	52	.678
Load X Automation	4.689	9.301	0.504	2	52	.607
Position X Load X Automation	9.856	9.301	1.060	2	52	.354

Table BB-11. Route Changes: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	10.439	9.198	1.135	1	52	.292
Load	25.556	9.198	2.778	1	52	.102
Automation	6.183	9.198	0.672	2	52	.515
Position X Load	4.178	9.198	0.454	1	52	.503
Position X Automation	6.951	9.198	0.756	2	52	.475
Load X Automation	5.545	9.198	0.603	2	52	.551
Position X Load X Automation	4.326	9.198	0.470	2	52	.627

Table BB-12. Interim Altitudes: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	5334.458	49.944	106.808	1	52	.000
Load	716.678	49.944	14.350	1	52	.000
Automation	35.445	49.944	0.710	2	52	.497
Position X Load	225.357	49.944	4.512	1	52	.038
Position X Automation	18.928	49.944	0.379	2	52	.686
Load X Automation	2.786	49.944	0.056	2	52	.946
Position X Load X Automation	24.448	49.944	0.490	2	52	.616

Table BB-13. Removal of Interim Altitudes: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	286.797	4.464	64.241	1	52	.000
Load	0.245	4.464	0.055	1	52	.816
Automation	14.379	4.464	3.221	2	52	.048
Position X Load	0.957	4.464	0.214	1	52	.645
Position X Automation	21.312	4.464	4.774	2	52	.012
Load X Automation	17.668	4.464	3.958	2	52	.025
Position X Load X Automation	31.075	4.464	6.960	2	52	.002

Table BB-14. Station Keeping: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Position	.190	0.810	4.257	4.257	70.951	3	50	.000
Load	.773	0.227	0.293	0.293	4.887	3	50	.005
Automation	.910	0.090	0.099	0.098	0.808	6	100	.566
Position X Load	.888	0.112	0.127	0.127	2.110	3	50	.111
Position X Automation	.887	0.113	0.127	0.125	1.032	6	100	.409
Load X Automation	.963	0.037	0.038	0.028	0.319	6	100	.926
Position X Load X Automation	.904	0.097	0.106	0.101	0.864	6	100	.524

Table BB-15. Handoffs Initiated: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	27021.953	350.172	77.168	1	52	.000
Load	487.935	350.172	1.393	1	52	.243
Automation	14.636	350.172	0.042	2	52	.959
Position X Load	1.257	350.172	0.004	1	52	.952
Position X Automation	267.619	350.172	0.764	2	52	.471
Load X Automation	36.886	350.172	0.105	2	52	.900
Position X Load X Automation	664.876	350.172	1.899	2	52	.160

Table BB-16. Handoffs Accepted: ANOVA Results.

	MS Effect	MS Error	F	df 1	df 2	p-level
Position	60480.059	329.186	183.726	1	52	.000
Load	4823.961	329.186	14.654	1	52	.000
Automation	405.776	329.186	1.233	2	52	.300
Position X Load	1144.213	329.186	3.476	1	52	.068
Position X Automation	43.276	329.186	0.131	2	52	.877
Load X Automation	116.685	329.186	0.354	2	52	.703
Position X Load X Automation	211.202	329.186	0.642	2	52	.531

Table BB-17. Datablock Movements: ANOVA Results.

	MS Effect	MS Error	F	df 1	df 2	p-level
Position	53321.758	893.799	59.657	1	52	.000
Load	2798.864	893.799	3.131	1	52	.083
Automation	1666.376	893.799	1.864	2	52	.165
Position X Load	925.341	893.799	1.035	1	52	.314
Position X Automation	2327.010	893.799	2.604	2	52	.084
Load X Automation	630.010	893.799	0.705	2	52	.499
Position X Load X Automation	318.709	893.799	0.357	2	52	.702

APPENDIX CC

Entry Questionnaire Results

Table CC -1. General Background Questions.

	Means	*SDs*
Use of vectoring for separation.	6.25	1.44
Use of vertical separation.	6.06	1.61
Use of speed control for separation.	5.06	1.95
Experience with video games.	5.13	3.01

Table CC-2. Importance of Aircraft Information.

Variable Label	Means	*SDs*
Current Aircraft Location	8.44	1.71
Current Altitude	8.25	1.57
Arrival Airport (within sector)	8.13	1.71
Most Recently Assigned Altitude	7.81	2.23
Current Heading	7.67	1.50
Density of Aircraft on Radar Display	7.63	1.45
Most Recently Assigned Heading	7.47	2.00
Aircraft Type	7.44	2.19
Entry Airspeed	7.31	2.69
Aircraft CallSign	7.19	3.45
Current Airspeed	7.19	2.04
Aircraft Waiting for Hand-off/Release	7.13	1.15
Most Recently Assigned Airspeed	6.75	2.08
Aircraft Holding/Spinning	6.75	2.38
Aircraft Near Exit Fix/Arrival Airport	6.69	2.15
Controller Ownership	6.50	1.86
Entry Altitude	6.50	2.55
Departure Airport (within sector)	5.69	2.65
Entry Fix	5.06	2.77
Exit Altitude	4.94	2.59
Exit Fix	4.56	2.90
Exit Airspeed	4.19	2.40
Aircraft Beacon Code	3.13	1.82

Table CC-3. Importance of Radarscope Information.

Variable Label	Means	*SDs*
Sector Boundaries	8.25	1.29
Restricted Area Boundaries	7.88	1.26
Filter Settings	6.56	2.25
Collision Alert	6.56	2.06
ILS Approaches	5.19	2.23
Obstructions	5.06	2.52
Airports	4.88	1.86
VORs	4.69	1.92
Fixes	4.44	2.00
Holding Patterns	4.19	2.48
System Clock	3.63	1.82
Future Aircraft List	2.13	1.89

APPENDIX DD

Post-Scenario Questionnaire Results

Table DD-1. Post-Scenario Questionnaire: Means and Standard Deviations.

	General Questions	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Coll.	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Automation 1	Simulation Realism	6.88	1.63	5.75	2.11	6.31	1.94	6.69	1.96	5.31	1.92	6.00	2.03	6.78	1.77	5.53	2.00	6.16	1.98
	Representation of Typical Workday	6.76	1.51	5.26	2.32	6.01	2.07	6.44	1.95	4.76	1.98	5.60	2.12	6.60	1.73	5.01	2.13	5.80	2.09
	How Well You Controlled Traffic	8.00	1.27	6.62	1.78	7.31	1.67	7.19	2.37	6.12	1.75	6.66	2.12	7.59	1.92	6.37	1.75	6.98	1.92
	Scenario Difficulty	6.42	2.83	5.35	2.41	5.88	2.64	6.92	2.73	5.86	2.09	6.39	2.45	6.67	2.75	5.60	2.24	6.13	2.54
	ATWIT Device Interference	2.12	1.30	2.22	1.26	2.17	1.26	2.59	1.39	2.68	1.60	2.64	1.48	2.36	1.35	2.45	1.44	2.40	1.38
	Oculometer Interference	2.17	0.94	2.41	0.84	2.29	0.88	3.30	2.22	3.29	2.12	3.29	2.14	2.73	1.77	2.85	1.65	2.79	1.70
	Simulation-Pilot Respondence	7.45	2.09	6.64	2.28	7.05	2.19	7.20	2.17	6.58	1.90	6.89	2.03	7.33	2.10	6.61	2.07	6.97	2.10
	DST Effectiveness Using TFPs	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX
	Competence Using DST	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX	XXX
Automation 2	Simulation Realism	6.19	1.76	5.82	1.72	6.00	1.72	6.50	1.71	6.32	1.62	6.41	1.64	6.35	1.71	6.07	1.66	6.21	1.68
	Representation of Typical Workday	5.85	1.88	5.20	1.82	5.52	1.85	5.98	1.82	5.57	1.69	5.77	1.74	5.91	1.82	5.39	1.74	5.65	1.79
	How Well You Controlled Traffic	8.19	1.52	7.19	1.22	7.69	1.45	7.19	2.37	6.94	1.61	7.06	2.00	7.69	2.02	7.06	1.41	7.37	1.76
	Scenario Difficulty	6.60	3.17	5.40	2.18	6.00	2.74	7.16	2.42	5.52	1.85	6.34	2.28	6.88	2.79	5.46	1.99	6.17	2.51
	ATWIT Device Interference	1.78	1.00	2.84	1.48	2.31	1.35	2.09	1.33	2.37	1.23	2.23	1.27	1.94	1.17	2.61	1.36	2.27	1.30
	Oculometer Interference	2.29	0.90	2.40	0.84	2.35	0.86	3.79	2.77	3.58	2.43	3.69	2.57	3.04	2.16	2.99	1.89	3.02	2.01
	Simulation-Pilot Respondence	7.84	1.59	6.78	1.78	7.31	1.74	7.84	1.50	7.29	1.76	7.57	1.64	7.84	1.52	7.04	1.76	7.44	1.68
	DST Effectiveness Using TFPs	4.79	1.30	4.81	1.31	4.80	1.28	4.52	2.21	4.62	1.99	4.57	2.07	4.66	1.79	4.71	1.66	4.68	1.71
	Competence Using DST	5.44	1.66	5.35	2.05	5.40	1.84	5.96	2.61	5.39	2.35	5.68	2.46	5.70	2.16	5.37	2.17	5.54	2.16
Automation 3	Simulation Realism	5.63	2.00	5.00	2.22	5.31	2.10	6.25	2.14	5.63	2.00	5.94	2.06	5.94	2.06	5.31	2.10	5.63	2.09
	Representation of Typical Workday	5.16	2.15	4.63	2.15	4.90	2.13	6.01	2.36	4.76	2.01	5.38	2.25	5.58	2.26	4.69	2.05	5.14	2.19
	How Well You Controlled Traffic	7.94	1.06	5.94	2.05	6.94	1.90	7.37	1.71	5.69	1.96	6.53	2.00	7.66	1.43	5.81	1.97	6.73	1.95
	Scenario Difficulty	5.97	2.90	4.42	2.74	5.19	2.89	5.92	3.02	4.61	2.69	5.26	2.89	5.94	2.91	4.51	2.68	5.23	2.87
	ATWIT Device Interference	1.87	0.76	4.06	2.99	2.97	2.42	2.06	1.33	3.19	2.03	2.62	1.78	1.97	1.07	3.62	2.56	2.80	2.12
	Oculometer Interference	2.47	0.76	2.29	0.90	2.38	0.82	3.41	2.17	3.61	2.73	3.51	2.43	2.94	1.67	2.95	2.11	2.94	1.89
	Simulation-Pilot Respondence	7.59	1.54	6.33	2.19	6.96	1.97	7.59	1.49	6.45	2.10	7.02	1.88	7.59	1.49	6.39	2.11	6.99	1.91
	DST Effectiveness Using TFPs	5.10	1.44	5.14	2.03	5.12	1.73	7.06	2.23	6.12	1.49	6.59	1.93	6.08	2.10	5.63	1.82	5.85	1.96
	Competence Using DST	5.29	1.84	4.73	1.87	5.01	1.85	6.40	2.40	5.24	2.14	5.82	2.31	5.84	2.18	4.99	1.99	5.42	2.12
Automation Collapsed	Simulation Realism	6.23	1.84	5.52	2.02	5.88	1.95	6.48	1.91	5.75	1.86	6.12	1.91	6.36	1.87	5.64	1.94	6.00	1.93
	Representation of Typical Workday	5.92	1.94	5.03	2.08	5.48	2.05	6.14	2.02	5.03	1.90	5.58	2.03	6.03	1.98	5.03	1.98	5.53	2.04
	How Well You Controlled Traffic	8.04	1.27	6.58	1.76	7.31	1.69	7.25	2.13	6.25	1.82	6.75	2.03	7.64	1.79	6.42	1.79	7.03	1.89
	Scenario Difficulty	6.33	2.92	5.05	2.45	5.69	2.75	6.66	2.73	5.33	2.26	6.00	2.58	6.50	2.81	5.19	2.35	5.84	2.67
	ATWIT Device Interference	1.92	1.03	3.04	2.16	2.48	1.77	2.25	1.34	2.75	1.65	2.50	1.52	2.09	1.20	2.89	1.92	2.49	1.65
	Oculometer Interference	2.31	0.86	2.37	0.84	2.34	0.85	3.50	2.36	3.49	2.39	3.50	2.36	2.90	1.86	2.93	1.87	2.92	1.86
	Simulation-Pilot Respondence	7.63	1.73	6.59	2.06	7.11	1.96	7.54	1.73	6.78	1.92	7.16	1.86	7.59	1.72	6.68	1.98	7.13	1.91
	DST Effectiveness Using TFPs	5.06	1.11	5.08	1.38	5.07	1.25	5.63	2.07	5.35	1.54	5.49	1.82	5.34	1.68	5.21	1.46	5.28	1.57
	Competence Using DST	5.40	1.40	5.18	1.60	5.29	1.50	5.94	2.04	5.37	1.79	5.65	1.93	5.67	1.76	5.27	1.69	5.47	1.74

Table DD-2. Realism: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p*-level
Position	.813	0.187	0.231	0.231	1.614	2	14	.234
Load	.607	0.393	0.649	0.649	4.541	2	14	.030
Automation	.629	0.371	0.589	0.589	1.768	4	12	.200
Position X Load	.928	0.072	0.078	0.078	.543	2	14	.593
Position X Automation	.701	0.299	0.426	0.426	1.277	4	12	.333
Load X Automation	.721	0.279	0.387	0.387	1.160	4	12	.376
Position X Load X Automation	.837	0.163	0.194	0.194	0.583	4	12	.681

Table DD-3. Realism: ANOVA Results.

	MS Effect	MS Error	F	df 1	df 2	*p*-level
Position	2.753	1.334	2.063	1	15	.171
Load	24.789	5.731	4.325	1	15	.055
Automation	6.609	1.530	4.318	2	30	.022
Position X Load	0.005	2.027	0.003	1	15	.961
Position X Automation	3.847	1.626	2.367	2	30	.111
Load X Automation	3.864	2.891	1.337	2	30	.278
Position X Load X Automation	0.193	2.248	0.086	2	30	.918

Table DD-4. Representativeness: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.570	1.793	0.318	1	15	.581
Load	48.459	7.128	6.798	1	15	.020
Automation	7.728	1.514	5.103	2	30	.012
Position X Load	0.570	1.922	0.296	1	15	.594
Position X Automation	3.402	1.734	1.962	2	30	.158
Load X Automation	4.691	2.156	2.176	2	30	.131
Position X Load X Automation	0.936	2.072	0.452	2	30	.641

Table DD-5. Participant Performance: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	15.188	3.008	5.049	1	15	.040
Load	72.535	2.010	36.080	1	15	.000
Automation	6.664	2.144	3.108	2	30	.059
Position X Load	2.518	2.120	1.188	1	15	.293
Position X Automation	0.297	1.869	0.159	2	30	.854
Load X Automation	5.939	1.983	2.995	2	30	.065
Position X Load X Automation	0.256	3.173	0.081	2	30	.923

Table DD-6. Difficulty: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	4.477	2.972	1.506	1	15	.239
Load	81.795	33.430	2.447	1	15	.139
Automation	18.212	4.510	4.038	2	30	.028
Position X Load	0.047	4.253	0.011	1	15	.918
Position X Automation	0.783	3.245	0.241	2	30	.787
Load X Automation	0.680	4.804	0.142	2	30	.869
Position X Load X Automation	0.475	6.114	0.078	2	30	.925

Table DD-7. Difficulty: ANOVA Results (N=6).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	4.014	4.847	0.828	1	5	.405
Load	21.125	12.758	16.470	1	5	.010
Automation	7.681	3.747	2.050	2	10	.179
Position X Load	6.125	7.492	0.818	1	5	.407
Position X Automation	1.347	7.381	0.183	2	10	.836
Load X Automation	1.042	7.475	0.139	2	10	.872
Position X Load X Automation	1.542	4.408	0.350	2	10	.713

Table DD-8. ATWIT Interference: ANOVA Results.

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	0.012	2.163	0.006	1	15	.942
Load	31.254	2.299	13.594	1	15	.002
Automation	4.753	1.505	3.158	2	30	.057
Position X Load	4.536	2.258	2.009	1	15	.177
Position X Automation	2.736	1.423	1.923	2	30	.164
Load X Automation	1.006	2.027	4.936	2	30	.014
Position X Load X Automation	1.214	1.629	0.746	2	30	.483

Table DD-9. ATWIT Interference: ANOVA Results (N=6).

	MS Effect	MS Error	*F*	df 1	df 2	*p*-level
Position	6.413	0.437	14.668	1	5	.012
Low Load	0.593	0.215	2.759	1	5	.158
High Load	8.051	0.359	22.426	1	5	.005
Load	12.077	0.371	32.545	1	5	.002
D Side	0.148	0.148	1.000	1	5	.363
R Side	1.384	0.404	25.725	1	5	.004
No Automation	0.131	0.056	2.353	1	5	.186
Limited Automation	0.253	0.329	0.770	1	5	.420
Full Automation	16.936	0.595	28.465	1	5	.003
Automation	2.936	1.827	1.607	2	10	.248
Low Load	0.698	0.630	1.109	2	10	.367
High Load	6.411	0.681	9.414	2	10	.005
Position x Load	19.519	1.284	15.198	1	5	.011
Position x Automation	6.505	2.683	2.425	2	10	.139
Load x Automation	11.282	0.794	14.203	2	10	.001
D side	0.694	0.494	1.404	2	10	.290
R side	15.475	1.173	13.197	2	10	.002
Position x Load x Automation	4.887	0.873	5.600	2	10	.023

Table DD-10. Oculometer Interference: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.001	1.263	0.001	1	15	.982
Automation	1.248	1.459	0.856	2	30	.435
Load X Automation	0.319	3.608	0.088	2	30	.916

Table DD -11. Responsiveness of Simulation Pilots: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Position	0.137	1.385	0.099	1	15	.757
Load	39.298	2.732	14.383	1	15	.002
Automation	4.485	2.117	2.119	2	30	.138
Position X Load	0.899	1.395	0.644	1	15	.435
Position X Automation	0.675	1.580	0.427	2	30	.656
Load X Automation	1.044	2.530	0.413	2	30	.666
Position X Load X Automation	0.169	2.177	0.078	2	30	.925

Table DD-12. Effectiveness of the DST in Resolving Conflicts When Using Trial Flight Plans: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Position	34.406	2.102	16.368	1	15	.001
Load	3.194	1.625	1.966	1	15	.181
Position X Load	3.852	3.261	1.181	1	15	.294

Table DD-13. Competence Felt Using the DST: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Position	9.595	8.146	1.178	1	15	.295
Load	11.234	1.207	9.307	1	15	.008
Automation	0.470	1.330	0.354	1	15	.561
Position X Load	2.301	1.821	1.263	1	15	.279
Position X Automation	2.258	1.819	1.241	1	15	.283
Load X Automation	2.215	1.992	1.112	1	15	.308
Position X Load X Automation	0.029	2.016	0.014	1	15	.906

APPENDIX EE

Subject Matter Expert Rating Forms/OTS Results

Table EE-1. Maintaining Safe and Efficient Traffic Flow: Means and Standard Deviations.

Safety and Efficiency	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	1) Separation and Resolving Potential Conflicts	5.92	1.48	5.37	1.70	5.64	1.60	5.92	1.48	5.37	1.70	5.64	1.60	5.92	1.46	5.37	1.68	5.64	1.58
	2) Sequencing Arrival/Departure Aircraft	6.41	1.30	5.79	1.66	6.10	1.50	6.41	1.30	5.79	1.66	6.10	1.50	6.41	1.28	5.79	1.64	6.10	1.49
	3) Using Control Instructions	6.47	0.88	5.72	1.36	6.09	1.19	6.47	0.88	5.72	1.36	6.09	1.19	6.47	0.87	5.72	1.34	6.09	1.18
Auto 2	1) Separation and Resolving Potential Conflicts	6.24	0.46	5.92	1.29	6.08	0.96	6.24	0.46	5.92	1.29	6.08	0.96	6.24	0.45	5.92	1.27	6.08	0.96
	2) Sequencing Arrival/Departure Aircraft	6.37	0.58	6.45	0.69	6.41	0.63	6.37	0.58	6.45	0.69	6.41	0.63	6.37	0.57	6.45	0.68	6.41	0.62
	3) Using Control Instructions	6.30	0.67	6.25	0.97	6.27	0.82	6.30	0.67	6.25	0.97	6.27	0.82	6.30	0.66	6.25	0.96	6.27	0.82
Auto 3	1) Separation and Resolving Potential Conflicts	6.36	0.89	5.74	1.34	6.05	1.16	6.36	0.89	5.74	1.34	6.05	1.16	6.36	0.88	5.74	1.32	6.05	1.15
	2) Sequencing Arrival/Departure Aircraft	6.68	0.57	6.10	1.44	6.39	1.12	6.68	0.57	6.10	1.44	6.39	1.12	6.68	0.56	6.10	1.41	6.39	1.11
	3) Using Control Instructions	6.48	0.60	6.28	0.85	6.38	0.73	6.48	0.60	6.28	0.85	6.38	0.73	6.48	0.59	6.28	0.84	6.38	0.73
Auto Collapsed	1) Separation and Resolving Potential Conflicts	6.17	1.03	5.67	1.44	5.92	1.27	6.17	1.03	5.67	1.44	5.92	1.27	6.17	1.02	5.67	1.44	5.92	1.27
	2) Sequencing Arrival/Departure Aircraft	6.49	0.88	6.11	1.33	6.30	1.14	6.49	0.88	6.11	1.33	6.30	1.14	6.49	0.87	6.11	1.32	6.30	1.13
	3) Using Control Instructions	6.42	0.72	6.08	1.09	6.25	0.93	6.42	0.72	6.08	1.09	6.25	0.93	6.42	0.71	6.08	1.09	6.25	0.93

Table EE-2. Prioritizing: Means and Standard Deviations.

Prioritizing	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	8) Taking Actions in Appropriate Order	6.07	1.18	5.45	1.37	5.76	1.30	6.07	1.18	5.45	1.37	5.76	1.30	6.07	1.16	5.45	1.35	5.76	1.29
	9) Preplanning Control Actions	4.91	1.00	4.10	1.39	4.50	1.26	4.91	1.00	4.10	1.39	4.50	1.26	4.91	0.98	4.10	1.37	4.50	1.25
	10) Handling Control Tasks for Several Aircraft	6.30	0.99	5.66	1.55	5.98	1.32	6.30	0.99	5.66	1.55	5.98	1.32	6.30	0.98	5.66	1.53	5.98	1.31
	11) Marking Flight Strips and Performing Tasks	3.62	2.62	3.12	2.21	3.37	2.40	3.62	2.62	3.12	2.21	3.37	2.40	3.62	2.58	3.12	2.17	3.37	2.38
Auto 2	8) Taking Actions in Appropriate Order	6.26	0.57	6.21	0.54	6.24	0.55	6.26	0.57	6.21	0.54	6.24	0.55	6.26	0.56	6.21	0.53	6.24	0.54
	9) Preplanning Control Actions	5.14	1.06	4.88	0.77	5.01	0.92	5.14	1.06	4.88	0.77	5.01	0.92	5.14	1.04	4.88	0.76	5.01	0.92
	10) Handling Control Tasks for Several Aircraft	6.30	0.67	6.12	0.58	6.21	0.62	6.30	0.67	6.12	0.58	6.21	0.62	6.30	0.66	6.12	0.57	6.21	0.62
	11) Marking Flight Strips and Performing Tasks	3.48	0.00	3.48	0.00	3.48	0.00	3.48	0.00	3.48	0.00	3.48	0.00	3.48	0.00	3.48	0.00	3.48	0.00
Auto 3	8) Taking Actions in Appropriate Order	6.51	0.62	5.95	0.93	6.23	0.83	6.51	0.62	5.95	0.93	6.23	0.83	6.51	0.61	5.95	0.91	6.23	0.82
	9) Preplanning Control Actions	5.21	0.80	4.35	0.94	4.78	0.96	5.21	0.80	4.35	0.94	4.78	0.96	5.21	0.78	4.35	0.92	4.78	0.95
	10) Handling Control Tasks for Several Aircraft	6.73	0.76	6.34	1.01	6.54	0.90	6.73	0.76	6.34	1.01	6.54	0.90	6.73	0.74	6.34	0.99	6.54	0.89
	11) Marking Flight Strips and Performing Tasks	3.70	0.88	3.48	0.00	3.59	0.62	3.70	0.88	3.48	0.00	3.59	0.62	3.70	0.87	3.48	0.00	3.59	0.62
Auto Collapsed	8) Taking Actions in Appropriate Order	6.28	0.84	5.87	1.03	6.08	0.96	6.28	0.84	5.87	1.03	6.08	0.96	6.28	0.84	5.87	1.03	6.08	0.96
	9) Preplanning Control Actions	5.09	0.95	4.44	1.10	4.76	1.07	5.09	0.95	4.44	1.10	4.76	1.07	5.09	0.94	4.44	1.09	4.76	1.07
	10) Handling Control Tasks for Several Aircraft	6.44	0.83	6.04	1.13	6.24	1.01	6.44	0.83	6.04	1.13	6.24	1.01	6.44	0.82	6.04	1.13	6.24	1.00
	11) Marking Flight Strips and Performing Tasks	3.60	1.57	3.36	1.26	3.48	1.42	3.60	1.57	3.36	1.26	3.48	1.42	3.60	1.56	3.36	1.25	3.48	1.42

Table EE-3. Providing Control Information: Means and Standard Deviations.

Providing Control Info.	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	12) Providing Essential ATC Information	5.34	1.07	4.66	1.27	5.00	1.21	5.34	1.07	4.66	1.27	5.00	1.21	5.34	1.05	4.66	1.25	5.00	1.20
	13) Providing Additional ATC Information	3.59	1.35	2.84	1.05	3.21	1.25	3.59	1.35	2.84	1.05	3.21	1.25	3.59	1.33	2.84	1.03	3.21	1.24
	14) Providing Coordination	5.92	0.84	4.30	1.80	5.11	1.61	5.92	0.84	4.30	1.80	5.11	1.61	5.92	0.83	4.30	1.77	5.11	1.60
Auto 2	12) Providing Essential ATC Information	5.55	0.60	5.19	0.62	5.37	0.63	5.55	0.60	5.19	0.62	5.37	0.63	5.55	0.59	5.19	0.61	5.37	0.62
	13) Providing Additional ATC Information	4.57	1.29	3.55	1.49	4.06	1.46	4.57	1.29	3.55	1.49	4.06	1.46	4.57	1.27	3.55	1.46	4.06	1.45
	14) Providing Coordination	5.63	0.86	5.43	1.06	5.53	0.95	5.63	0.86	5.43	1.06	5.53	0.95	5.63	0.84	5.43	1.04	5.53	0.94
Auto 3	12) Providing Essential ATC Information	5.73	0.55	5.03	1.32	5.38	1.06	5.73	0.55	5.03	1.32	5.38	1.06	5.73	0.54	5.03	1.30	5.38	1.05
	13) Providing Additional ATC Information	4.20	1.31	3.53	1.45	3.86	1.40	4.20	1.31	3.53	1.45	3.86	1.40	4.20	1.29	3.53	1.43	3.86	1.39
	14) Providing Coordination	6.07	1.27	4.86	1.61	5.46	1.55	6.07	1.27	4.86	1.61	5.46	1.55	6.07	1.25	4.86	1.59	5.46	1.54
Auto Collapsed	12) Providing Essential ATC Information	5.54	0.78	4.96	1.11	5.25	1.00	5.54	0.78	4.96	1.11	5.25	1.00	5.54	0.77	4.96	1.11	5.25	1.00
	13) Providing Additional ATC Information	4.12	1.35	3.31	1.36	3.71	1.41	4.12	1.35	3.31	1.36	3.71	1.41	4.12	1.35	3.31	1.35	3.71	1.40
	14) Providing Coordination	5.87	1.00	4.86	1.56	5.37	1.40	5.87	1.00	4.86	1.56	5.37	1.40	5.87	1.00	4.86	1.55	5.37	1.40

Table EE-4. Technical Knowledge: Means and Standard Deviations.

Technical Knowledge	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	15) Knowledge of LOAs and SOPs	6.78	0.65	6.03	1.32	6.40	1.09	6.78	0.65	6.03	1.32	6.40	1.09	6.78	0.64	6.03	1.29	6.40	1.08
	16) Knowledge of Capabilities and Limitations	6.72	0.85	6.35	1.01	6.53	0.94	6.72	0.85	6.35	1.01	6.53	0.94	6.72	0.84	6.35	0.99	6.53	0.93
	17) Effective Use of Equipment	6.44	0.73	5.69	1.40	6.06	1.16	6.44	0.73	5.69	1.40	6.06	1.16	6.44	0.71	5.69	1.38	6.06	1.15
Auto 2	15) Knowledge of LOAs and SOPs	6.88	0.45	6.93	0.13	6.90	0.33	6.88	0.45	6.93	0.13	6.90	0.33	6.88	0.44	6.93	0.13	6.90	0.33
	16) Knowledge of Capabilities and Limitations	6.71	0.43	7.00	0.29	6.85	0.39	6.71	0.43	7.00	0.29	6.85	0.39	6.71	0.42	7.00	0.28	6.85	0.39
	17) Effective Use of Equipment	6.00	0.82	5.69	1.08	5.85	0.95	6.00	0.82	5.69	1.08	5.85	0.95	6.00	0.80	5.69	1.06	5.85	0.95
Auto 3	15) Knowledge of LOAs and SOPs	6.82	0.62	6.78	0.54	6.80	0.57	6.82	0.62	6.78	0.54	6.80	0.57	6.82	0.61	6.78	0.53	6.80	0.57
	16) Knowledge of Capabilities and Limitations	7.02	0.28	6.78	0.65	6.90	0.51	7.02	0.28	6.78	0.65	6.90	0.51	7.02	0.27	6.78	0.64	6.90	0.50
	17) Effective Use of Equipment	6.50	0.73	5.75	1.44	6.13	1.18	6.50	0.73	5.75	1.44	6.13	1.18	6.50	0.72	5.75	1.41	6.13	1.17
Auto Collapsed	15) Knowledge of LOAs and SOPs	6.83	0.57	6.58	0.90	6.70	0.76	6.83	0.57	6.58	0.90	6.70	0.76	6.83	0.57	6.58	0.90	6.70	0.76
	16) Knowledge of Capabilities and Limitations	6.81	0.58	6.71	0.75	6.76	0.67	6.81	0.58	6.71	0.75	6.76	0.67	6.81	0.58	6.71	0.75	6.76	0.67
	17) Effective Use of Equipment	6.31	0.78	5.71	1.29	6.01	1.10	6.31	0.78	5.71	1.29	6.01	1.10	6.31	0.77	5.71	1.28	6.01	1.10

Table EE-5. Communicating: Means and Standard Deviations.

Comm.	Question	R Side Controller Position						D Side Controller Position						Position Collapsed				Load/Pos. Collapsed	
		Low Load		High Load		Load Coll.		Low Load		High Load		Load Coll.		Low Load		High Load			
		Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs	Means	SDs
Auto 1	18) Using Proper Phraseology	6.77	0.65	6.27	0.92	6.52	0.83	6.77	0.65	6.27	0.92	6.52	0.83	6.77	0.64	6.27	0.91	6.52	0.82
	19) Communicating Clearly and Efficiently	6.78	0.65	6.22	1.22	6.50	1.00	6.78	0.65	6.22	1.22	6.50	1.00	6.78	0.64	6.22	1.20	6.50	1.00
	20) Listening for Pilot Readbacks and Requests	5.88	0.81	5.44	1.15	5.66	1.00	5.88	0.81	5.44	1.15	5.66	1.00	5.88	0.79	5.44	1.13	5.66	1.00
Auto 2	18) Using Proper Phraseology	6.63	0.58	6.84	0.45	6.73	0.52	6.63	0.58	6.84	0.45	6.73	0.52	6.63	0.57	6.84	0.45	6.73	0.52
	19) Communicating Clearly and Efficiently	6.71	0.43	6.99	0.47	6.85	0.46	6.71	0.43	6.99	0.47	6.85	0.46	6.71	0.42	6.99	0.46	6.85	0.46
	20) Listening for Pilot Readbacks and Requests	6.19	0.54	6.19	0.75	6.19	0.64	6.19	0.54	6.19	0.75	6.19	0.64	6.19	0.53	6.19	0.74	6.19	0.64
Auto 3	18) Using Proper Phraseology	6.88	0.45	6.71	0.68	6.79	0.57	6.88	0.45	6.71	0.68	6.79	0.57	6.88	0.44	6.71	0.67	6.79	0.57
	19) Communicating Clearly and Efficiently	7.02	0.28	6.91	0.45	6.96	0.37	7.02	0.28	6.91	0.45	6.96	0.37	7.02	0.27	6.91	0.44	6.96	0.37
	20) Listening for Pilot Readbacks and Requests	6.25	0.58	6.13	1.02	6.19	0.82	6.25	0.58	6.13	1.02	6.19	0.82	6.25	0.57	6.13	1.01	6.19	0.81
Auto Collapsed	18) Using Proper Phraseology	6.76	0.57	6.61	0.74	6.68	0.66	6.76	0.57	6.61	0.74	6.68	0.66	6.76	0.56	6.61	0.73	6.68	0.66
	19) Communicating Clearly and Efficiently	6.84	0.49	6.71	0.85	6.77	0.70	6.84	0.49	6.71	0.85	6.77	0.70	6.84	0.48	6.71	0.85	6.77	0.69
	20) Listening for Pilot Readbacks and Requests	6.11	0.66	5.92	1.03	6.01	0.86	6.11	0.66	5.92	1.03	6.01	0.86	6.11	0.66	5.92	1.02	6.01	0.86

Table EE-6. Maintaining Safe and Efficient Traffic Flow: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Load	.160	0.840	5.234	5.234	8.723	3	5	.020
Automation	.156	0.844	5.391	5.391	1.797	6	2	.400
Load X Automation	.025	0.975	38.269	38.269	12.756	6	2	.074

Table EE -7. Maintaining Separation and Resolving Potential Conflicts: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	2.981	0.480	6.213	1	7	.041
Automation	0.941	0.614	1.534	2	14	.250
Load X Automation	0.102	0.844	0.121	2	14	.887

Table EE -8. Sequencing Arrival and Departure Aircraft Efficiently: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	1.688	0.164	1.309	1	7	.015
Automation	0.481	0.309	1.558	2	14	.245
Load X Automation	0.626	0.605	1.036	2	14	.381

Table EE -9. Using Control Instructions Effectively/Efficiently: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	1.333	0.393	3.394	1	7	.108
Automation	0.341	0.393	0.867	2	14	.442
Load X Automation	0.545	0.517	1.055	2	14	.374

Table EE -10. Prioritizing: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Load	.173	0.827	4.773	4.773	7.956	3	5	.024
Automation	.186	0.814	4.362	4.362	1.454	6	2	.462
Load X Automation	.107	0.893	8.353	8.353	2.784	6	2	.288

Table EE-11. Taking Actions in an Appropriate Order of Importance: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	2.083	0.262	7.955	1	7	.026
Automation	1.196	0.235	5.081	2	14	.022
Load X Automation	0.390	0.418	0.932	2	14	.417

Table EE-12. Preplanning Control Actions: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	5.005	0.160	31.288	1	7	.001
Automation	1.038	0.386	2.690	2	14	.103
Load X Automation	0.438	0.640	0.684	2	14	.521

Table EE-13. Handling Control Tasks For Several Aircraft: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	1.928	0.251	7.671	1	7	.028
Automation	1.278	0.510	2.505	2	14	.117
Load X Automation	0.219	0.291	0.752	2	14	.490

Table EE-14. Marking Flight Strips While Performing Other Tasks: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	1.000	00.689	1.452	1	7	.267

Table EE-15. Providing Control Information: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Load	.123	0.877	7.134	7.134	11.890	3	5	.010
Automation	.061	0.939	15.277	15.277	5.092	6	2	.173
Load X Automation	.072	0.928	12.981	12.981	4.327	6	2	.200

Table EE-16. Providing Essential ATC Information: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	4.083	0.083	49.000	1	7	.000
Automation	0.751	0.766	0.980	2	14	.400
Load X Automation	0.151	0.397	0.379	2	14	.691

Table EE-17. Providing Additional ATC Information: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	7.922	0.458	17.312	1	7	.004
Automation	3.143	0.567	5.540	2	14	.017
Load X Automation	0.133	0.550	0.242	2	14	.788

Table EE-18. Providing Coordination: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	12.319	1.134	1.861	1	7	.013
Automation	0.820	1.093	0.750	2	14	.490
Load X Automation	2.136	.684	3.120	2	14	.076

Table EE-19. Technical Knowledge: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Load	.146	0.854	5.859	5.859	9.764	3	5	.016
Automation	.068	0.932	13.798	13.798	4.599	6	2	.189
Load X Automation	.134	0.866	6.478	6.478	2.159	6	2	.350

Table EE-20. Showing Knowledge of LOAs and SOPS: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.750	0.238	3.150	1	7	.119
Automation	1.125	0.312	3.610	2	14	.054
Load X Automation	0.758	0.332	2.285	2	14	.138

Table EE-21. Showing Knowledge of Aircraft Capabilities and Limitations: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.130	0.094	1.378	1	7	.279
Automation	0.642	0.204	3.150	2	14	.074
Load X Automation	0.504	0.157	3.215	2	14	.071

Table EE-22. Showing Effective Use of Equipment: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	4.376	0.142	3.775	1	7	.001
Automation	0.347	0.731	0.475	2	14	.632
Load X Automation	0.256	0.659	0.389	2	14	.685

Table EE-23. Communicating: MANOVA Results.

	Wilks Λ	Pillai's	Hotelling	Roy's	*F*	df 1	df 2	*p-level*
Load	.821	0.179	0.219	0.219	0.364	3	5	.782
Automation	.211	0.789	3.741	3.741	1.247	6	2	.509
Load X Automation	.254	0.746	2.934	2.934	0.978	6	2	.585

Table EE-24. Using Proper Phraseology: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.279	0.272	1.025	1	7	.345
Automation	0.324	0.201	1.611	2	14	.235
Load X Automation	0.505	0.118	4.278	2	14	.035

Table EE-25. Communicating Clearly and Efficiently: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.202	0.263	0.769	1	7	.410
Automation	0.924	0.218	4.243	2	14	.036
Load X Automation	0.718	0.136	5.293	2	14	.019

Table EE-26. Listening for Pilot Readbacks and Requests: ANOVA Results.

	MS Effect	Ms Error	*F*	df 1	df 2	*p-level*
Load	0.421	0.302	1.393	1	7	.277
Automation	1.510	0.297	5.077	2	14	.022
Load X Automation	0.206	0.424	0.485	2	14	.626

APPENDIX FF

Exit Questionnaire

Table FF-1. Experiment Ratings.

	Mean	SD
Genera Center Airspace hands-on training	7.63	1.78
Decision Support Tool hands-on training	7.06	2.11