2003

Effects of a supersector and color-coded aircraft altitude on ATC performance

Liza Tam
San Jose State University

Follow this and additional works at: https://scholarworks.sjsu.edu/etd_theses

Recommended Citation

Tam, Liza, "Effects of a supersector and color-coded aircraft altitude on ATC performance" (2003).
Master's Theses. 2427.
DOI: <https://doi.org/10.31979/etd.vpj8-psmr>
https://scholarworks.sjsu.edu/etd_theses/2427

This Thesis is brought to you for free and open access by the Master's Theses and Graduate Research at SJSU ScholarWorks. It has been accepted for inclusion in Master's Theses by an authorized administrator of SJSU ScholarWorks. For more information, please contact scholarworks@sjsu.edu.

**EFFECTS OF A SUPERSECTOR AND COLOR-CODED
AIRCRAFT ALTITUDE ON ATC PERFORMANCE**

A Thesis

Presented to

The Faculty of the Department of Information Systems Engineering

San José State University

In Partial Fulfillment

of the Requirements for the Degree

Master of Science in Human Factors/Ergonomics

by

Liza Tam

May 2003

UMI Number: 1415739

UMI Microform 1415739

Copyright 2003 by ProQuest Information and Learning Company.

All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

© 2003

Liza Tam

ALL RIGHTS RESERVED

APPROVED FOR THE
DEPARTMENT OF INDUSTRIAL SYSTEMS ENGINEERING

Dr. Kevin M. Corker

Dr. Kevin Jordan

Dr. Dwight Miller, Sandia National Laboratories

APPROVED FOR THE UNIVERSITY

ABSTRACT

EFFECTS OF A SUPERSECTOR AND COLOR-CODED AIRCRAFT ALTITUDE ON ATC PERFORMANCE

by Liza C. Tam

The increasing demand on airspace capacity requires researchers to examine methods of accommodating a greater number of aircraft operations within a given airspace without sacrificing safety or overburdening air traffic controllers. We examined an air traffic management (ATM) concept called *supersector* operations. This concept redistributes the roles and responsibilities between a tactical controller and a planning controller for management of traffic flow and intersector operations. We compared the effects of delivering aircraft altitude information to the tactical controller through the datablock (alpha-numeric) alone and when complemented by the color saturation of the aircraft on the radar display. We also compared the performance of controllers working in the supersector versus a current-day sector. We found the supersector operational concept to be beneficial under high traffic density conditions. However, the color-saturation coding of altitude alone did not yield statistically significant results. Avenues for future research are proposed.

ACKNOWLEDGMENTS

I would like to express my deep appreciation to Dr. Kevin Corker for his unfailing guidance and encouragement. Each time this project stalled, your sharp intelligence, optimism, and insight kept me focused and energized. My gratitude also extends to Anthony Tang. Without Anthony's dedication, patience, and skill, this computer simulation would not have been possible. I thank Samantha Starr for her generosity and assistance, and Drs. Kevin Jordan and Dwight Miller for their insightful and invaluable suggestions. I am grateful to Dr. Vu Duong at the EUROCONTROL Experimental Center for endowing the Human Automation and Integration Laboratory (HAIL) at San José State University with the grant that funded this research.

I would like to thank my parents for teaching me the importance of education and persistence, and Betsy and Conrad Josias for their consistent interest and encouragement. Finally, I would like to thank my husband, Andrew, for his suggestions, patience, and endless support.

TABLE OF CONTENTS

INTRODUCTION.....	1
AIR TRAFFIC GROWTH IN EUROPE.....	4
THE AIR TRAFFIC CONTROLLER & RADAR INFORMATION	5
A NEW OPERATIONAL CONCEPT: A SUPERSECTOR.....	6
AIRSPACE COMPLEXITY	11
<i>Airspace Complexity and Cognition.....</i>	<i>11</i>
<i>Airspace Complexity Metrics</i>	<i>12</i>
INFORMATION AND COMPUTER DISPLAYS.....	14
<i>Data Delivery versus Information Delivery.....</i>	<i>14</i>
<i>Color</i>	<i>16</i>
HUMAN PERFORMANCE METRICS	21
WORKING MEMORY: THE VISUO-SPATIAL SKETCHPAD.....	22
HYPOTHESES	23
LITERATURE REVIEW.....	25
<i>Human-Computer Interaction.....</i>	<i>25</i>
<i>Color-Coding of Information</i>	<i>26</i>
<i>Airspace Complexity.....</i>	<i>30</i>
<i>Complexity Metrics.....</i>	<i>30</i>

<i>Sector Design</i>	33
<i>Human Performance Metrics</i>	34
<i>Summary</i>	35
METHOD	37
<i>Participants</i>	37
<i>Simulation Program</i>	37
<i>Equipment</i>	38
<i>Map Display</i>	39
<i>Aircraft Display</i>	42
<i>Trajectory Information</i>	45
<i>Communications</i>	46
<i>Traffic</i>	46
<i>Procedure</i>	47
<i>Dependent Variables</i>	49
<i>Independent Variables</i>	51
RESULTS	54
<i>Actual Conflicts</i>	54
<i>Potential Conflicts</i>	55
<i>Conflict Resolution Performance</i>	57
<i>Number of Actions Taken</i>	61
<i>Subjective Workload Ratings</i>	62
DISCUSSION	64
<i>Actual Conflicts</i>	65
<i>Potential Conflicts</i>	66
<i>Time to Conflict</i>	67

<i>Subjective Workload and Performance Metrics.....</i>	<i>68</i>
<i>Color coding of Altitude</i>	<i>70</i>
<i>Airspace</i>	<i>72</i>
<i>Lessons for displaying colors.....</i>	<i>73</i>
<i>Participants.....</i>	<i>74</i>
CONCLUSION	75
<i>Modes of Control.....</i>	<i>75</i>
<i>Participants.....</i>	<i>77</i>
REFERENCES	81
APPENDIX A <u>Training Document Used for Briefing Participants</u>.....	93
APPENDIX B <u>Participant Informed Consent Form</u>	110
APPENDIX C <u>Graphs Illustrating Subjective Workload Ratings as a Function of Time in Each Scenario</u>	112

LIST OF FIGURES

<i>Figure 1.</i> Datablocks as they currently appear on controllers' radar screens at air traffic control centers. Each aircraft is represented by a short green line.	6
<i>Figure 2.</i> An outline of the supersector shape and the trajectory-based airways. The green arrows represent the north and eastbound airways while the blue arrows represent the south and westbound airways.....	9
<i>Figure 3.</i> A color – saturation coding scheme is used to display aircraft altitudes. As aircraft climb to higher altitudes, the color saturation of the aircraft displayed on the air traffic controller's radar screen decreases.	20
<i>Figure 4.</i> A screenshot of the supersector airspace as seen by participants.	40
<i>Figure 5.</i> A screen shot of the Uniform Juliet airspace as seen by participants on their screens during the simulation.	41
<i>Figure 6.</i> Datablock on the radar screen display as shown during the experiment.	45
<i>Figure 7.</i> Seating arrangement of participants during data collection. Data was collected from two participants simultaneously.....	48
<i>Figure 8.</i> A diagram of the 2 x 2 x 2 within-subject experimental design that includes two levels of traffic intensity, two types of airspace operations, tested with and without the presence or color-saturation coding.....	53
<i>Figure 9.</i> Mean number of actual conflicts as a function of traffic intensity. Performance in the supersector resulted in zero actual conflicts.....	55

<i>Figure 10.</i> Mean number of actual and potential conflicts as a function of traffic intensity.....	56
<i>Figure 11.</i> Mean number of potential conflicts in the supersector airspace under both color-coding conditions.	57
<i>Figure 12.</i> Conflict resolution performance counted as the length of time remaining, when a conflict was averted. Performance was measured as the mean amount of time remaining before a conflict, at the time a conflict was resolved.	59
<i>Figure 13.</i> The effect of color-saturation coding on conflict resolution performance in the supersector airspace. The error bars at each point represent the standard error.	60
<i>Figure 14.</i> Graphs display a 3-way interaction of traffic levels, types of airspace and two levels of saturation coding. Performance was measured as a function of the number of actions taken.	61
<i>Figure 15.</i> Mean number of actions taken in high traffic scenarios, as a function of the type of airspace operations.	62
<i>Figure 16.</i> Mean subjective workload ratings across all conditions. Workload was rated on a scale of 1-10, where 1 was low and 10 was high.	63

Introduction

As the airline industry becomes increasingly competitive and the volume of air cargo continues to rise, air traffic density over European skies is set to expand to unprecedented levels. Each year, the number of aircraft operations over Europe continues to rise (EUROCONTROL, 2002), while aviation authorities search for workable technologies and procedures to support this growth (Boeing, 2002; Duong, Hoffman, & Nicolaon, 1997; EUROCONTROL, 2002; Hoekstra, Ruigrok, & Van Gent, 2000; Johnson, Battiste, & Belcher, 1998).

Currently, the distribution of responsibility in air traffic management is based on the division of the airspace into sectors. A controller, or a team of controllers, manages each sector and as part of the air traffic management (ATM) process, pilots are required to communicate with air traffic controllers. To fly an aircraft safely from departure to destination requires navigating through numerous sectors and switching between airspace managers (called handoffs); the result is a journey composed of a complex amalgam of technological employment and procedural organization both on the ground and in the cockpit.

Perhaps, the greatest hurdle to accommodating more aircraft operations is the ability of the air traffic controller to safely manage a growing number of aircraft in a finite space. Even though sophisticated communication, satellite, and navigation technologies supply information to controllers and pilots, ultimately, it is the human operator – the air traffic controller – who has to take this information and make safety-

critical decisions in situations that often require urgency. The effort of handling traffic depends on at least two key factors: airspace complexity and the information available to the controller on the air traffic display.

In an effort to safely increase airspace capacity, many procedural changes and operating concepts (Boeing, 2002; Duong et al., 1997; Hoekstra et al., 2000; Van Gent, Hoekstra, & Ruigrok, 1998) have been proposed. These concepts, through new technologies or operating procedures, aim to reduce or redistribute controller responsibilities. Along these lines, the Innovative Research Unit at EUROCONTROL Experimental Center (EEC) has proposed an operational concept that restructures both the operating procedures and architecture of the current airspace (EEC, 2002). This concept, called the *supersector*, proposes simplifying the array of airways and restructuring air traffic management operating procedures. In particular, the roles and responsibilities of controllers would become time based, the size of each sector would increase significantly, and the sector would be shared.

The fundamental role of an air traffic controller is to safely and efficiently manage aircraft while traversing from origin to destination. The management of safe and efficient airspace relies on coordinated efforts among airspace actors (i.e., pilots, airspace management, airlines and air traffic controllers) who understand each other's roles and responsibilities. Thus, it is vital that any evaluation of new operational concepts examines the effects of change on operator performance, behavior, safety, and efficiency.

Changing the roles and responsibilities of controllers have been shown to impact controllers' levels of workload, situation awareness, and stress and is most evident when

controllers engage in activities such as visual search, prioritizing tasks, and conflict detection (Endsley, 1997; Galster, Duley, Masalonis, & Parasuraman, 2001). To support the controller in decision making, two elements of information delivery need to be defined. First, what information should be displayed? To enable effective decision making, researchers need to establish the informational needs, such as pilot intentions, aircraft position, weather, route information, or incoming sector traffic of the job requirements. Second, to minimize controllers' cognitive workload, the delivery of information should afford efficient, yet accurate cognition and decision making. Hence, researchers have to answer the question of 'how to display' information.

To address both of these issues, a parallel approach is proposed. The redesign of the airspace structure and operating procedures address the need to increase the airspace capacity while the integration of task-critical information addresses the need for efficient information delivery to enhance safety. This research evaluates a proposed method of improving the delivery of time-critical data to aid the decision making and performance of air traffic controllers under the new operating concept put forward by the EEC. This evaluation is not intended as an operational assessment of the supersector concept. Rather, this experiment only examines the performance of the tactical controller working under supersector procedures and managing traffic under normal conditions.

The impact of operational change is reflected not only by operational results, such as growth in the number of aircraft a sector supports or the type of technologies integrated into a system, but also by the effects on the human operators in that system including their perceptual, physical and cognitive workloads. At the core of airspace

operations is the human whose performance is affected by stress, workload, and temporal constraints. The ability of the human to manage complex supervisory tasks and engage in safety-critical decision making depends on the efficiency of information delivery and the complexity of the system. It is therefore important that the requirements of the human operator be given adequate consideration in the airspace design process, especially when safety and efficiency have to be maintained as the system's two most fundamental goals.

Air Traffic Growth In Europe

Based on conservative estimates, the number of instrument flight rules (IFR) air operations over France is forecast (EUROCONTROL, 2002) to grow at an average rate of 1.9% per year between 2001 – 2009, resulting in a 15.9% growth in air traffic over 8 years. Forecasts, based on a high growth scenario, estimate the annual number of IFR flights during this period to grow at an average rate of 3.8% per year. This amounts to a 34.9% increase in aircraft operations compared to the 2,502,242 flights that operated over France during 2001. The projected growth in flight operations translates directly into greater workload for air traffic controllers, who must manage significantly more traffic within the same volume of airspace.

Not surprisingly, air traffic growth is related to economic growth. EUROCONTROL's STATFOR statistical forecasting office predicts future European air traffic volume partly on the economic performance and political climate of each region. Concepts that enable the air traffic managements system to absorb more flight operations safely and efficiently are therefore strongly encouraged, since the ability to increase the

number of flight operations facilitates the growth of cargo transport, exports, tourism and economic growth.

The Air Traffic Controller & Radar Information

The fundamental job of the air traffic controller is to keep traffic legally separated and operating safely and efficiently in an en route, oceanic, approach or terminal area environment. Currently, the allocation of responsibility is divided between the various air traffic management authorities based on whether aircraft are en route, in the oceanic environment, on approach or departure to and from an airport, or in the airport terminal area.

To maintain legal separation, aircraft operating at altitudes between flight levels 19,000 to 28,000 ft must maintain a vertical separation of 1000 ft; above 28,000 ft, aircraft must remain vertically separated by at least 2000 ft. Laterally, aircraft must maintain a 5 nm separation.

A controller's mental picture of the airspace is primarily based on the information on the radar display. The text information that pertains to aircraft altitude is critical for maintaining legal separation because aircraft are displayed on the radar screen in a two-dimensional plan view. In today's control centers, controllers source aircraft altitude information from the datablock – a few lines of text – linked to each aircraft on the radar screen (see Figure 1). To decide on the likelihood of conflicts arising, perhaps from aircraft climbing through several flight levels or crossing traffic, controllers must first read the text information from at least two datablocks and then compare them.

Figure 1. Datablocks as they currently appear on controllers' radar screens at air traffic control centers. Each aircraft is represented by a short green line.

Second, the controller may also have to do mental arithmetic to ensure legal separation while scanning the radar screen for other possible incursions. Scanning, reading, and then mentally calculating differences in aircraft altitude takes time as well as cognitive effort. Moreover, it leaves the controller vulnerable to errors when reading numbers or doing mental arithmetic.

A New Operational Concept: A Supersector

Current operations require communications between pilots and controllers as they enter and exit any sector. As the controller's workload in a sector reaches capacity, the size of the sector is reduced and divided to enable the workload to be shared. What results is a cyclical reduction of sector size due to an increasing workload. As the number of sectors increases (to make the airspace more manageable), so does the frequency of radio or verbal communications due to the increased regularity of hand offs. A hand off is the process where the controller in one sector hands over responsibility of an aircraft to

a controller in an adjacent sector. Eventually, the reduction in sector size reaches its physical limitations (rather than the performance limits of controller and pilot), making it impossible to reduce the size of a sector any further. As a result, the distribution of workload and sector design needs to be reconsidered if the airspace is to safely and efficiently absorb more aircraft operations in the near future.

The operational concept proposed by the EEC, dubbed the supersector, is intended to operate between flight levels 19,500 to 34,000 ft and focuses on the area over northwest of France. It integrates the airspace of six to eight current day sectors (see Figure 2). Investigation of the viability of the supersector is an initial exploration of the EEC's SectorLess – 1 concept (EEC, 2002). The supersector has been proposed by the EEC to address the transition from the current sector design of the airspace to an operational concept centered on trajectory-based individual control.

The supersector's airways were restructured to consolidate the flow of traffic, and are based on needs of the current traffic patterns. Each airway has traffic that flows in only one direction, with a set of pre-assigned altitudes for each direction. Three major airports—Charles de Gaulle, Orly, and Le Bourget—are inside the boundaries of the supersector. As such, there is a regular stream of climbing and descending traffic. In terms of the total volume, a significant proportion of traffic is from overflights.

An innovation of the supersector is the redistribution of ATM roles and responsibility; it is a concept that significantly differs from that of current ATM. Under the supersector, ATM consists of two 4-person air traffic control (ATC) teams who share the responsibility of traffic management. One team is assigned the task of managing the

north and east bound traffic, while the second team manages the south and west bound traffic. In each team, three members' roles are based on the expected arrival times of incoming aircraft, which are set at 15, 10, and 5 minutes out from the sector boundaries. The fourth team member, the tactical controller, would manage the traffic within the supersector. The two teams would be coordinated by a supervisor, resulting in a total of nine people managing the traffic not only within the supersector but also engaging in the task of planning for traffic arrival and exit both to and from the supersector (EEC, 2002).

Figure 2. An outline of the supersector shape and the trajectory-based airways. The green arrows represent the north and eastbound airways while the blue arrows represent the south and westbound airways.

Converting from the current day operational procedures to an, as yet, unproven airspace concept, is potentially rife with unforeseen consequences. Research and testing

of new operational concepts under different environmental and operating conditions are therefore essential. To aid operators—both pilots and controllers—in coping with fluctuating workloads and decision making responsibilities, the airspace system is becoming increasingly reliant on automation and decision-support tools. However, despite the sophisticated technological capabilities, decision-support tools can have a deleterious impact on operators' performance. This is because decision-support tools and their modes of information delivery are specifically designed to support the cognitive and information processes of the controller when performing tasks under particular working environments. Moreover, decision-support tools can only provide limited aid when the stress of decision making is exacerbated by the physical limitations of the airspace. The design of information displays—in this case, the delivery of altitude information via color and text—should therefore be developed in tandem with the procedures, demands, and operational needs of the operators who maneuver and manage the sector.

As such, when assessing the viability of the supersector as a method of alleviating the impact of air traffic growth on controllers, effort should be directed toward identifying the effects of augmented roles and responsibilities on the working practices of the human operator. For this reason, the examination of the supersector concept herein focuses on the tactical controller, who, we postulate, would face the most notable change—the sharing of a large sector with another controller—when adapting to the supersector's working procedures. This contrasts with the current practice of a single controller managing all the traffic in a much smaller airspace.

Airspace Complexity

The purpose of creating one-directional airways and integrating six to eight sectors is to decrease the complexity of the airspace and create a greater degree of predictability in the traffic pattern. By expanding the airspace governed by two ATC teams, it is envisioned that greater autonomy—in terms of both physical and procedural autonomy—will be available to the controller for rerouting aircraft and avoiding conflicts. The consolidation of airways in the supersector is intended to reduce the complexity of traffic distribution by consigning large volumes of aircraft to a handful of airways and providing greater predictability in the traffic pattern.

Airspace complexity is a construct that is difficult to compare across different sectors. No two sectors are completely alike. Moreover, since so many factors (e.g., aircraft density, geographical variances, weather patterns, traffic configurations, workload demands, and temporal changes of traffic distribution) impact the controller's ability to make decisions and manage the workload, it is difficult to gauge and compare complexity levels between sectors. The factors that make one sector particularly complex, such as a wide range of aircraft speeds, may not be a factor in another sector that is considered to be equivalently complex. The other sector may be complex due to the high frequency of climbs and descents or frequency of exits and entries.

Airspace Complexity and Cognition

Increasing the predictability of traffic movements enables the controller to quickly develop a *default knowledge* (Leroux, 1998), based on the controller's experience with

events in a particular time, place, and context. Default knowledge is the basis by “which controllers treat situations although they logically lack of [*sic*] information to do so” (Bresolle, Benhacene, Boudes, & Parise, 2000, p. 2) and develops through the repeated experience of observing traffic and movements in the sector, allowing the controller to develop expectations for movements and events in the airspace. Moreover, default knowledge becomes an invaluable cognitive resource when a controller needs to engage in timely decision making. Default knowledge enables the controller to make an educated guess with respect to the probable intent and movement of aircraft. Its acquisition results in the development of *sentry parameters* (Leroux, 1998), the cognitive alerting mechanism that allows controllers to differentiate between normal and unusual events. As such, the controller becomes more sensitized to unusual events and movements.

Airspace Complexity Metrics

Airspace complexity has been measured in many ways. Complexity metrics have been based on the traffic density, aircraft maneuvers, structure of the sector, or the mental and physical workload imposed on the controller. The design and procedures proposed in the supersector concept address the issue of airspace complexity in terms of each of these aspects.

Delahaye and Puechmorel (2000) and Histon and Hansman (2001) measured airspace complexity with physical factors, such as the type and number of aircraft in a sector, the shape of a sector, the array of trajectories, the number of climbs and descents, entries and exits, and the structure of sectors. Alternatively, Auley, Lenzi, and Lepadatu

(2001) and Morrow-Magyarits and Kopardekar (2000) measured complexity by assessing the cognitive and physical workloads imposed on controllers. Despite the lack of uniformity in measuring airspace complexity, it is generally agreed that traffic complexity is positively correlated with controller workload. Therefore, efforts to reduce controller workload should logically focus on reducing traffic and airspace complexity. The rationale is that a reduction in airspace complexity – and the subsequent workload reduction – would enable controllers to accommodate a greater number of aircraft operations.

If the supersector manages to reduce the complexity of the workload and airspace, then the effectiveness of this new operational concept has to be quantified. For this purpose, the following are used to compare performance in the supersector against performance in a current sector under today's operating procedures:

- The number of extra aircraft operations safely absorbed by the supersector
- The degree of change in controllers' cognitive and physical workload

The experiment discussed herein does not aim to quantify or compare the complexity of the supersector against a current sector, but rather, it aims to gauge the potential of the supersector by comparing the performance of tactical controllers working in each of the two airspaces. This comparison is made by manipulating the traffic intensity moving through the sector per unit time and then comparing controllers' performances between the supersector and current-day airspace based on two traffic-intensity levels.

Another factor affecting the level of workload is the mode of information delivery on the controller's display. The efficient delivery of information can significantly aid controllers in reducing both their mental and physical workload (Luder, 1984; Mulligan, 2001), thereby reducing the task complexity. Therefore, a method of delivering time-critical information will also be tested for efficacy and efficiency of information delivery.

Information and Computer Displays

Advancements in computer display technology provide an opportunity to increase the efficiency of information delivery to operators whose responsibilities demand both speed and accuracy. The greater availability of colors and the improved resolution of graphics allow display designers to exploit technological capabilities and deliver decision-support information more efficiently and effectively. However, the difficulty of designing appropriate and effective displays is compounded by the dynamic nature of air traffic control. Additionally, the limited screen real estate and the demand for timely information result in a display environment where screen clutter can become a real risk to safe operations. Therefore, the efficient delivery of information becomes an important variable when there is a demand for a radar screen that is information rich with a reduced tendency for clutter.

Data Delivery versus Information Delivery

An important issue in display design involves the distinction between raw data and meaningful information. To be effective, operators must search and process information before making a response. The delivery of raw data leaves the operator with the

responsibility of interpreting the information. Alternatively, the delivery of information—in a meaningful format—reduces both the cognitive load on an operator and the time needed for cognition and response.

Features such as shape and color can be used as an information-rich source of data. By making visual features meaningful, information can be processed more rapidly, while enabling the operator to maintain vigilance over the system. The use of information input from the environment (such as color, shape, size) for decision making is called “bottom-up processing” (Wickens & Holland, 2000). Bottom-up processing refers to perceptually based cognition, rather than cognition from previous knowledge. If cognition is based on previous knowledge, such as knowledge of the performance characteristics on an aircraft, then this is called “top-down processing” (Wickens & Holland, 2000).

Despite the volume of literature that addresses the issues of human-machine interaction (Bennett, Toms, & Woods, 1993; Raskin, 2000; Silverstein, 1987; Stokes, Wickens, & Kite, 1990), there is no single design solution that works for every situation. Humans rely on a variety of cognitive processing resources depending on their level of expertise, the types of tasks being performed, the complexity and number of actions to be completed, the type of attention required (e.g., focused or selective), and whether the tasks are accomplished serially or simultaneously. Other factors to consider when selecting one display design over another include technological constraints, the environmental characteristics of the workplace, and the accuracy required for the task. Therefore, the design of any information display has to be customized to the task at hand.

Color

The perception of color is based on the interaction of light waves with chemicals in the visual system. Humans can generally identify nine colors on an absolute basis and twenty-four colors when properties such as saturation, luminescence and value are manipulated (Stokes et al., 1990). Hence, the wide range of perceptually distinguishable colors presents a strong option for enhancing the efficiency of information delivery, potentially adding another dimension to the two-dimensional plan display currently found on ATC radar screens.

Due to perceptual adaptation, and at times, difficulties in differentiating between colors, color-coding should only be used as a redundant feature. That is, color should only be used to reinforce information that is already present in another display format. It should never be the sole mode of information delivery. Redundant color coding in the aviation domain is a well-documented area of research, although much of this work has focused on its application in the design of cockpits displays (Andre, Wickens, Moorman, & Boschelli, 1991; Beringer, Allen, Kozak, & Young, 1993; Luder, 1984; Reising & Aretz, 1989).

A review of a wide collection of color-coding research by Christ (1975) found that color coding has been used in displays with mixed success and the use of color to enhance ATC performance has not been consistently proven. Studies by both Schmit and Tullis, as cited by Narborough - Hall (1985), found no improvement in performance with redundant color coding in ATC displays. Color can at times interfere with search and object comparison tasks due to visual adaptation to the color display from long-term

exposure. Additionally, highly saturated colors can appear to float when displayed in a high-contrast color scheme, especially when displayed with a color opposite on the color wheel. This phenomenon is called *chromostereopsis*. The different dimensions of coding color (e.g., chromaticity, brightness, and luminance) can produce mixed results. Varying the levels of a color property (e.g., chromaticity) may aid visual search, while combining the levels of two or more color properties may have little or a detrimental affect depending on the task and the number of distracters on the screen (Nagy & Sanchez, 1992). Therefore, a judicious approach should be taken when using color in display designs.

Generally, redundant color-coding provides a salient and efficient method of aiding visual search. In the aviation context, where information displays are information rich, complex, and rather dynamic, the redundant color coding of information has been commonly found to enhance performance, if only to a limited degree, (Andre et al., 1991; Johnston, Horlitz, & Edminston, 1993; Mulligan, 2001) by reducing visual search and conflict detection time.

Color also provides a means of quickly grouping information with little cognitive effort, since the perception of color is a physiological rather than cognitive phenomenon. Hence, the bottom-up processing of information results in a reduction in the cognitive effort and mental workload required for decision making. As well as being a coding mechanism that requires little cognitive effort, color is a means of grouping information. Humans have a tendency to group objects based on common visual features. This is the central thesis of Gestalt theory. Gestalt theory also explains the tendency for humans to

perceptually organize objects with a common visual feature, such as color. One of the laws of Gestalt psychology, known as the *law of common fate*, states that “things that are moving in the same direction appear to be grouped together” (Goldstein, 1999, p.183). Therefore, with each airway containing aircraft moving in a single trajectory, it is believed that will color provide a strong grouping mechanism by which controllers can identify *my own aircraft*. In the experiment proposed herein, all aircraft *owned*, or for which a controller has responsibility, will be displayed in a single hue, either blue or green. In this case, it is the hue that forms the basis of grouping aircraft management responsibility.

The selection of color codes is important for facilitating learning and communication. Therefore, color codes should be meaningful and easy to learn. Bennett and Flach (1992) discussed the importance of graphical displays that facilitate the semantic mapping of information. In this respect, they cited Gibson’s ecological optics which asserted that “skilled performance is possible when the geometry of optic flow fields is specific to properties of surfaces and movements of an observer” (Bennett et al., 1993, p. 530). In this experiment proposed, aircraft altitude will be coded by varying levels of color saturation.

Saturation is the purity of a color relative to its brightness. It is inversely related to the amount of white in a color. In this experiment, coding is based on a system whereby saturation decreases (colors appear lighter) as an aircraft climbs to higher altitudes. Semantically, controllers can draw meaning from the color scheme in two ways. First, under the green hue, the aircraft at lower altitudes are displayed in a highly saturated

green (dark green) while aircraft at higher altitudes are displayed in increasingly less saturated hues (lighter shades of green). Alternatively, controllers can bring meaning to this system by associating the dark green aircraft to the dark green of foliage and landscapes. As the altitude increases, the aircraft head towards the clouds. This is reflected by an aircraft display that is lighter. Similarly, in traffic sets colored in a blue hue, low flying aircraft are displayed in a dark blue while higher aircraft are displayed with a reduced color saturation, in a pale blue similar to clouds (see Figure 3).

Figure 3. A color – saturation coding scheme is used to display aircraft altitudes. As aircraft climb to higher altitudes, the color saturation of the aircraft displayed on the air traffic controller’s radar screen decreases.

Human Performance Metrics

Human performance has been measured both subjectively and objectively (Galster et al., 2001; Manning, Mills, Fox, Pfleiderer, & Moglika, 2001; Stager, Ho, & Garbutt, 2001). The impact of new airspace operational concepts on the human operator has been the topic of much research (Corker, Gore, Fleming, & Lane, 2000; Endsley, 1997; Endsley, Mogford, Allendoefer, Snyder, & Stein, 1997; Hilburn, Bakker, Pekala, & Parasuraman, 1997; O'Hare & Roscoe, 1990; Remington & Johnston, 1995) since human performance and safety are invariably intertwined. Studies that examine the impact of new ATM concepts and technologies on the human operator tend to focus on situation awareness, stress, workload, actions and human-machine interactions of the operators.

Workload is a concept without a concrete definition, but generally refers to the cognitive effort required to perform a task. However, Manning, Mills, Fox, Pfleiderer, and Moglika (2000) integrated *taskload*, the physical effort of an operator, into a workload metric called the Performance and Objective Workload Evaluation Research (POWER). New operational concepts have been known to impact the workload of controllers because some concepts that are meant to increase autonomy for pilots—such as self-separation in free flight—result in a marked decrease in air traffic predictability and controller situation awareness. As such, the controllers, who are still expected to maintain supervisory control, may be left with little intent information. This results in the escalation of stress, anxiety, and cognitive effort. Therefore, the impact on the controller has to be assessed when proposing to introduce an operational concept capable of

physically absorbing more traffic. To maintain safety, it is desirable to have human operators who are not overloaded with task demands.

Measures of workload have been collected using objective and subjective methods. Subjective methods rely on the perceived cognitive effort of the subject whereas objective methods ascertain affect via performance measures such as accuracy. However, none of these methods are considered perfect as they each have limitations. Objective performance measurements (e.g., the number of errors, response times) may not correlate to any improvements perceived by the participant.

To create a balanced data collection, this experiment will collect both subjective and objective measurements of performance. To avoid memory effects, controllers will be asked to estimate their workload at frequent intervals. To buttress these results, objective performance measurements will also be collected. In this case, activity level will be quantified by the number of actions taken. To supplement the subjective workload analysis, the performance of the participant, measured in terms of the number of actual and potential conflicts and conflict resolution times, will be used to compare participant performance under the different operating conditions.

Working Memory: The Visuo-Spatial Sketchpad

Baddeley (1995) explained the construct of short-term memory with his model of working memory, a three-part system that stores short-term memory. These three parts are composed of the visuo-spatial sketchpad, the phonological loop and the central executive. In Baddeley's model, the information from the visuo-spatial sketchpad and

the phonological loop is integrated in the central executive. This model of working memory is important to the design of information displays as it underscores the significance of information modalities in aiding cognition. In this model, working memory would result in more efficient cognition if the modality of the information for both encoding and response is the same.

Wickens' Multiple Resource Theory (MRT) (Wickens & Holland, 2000) buttresses this argument. MRT attempts to explain how humans can process and respond to information in several modes simultaneously. The experimental results by Wickens, Sandry, and Vidulich (1983) found that, indeed, visual information for spatial tasks and auditory information for tasks requiring a speech response generated superior performance. The potential application of this knowledge to the design of complex displays is broad. Considering that air traffic controllers perform safety-critical tasks, displays that enable faster yet accurate decision making can make an important contribution towards the task of reducing controller workload.

Hypotheses

Experiments examining the effects of saturation-coded aircraft altitude on controller performance will present two levels of airspace complexity—medium and high—reflecting operations in the real world where variations in traffic densities and trajectories occur regularly.

The experimental design has the following aims:

Aim 1: Measure controllers' performance under two levels of air traffic complexity operating in both the supersector and current day airspace.

Aim 2: Examine the effect of saturation-coded altitude display on the conflict detection and subjective mental workload of en route air traffic controllers working in both the supersector and current day airspace.

The following hypotheses will be tested with a human-in-the-loop simulation experiment:

1. Operations under a supersector will reduce controller mental workload, yield earlier conflict resolution and fewer potential conflicts when compared to the current - day sectorized airspace.
2. Color-saturation coding of altitude results in earlier conflict resolution, fewer conflicts and a reduction in controller workload than when altitude information were only presented as alphanumeric text in a data block.

Literature Review

Human-Computer Interaction

There is no single rule for designing information displays; inevitably, the most appropriate design depends on the type of task being performed, how information will be used in decision making, the urgency and accuracy of the informational needs, and the physical characteristics of the working environment (Bennett et al., 1993; Christ, 1975; Hess, Detweiler, & Ellis, 1999; Johnston et al., 1993). The role of information displays is to aid the user in decision making and the effectiveness of any display depends on whether the information is used for a single task, several tasks, or a sequence of tasks.

Williams (2000) examined the performance of pilots in navigating to the nearest airport with navigation information presented in each of three formats: a pure text display, a text display enhanced with graphical orientation markers, and a pure graphical display. He found that graphical displays produced superior performance when compared to the text displays. Moreover, he concluded that the trade-off between graphical and textual displays and the selection of one information code over the other is entirely task dependent. The selection of graphics—instead of a text display—is dependent on the type of task, the information required, and the response modality of the operator.

Additionally, Williams' conclusions support the theory proposed by Wickens, Sandry, and Vidulich (1983): that tasks requiring visual-spatial cognition should be supported with information presented in the same cognitive mode. They asserted that the

mode of information input (encoding) and information output (response) should use the same cognitive resources. Wickens et al. (1983) carried out an experiment that examined performance with different encoding and response modalities for a flight task. Their results found the best performances occurred when a spatial task requiring a manual response was displayed visually, and also when a verbal task that required a speech response was presented aurally.

Scholten (1969) was one of the first to explore the modes of displaying aircraft height on ATC radar screens. He found that 2-dimensional (2-D) representations were preferable to 3-dimensional (3-D) representations due to the difficulties in accurately estimating horizontal distances and vertical differences in a 3-D display when the aircraft were some distance apart. The design of 3-D displays relies on binocular parallax; a concept that has significant disadvantage when presented on a 2-D surface in that objects made to look three dimensional on a two-dimensional surface can present the viewer with ambiguous or illusory depth cues. In scenarios where accuracy in estimating distances is important, Scholten concluded that 3-D displays should not be used. Instead, he proposed color-coding aircraft altitude so as to add a third dimension to a two-dimensional display without the disadvantages of a three-dimensional display.

Color-Coding of Information

Surprisingly, there has been little research in the use of color coding to represent altitude. The most common use of color coding in information displays is found in a cartographic context, where color is used to code either different types of terrain or changes in the height of terrain. In ATC radar displays, color-coding is most commonly

used to indicate the different states of an aircraft. For example, in the Operational Display and Input Development (ODID) display developed by EUROCONTROL, the datablocks change color if an aircraft is approaching the controller's sector, is already in the sector, or has just left the sector. Scholten (1969) noted that some of the earliest research into color-coding altitude was conducted in 1947, and again in 1960 by Schneider (1960), who suggested the use of five colors to represent altitude. Scholten also proposed a display system whereby the altitude as well as the speed vector would be represented by a color code. This display has the potential of producing 25 (5^2) different color combinations, an excessive use of color, especially when considering that good human-machine interface (HMI) design principles aim to limit the number of colors used for information coding (Stokes et al., 1990). Additionally, it is believed that there is a lack of an intuitive relationship between speed and color to make this display format useful.

Research into performance with color-coded information (Christ, 1975; Johnston et al., 1993; Mulligan, 2001) found significant reductions in reaction time and error rates when comparing controller performance with information delivery in a text format. Johnston, Horlitz, and Edmiston (1993) found significant improvements in search time when the altitude of aircraft was coded in various colors. In this experiment, each aircraft displayed potentially three different colors. Aircraft were coded in colors that reflected the initial, current and assigned altitude. The potential problem with this display format is that too many colors can be employed, and information coded in each of the number of different color bands may impose a heavy load on working memory. Additionally, the

experiment by Johnston et al. (1993) used naïve participants rather than trained air traffic controllers. One could argue that due to this participant pool, the bias toward color as a method of coding altitude was therefore amplified since the participants were not trained in any other method of sourcing altitude information. As such, these results cannot be used to predict any improvements in performance from a population of trained air traffic controllers.

Both studies by Johnston et al., and Scholten used a color coding scheme that was based on the entire color spectrum. Therefore, the altitude associated with each color still had to be learned by the controller. The inherent weakness in using the entire color spectrum is that the operator has to learn the order of the colors as well as their proximity on the spectrum. There is a lack of meaning that the controller can draw from the display other than learning that two similar colors (e.g., red and orange) indicate the aircraft are in close proximity.

Color coding has also been used for the delivery of tactical information in combat aircraft, as in the cockpit of the F-22 Raptor. It has been tested in cockpits for presenting nearby traffic via the Cockpit Display of Traffic Information (CDTI) (Beringer et al., 1993), for presenting military pilots with horizontal-situation displays (Derefeldt et al., 1999), and even in weather information for commercial aircraft (Scanlon, 1994). EUROCONTROL has been using color to code the status of flights in their ODID displays for air traffic controllers.

On the whole, research into color-coding radar display information, as a means of reducing controller workload, appears to be promising. In an environment where

response rate and accuracy are both significant performance parameters, a color display of information could be particularly advantageous for enhancing human performance. By encoding aircraft displays with color, information redundancy is provided, and the reliance on textually based information processing is reduced. Consequently, this means the frequency and effort required for searching, reading, processing, and decision making—the total mental workload—is also, hypothetically, diminished.

Representing a spatial concept with pictorial cues is consistent with the encoding-compatibility principle and also supports the working memory model proposed by Baddeley (1995). By linking color to a spatial semantic, the processing of information is compatible with the memory mode. Conversely, altitude information displayed as text requires cognitive processing in the phonological loop, a modality inconsistent with the spatial context of the information.

Hess, Detweiler, and Ellis (1999) explored the link between display layout and the ability to keep track of rapidly changing information. Their experimental design examined the role of verbal, spatial and visual-spatial icons in supporting monitoring tasks as a function of the number of objects and attributes associated with the objects. Their experiment examined the interaction between display layout and the spatial nature of the information and highlighted the cognitive cost of monitoring more information. As the number of objects to be monitored increased, response time grew accordingly. Moreover, displays that provided redundant information aided the development of information coding strategies or mnemonics. Their research highlighted the role that display design can have in reducing the load on working memory and underscores the

need to provide redundant information so to improve response time and performance in complex displays.

Airspace Complexity

The difficulty in establishing complexity metrics lies in the tendency of traffic patterns to be affected by a myriad of factors. The variables that impact traffic complexity are multidimensional; as such, complexity becomes a complicated concept to quantify. Nevertheless, the assessment of new ATC concepts and decision-aiding tools must occur in the context of varying airspace complexity so that evaluation of new tools and operating procedures are tested with ecological validity. It is possible that a procedure or decision-aiding tool that significantly improves performance under low airspace complexity conditions has either no effect or an adverse effect under high airspace complexity. The alternative may also be true. The difficulty in making complexity metrics comparable across a broad array of research and experiments is due to researchers defining airspace complexity as a function of a vast array of elements. The impact of sector design on performance and airspace complexity will also be discussed.

Complexity Metrics

There is no single definition or metric of airspace complexity, due to the many elements (such as traffic trajectories, speed, volume, controller mental and physical workload, traffic volume and entry rate, capacity, and different types of aircraft in an airspace) that contribute to variations in complexity. Morrow-Magyarits and Kopardekar

(2000) defined dynamic density (DD) as being analogous to air traffic complexity whereby DD measures ATC workload as a function of the complexity of traffic patterns.

In the United States, the current method of measuring airspace complexity is generally regarded as being too simplistic (Histon & Hansman, 2001; Mills, 1998; Pawlak, Brinton, Crouch, & Lancaster, 1996), as it is only a measure of the number of aircraft expected to occupy a sector. So, rather than being a measure of airspace complexity, a more accurate description of this metric is 'airspace density.' By using airspace density metrics as an estimate for complexity, the impact of factors such as dynamics, the number of trajectories in which the traffic is flowing, or the mental and physical workload imposed on the controller is ignored.

Airspace complexity in Europe is not measured by any single formula. Each control center has its own methods of deciding when a sector is "excessively" complex. The difficulty in making absolute, quantifiable, complexity metrics across multiple sectors are exacerbated by the different elements (such as geographic location and layout of the airport, number of available runways, types of aircraft in an airspace) that impact each area and thus, make comparing different airspace unscientific and inappropriate.

Pawlak, Brinton, Crouch, and Lancaster (1996) proposed a technique of measuring and evaluating airspace complexity, based on the integration of mental and physical workload measurements. In this approach, complexity measurements are a function of the required cognitive workload. Their research initiated from an investigation of airspace complexity metrics under the context of free flight operations.

Auley, Lenzi and Lepadatu (2001) created the methodology called CAPAN, which measures capacity as a function of controllers' workload rather than as a function of the number of aircraft or traffic volume, as does NASA's dynamic density model (Morrow-Magyarits & Kopardekar, 2000).

Histon and Hansman (2001) focused on the structural, operational and aircraft distribution elements that contribute to controller workload, such as flow structures and aircraft characteristics. They proposed a complexity metric algorithm based on the interaction of traffic structure, load and operations.

Alternatively, Delahaye and Puechmorel (2000) approached complexity metrics from a more geometrical perspective. Instead of viewing the airspace complexity as a function of multiple factors, they examined the geometrical properties of the traffic pattern. A complexity algorithm was calculated based on the geometrical properties as a function of time. They also measured the evolution of traffic patterns as a dynamic system. Integrating the metrics from geometric and dynamic properties, Delahaye and Puechmorel applied the data into the Kolmogorov entropy, an algorithm for testing and predicting events with degrees of complexity. They concluded that despite having geometric and dynamic data with respect to the movements of aircraft, it is very difficult to build a model to predict airspace complexity when both quantitative (e.g., vectors, number of aircraft) and qualitative (e.g., mental workload) factors need to be integrated.

The different approaches toward defining airspace complexity only highlights the difficulties of empirically classifying complexity with one simple metric, especially since airspace elements and controller workloads vary vastly. However, it necessary to

compare what is seemingly, equivocally complex traffic so to understand how various information displays or working methods may better assist human operators in managing their workload.

Sector Design

The design of a sector has a direct influence upon the complexity and mental workload required for managing an airspace. Rodgers, Mogford, and Mogford (1998) examined the relationship between operational errors, sector design, and sector complexity. Their analysis of operational errors in the Atlanta Air Route Traffic Control Center (ARTCC) sector found evidence to indicate the possibility of a link between sector complexity, reductions in situation awareness, and operational errors. The overall operational error (OE) rates could be predicted by statistical analysis of sector characteristics and flow variables, although there was low accuracy in their calculations. Sector design principles that tend to alleviate controllers' mental workload were also discussed. For example, designing a sector so that the shape is parallel with the direction of traffic flow has been shown to reduce controllers' mental workload. Increases in sector complexity—which may be associated with the design of the sector—may also have the effect of reducing situation awareness and increasing operational errors. Correlations and interactions among these elements warrant further investigation.

Although the design of the sector has been shown to impact the rate of OEs and cognitive workload, the level of complexity and traffic load are also relevant factors. Rodgers, Mogford, & Mogford (1998) cited the research of Stager and Hameluck and Stager, Hameluck, and Jubis in discussing the higher rates of OEs that tend to occur

under conditions of low to moderate workload. Reasons given for these errors include lapses in attention and judgment and problems with communications.

Human Performance Metrics

Subjective mental workload is a measurement of cognitive effort as estimated by the participant; hence the term “subjective.” Accurate measurements can be skewed by the timing of measurements—whether it is taken during or after a task—and how recently the task was performed. Procedures for collecting subjective mental workload metrics include the Air Traffic Workload Input Technique (ATWIT) and NASA Task Load Index (NASA TLX.) The ATWIT methodology requires the participant to estimate the level of their mental workload during the experimental procedure by entering a value on a numeric scale via a keyboard. Conversely, NASA-TLX interviews the participants after the experiment. Both methods have their strengths and weaknesses.

By asking participants to respond with estimations of workload during a task, it could be argued that a secondary task is created, thereby elevating the mental workload from that experienced by solely working on the primary task. NASA-TLX has the disadvantage of being employed at the end of an experiment and therefore risks having the data tainted by recency effects.

Objective measurements of performance may include reaction time, the number of errors made, and the number of actions taken. To objectively validate any changes in performance from operational changes or technological applications, methods of measuring performance and activity levels (i.e., number of actions taken) are used to complement subjective workload metrics. As such, there may be contradictions

between estimations of subjective mental workload and data collected from measuring the objective performance and activity levels. Manning et al. (2000) developed the POWER method to objectively measure controller workload. The POWER method adapted techniques from both ATWIT and NASA TLX but included objective measurements such as a count of the total number of conflict alerts, handoffs, altitude changes, and data entry errors. Their study examined the relationship of sector complexity, performance, and controller workload. While participants observed traffic samples from the Kansas ARTCC, they estimated their workload and performance. Their results found a correlation between workload and sector complexity. However, the relationship between sector complexity and controller performance was not strongly evident.

Summary

To increase the capacity of the airspace, it will be critical to reduce the mental workload of the air traffic controller. Since metrics of complexity and capacity are centered on the workload and capability of the controller in coping with fluctuating complexity and volumes of traffic, then reducing mental workload will be a critical move towards maintaining a level of performance that ensures safety while increasing capacity.

Innovations in airspace design have to be developed with considerations to human-machine interface issues, as information delivery, speed, and accuracy of information processing will be integral parameters for the success and safety of any new operating system. Research into color coding spatial information has produced some positive

results and deserves further investigation as a method of reducing controllers' mental workload and improving performance and situation awareness.

This research proposes to study the effect of color-saturation coding aircraft altitude on the performance and mental workload of en route air traffic controllers operating under the procedures defined by the EEC for supersector operations. Altitude information in a saturation-coded format is intended to reduce the likelihood that the information will be misread. The display of aircraft using color saturation levels is capable of conveying information about potential conflicts without the need to read and process the text information in each and every datablock. By presenting aircraft altitude via color saturation, two pieces of information are integrated and delivered simultaneously. Information about altitude differences can be quickly interpreted by comparing the saturation of two or more aircraft. Information about who has responsibility for an aircraft is contained by the hue of the display.

Method

Participants

Ten air traffic controllers were recruited from the National Aviation Research Institute (NARI). Each data collection session lasted for one day and two controllers were tested simultaneously in each of the one day sessions. We recruited three female and seven male air traffic controllers, aged 35 – 66 years old, with 9 to 35 years of active radar experience. The mean age of the participants was 45. The participants had a mean of 15.5 years of experience working in a radar capacity. Seven participants were actively working as air traffic controllers at Oakland Air Route Traffic Control Center (ARTCC), while three participants were retired controllers. No participants had color vision problems since all applicants are tested for color vision deficiencies upon applying to the Federal Aviation Administration (FAA) for employment as an air traffic controller. The selection criteria required all participants to have at least 2 years working experience as an en route controller in the radar capacity. All the participants were paid for their participation.

Simulation Program

The simulation program was written in C++ using graphics from Open GL (Graphics Library by Silicon Graphics). Flight records (flight name, aircraft type, speed, altitude, flight path) were written in NASA's Pseudo Aircraft System (PAS) format.

The simulation displayed aircraft in a plan display airspace, with the aircraft moving at a climb, cruise, and descend rates similar to real aircraft movements. Waypoints, airports, and, sector boundaries were displayed on the map. The program was written so that every 7 s, the computer scanned for potentially conflicting aircraft pairs up to 5 min the future. These conflict pairs were logged and used for data analysis.

Equipment

Five personal computers were dedicated to this experiment. Four of these were used as work stations while the fifth computer was used as a server. Each computer was a Pentium III ® PC, operating on a Windows 2000 ® platform with 256 MB of random access memory (RAM). We used video accelerator cards manufactured by ATI Graphics, namely the Radeon 7000 ® with 64MB. Each controller was linked to a single pseudo pilot through headphones and the computer server (which displayed the same traffic in real time). Microsoft Sidewinder ® version 1.0 software was used for voice communications. Voice compression on this software was at 32 kilobytes per second, which simulated a voice quality similar to that of land-line telephones.

The traffic was displayed on 17 inch (15 inch viewable) Dell ® Ultrascan P991 computer monitors set at a resolution of 1024 x 768 pixels. The monitors ran on 16-bit, high color display. Each monitor was calibrated and placed on a brightness setting of 35 and a contrast setting of 81. The keyboard commands were entered through a Dell ® QWERTY keyboard and mouse. The participants used a Kensington ® 4-button trackball as a pointing device.

Map Display

Two different maps were used, which represented two different types of airspace. All colors were defined as Red Green Blue (RGB) values when programming the computer simulation. In the supersector airspace, the display was presented on a dark gray background (defined as 35/255, 35/255, 35/255) while the sector boundary was marked by a thin gray line. The color of the supersector boundary was defined as 80/255, 80/255, 80/255. Airways, which were 15nm wide, were displayed on the map in dark gray (45/255, 45/255, 45/255) (see Figure 4). To increase traffic intensity, we increased the number of flights through both the supersector and the current day sector. However, we endeavored to maintain fidelity by keeping to the origin and destination routes of the original traffic samples.

Figure 4. A screenshot of the supersector airspace as seen by participants.

To enhance realism, the current day airspace was based on a sector in southeastern France, the Uniform Juliet (*UJ*) sector. This sector has an area approximately one sixth the size of the supersector. In the *UJ* airspace, the sector boundary was defined with the color 229/255, 229/255, 229/255. Due to the small size of the *UJ* sector, no airways were displayed. However, the background color was the same in both types of airspace. (See Figure 5).

Aircraft Display

Aircraft were represented on the screen by solid-colored triangles measuring 6 mm in length and 4 mm in width. The screen display of each aircraft was called a *target*. Each target was equal to 40 min of visual arc on the long axis when the participant was seated 50 cm away from the screen. The apex of the target pointed to the trajectory of the aircraft's heading. To increase the salience of the target trajectory, controllers generally had the vector of every aircraft displayed at all times so that its position 1 min ahead was always visible. Therefore, the vector line extending from the top of each target pointed to the aircraft's immediate trajectory. Vector lines were displayed in the color of the target. In the scenarios where saturation coding was used, the color of the target was determined by its current altitude.

Aircraft color was presented in two conditions: with and without color-saturation variances as a means of coding aircraft altitude. With color-saturation coding, as aircraft altitude increased, the saturation of the color decreased (i.e., became lighter). The RGB values of the colors were selected from the Munsell color system (Macbeth, 2001) so that all colors were spaced by a just-noticeable-difference, one step apart.

Under the color-saturation coding condition, all aircraft in the blue hue had five values ranging from 4 to 8. These five blue values had the following RGB definitions (from least to greatest saturation):

1. 33/255, 105/255, 132/255
2. 57/255, 30/255, 156/255
3. 82/255, 158/255, 181/255

4. 107 /255, 186 /255, 206/255
5. 181/255, 239 /255, 255/255

Under color-saturation coding, the aircraft displayed in the green hue had six values ranging from 4 to 9. These values were defined by the following RGB definitions (from least to greatest saturation):

1. 0/255, 126/255, 55/255
2. 50/255, 152 /255, 77/255
3. 84/255, 177/255, 99/255
4. 113/255, 203/255, 121/255
5. 138/255, 229/255, 143/255
6. 149/255, 255/255, 172/255

Five levels of color saturation were available under the blue hue and six under the green hue since the EEC operating concept assigned five flight levels to southbound aircraft and six flight levels to northbound flights (EEC, 2002).

When managing traffic without color-saturation coding in the supersector airspace, the participants identified the aircraft under their responsibility by the hue of the targets. The targets were displayed with a single value of blue and a single value of green. That is, the varying saturation levels were absent but the blue and green hue of each traffic set remained. Under this condition (supersector airspace and without color coding of

altitude), the blue targets and green targets were displayed with the 3rd and 4th level of color saturation.

When participants were managing traffic in the UJ sector, all aircraft were presented in the green hue, following the current standard practice. Therefore, with color-coding present, the above six green saturation values were used. Without the presence of color-saturation coding, all targets were displayed with the 4th green value, defined by RGB values as 113/255, 203/255, 121/255.

A J-ring is a thin circle, displayed in the same color as the aircraft, showing the five-mile radius boundary around an aircraft. To aid controllers in making decisions about distance—in particular the five mile radius—the J-ring tool was made available. This served not only as a decision-support aid for detecting potential conflicts, but also provided a sense of scale and distance for the participants. In current operations, controllers at the ARTCC can turn on J-rings when managing traffic. Participants could turn the J-rings on and off by hovering the cursor over the target and hitting the 'J' key.

Datablocks were always displayed with white text and linked to all targets by a thin, white leader line. Participants were able to turn datablocks on and off when aircraft were not in their sector. Datablocks showed the aircraft's flight number and current altitude. When an aircraft was climbing or descending, an arrow would appear on the second line between the current aircraft altitude and its assigned, final altitude. The third line of the datablock displayed each aircraft's computer identity number (CID) and ground speed in knots. (See Figure 6).

Figure 6. Datablock on the radar screen display as shown during the experiment.

When aircraft appeared on the screen outside the sector, the call sign of the datablock flashed every 2 seconds. This indicated to the participants that they had not accepted control of the aircraft. Upon clicking on the aircraft target, participants could assume they had accepted control or the handoff from an adjacent sector. Accordingly, the pseudo pilots would also assume that they were the pilots of the now-accepted aircraft.

Trajectory Information

Participants could access information on aircraft history by bringing up a path history showing the last 1, 2, 5, and 8 refreshes. The radar display was refreshed every 7 seconds. To access information about the anticipated aircraft trajectory, participants could display the trajectory path up to 1, 2, 4, and 8 min into the future. Route information was available when the cursor hovered over the target and the controller hit

the 'R' key. A line showing the aircraft's flight plan would appear and it then disappeared after 3 seconds.

Communications

The participants were also told they could reroute aircraft, by giving pilots "fly-direct" instructions. In this instance, pilots were asked to fly direct to a waypoint along their flight plan, skipping earlier waypoints along their flight plan. Additionally, controllers could also ask pilots to temporarily vector for traffic. Under these circumstances, the vector lines, which controllers generally left displayed at all times, would be displayed in yellow. This acted as a visual cue so that at a later point, controllers would be reminded that an aircraft needs to be vectored back to its original flight path. Finally, participants could direct the pseudo pilots to change the altitude of an aircraft, either temporarily or permanently, and adjust aircraft speeds.

Traffic

The airways were developed by the EEC Innovative Research team based on the demands of current and anticipated traffic patterns. The supersector traffic was sourced from the EEC. It was a real Parisian traffic sample that passed through northwestern France on September 5th, 2001. The traffic in the UJ sector was also supplied by the EEC and consisted of typical airlines and routes for flights that pass through the UJ sector. In both airspace conditions, the traffic intensities were amplified after receiving the samples from the EEC. Pilot studies confirmed that the two traffic intensity levels were adequately different.

Procedure

In each session, we briefed and collected data from two controllers simultaneously. In the supersector scenarios, the participants were required to interact and share responsibility for the traffic in the sector, as described in the EEC's operational concept document (EEC, 2002). When the controllers managed traffic in the UJ airspace, they operated independently of one another since they were managing the same traffic and airspace scenarios simultaneously; this could be likened to two controllers managing the same traffic in a "parallel universe" situation. The pseudo pilots were the author and another assistant trained in using the simulation program and 'piloting' the aircraft on the screen.

Each participant was linked to one pseudo pilot through a microphone and a headset. Each pseudo pilot was given the task of 'piloting' all the aircraft assigned to each controller. Traffic assignment under the supersector condition was based on the color of the aircraft. Therefore, one participant controlled the green-hued aircraft (with six levels of color saturation when color coding was turned on) and the second participant controlled the blue-hued aircraft (with five levels of color saturation when color coding was turned on). When managing the UJ airspace, the participants were responsible for managing all aircraft transitioning through the sector.

The experiment was carried out in a darkened cubicle to prevent ambient light from reflecting on the surface of the monitor. It also created lighting conditions similar to the working environment at the en route air traffic control center. Moreover, the darkened cubicle enabled the participants to better distinguish between the various color saturation

levels of the targets on the screen. The participants were seated adjacent to one another since they needed to consult with each other when managing the traffic in the supersector (see Figure 7).

Figure 7. Seating arrangement of participants during data collection. Data was collected from two participants simultaneously.

Participants were briefed for 30 min with a training document (see Appendix A) and were asked to sign a participant consent form (see Appendix B). After the briefing, they were given two 45 min training sessions on each of the two types of airspace, with the color-saturation coding of aircraft altitude turned on. During this period, the participants familiarized themselves with the keyboard commands for amending datablocks, making altitude and speed changes, the screen and airspace displays, and the quality of sound on the headsets. The participants first engaged in a 45 min practice scenario of the supersector airspace with color-saturation coding of the aircraft altitude turned on. They were then given a 10 min break and then began the 45 min practice

scenario in the UJ airspace, again with the color-saturation coding of the aircraft altitude turned on. The color-coding condition was turned on during the practice sessions to give the participants time to learn the levels of color saturation assigned to each altitude. After the two practice sessions, the participants had a 15 min break before data collection began.

Four traffic scenarios were used in the experiment. Each scenario differed based on the two types of airspace and the two levels of traffic intensity. Over the course of the data collection period, the participants were given each scenario twice, once with and once without the color-saturation coding of altitude. However, they were not told that they were to manage the same traffic scenario twice under each color-coding condition. The scenarios were run in random order so as to avoid order effects.

Each data collection scenario ran for 35 min with a 5 min break between each scenario. The participants engaged in three scenarios before given a 35 min lunch break. They then returned and participated in another two scenarios before given a 10 min break. Finally, they managed traffic in another three scenarios. After the data collection, the participants sat down in a debriefing session to discuss their experience with the color-saturation coding of aircraft altitude, working in the supersector, the workload, and their overall experience with the traffic samples. Since the experiment was a $2 \times 2 \times 2$ within-subjects design, all participants managed traffic under all eight conditions. In total, the briefing and data collection, with the rest breaks, took 9 hr to complete.

Dependent Variables

The participants' performances were measured with the following variables:

- Subjective workload ratings
- Time before a conflict that an evasive action occurred
- Number of aircraft within 5 min of a potential conflict
- Number of aircraft that actually conflicted
- Number of actions taken in each scenario

To gauge participants' perception of workload, a small box appeared on the top right hand corner of the screen every 3 min during each scenario. In the box was the flashing text "ENTER A SCORE HERE" and participants were given 20 s to respond to this request. A workload rating between 1 (*low*) and 9 (*high*) could be entered with the F1-F9 keys on the keyboard. A rating of 10 could not be manually entered, since this implied the participants were at their maximum working capacity. If they really were working at their maximum level, then they would not have the time to enter a score in the box. Therefore, if there was no response within 20 seconds, the computer assumed the participants were too busy to enter a score and generated a default score of 10.

Since it is extremely difficult, if not subjective, to decide when to start counting a reaction time to a potential conflict, it was decided that a reverse measurement would be taken. This measurement would time how long the participant had until conflict, and was defined as an incursion of legal separation. Every 7 s, the computer scanned for aircraft pairs that would conflict should neither aircraft change vector. It then counted down until either the potential conflict was averted, or a conflict occurred. The data therefore recorded how many conflicts occurred as how long a controller took to prevent a conflict.

This was used as a basis for comparing performance under different types of airspace, traffic and color-coding conditions.

The number of actions taken by each controller was recorded as an objective measure of controller activity. This included the number of mouse clicks, the number of times datablocks had to be dragged, the frequency of checking aircraft routes and vectors, zooming and panning over the map as well as amending the aircraft datablocks.

Independent Variables

Participants' performances were manipulated using three independent variables, each at two levels, with a 2 x 2 x 2 fully crossed, repeated measures experimental design. Participants handled aircraft in two types of airspace operations, the supersector and the current day airspace, represented by the UJ sector. Operations in the supersector airspace required the participants to cooperate, and share one large sector and divide responsibility based on aircraft trajectory. This airspace included various traffic crossing points. Each controller handled the traffic heading in one direction. In contrast, the UJ sector was approximately one sixth of the size of the supersector and each controller managed all of the traffic traversing the sector.

Color was the second independent variable, tested under two conditions, with and without saturation coding. Color-saturation coding was used to code the different aircraft altitudes. Aircraft were designated altitudes depending on their heading, where north bound aircraft were given altitudes of 210, 230, 250, 270, 290, 330 (in 100s of feet) and south bound aircraft were assigned 220, 240, 260, 280, 310 (in 100s of feet). The assignment of flight levels for particular trajectories is a standard procedure for ATM

authorities. The northbound aircraft were displayed in a green hue and the southbound aircraft were displayed in a blue hue. As mentioned earlier, all targets in the UJ airspace were displayed in a green hue.

The final independent variable was traffic load. Two levels of traffic intensity, high and medium, were used to test for differences in performance. Under the supersector, high traffic condition, 82 flights traversed the sector while 67 flights occurred during the scenarios with a medium traffic level. In the UJ scenarios, 92 flights were created for the high traffic scenario while 82 flights were written for the medium traffic set. Despite the difference in numbers, the traffic intensities were tested by retired controllers and were regarded as being equivocal. The purpose of having two traffic intensity levels was to test if changes in performance, from either the color coding or type of airspace, were sensitive to traffic-intensity conditions. Under high traffic conditions in the supersector, approximately 70 aircraft entered the supersector over 35 min. Under the medium traffic condition, this volume was reduced to 45 flights. In the UJ airspace, high traffic consisted of approximately 40 aircraft entering the airspace over 35 min while the low traffic had approximately 25 aircraft in 35 min.

Hence, the experiment design was a fully factored $2 \times 2 \times 2$ within-subject design (see Figure 8) that required the participants to engage in all eight conditions. Each condition was represented by a 35 min scenario where performance from the last 30 minutes of each scenario was used as data. The first 5 min of each scenario was regarded as a warm up time for the participants so that they could gain an awareness of the situation.

Figure 8. A diagram of the 2 x 2 x 2 within-subject experimental design that includes two levels of traffic intensity, two types of airspace operations, tested with and without the presence or color-saturation coding.

Results

Assessment of controller performance was based on the number of conflicts and potential conflicts that occurred in each scenario. To gain further insight into possible variations in the participants' strategies or difficulties in managing traffic, we also recorded when a controller averted a conflict. This provided information about how early conflicts were being resolved and not simply how many potential conflicts resulted. Additionally, an objective analysis of controller activity was based on the number of actions (i.e., number of keystrokes, mouse clicks) each participant made. These data were compared to the controllers' subjective workload measurements. A 2 x 2 x 2 within subjects ANOVA with unweighted means was used to analyze the results.

Actual Conflicts

Legal separation is defined as a minimum of 5nm lateral separation and 1000 ft vertical separation, when at an altitude beneath 29,000 ft. Above an altitude of 29,000 ft, a 2000 feet vertical separation is required.

The number of actual conflicts and potential conflicts in each scenario were added together for the analysis since performance in the supersector, under all conditions, resulted in zero conflicts by all of the participants (although there were several potential conflicts). In contrast, performance in the UJ sector resulted in actual conflicts (see Figure 9) ranging from a mean of 0.8 to 4.6 that increased as a function of traffic intensity. Color-saturation coding had no significant effect on the number of conflicts across all conditions, $F(1, 9) = 2.561, p = .144$.

Under medium traffic, the supersector yielded significantly fewer conflicts than the UJ airspace, $F(1, 9) = 18.00, p = .002$. This was amplified under the high traffic condition where the supersector yielded significantly fewer conflicts, $F(1, 9) = 44.655, p < .001$.

Figure 9. Mean number of actual conflicts as a function of traffic intensity.

Performance in the supersector resulted in zero actual conflicts.

Potential Conflicts

A potential conflict was defined as a situation where aircraft are within five minutes (based on current trajectories) of violating legal separation standards. Overall, operations in the supersector resulted in significantly fewer potential conflicts, $F(1, 9) = 252.203, p < .001$. There was also a significant two-way interaction between airspace and traffic levels, $F(1, 9) = 102.402, p < .001$. The mean number of potential conflicts in the supersector ranged from 0.4 to 1.2, depending on the traffic level and whether color-

saturation coding was turned on. In contrast, the mean number of actual and potential conflicts when operating in the UJ sector ranged from 8 to 18 (see Figure 10). This suggests that traffic levels had a differential and significantly larger impact upon the UJ airspace than it did on supersector operations.

Figure 10. Mean number of actual and potential conflicts as a function of traffic intensity.

As expected, operations under high traffic conditions resulted in significantly more potential conflicts when compared to operations with a medium traffic density, $F(1, 9) = 234.305, p < .001$. Although the number of potential conflicts in the supersector grew from 0.4 (under medium traffic without color coding) and 0.9 (under medium traffic with color coding) to 1.2, operationally this was a small increase. However, statistically, the

difference was significant, $F(1, 9) = 5.211, p = .048$, (see Figure 11). Color-saturation coding did not significantly affect performance in either the supersector or the current day sector. Moreover, it did not affect performance based on the different traffic levels.

Figure 11. Mean number of potential conflicts in the supersector airspace under both color-coding conditions.

Conflict Resolution Performance

When aircraft were within five minutes of a potential conflict, the computer counted down until the conflict was resolved or the conflict occurred. The 5 minute ceiling on the controller intervention time was chosen because this was regarded as a situation where traffic was dangerously close. Overall, potential conflicts were resolved significantly earlier in the supersector airspace than in operations under the UJ sector, $F(1, 9) = 7.224, p = .025$, (see Figure 12). The calculation of time-to-conflict enabled us

to measure of how long before a potential conflict, a situation was resolved. Therefore, an earlier or longer time-to-conflict is more desirable when managing traffic.

Based on traffic levels, it was found that under medium traffic, the supersector provided significantly better performance, $F(1, 9) = 15.601, p = .003$. However, under high traffic, the type of airspace did not have any significant impact on conflict resolution time, $F(1, 9) = 1.588, p = .239$. This may have been due to the high standard error present in the data (see Figure 13). The conflict resolution times under the supersector and high-traffic conditions ranged from a mean of 14 s to 300 s while the range under the UJ and high traffic conditions were much smaller, from a mean of 76 s to 151 s.

Figure 12. Conflict resolution performance counted as the length of time remaining, when a conflict was averted. Performance was measured as the mean amount of time remaining before a conflict, at the time a conflict was resolved.

Color-saturation coding had neither a significant effect on the time-to-conflict under either airspace condition nor in either traffic conditions. Performance under the no-color-saturation coding condition was similar to performance when color-saturation coding was present. (See Figure 13).

Figure 13. The effect of color-saturation coding on conflict resolution performance in the supersector airspace. The error bars at each point represent the standard error.

In operations without color-saturation coding, the type of airspace had a moderate effect on performance, $F(1, 9) = 4.889, p = .054$, while the level of traffic significantly affected time-to-conflict performance, $F(1, 9) = 11.418, p = .008$. However, with color-saturation coding, the supersector produced significantly better performance, $F(1, 9) = 6.203, p = .034$, than in the UJ airspace while traffic levels affected performance at a significance level of $F(1, 9) = 5.979, p = .037$. No two- or three-way interaction was evident.

Number of Actions Taken

The number of actions that each controller took was aggregated and compared for significant differences. These data were collected as a means of objectively gauging controller activity levels. There were no main effects, but a significant three-way interaction was found, $F(1, 9) = 5.445, p = .044$, (see Figure 14).

Figure 14. Graphs display a 3-way interaction of traffic levels, types of airspace and two levels of saturation coding. Performance was measured as a function of the number of actions taken.

Comparison of data under the high traffic conditions found that participants significantly reduced their number of actions when working in the supersector, compared to working in the current day sector, $F(1, 9) = 27.995, p < .001$, (see Figure 15). Under the supersector and high traffic conditions, participants made a mean of 364.9 and 351.9 actions, with and without color coding, respectively. When operating in the UJ airspace, the participants made a mean of 445.4 and 412.9 actions, with and without color-coding

respectively. However, under medium traffic conditions, there was no significant difference in performance based on either the traffic conditions or the color-saturation coding condition. Color-saturation coding was not a significant variable.

Figure 15. Mean number of actions taken in high traffic scenarios, as a function of the type of airspace operations.

Subjective Workload Ratings

Subjective workload ratings were based on estimations made by the controllers every 3 min during each scenario. Operations in the supersector, when compared to the UJ sector, yielded a significantly lower subjective workload rating, $F(1, 9) = 19.136, p = .002$. As expected, medium traffic resulted in a significantly lower workload rating, $F(1, 9) = 27.692, p = .001$, than operations under a high traffic condition. A two-way interaction between airspace and traffic level was also found to be significant, $F(1, 9) =$

12.193, $p = .007$, (see Figure 16). This suggests that the impact of workload was different across the two airspace operating conditions.

The data varied widely between each participant and within each scenario. As a result, relationships between the subjective workload ratings and the objective performance measurements cannot be strongly inferred. The results of the subjective workload ratings can be found in Appendix C, where the workload ratings have been graphed as a function of time during each of the eight scenarios.

Figure 16. Mean subjective workload ratings across all conditions. Workload was rated on a scale of 1-10, where 1 was low and 10 was high.

Discussion

The first experimental hypotheses predicted that operations under a supersector would reduce controller mental workload, yield earlier conflict resolution, and fewer potential conflicts when compared to the current sectorized airspace. The results have strongly supported this. The absence of actual conflicts in the supersector plus the significantly fewer potential conflicts in the supersector, $F(1, 9) = 252.203, p < .001$, support the notion that the supersector may provide the benefit of absorbing the larger growth in traffic without negating the performance of the tactical controller. The significantly earlier conflict resolution performance in the supersector also buttresses this hypothesis, $F(1, 9) = 7.224, p = .025$. The data from the subjective workload ratings also support the supersector concept. Controllers working in the supersector responded with significantly lower workload ratings, $F(1, 9) = 19.136, p = .002$, when managing traffic in the supersector.

The second hypothesis, which predicted that color-saturation coding of altitude would result in earlier conflict resolution, fewer conflicts, and a reduction in controllers' mental workload was not strongly supported by the data. However, considering the three-way interaction when measuring the number of actions taken and the large standard error when comparing conflict resolution performance with and without color under the supersector condition (see Figure 13), the data suggests that color-saturation coding as a method of coding altitude, may deserve further attention.

The difference in performance between the supersector and the UJ airspace may be explained by the structure and physical characteristics of the two airspaces as well as the redistribution of responsibilities. Although investigations to explore the viability of the supersector are, for now, in a preliminary stage, the experimental results here point to the advantages of this operational concept to the tactical controller. The most consistent finding was the marked and significant improvement in performance with the supersector compared to performance with the current-day airspace. Inferred reasons for the absence of statistical differences with color-saturation coding and results will be discussed in terms of performance measurements.

Actual Conflicts

The absence of actual conflicts under the supersector operations suggests that tactical controllers may find this operational concept to be workable. One explanation for the improved performance is the division of responsibility for traffic in a larger airspace may have resulted from different demands on the participants' visual search and conflict detection performance. By dividing the traffic management responsibilities into two groups via hues, participants reported that it was easy to quickly identify "my own" aircraft. Additionally, the one-directional airways meant that only one controller managed each airway. Negotiations with the other controller was mainly at crossing points and times when a controller wanted to use an altitude assigned to traffic moving in the opposite direction.

Potential Conflicts

The larger area of the supersector may have enabled the participants to vector and reroute aircraft direct to waypoints and reduce the traffic volume in the congested areas of the airspace. The larger airspace may have enabled the controllers to have more alternatives when looking for strategies to manage traffic and provided a longer look-ahead time in which to detect and resolve conflicts. The greater look-ahead time, combined with the larger area may have enabled the participants to better strategize and prioritize tasks. The small size of the UJ sector, plus the short look-ahead and transition time, may explain why performance in the UJ sector resulted in more conflicts, and closer potential conflicts.

Interestingly, the ability to detect and resolve conflicts evolved differently as a function of airspace type and traffic intensity. We found that as traffic intensity grew, operations in the UJ airspace resulted in a two-fold increase in the number of actual conflicts and potential conflicts, rising from a mean of 8.2 under medium traffic to 18 under high traffic. However, in the supersector airspace, the mean number of potential conflicts did not fluctuate noticeably with growing traffic intensity. Although the number of potential conflicts in the supersector rose significantly with higher traffic intensity, operationally, this rise (from a mean of 0.4 and 0.9 to 1.2) was not large (see Figure 11). This demonstrated that conflict detection and resolution in the supersector airspace did not deteriorate greatly despite a significant rise in the intensity of traffic and reflects an operational concept that may absorb a significantly greater number of aircraft without the expected rise in conflict pairs.

Although color-coding did not have a statistically significant effect on performance, it resulted in weaker conflict detection performance under the supersector, medium traffic condition. However, this effect dissipated as traffic intensity grew as color-saturation coding had no significant effect on conflict resolution time under high traffic conditions. The weaker results under the medium traffic condition may support the findings of Stager and Hameluck, (1990) and Stager, Hameluck, and Jubis, (1989), whereby the number of operational errors tend to increase under low to moderate workload. The redundant color coding may have offered a false sense of security, which amplified the tendency to be less vigilant under moderate workload conditions. This possibility warrants further investigation.

Time to Conflict

The supersector airspace produced significantly earlier conflict resolution compared to the UJ airspace. However, under high traffic, the supersector did not produce significantly better performance than operations in the UJ airspace. This was likely due to the large size of the standard error under the supersector condition. Nonetheless, the means of the performances reflected a consistently earlier conflict resolution time in the supersector condition during high traffic.

Under supersector operating conditions, performance differed between the high and medium traffic intensity levels. Conflict resolution, under high traffic intensity, occurred 20 s earlier with color-saturation coding. However, since color-saturation coding produced weaker performance under medium traffic conditions, any effect of color may have been statistically negated. Figure 13 showed a set of crossed data that hinted at an

interaction between traffic intensity levels and color coding in the supersector airspace. However, this interaction was not significant.

The lack of difference between the color conditions in the UJ airspace suggests that the small size and short look-ahead time of this airspace may have left the controller in a time-critical decision making state, to the extent that the introduction of a decision-support aid may not yield much improvement in performance. Alternatively, this may be explained by the short training time given to the participants and thereby, the lack of learning effects. With prolonged exposure to the color-saturation coding of altitude information, differences in performance may have eventually surfaced.

Subjective Workload and Performance Metrics

There were differences between the subjective and objective measurements of performances. The subjective workload measures found a two-way interaction between the type of airspace and traffic levels. Under medium traffic intensity, there were no significant differences between any variables. However, as traffic intensity grew, subjective workload in the supersector airspace increased only moderately (by approximately 0.4 points), while under the UJ airspace, subjective workload ratings escalated sharply (by approximately 1.8 points). Hence, differences between operational concepts and their effects on controllers' subjective workload did not become apparent until the traffic intensity heightened. As such, under high traffic intensity scenarios, the supersector only marginally increased controllers' subjective workload ratings. Color-saturation coding did not significantly impact controllers' workload ratings.

To objectively measure the participants' level of activity, we tallied the number of actions taken in each scenario. This tally included tasks such as amending datablocks for altitude or speed changes, moving datablocks on the screen, checking routes, turning on J-rings, turning vector displays, panning, and zooming. The results found a three-way interaction between traffic intensity, the type of airspace, and color coding.

With color-saturation coding, there was a two-way interaction between the type of airspace and traffic levels. As expected, in the UJ airspace, the transition from a medium- to high-traffic intensity resulted in a sharp rise in the number of actions taken. However, the reverse happened in the supersector airspace; there was a noticeable drop in the number of actions when traffic intensity increased. This was a surprising result, considering there were more planes in the airspace and more conflicts to resolve. Since the only difference between the two types of scenarios was the presence of color coding, then we may conclude that the controllers relied heavily on the color coding when working with a display featuring a higher density of targets and datablocks. Without the color coding, the number of actions taken in the supersector condition remained the same, irrespective of the traffic intensity (see Figure 14).

Since the area of the supersector is six to eight times greater than the UJ airspace, there were significantly more targets and datablocks to keep track of within the same sized screen display. Additionally, the controllers could not afford to zoom in closer because they risked losing the displays of aircraft on the edge of the sector. The only aid they had for detecting potential conflicts was the saturation coding. As such, it is

possible that visual search, rather than physical actions, became the new strategy for managing traffic.

In the supersector scenarios where color-saturation coding was not available, the number of actions taken under both traffic levels was similar (see Figure 14). Under medium traffic with color, the range of the data spanned from 220 to 904 actions, with a standard error of 61.78. Without color, the minimum tally was 199, the maximum tally was 799 with a standard error of 56.88.

The rise in the number of actions in the UJ airspace underscores the likely transition toward a more reactionary level of control as traffic intensity rose. The number of actions rose with an increasing number of targets since the datablocks needed to be kept uncluttered while information was being sourced and amended. Therefore, the increasing need to react to the most salient and available information was reflected by the escalation in actions taken with increasing traffic.

The first experimental hypotheses proposed that operations under a supersector will reduce controller mental workload, result in fewer potential conflicts and yield earlier conflict resolution when compared to the current, sectorized airspace. The results here strongly support this hypothesis. The supersector not only reduced the number of potential conflicts but also yielded zero actual conflicts. Hence, the hypothesis was upheld.

Color coding of Altitude

Since the participants were all experienced air traffic controllers, with an average of 15.5 years of experience, searching for altitude information from the datablocks could be

considered to have been an over-learned response. The tendency to look for altitude information from the data blocks probably entrenched despite being given 2 hr of training and practice at the beginning of experimental trials, with the color-saturation coding and being encouraged to increase their awareness of the color-saturation coding. Since the majority of the participants were active controllers, we had no intention of diverting them from their usual and trained reactions. Additionally, with the logistical constraints of the experiment and our resources, the best result we could have expected was to have the participants make a conscious effort to use this new display concept. The two 45-minute practice session with the color coding probably had a lack of effect in initially training the participants to actively use the various color-saturation levels. However, the participants reported that as the experiment carried on, they started to use the colors to ascertain the proximity of one aircraft's altitude to any nearby plane. Therefore, when two aircraft with different shades came together, the controller was confident that evasive maneuvers were not required. However, confidence in the saturation coding system waned as aircraft approached higher altitudes. This was due to difficulties in differentiating between colors bearing little saturation; the targets with higher altitudes started to blend into a near white color. One explanation may be the prolonged exposure to the high contrast color on the screen, despite the fact that the various levels of saturation coding were chosen, based on their just noticeable differences.

Debriefings with the participants indicated an interest in the color-saturation coding display, but all agreed that the years of training and on-the-job experience had biased their reactions to the altitude information. All participants found themselves immediately

looking for altitude information mainly from the datablocks. Under high traffic intensity, many reported using the colors to compare for altitude differences but mainly, the participants reported using the color display for identifying which aircraft were assigned to their control when managing traffic in the supersector.

The second hypothesis was that the color-saturation coding of altitude would result in earlier conflict resolution, fewer conflicts and a reduction in controller workload than if altitude information were only presented as alphanumeric text in a data block. Due to the factors noted above, color was found not to be a significant variable. However, based on the three-way interaction found in the number of actions taken, the color-saturation coding of altitude deserves further attention. More salient color differentiation and a participant pool which is less inclined to source altitude information from datablocks, may eventually prove this hypothesis to be correct. Alternatively, participants may be given more prolonged training with the color-saturation coding system.

Airspace

Participants generally regarded the division of responsibility in the supersector airspace with a positive attitude and the traffic conflicts presented in each scenario were viewed as being realistic. The airways provided predictability for the participants and made risky areas, such as airway crossing points, more salient than the traffic patterns in the UJ airspace. Although the traffic flow in the UJ airspace did follow a set of routes, the strikingly smaller area and short sector transition time may have prevented planning and strategic methods of avoiding conflict situations. Essentially, the participants noticed

that they engaged in two types of supervisory control when managing traffic in the different types of airspace.

Lessons for displaying colors

The various levels of saturation coding were easy to differentiate when members of an aircraft pair were at vastly different altitudes. The participants could easily decide that an aircraft pair was at different altitudes, even when their hues were different, by comparing their saturation levels. However, participants reported difficulties in differentiating aircraft that were at higher altitudes. That is, they were not confident that there was vertical separation when the colors of both targets were displayed in a near white color. Participants found that the color coding of altitude could potentially be a useful decision-making aid which increased the efficiency of their visual search and conflict detection performance. In this experiment, however, controllers found they primarily sourced altitude information from the datablocks.

Some controllers suggested that color-coded datablocks may actually have been more helpful since most of their visual search focused on comparing the information from datablocks rather than the targets. However, for display purposes, this may prove problematic since the color only displays one element of the datablock information—the altitude—and the comparison of datablock colors for possible conflicts could result in confusion and errors when comparing data blocks for information other than altitude.

Future use of saturation coding should provide more salient color differentiation, particularly for those targets with little color saturation. Alternatively, when the display

allows for fewer than five saturation levels, selection of colors should vie for a range focusing on colors with high to moderate saturation levels.

Participants

Although both the airspaces in the experiment were set over France, the participants were air traffic controllers from Oakland ARTCC in California, USA. We believe this strengthened the validity of our data.

If the participants had been French controllers, then expectations—default knowledge—may have interfered with the behavior and outcomes of the data. Since our participants had no default knowledge with regard to the air traffic over France, then the results were not biased with the participants' prior experience. French air traffic controllers have a set of expectations for traffic patterns based on the context of sector and time. Using American controllers introduced learning effects into the data. However, this avoided the possibility of confounding the data with experience and expectations for traffic patterns. Hence, the use of American controllers produced an unbiased participant pool.

Conclusion

Overall, the most noticeable differences in performance were in scenarios where the supersector was crossed with high traffic. It was these scenarios that tested the controllers' ability to manage a heightened workload with the redistribution of roles and responsibilities. The improvement in performance measured the potential of the supersector in not only reducing the workload of the controller as a means of reducing airspace complexity, but also its impact on airspace safety and efficiency. The results suggest that the supersector may enable aircraft to fly more fuel- and time-efficient routes; moreover, the design of the airspace and procedural organization may enable the controllers to be more efficient in prioritizing workloads due to the longer event horizon and larger airspace. Additionally, this information provided insight into the levels of supervisory control faced by air traffic controllers in each type of airspace. The salient differences found herein serve to magnify the urgency of solutions to manage the growth in air traffic in the near future.

Modes of Control

The different approaches for managing air traffic during this experiment may be viewed from the context of Hollnagel's model of control (Hollnagel, 1993). This is a model of supervisory control that may provide a method of interpreting the experimental data. Hollnagel defines four levels of control, where each level is based on the supervisor's ability to view the entire situation and make strategic decisions. With increasing workload, reduced availability of information, and time demands, the level of

control is reduced. As such, only two levels of control will be discussed with respect to the management of traffic in the supersector and the UJ airspace.

These are *opportunistic control* and *strategic control*. The situation where a supervisor is unable to strategize in a medium-, or long-term context, and where each action is decided by the current situation is called opportunistic control. Under opportunistic control, information and decisions are drawn from salient information rather than planned goals. Strategic control is one where the operator is able to draw information from a broad range of sources with a wide event horizon. This type of control allows the supervisor to prioritize tasks and plan actions well ahead of time.

In the UJ sector, we saw that heightened traffic intensity resulted in the marked increase in conflicts. The controllers had access to available information and there was predictability in the movements and intentions of aircraft, but the narrow event horizon and perhaps, the lack of time to formulate any long-term strategy, may lead us to interpret operations in the UJ airspace as being opportunistic control.

In contrast, operations in the supersector resulted in only a small rise in potential conflicts and zero actual conflicts. The earlier conflict resolution in the supersector and the long look-ahead time suggests the controllers likely had a medium- to long-term event horizon and adequate information to formulate a series of conflict resolution and task prioritization plans, thereby engaging in a strategic level of control.

This marked difference in the levels of control is of significant importance for formulating and assessing operational concepts. Enabling controllers to engage in strategic control has numerous advantages, some of which include the ability to delegate

tasks well ahead of time, the ability to better organize traffic patterns, workloads, responsibilities, and aircraft movements. As such, the term “efficiency”, as a goal of a new operational concept, expands in its definition to cover both the efficient maneuvering of aircraft and the efficient organization of the controllers’ workload. Based on the modes of control, the supersector seems to provide a level of strategic control for the tactical controller that is two “levels” above that of the current sector design. However, these claims have to be qualified by further investigations of the concept and the integration of the other controllers’ roles in the working of the supersector on a more holistic scale. The experiment here only tested the performance of the tactical controller working under supersector procedures.

Participants

The potential of color saturation to code aircraft altitude was tempered by the automatic tendency of experienced controllers to look to the datablock for altitude information. Future studies investigating the redundant color coding of datablock information should look to recruiting controllers with adequate ATC training but little live traffic experience. This would minimize the tendency for automatic behavior generated by years of experience in controlling traffic with the current ATC radar screen display.

Although participants suggested color coding both targets and datablocks, this should be approached with caution. The color coding of datablocks that contain information along several dimensions (speed, altitude, intent to climb or descend) presents a situation where errors in reading and comparing information is quite possible.

Additionally, the excessive use of color can be a serious distraction when engaging in time- and safety-critical visual search tasks.

It should be noted that this experiment only tested the performance of the tactical controller in the supersector concept. The ability of the supersector in reducing airspace complexity and controller workload has to be further tested on a more holistic scale. Interactions between the different controller roles, communications about weather, intent, and other types of information flow, have not been tested nor stressed in the context of supersector operations. Additionally, events such as emergencies, bad weather, and rerouting procedures were not incorporated within this experiment's traffic scenarios. As such, the case for the supersector operations can only be made under the context of normal operating conditions and based only on the perceived benefits to the tactical controller. To assess the viability of the supersector as an operational concept, future research needs to integrate the various layers of ATC with aircraft operations. An operational assessment should then test the coordination of this airspace operational concept with unexpected events, varying weather conditions and different traffic complexities.

This experiment found the supersector to be advantageous for the tactical controller in that the division of responsibility and the color coding of aircraft delegation improved conflict detection and resolution performance. Additionally, tactical controllers experienced a significant decrease in subjective workload. However, to better understand the interactions of the operators in their various roles, scenarios need to be developed so that the potential supersector—as an entire operating concept—may be examined more

thoroughly. It is possible that a controller in another role, perhaps the approach controller who lines up the aircraft before entry into the airspace, may experience a negative effect in performance due to some unforeseen factor.

The time and logistical constraints resulted in the recruitment of a small number of participants for this study. Further studies that incorporate a greater number of participants may achieve more statistically powerful data with which to assess the viability of the workload, the supersector and in particular, the potential of color-saturation as a means of coding altitude information.

Additionally, future research should also examine the color-coding of altitude information with a greater degree of differentiation between color-saturation levels. Debriefings with the participants found that several participants had difficulty in identifying differences in color saturation when two targets had little saturation, despite that the colors selected for the experiment were chosen for their just-noticeable-difference. This resulted in situations where participants repeatedly checked on two aircraft in near proximity only to find that they were indeed adequately separated. A more salient difference in color saturation between all altitude levels would have prevented this response.

This experiment found strong evidence to suggest that the procedures and design of the supersector have the potential to reduce airspace complexity and controller workload. The striking difference in performance between the supersector and the current sector airspace provides a strong argument for further investigation into this operational concept. Although the color-saturation coding of altitude did not result in a significant

difference in performance, the steps discussed previously could prove this to be otherwise. Further investigations into the viability of the supersector should examine the integration of decision-aiding displays to support the performance of the operators in their various air traffic management roles.

References

- Andre, A. D., Wickens, C. D., Moorman, L., & Boschelli, M. M. (1991). Display Formatting Techniques for Improving Situation Awareness in the Aircraft Cockpit. *International Journal of Aviation Psychology*, 1(3), 205 - 218.
- Auley, G. M., Lenzi, J. M., & Lepadatu, C. (2001). *Presentation to Skyguide: Final Results Capacity Analysis (Phase I) Geneva ACC - Geneva - Zurich ACC*.
Bretigny-sur-Orge, France: EUROCONTROL Experimental Centre.
- Baddeley, A. D. (1995). Working Memory or Working Attention. In A. Baddeley & L. Weiskrantz (Eds.), *Attention: Selection, Awareness and Control*. Oxford: Oxford University Press.
- Bennett, K. B., Toms, M. L., & Woods, D. D. (1993). Emergent Features and Graphical Elements: Designing More Effective Configural Displays. *Human Factors*, 35(1), 71 - 97.
- Bennett, K. B., & Flach, J. M. (1992). Graphical Displays: Implications for divided attention, focused attention, and problem solving. *Human Factors*, 34, 513 – 533.

Beringer, D. B., Allen, R., Kozak, K. A., & Young, G. (1993). *Responses of Pilots and non-pilots to Color-coded Altitude Information in a Cockpit Display of Traffic Information*. Paper presented at the Human Factors and Ergonomics Society 37th Annual Meeting, Santa Monica.

Boeing. (2002). *Air Traffic Management Working Together Team Phase 1: System Performance Requirements Document* (D780-10010-1). Retrived February 12th, 2002, from <http://www.boeing.com/atm/pdf/sprd.pdf>

Bresolle, M. C., Benhacene, R., Boudes, N., & Parise, R. (2000, June 13th-16th, 2000). *Advanced Decision Aids for Air Traffic Controllers: Understanding Different Working Methods from a Cognitive Point of View*. Paper presented at the 3rd USA/Europe Air Traffic Management R&D Seminar, Napoli, Italy.

Christ, R. E. (1975). Review and Analysis of Color Coding Research for Visual Displays. *Human Factors*, 17(6), 542 - 570.

Corker, K., Gore, B., Fleming, K., & Lane, J. (2000). *Free Flight and the Context of Control: Experiments and modeling to determine the impact of distributed air-*

ground air traffic management on safety and procedures. Paper presented at the

3rd USA-Europe Air Traffic Management R&D Seminar, Napoli, Italy.

Delahaye, D. L. C., & Puechmorel, S. E. (2000, June 13th - 16th). *Air Traffic*

Complexity: Towards Intrinsic Metrics. Paper presented at the 3rd USA/Europe

Air Traffic Management R&D Seminar, Napoli, Italy.

Derefeldt, G., Skinnars, Ö., Alfredson, J., Eriksson, L., Andersson, P., Westlund, J.,

Berggrund, U., Holmberg, J., & Santesson, R. (1999). Improvement of Tactical

Situation Awareness With Color-Coded Horizontal-Situation Displays in Combat

Aircraft. *Displays*, 20, 171-184.

Duong, V., Hoffman, E., & Nicolaon, J. P. (1997, June 17-20). *Initial Results of*

Investigation into Autonomous Aircraft Concept (FREER-1). Paper presented at

the 1st USA/Europe Air Traffic Management R&D Seminar, Saclay, France.

EEC. (2002). *Supersector (SectorLess-1) Operational Concept Document* (Document

0.1). Bretigny-sur-Orge: EUROCONTROL Innovative Research.

Endsley, M. R. (1997). *Situation Awareness, Automation and Free Flight*. Paper

presented at the 1st USA-Europe Air Traffic Management R&D Seminar, Saclay,
France.

Endsley, M. R., Mogford, R. H., Allendoefer, K. R., Snyder, M. D., & Stein, E. (1997).

*Effect of free flight conditions on controller performance, workload and situation
awareness* (DOT/FAA/CT-TN97/12). Atlantic City, N.J.: Federal Aviation
Administration.

EUROCONTROL. (2002, 12 March, 2002). *Air Traffic Statistics & Forecasts*

*(STATFOR): Medium-Term Forecasts of Annual Number of IFR Flights (2002-
2009). Volume 2*. STATFOR. Retrieved, from the World Wide Web:

[http://www.eurocontrol.int/statfor/forecasts/MTFReport%20Feb02%20vol2%20v
10.pdf](http://www.eurocontrol.int/statfor/forecasts/MTFReport%20Feb02%20vol2%20v10.pdf)

Galster, S. M., Duley, J. A., Masalonis, A. J., & Parasuraman, R. (2001). Air Traffic

Controller Performance and Workload Under Mature Free Flight: Conflict

Detection and Resolution of Aircraft Self-Separation. *International Journal of Aviation Psychology*, 11(1), 71-93.

Goldstein, E. B. (1999). *Sensation and Perception* (5th ed.). Pacific Grove: Brooks/Cole Publishing Company.

Hess, S. M., Detweiler, M. C., & Ellis, R. D. (1999). The Utility of Display Space on Keeping Track of Rapidly Changing Information. *Human Factors*, 41, 257-281.

Hilburn, B. G., Bakker, M. W. P., Pekala, W. D., & Parasuraman, R. (1997). *The Effect of Free Flight on Controller Mental Workload, Monitoring and System Performance*. Amsterdam, The Netherlands: Confederation of European Aerorspace Societies.

Histon, J., & Hansman, R. J. (2001, April 5-6). *Introducing Structural Considerations into ATC Complexity Metrics*. Paper presented at the Joint University Program (JUP) for Air Transportation Research Meeting, Boston, MA.

Hoekstra, J. M., Ruigrok, R. C. J., & Van Gent, R. N. H. W. (2000, June 13-16). *Free*

Flight in a Crowded Airspace? Paper presented at the 3rd USA/Europe ATM

R&D Seminar, Napoli, Italy.

Hollnagel, E. (1993). *Human Reliability Analysis: Context and Control*. London:

Academic Press.

Johnson, W., Battiste, V., & Belcher, S. (1998). *Analysis of CDTI Requirements for Free*

Flight Self-Separation. Moffett Field, CA: NASA Ames.

Johnston, J. C., Horlitz, K. L., & Edminston, R. W. (1993). *Improving Situation*

Awareness Displays for Air Traffic Controllers. Paper presented at the

Proceedings of the Seventh International Symposium on Aviation Psychology,

Columbus, Ohio.

Leroux, M. (1998, 1-4th December, 1998). *In Depth Evaluation of ERATO Tools:*

Towards an Operational Prototype. Paper presented at the 2nd USA/Europe Air

Traffic Management R&D Seminar, Orlando, FL.

Luder, C. (1984). Redundant Color Coding on Airborne CRT Displays. *Human Factors*, 26(1), 19 - 32.

Macbeth, G. (2001). *Munsell Conversion Program*. Retrieved 10/10/01, 2001, from the World Wide Web: <http://www.munsell.com>

Manning, C., Mills, S. H., Fox, C., Pfliegerer, E., & Moglika, H. (2000a, June 13 - 16).

Investigating the validity of Performance and Objective Workload Evaluation

(POWER). Paper presented at the 3rd USA/Europe Air Traffic Management R&D

Seminar, Napoli, Italy.

Manning, C., Mills, S. H., Fox, C., Pfliegerer, E., & Moglika, H. (2000b, June 13 -16).

Investigating the Validity of Performance and Objective Workload Evaluation

Research (POWER). Retrieved July 23, 2001, from the World Wide Web: Ftp at:

<http://atmseminar2000.eurocontrol.fr/acceptedpapers/pdf/paper66.pdf>

Manning, C., Mills, S. H., Fox, C., Pfliegerer, E., & Moglika, H. (2001, December 3-7).

The Relationship Between Air Traffic Control Communication Events and

Measure of Controller Taskload and Workload. Paper presented at the 4th

USA/Europe Air Traffic Management R&D Seminar, Santa Fe, NM.

Mills, S. H. (1998). *The Combination of Flight Count and Control Time as New Metric of*

Air Traffic Control Activity (Final Report DOT/FAA/AM-98/15). Oklahoma City,

OK: Federal Aviation Administration.

Morrow-Magyarits, S., & Kopardekar, P. (2000). *Dynamic Density: The Measure of Air*

Traffic Control Sector Complexity: Phase I (Draft Report for the FAA Report No.

ASD-130, ATP-400): FAA and NASA Ames Research Center.

Mulligan, J. (2001, January 21 - 26). *Scan-path Analysis of Air Traffic Control Displays.*

Paper presented at the Twenty-Sixth Annual Interdisciplinary Conference,

Jackson Hole, Wyoming.

Nagy, A. L., & Sanchez, R. R. (1992). Chromaticity and Luminance as Coding

Dimensions in Visual Search. *Human Factors*, 34(5), 601 - 614.

Narborough - Hall, C. S. (1985). Recommendations for Applying Colour Coding to Air

Traffic Control Displays. *Displays*, July, 131 - 137.

O'Hare, D., & Roscoe, S. (1990). *Flightdeck Performance: The Human Factor*. Ames,

IA.: Iowa State University Press.

Pawlak, W. S., Brinton, C. R., Crouch, K., & Lancaster, K. M. (1996). *A Framework for*

the Evaluation of Air Traffic Control Complexity. Paper presented at the AIAA

Guidance Navigation and Control Conference, San Diego, CA.

Raskin, J. (2000). *The Humane Interface* (1st ed.). Reading, MA.: Addison - Wesley.

Reising, J. M., & Aretz, A. J. (1989). Color Computer Graphics in Military Cockpits. In

H. J. Durrett (Ed.), *Color and the Computer* (pp. 151 - 169). Boston, MA.:

Academic Press, Inc.

Remington, R., & Johnston, J. C. (1995). *Simulation Experiments Investigatin Controller*

Workload under Free Flight Conditions. Retrieved 1st February, 2001, from the

World Wide Web: <http://visions.arc.nasa.gov/HST/Brief/Attention.C>.

/Controller.W.html

Rodgers, M., Mogford, R., & Mogford, L. (1998). *The Relationship of Sector*

Characteristics to Operational Errors (Final Report DOT/FAA/AM-98/14).

Oklahoma City: William J. Hughes Technical Center.

Scanlon, C. H. (1994, April 13 - 15). *Cockpit Graphical Weather Information Shown to*

Engage Efficiency, Safety and Situation Awareness. Paper presented at the 39th

Corporate Aviation Safety Seminar (CASS), St. Louis, MO.

Scholten, C. G. H. (1969). *Elements for a New Departure in Air Traffic Control*.

Unpublished doctoral dissertation., University of Eindhoven, Eindhoven, The

Netherlands.

Silverstein, L. D. (1987). Human Factors for Color Display Systems: Concepts, Methods,

and Research. In H. J. Durrett (Ed.), *Color and the Computer* (pp. 27 - 61).

Boston, MA.: Academic Press Inc.

Stager, P. & Hameluck, D. (1990). Ergonomics in Air Traffic Control. *Ergonomics*, 33,

493-499.

Stager, P., Hameluck, D., & Jubis, R. (1989). *Underlying factors in Air Traffic Control*

Incidents. Paper presented at the 33rd Annual Meeting of the Human Factors

Society, Santa Monica, CA.

Stager, P., Ho, G.W., Garbutt, J.M. (2001). *An On-Line Measure of Controller Workload*.

Paper presented at the 11th Symposium on Aviation Psychology, Columbus,

Ohio.

Stokes, A., Wickens, C., & Kite, K. (1990). *Display Technology*. Warrendale, PA:

Society of Automotive Engineers Inc.

Van Gent, R. N. H. W., Hoekstra, J. M., & Ruigrok, R. C. J. (1998, May 27 -29). *Free*

Flight with Airborne Separation Assurance. Paper presented at the HCI Aero,

Montreal, Canada.

Wickens, C. D., & Holland, J. (2000). *Engineering Psychology and Human Performance*

(3rd ed.). New York: Harper Collins.

Wickens, C. D., Sandry, D. L., & Vidulich, M. (1983). Compatibility and Resource

Competition Between Modalities of Input, Output and Central Processing: Testing

in a Model of Complex Task Performance. *Human Factors*, 25(2), 241 - 248.

Williams, K. W. (2000). *Comparing Text and Graphics in Navigation Display Design*

(Report No. DOT/FAA/AM-00/8). Oklahoma City, OK.: Federal Aviation

Administraion Civil Aeromedical Institute.

Appendix A

Training Document Used for Briefing Participants

The Effects of Color Saturation - Coding of Aircraft Altitude on the Performance of Air Traffic Controllers.

Human Automation Integration Laboratory (HAIL)

San Jose State University

Point of Contact:

Liza Tam: E-mail: ltam1@email.sjsu.edu

Telephone: 408 924 3808

Dr. Kevin . M. Corker: E-mail: kcorker@email.sjsu.edu

Telephone: 408 924 3988

INTRODUCTION

Thank you for participating in this research experiment. Your efforts are appreciated, and the results of which we hope, will be useful in the development of future air traffic control operating concepts and information displays that will better assist controllers in their various roles.

Solutions to meet the unrelenting demand for airspace have been the subject of much research. This experiment tests the potential of a newly proposed airspace operating concept and a proposed method of reducing the mental workload of the controllers managing this airspace.

PURPOSE

Due to the growing volume of air traffic, sectors in Europe are commonly 'resectored' (or subdivided) so that the workload can be offloaded from a single controller to two or more controllers. What results is an area of very small sectors. The constant increase in the number of sectors eventually becomes unmanageable due to the growing frequency of handoffs and voice communications between pilots and controllers. To address the issues presented by constant resectorization, EUROCONTROL Experimental Center (EEC) has proposed an airspace whereby 6-8 current sectors are integrated into one greater sector. This is referred to as a 'Supersector.' The purpose of this study is to evaluate the viability of such a concept, and examine methods whereby the delivery of information to controllers can aid in reducing their mental workload.

This experiment examines the potential of using color saturation to represent aircraft altitude on a radar screen. In this experiment, improvements in performance are measured by

reduction in the number of conflicts and reaction time of controllers operating under varying workload level.

SIMULATION

The simulation examines mental workload and performance under 3 sets of conditions:

- Type of Airspace
- Presence/Absence of color saturation-coding of aircraft altitude
- Level of Traffic Density

Airspace

The experiment will present the controller with two airspace designs. One will be the proposed supersector. The boundaries will be displayed on the screen with a thin grey border. The operational routes in the supersector are marked in dark grey. The second airspace will be a replication of the sector in south-east France and is referred to as 'Uniform Juliet.' When managing Uniform Juliet, you will be presented with aircraft operating in a pattern similar to what can be currently found in today's operations. In other words, it is a simulation of a current, classical sector.

The simulation investigates performance under the standard classical sector conditions, compared to the supersector conditions. Under each of these operating conditions, controllers will also be asked to control traffic with high versus medium traffic density and altitude color saturations coding versus the absence of altitude color saturation coding.

Your sector can be seen on the screen, defined by a bordered area. The altitude levels you will be controlling are FL190 – FL340.

When controlling the **classical sector**, you will be responsible for **EVERY** aircraft passing through that sector, as is the current practice.

When controlling traffic in the **supersector**, your responsibility is **shared** between you and your adjacent controller.

Saturation Coding of Altitude

Color saturation refers to the intensity of a color. A color with little or no white is highly saturated while a color with a lot of white has low saturation. For example:

High

Low

The Supersector operates between flight levels: 190 – 340 (see diagram 1 below).

Between 190 – 290, aircraft are separated by a minimum of 1000 feet vertically and 5 nm laterally. Above 29 000 feet, aircraft are separated by 2000ft vertically while maintaining the 5nm lateral separation.

In this experiment, under supersector operations, aircraft heading north (green aircraft) use odd level altitudes and above FL290, use FL330. For aircraft heading south (blue aircraft), even level altitudes are used. When aircraft are above FL290, FL310 is used.

Diagram 1. The various color - saturation codes are shown with the aircraft altitudes

In our experiment, the Supersector traffic is divided between north bound and southbound traffic. This is the basis of traffic management roles. The north and east bound traffic are coded as green while south and west bound traffic are coded as blue. One controller manages all the blue traffic while the other controller manages all the green traffic. Each route has traffic that travels in only one direction.

Waypoints & Airports

Airports are marked by small yellow triangles while waypoints are displayed as small white triangles. Rolling the cursor over either airports or waypoints will display its name.

Data Blocks

Each aircraft will appear on the screen from outside the managed sector. The call sign will be flashing. To accept an aircraft into your sector, the controller needs to roll the cursor over the target (the aircraft) and click the left mouse button.

Aircraft will appear like the one below:

Roles

In the supersector, there are two controllers managing the traffic. Each controller manages the traffic moving in one direction.

- One controller manages the blue aircraft (south and west bound).
- The second controller manages the green aircraft (north and east bound).

Each route has traffic moving in one direction, although they may split into different routes as they travel through the supersector. There are several crossing points in the sector and therefore, you may find it necessary to consult each other.

These routes are only samples. The actual routes are shown on the screen.

When managing traffic in the classical sector (Uniform Juliet), you will both be controlling the same traffic simultaneously. As such, you will be linked with the pseudo pilot who will be 'piloting' all the flights on the screen. In this case, each controller works alone since each simulation runs separately and in parallel.

PROCEDURE

During each simulation, you will be required to do two tasks:

- Manage all the traffic entering your sector
- Enter a workload rating every 3 minutes

Each scenario will take 35 minutes. The first 5 minutes are for 'warming up' and observing the state of the traffic. From the 6th minute to the 30th minute, the data collection will occur. At the 31st minute, the scenario will stop and a new scenario will then be loaded for another session. Below is our estimated timetable.

8am -10:15am: Briefing

10:15 – 10:30: Break

10:30 – 11:05: Scenario 1

11:10 – 11:45: Scenario 2

11:50 – 12:25: Scenario 3

12:25 – 1:00: Lunch

1:05 – 1:40: Scenario 4

1:45 – 2:20: Scenario 5

2:20 – 2:30: Break

2:35 – 3:10: Scenario 6

3:45 – 4:20: Scenario 7

4:25 – 4:40: Scenario 8

4:45 – 5:15: Debriefing

Workload Recording

Every 3 minutes, you will see a box appear in the top right hand corner of the screen, titled 'Workload'. The purpose of this is to collect data on what you, the controller, estimate to be a measure of your mental workload at that point in time.

Measurement is on a scale of 1- 10 where 1= virtually no workload and 10 = extremely high workload. You enter a value by pressing one of the F keys. So, to enter a workload rating of 4 out of 10, you would press the F4 key.

The program allows you to enter a value between 1-9 since a 10 implies that you are at your maximum threshold. If you were operating at your maximum threshold, you would not have the time to respond. Therefore, when you don't respond to the box within 20 seconds, the box will disappear and a score of 10 will be recorded.

Managing Traffic

Between FL190 – FL290, traffic have to maintain vertical separation of 1000 feet. Above FL290, all aircraft have to maintain a separation of 2000 feet. At all times, aircraft must maintain a lateral separation of 5nm.

COMMANDS

The following keyboard commands are available for controlling traffic. As with real life, verbal instructions are given to the pseudo pilot and then the commands are entered via the keyboard. The examples below use the adjacent datablock:

AFR728
260
921 310

Altitude change

Enter 'A' immediately followed by required flight level. Leave a space and then type either the flight CID, callsign or click on the target.

So, if we were to instruct the above aircraft to descend to FL240, the following commands could be used:

A240 AFR728 or A240 921 or A240(SPACE)(click on the target)(ENTER)

During descent, the second line will display: 253↓240

The datablock will have the second line displayed as: 240 once the aircraft has leveled off at FL240. The first number on the second line always shows the actual altitude.

Clearing aircraft to assigned altitudes

If the above aircraft originally requests 220, but the controller can only give it 240, the controller will enter the commands

A220 AFR728

C240 AFR728

Once the aircraft have reached 240, then the second line of the datablock will display 240T220. This means the aircraft has been assigned a temporary altitude of 240 but requests to return to 220 later. If you need to make a permanent change to the aircraft altitude, then the

pseudopilot will enter a command to clear it to the new altitude, and the datablock's second line will then simply display 220. The air traffic controller must then give the aircraft total clearance by entering:

CAFR728

Some aircraft will be taking off from airports inside your sector and they will have altitudes to which they have already been cleared. Alternatively, you may wish to clear aircraft that are flying at temporary altitudes to fly at their previously assigned altitudes. If you wish to clear an aircraft to its preassigned altitude, you enter the "C" command followed by the call sign, CID or clicking on the target.

e.g. CAFR728 or C921 or C(click on the target)(ENTER)

Speed change

Enter 'S' immediately followed by required speed in knots. Leave a space and then type either the flight CID, callsign or click on the target.

So, if we were to instruct the above aircraft to speed up to 340 knots, the following commands could be used:

S340 AFR728 or S340 921 or S340(SPACE)(click on the target)(ENTER)

Altitude & Speed Change in one command line

Both altitude and speed instructions can be entered with one command. So, if you require AFR728 to climb to FL280 and speed up to 340 knots, you would enter one of the following:

A280 S340 AFR728 or A280 S340 921 or A280 S340(SPACE)(click on target)(ENTER)

Heading Change

Enter 'H' immediately followed by required degrees. Leave a space and then type either the flight CID, callsign or click on the target.

So, if we were to instruct the above aircraft to change vector by 2 degrees the following commands could be used:

H002 AFR728 *or* H002 921 *or* H002(SPACE)(click on the target)(ENTER)

Hand offs

Enter 'H' immediately followed by the flight CID or call sign.

So, if we were to hand off AFR728, the following commands could be used:

HAFR728 *or* H921 *or* H(click on the target)(ENTER)

An aircraft that has been handed off will have an "O" displayed at the end of the third line.

This remains displayed for 5 minutes.

E.g.	AFR728
	280
	921 310 O

Acceptance of aircraft

Aircraft approaching your sector will have a flashing call sign. To accept these aircraft, click on the target. The call sign will stop flashing and you have control of that flight.

TOOLS

Zoom

You can zoom in and out by using the – and + keys located in the top right hand corner of the keyboard.

Pan

Panning is done via the arrow keys (← ↑↓→) in the bottom right hand of the keyboard.

Vector

The location of aircraft 1,2,4, & 8 minutes ahead can be found through the ‘v’ command.

Hitting ‘v’ repeatedly will result in a vector line extending from each aircraft.

Hitting it 5 times (1,2,4,8, nil) will hide the vector line from display.

History

Hitting the “H” key and enter will display lines behind aircraft, showing you a history of their movements. Like the vector ‘V’ key, repeatedly pressing ‘H’ will give you an extended history. But unlike the vector lines, history lines are over 1,2,3,4 & 5 last updates.

J-ring

Rolling the cursor over any target and pressing the 'J' key will result in the display of a J-ring for that particular aircraft. To turn it off, again, roll the cursor of the target and press the J key.

Route

To see the route of any particular aircraft, roll the cursor over the target and press the 'R' key. The route will display for 3 seconds and then disappear. It can be toggled if the controller only wants to view the route for half a second by pressing the 'R' key again.

Turning off Data Blocks

You can turn off the datablocks of any aircraft NOT in the sector. This is designed to reduce the clutter on the display. You cannot turn off the datablocks of aircraft in the sector.

To turn off the datablock from each aircraft, roll the cursor over the target and press the 'S' key. The target remains visible but the datablock disappears. You can turn it on again by repeating this.

DATA COLLECTION

We are collecting data about your performance in each scenario to assess the operational concept and the type of computer display that best assists controllers under varying workloads.

Each participant will remain anonymous and the data will only be used to assess the potential of using color saturation to display aircraft altitude.

Each participant's performance will be measured based on:

1. Number of conflicts
2. Time when action was taken to prevent a conflict
3. Number of operational errors (e.g. placing aircraft on conflict headings when initially, they were not headed for a conflict.)
4. Subjective workload rating

Appendix B

Participant Informed Consent Form

INFORMED CONSENT

INDIVIDUAL'S CONSENT TO VOLUNTARY PARTICIPATE IN RESEARCH ON AIR TRAFFIC CONTROLLER INTERACTION WITH DECISION AIDING AND ALERTING SYSTEMS.

The research into air traffic controller interaction with decision aiding and alerting systems is sponsored by the EUROCONTROL Experimental Center (EEC).

Risks & Benefits

Through this effort, EUROCONTROL expects to gain insight into decision aiding and to develop a set of human performance guidelines to support future ATM system design and methods of reducing air traffic controllers' mental workload.

Fatigue is a possible risk of this experiment. However, regular breaks have been scheduled throughout the data collection period.

Terms of Compensation

Participants will be paid \$400 for fully participating in the experiment and \$0.33 for each mile traveled to and from SJSU plus parking expenses. Should the participant wish to withdraw from the study during the experiment, their relationship with EEC and SJSU will not be affected and they will not lose any services. However, the participant shall only be compensated for their mileage and parking expenses.

Research

The participants will be asked to partake in an airspace simulation on the computers at San Jose State University. The research will be conducted by the San Jose State University under the direction of the principal investigator Dr. Kevin M. Corker & Liza Tam. The research shall consist of set of observational studies at San Jose State University's Human Automation Integration Laboratory (HAIL).

I _____, understand that this study entitled "Effects of Color Saturation-coded Altitude Display on Performance of Air Traffic Controllers" is sponsored by EUROCONTROL and is being directed by SJSU and Dr. Kevin M. Corker.

I have been recruited to volunteer as a subject in the project named above. The purpose of this effort is to determine the process of decision aiding in air traffic management. The observation period will last approximately 10 hrs during which time question for clarification may be asked on a not-to-interfere basis.

The approach being will be data collection on workload and conflict detection undertaken as air traffic controllers operate newly installed aircraft display systems, specifically the display of color-coded aircraft over the European airspace.

I understand that my participation in this study is completely voluntary. I am participating because I want to. Liza Tam & Dr. Kevin Corker has adequately answered all questions I have about this study, my participation, and the procedures involved. I understand that Liza Tam & Dr. Corker 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 participation, I will be informed.

I understand that records of this study will be kept confidential, and that I will not be identifiable by name or description in any reports or publications about this study. All records and interviews shall be classified and confidential. All data collected from this study shall be kept in a locked cabinet, accessible only by Liza Tam and Kevin Corker.

I have read this consent document and freely consent to participate in this study under the conditions described. I have received a copy of this consent form. I understand that I may withdraw from this study at any time without penalty.

If I have questions about this study or need further information, I will contact Kevin M. Corker, the Principal Investigator, at (408) 924 3988 or via email at kcorker@email.sjsu.edu

If I have any comments or complaints about the conduct of this experiment, I may contact:

Dr Nabil Ibrahim
Associate Vice President
Graduate Studies & Research
Phone: 408-924-2427
Email: nibrahim@sjsu.edu
Fax: 408-924-2477

Research Subject _____

Date _____

Investigator _____

Date _____

Witness _____

Date _____

Appendix C

Graphs Illustrating Subjective Workload Ratings as a Function of Time in each Scenario

Figure C1. Subjective workload rating in the supersector airspace, with color -saturation coding and under high-traffic conditions.

Figure C2. Subjective workload rating in the supersector airspace, with color-saturation coding and under medium-traffic conditions.

Figure C3. Subjective workload rating in the supersector airspace, without color-saturation coding and under high-traffic conditions.

Figure C4. Subjective workload rating in the supersector airspace, without color-saturation coding and under medium-traffic conditions.

Figure C5. Subjective workload rating in the UJ airspace, with color-saturation coding and under high-traffic conditions.

Figure C6. Subjective workload rating in the UJ airspace, with color-saturation coding and under medium-traffic conditions.

Figure C7. Subjective workload rating in the UJ sector, without color-saturation coding and under high-traffic conditions.

Figure C8. Subjective workload rating in the UJ sector, without color-saturation coding and under medium-traffic conditions.