Design
of
Collaborative Systems
for
Modern Cockpits

by

Paul McKay

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Systems Design Engineering

Waterloo, Ontario, Canada, 2009

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

One of the most significant developments in cockpit technology over the past several years is the emergence of a new cockpit architecture that uses cursor control devices and keyboards for interaction with individual and shared displays. This architecture has allowed for the design of cockpit interfaces with many advantages compared to traditional designs. However, there are a number of challenges associated with these new cockpits that should be addressed so that pilots will be able to take full advantage of the performance improvements available from the new designs.

This thesis describes three of the major challenges associated with the new architecture: supporting awareness, assisting interruption recovery, and mitigating interaction conflicts. It also describes the analysis process used to identify these challenges and proposes an interface augmentation with the potential to address them. The proposed design uses visualizations of the history of operator interactions with the interface to provide cues to the pilots about where each of them has been (and is currently) interacting. This interaction data includes both visual (sourced from a gaze tracking system) and input (from the keyboard or cursor control device) information, and was communicated on the interface using dynamic borders around the relevant areas of the interface. This augmentation aimed to address the three identified challenges by providing pilots with: improved awareness of each other's actions, visual cues of where they were working prior to an interruption and what has changed since, and clear indications of where each is working to allow them to avoid conflicts.

A two-stage evaluation process was used to determine the utility of the interface concept in a cockpit context by developing a non-interactive video prototype and showing it to pilots. The results of the evaluation indicated that the design has sufficient potential to warrant further study, as evaluation in higher fidelity environments would help provide further evidence of its potential utility for live cockpit operations. Therefore, future work should include the development and evaluation of a fully interactive prototype for live cockpit operations, as well as further examination of the design concept's potential for use as a training tool.

Acknowledgements

This thesis would not have been possible without the assistance and hard work of a number of people at both the University of Waterloo and at Esterline|CMC Electronics Inc. First and foremost, the guidance, support, and supervision of Dr. Stacey Scott helped me develop as a graduate student and critical thinker and ensured that my research stayed on track and within scope. I am very grateful to have had the opportunity to work with her over the past two years. Also, Dr. Jonathan Histon (UW) and Mr. Gerard Torenvliet (CMC) were instrumental in helping to plan the course of this research and providing suggestions that helped to define the methodology. On the technical side, the evaluation process would not have been possible without the contributions of Mr. Michael Perlin (CMC) during the data collection phase and Mr. James Wallace (UW) during the experimental study.

I would also like to acknowledge the support of my other colleagues at both UW and CMC. Their support and friendship was an enormous help in keeping me interested and motivated while working on this research.

Finally, to my family: Mom and Dad, your love and guidance has made me what I am today, and I can't thank you enough for always supporting me in whatever I choose to do. Ann, you're always there when I need someone to talk to, and I couldn't ask for a better sister. Steph, you're amazing and I can't believe how lucky I am to have you.

Table of Contents

Appendices

List of Figures

List of Tables

List of Acronyms

ASL – Applied Science Laboratories.

ATIS – Automatic Terminal Information System.

CSCW – Computer Supported Cooperative Work.

CYKF – Region of Waterloo International Airport.

CYQG – Windsor International Airport.

FO – First Officer.

HACT – Human Automation Collaboration Taxonomy.

HCI – Human Computer Interaction.

IFR – Instrument Flight Rules.

ILS – Instrument Landing System.

KORD – Chicago O'Hare International Airport.

LCD – Liquid Crystal Display.

NASA – National Aeronautics and Space Administration.

NTSB – National Transportation Safety Board.

OSD – Operational Sequence Diagram.

PFD – Primary Flight Display.

PIC – Pilot-In-Command.

SA – Situation Awareness.

SAGAT – Situation Awareness Global Assessment Technique.

SART – Situation Awareness Rating Technique.

SA-SWORD – Situation Awareness – Subjective Workload Dominance.

SDG – Single Display Groupware.

STAR – Standard Terminal Arrival Route.

SV – Sheridan-Verplank.

TCAS – Traffic Alert and Collision Avoidance System.

VHF – Very High Frequency.

Chapter 1

Introduction

Human collaborative work has long been a research focus in a wide variety of domains, including psychology, aviation, and the military. The aviation domain, in particular, has been a hub for significant work (Foushee et al., 1986, Wiener et al., 1993) over the past several decades aimed at studying and improving collaboration. During this time, advances in cockpit computer technology, particularly automation, have led to a steady increase in the level of human-machine collaboration in the cockpit (an environment that once involved primarily human-human collaboration). More recently, the idea of using computer systems to assist collaboration in various other domains has become an important area of research in human computer interaction, generally known as computer supported cooperative work (CSCW). Traditional CSCW research has focused on distributed systems such as email, chat, and web-based systems (e.g., Grudin, 1994, Gutwin and Greenberg, 1998, Yankelovich et al., 2004), but technology advances in the last decade have driven the expansion of CSCW research into the development of collaborative technologies for co-located environments. These technologies often involve the use of shared displays, which can include single display (Stewart et al., 1999, Tse et al., 2004, Zanella and Greenberg, 2001) or multi-display (Dietz and Leigh, 2001, Han, 2005, Johanson et al., 2002, Rekimoto, 1997, Streitz et al., 1999) systems with a variety of different interaction techniques. In recent years, shared displays have started to gain acceptance as an effective element of cockpit interfaces. This thesis examines some of the challenges associated with supporting cockpit collaboration using shared displays, proposes a new display augmentation concept to help mitigate these challenges, and discusses the development and testing of a prototype implementation of this concept.

1.1 Motivation

1.1.1 Cockpit Technology Advances

During the past several decades, cockpit technology has been steadily developing, with notable advances being made in both the cockpit systems themselves and the human-machine interfaces that are used to control them. Older aircraft used primarily mechanical and simple electrical systems that were monitored and controlled using analog instruments and manual controls. In the late 1970s and early 1980s, aircraft such as the McDonnell Douglas MD-80 and Boeing 757 became the vanguard of a widespread deployment of integrated digital cockpit systems (glass cockpits), which consolidated a wide variety of individual instruments into a few integrated, multi-function displays that allow pilots to easily locate and understand important flight information. Along with the development in cockpit displays, these aircraft use significantly more advanced flight controls; in particular, the level of automation in the aircraft systems has increased dramatically, leading to a commensurate growth in the need to interact with the computers driving the displays and automation functions.

Data input into these flight computers has traditionally and successfully been accomplished using entirely keyboard, button, and knob-based interaction with a primarily text-based interface. However, emerging cockpit configurations are based on the concept of a “Windows-style” interface that uses an on-screen cursor with a cursor control device and a keyboard for input. Many of these systems propose a layout similar to the one shown in Figure 1-1, in which there is a display on each side of the aircraft (one or more screens for each pilot) and a center display (one or more screens) that is shared by the pilots. The introduction of systems of this type allows for the design of cockpits that use the shared display as a focus for collaboration, both between the two pilots and between the pilots and the automation.

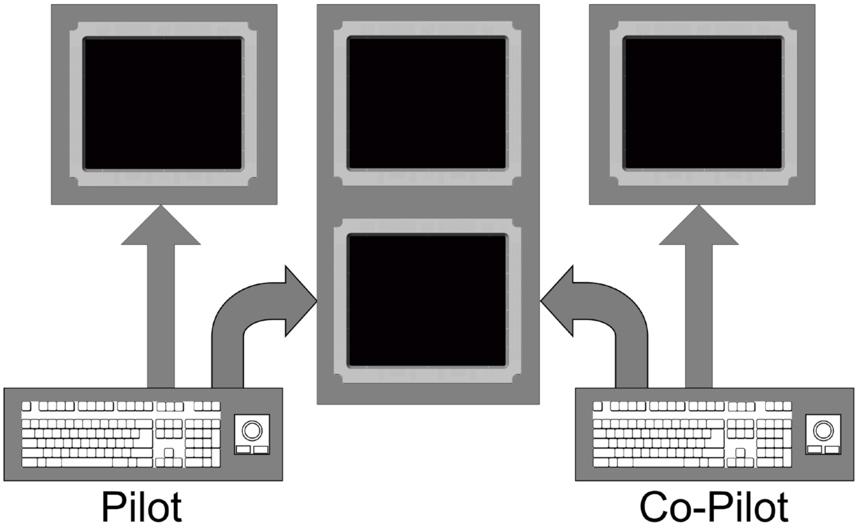


Figure 1-1 – Emerging Cockpit Architecture

1.1.2 New Cockpit Architecture Advantages

There are a number of advantages associated with this new architecture, foremost among which is the increased display area available compared to a conventional glass cockpit. Because the shared display uses two or more relatively large LCD screens, more ecologically appropriate graphical designs can be used for the display and control of the functions located there. Foremost among these functions are flight planning, flight management, and navigation, and the large shared display space allows for the provision of an interactive map that can be used in carrying out these functions. Allowing the pilots to work with their flight plan directly on a shared map, instead of individually on a waypoint list, helps to provide a much better contextual understanding of the position of the aircraft and the status of the flight. Improving pilot awareness of the navigation status of the aircraft has important potential benefits to both safety and efficiency; examples of such benefits could include reducing the likelihood for controlled flight into terrain accidents and improving pilot selection of optimal alternate routes when avoiding inclement weather. Additionally, the shared display provides a common access point to secondary functions, including aircraft systems management, navigation sensor management, communications management, maintenance information, and checklists, and the crew's interaction with all of these functions can be improved using ecologically designed graphical interfaces.

1.1.3 New Cockpit Architecture Challenges

The wide range of functions being monitored and controlled on the shared display also leads to one of the major drawbacks of this architecture: the potential for situations to arise that create conflicts between the pilot and co-pilot interacting with the system. Such "interaction conflicts" can be defined as situations in which one user's interaction with the system interferes with another user's interaction. If not addressed, these conflicts have the potential to be a significant safety concern. In addition to the potential for interaction conflicts, the interaction style of the system (cursor control and keyboard input) removes some of the peripheral collaborative cues that existed in older generations of aircraft. For example, when changing a radio frequency on older aircraft, pilots had to physically reach out to the radio control on the instrument panel and turn a knob (a movement which could be relatively easily detected in the peripheral vision of the second crewmember). Pilots of aircraft using the emerging cockpit architecture now accomplish this task (and many similar tasks) with input on the keyboard and/or cursor control device. This makes it much more difficult for crewmembers to maintain a sense of "peripheral awareness" of each other's actions. This reduction in peripheral awareness has the potential to negatively affect crew collaboration and situation awareness, which are both important in maintaining flight safety.

1.1.4 Additional Cockpit Design Challenges

In addition to the specific issues associated with the emerging cockpit architecture, any cockpit design must account for the complex, time-critical, and high workload nature of the aviation domain. One particular issue associated with any domain of this type is the potential for interruptions to occur, temporarily distracting an operator from their task. These interruptions tend to have a detrimental effect on task performance, and have been cited as contributing factors in several aviation incidents and accidents (e.g., Dismukes et al., 1998, NTSB, 1988). Thus, assisting the operator in resuming their interrupted tasks is an important design consideration for any aviation interface.

1.1.5 Problem Statement

Aviation is a complex domain that involves many time-critical and life-critical tasks. The emergence of a new cockpit architecture that uses a shared display (as shown in Figure 1-1) presents an opportunity to improve collaboration between cockpit crew members, with a resulting improvement in operational performance and safety. This new architecture does, however, have a number of drawbacks that need to be addressed before systems of this type can reach their full potential.

1.2 Research Objectives and Approach

In order to improve collaboration in two-pilot crews, increasing both performance and safety, this thesis aimed to generate an answer to the following general research question:

How can a cockpit using the emerging architecture shown in Figure 1-1 be designed to help mitigate interaction conflicts, improve crewmember awareness, and assist operator interruption recovery?

The approach to answering this question can be described in terms of three more specific objectives, as follows:

- **Objective 1 – Examine current practices for effective collaboration in cockpit environments.** A multi stage analysis process was used to help develop an understanding of how pilots collaborate in current cockpits. The analysis process included a literature review of shared display and cockpit research (described in Chapter 2), as well as the development of an operational sequence model for a representative aviation task, the development of the operational sequence model into action and information requirements, and informal interviews with experienced pilots (described in Chapter 3). The product of this analysis process was a set of design challenges that should be addressed in the design of a collaborative cockpit interface.

- **Objective 2 – Design a collaborative interface for use on individual and shared displays in modern (new architecture) cockpits.** Based on the understanding of cockpit operations and interface design challenges obtained in meeting Objective 1, an augmented cockpit interface was designed to address the challenges particularly relevant to collaboration. Input and gaze tracking data were combined in order to augment an existing cockpit interface with integrated visualizations of current and past operator interactions. An in-depth description of the augmented design, including a detailed discussion of the principles and processes used to create it, is presented in Chapter 4.

- **Objective 3 – Evaluate the utility of the improved interface for addressing the identified design challenges.** A non-interactive prototype of the cockpit interface augmentation was evaluated in a human-participant experimental study that was conducted using a novel two-phase methodology (described in Chapter 5). The results of the study and a discussion of their implications are found in Chapter 6 and Chapter 7, respectively.

1.3 Thesis Organization

This thesis is organized into seven chapters, as follows:

- **Chapter 1, Introduction –** introduces the motivation and research objectives of this thesis.

- **Chapter 2, Background –** contains a review of research related to cockpit design, awareness, interruption recovery, shared display systems, and automation.

- **Chapter 3, Domain Analysis –** describes the analysis process used to further develop an understanding of the task domain and the context of the design problem, including the development of an operational sequence model and a framework for studying interaction conflicts.

- **Chapter 4, Interface Design –** describes an augmented cockpit interface designed to promote awareness, assist interruption recovery, and mitigate interaction conflicts, and explains the design process and rationale.
- **Chapter 5, Evaluation Methodology –** describes the two-stage evaluation methodology used to evaluate a non-interactive prototype of the proposed design.
- **Chapter 6, Results –** presents the quantitative and qualitative results of the evaluation process described in Chapter 5.
- **Chapter 7, Discussion** – examines the results of the evaluation process in terms of the hypotheses, discusses some additional insights provided by these results, and makes recommendations for improvement of both the interface design and the evaluation process.
- **Chapter 8, Conclusion –** discusses how well the research objectives were met and presents some recommendations for future work.

Chapter 2

Background

This chapter contains a review of research literature relevant to understanding the aviation domain and the design challenges associated with new architecture cockpits. This review begins with a discussion of existing cockpit procedures and work practices, which is followed with a description of one of the most popular metrics for pilot performance (situation awareness, or SA) and how it relates to other forms of awareness from different domains. The current state of research in both interruption recovery and shared display systems is then discussed, including some analysis of the limitations of current work. Finally, the chapter ends with a discussion of the interaction between humans and automation and some of the limitations associated with current automation strategy.

2.1 Cockpit Research

When designing any cockpit display or control, an important part of the design process is to examine existing cockpit procedures and work practices. This is particularly true when considering cockpits designed for two or more crew members, because the interaction of these crew members with each other and with the cockpit systems form the building blocks of a larger system of distributed cognition (Hutchins and Klausen, 1996). The concept of distributed cognition was developed to address the fact that the performance of a system made up of multiple people interacting with technology cannot be easily modeled or predicted by examining the individual skills or capabilities of any one part of the system (Perry, 2003). Research examining the performance of distributed cognitive systems in aviation (Hutchins, 1995, Hutchins and Klausen, 1996), ship navigation (Hutchins, 1990), and firefighting (Toups and Kerne, 2007) has shown that the distribution of information, information storage, and cognitive work among the different parts of a system allows for more robust overall system performance. For example, Hutchins and Klausen (1996) showed that effective performance of a cockpit task involving multiple crew members does not require each crew member to perform perfectly because errors can be detected and corrected by the others. In another examination of distributed cognition in a time and life-critical domain, Toups and Kerne (2007) demonstrate that when individual firefighters are aware of the overall firefighting situation as well as of the specific actions of the other firefighters, the coordination and performance of the overall firefighting team is improved.

Research into existing cockpit procedures and practices is also important when attempting to design a cockpit system that addresses specific limitations of current systems. While much of this specific information is available from subject matter experts, ethnographic studies that have carefully examined the interactions between members of the flight crew and between the flight crew and the cockpit have revealed work practice information that is difficult for experts to articulate. Some examples of studies of this type include investigations of the use of cockpit elements as memory aids (Hutchins, 1995) and the use of paper in the cockpit (Nomura et al., 2006). Knowledge obtained from these studies provides important insights into some of the design challenges associated with the emerging cockpit architecture. For example, Nomura et al.'s (2006) research on paper use in the cockpit showed that the pilot flying and pilot not flying kept much of the same information easily available (such as approach plates, airport maps, departure and arrival procedures), but some specific pieces of data were only used by one or the other (such as crosswind tables, circling charts, and V speeds[1]). The incorporation of any of this information into the shared display has design implications related to both interaction conflicts and crew awareness. For instance, when information is used by only one pilot or the other, it will be important to ensure that the pilot who needs the information can access it when required without affecting (i.e., causing an interaction conflict with) the ongoing tasks

---

[1] V speeds are reference speeds specific to an aircraft type, labeled as Vx (where x can be a letter, number, or set of letters and numbers). For example, V1 is the maximum speed at which pilots can abort a takeoff and still be able to stop safely without leaving the runway

of the other pilot. When information is needed by both pilots, it will be important to present it in a way that not only promotes individual understanding, but also helps the pilots to understand whether his crew member has also accessed and understood the information.

2.2 Awareness

In the aviation domain, the concept of situation awareness has become very popular over the past two decades as a metric for and predictor of pilot performance (Endsley, 1999). The commonly accepted definition of SA as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" was popularized by Endsley (1988) and has been widely used since. This definition separates SA into three distinct but related processes (or levels):

i) Perception: basic perception of important information in the task environment.

ii) Comprehension: integration of this information to form an understanding of the current situation.

iii) Projection: anticipation of future events based on an understanding of the current situation.

The concept of SA is most commonly used to help in explaining individual task performance by considering SA as an element of performance that is separate from, but closely related to, decision making (as shown in Figure 2-1).

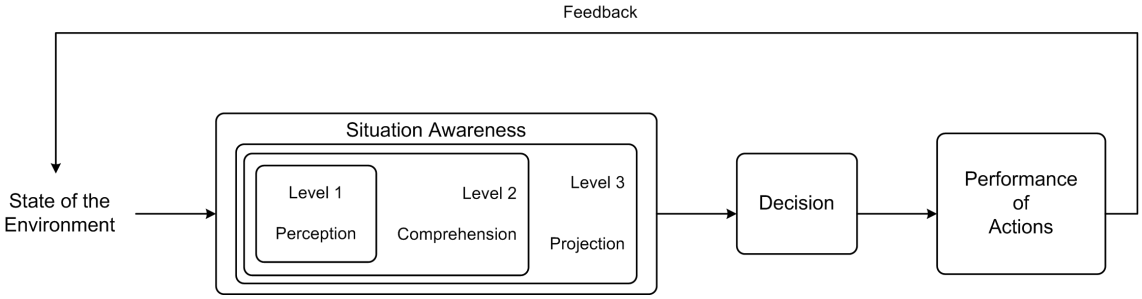


Figure 2-1 – Situation Awareness model (adapted from Figure 2 of (Endsley, 2000))

To this end, several different methods have been developed for measuring operator SA, including the Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 1987), the Situation Awareness Rating Technique (SART) (Taylor, 1990), and the Situation Awareness – Subjective Workload Dominance (SA-SWORD) technique (Vidulich and Hughes, 1991), which have all seen use in evaluation of air crew or cockpit design performance. For example, in a NASA research program aimed at developing synthetic vision systems for cockpits, Kramer et al. (2004) used SA-SWORD as one of their major metrics for comparison of different designs.

The concept of SA has also seen some application to the analysis of team performance. Salas et al. (1995) described team SA as a product of the individual SA of the team members and the communication and teamwork processes that exist between them, and Cooke et al. (2001) developed this concept into a more detailed framework for team SA and discussed some potential methods of assessing SA in team environments. However, these applications of "conventional" SA to team situations are all limited by the fact that SA was originally developed in the context of individual performance and does not specifically address group environments.

In other work domains, the term "awareness" has been used with many other modifiers, and some of these other "types" of awareness were developed specifically to address collaborative settings. Some examples from CSCW include peripheral awareness (awareness information sourced from an operator's peripheral attention (Cadiz et al., 2002)) and group awareness ("the up-to-the-moment understanding of others' activities in a shared space" (J. Hill and Gutwin, 2004)). In the cockpit environment, there are a variety of ways in which a pilot can obtain awareness information. For example, in older two-pilot cockpits, peripheral awareness information was available simply from peripheral vision, by observing that a co-pilot was reaching to adjust something on the instrument panel. However, in software systems (such as modern glass cockpits), this "natural" form of peripheral awareness information is often lost, which has led to research in providing analogous information on a computer display (Cadiz et al., 2002).

An important limitation of existing approaches to providing on-screen awareness information is that (in most cases) this information is displayed separately from primary task information. In addition to requiring dedicated screen real estate, this approach increases the cognitive burden on the user by forcing them to interpret how the awareness information relates to the primary task. Group awareness information can also be found in a number of different forms, including some notable research into the development of software "widgets" (such as buttons and menus) that provide real-time information about collaborators' actions. For example, in the suite of multi-user interface components developed by Hill and Gutwin (2004), a button shows one user's click action to all other users by showing normal button "click" feedback and highlighting it with a colour and label. Widgets of this type are useful at providing information about the current situation, but are less useful in providing an overall context for how the situation has developed because they do not show any historical record of actions.

There has, however, been some research into providing such contextual information; notably, Hill et al. (1992) suggested creating a computational analogy to paper document "wear" (e.g. dog-earing[2], annotations, etc.) to provide information about the history of collaborative work on an electronic document. For example, in their "EditWear" shared editor software application, visual traces of user activity in an electronic document are indicated via a dynamic histogram embedded into the scrollbar of the document window (see Figure 2-2).

---

[2] Dog-earing is the practice of folding down a corner of a page, typically the upper right, to mark a location in a paper document.

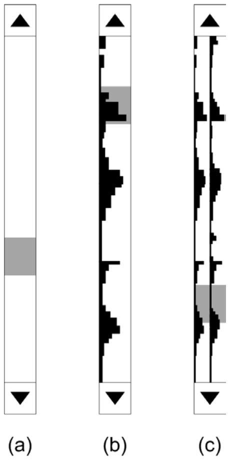


Figure 2-2 – "EditWear" scrollbar concept (adapted from Figure 1 of (W. C. Hill et al., 1992))

The size of each mark in these histograms is proportional to the amount of viewing or editing that has taken place at that point in the document. Figure 2-2 (a) shows a normal scroll bar, (b) shows a scroll bar with a histogram showing document editing activity, and (c) shows a scroll bar with two histograms, one each for editing and viewing activity information. Placing this information in the scrollbar of the document allows users reviewing the document to easily skip to locations that have been heavily read or edited by simply moving the scrollbar to the appropriate location on the histogram. This particular implementation is limited in that it was only designed to support asynchronous collaboration; however, it is likely that the concept of displaying traces of user activity on an interface has the potential to be useful in synchronous collaboration as well.

2.3 Interruption Recovery

The potential for interruptions to occur and distract an operator from their task is an issue in many work domains, but is of particular relevance when examining complex domains. A sizeable body of research has been published examining interruptions in a variety of tasks, with results almost invariably showing that interruptions have a significant effect on task performance (Altmann and Trafton, 2004, Bailey et al., 2001, Loukopoulos et al., 2001, McFarlane, 1999, Scott et al., 2006, Scott et al., 2008, Smallman and St. John, 2003, St. John et al., 2005, Trafton et al., 2005, Trafton et al., 2003). In a study examining interruptions in relatively simple tasks such as addition, counting, and reading comprehension, Bailey et al. (2001) found that interrupting a user decreased their task performance and increased their level of annoyance with the task, and that the magnitude of these effects was related to their perceived mental workload at the time of the interruption. These detrimental effects can also be observed in more complex tasks; McFarlane (1999) discusses how interruptions are more likely in tasks involving automated systems (including aviation) and demonstrates that there is decreased performance on such tasks when interruptions occur.

In a time and life-critical domain such as aviation, interruptions and the resulting negative effects can have disastrous consequences. For example, in 1987, Northwest Airlines Flight 255 crashed on takeoff after the pilots were interrupted from their normal pre-flight routine and failed to set the flaps (NTSB, 1988). In a study specifically examining interruptions in an aviation setting, Loukopoulos et al. (2001) found that aircrew performance on flows and checklists was noticeably affected when the crew members were interrupted while carrying out these tasks.

Many researchers have proposed and studied methods for mitigating the effects of interruptions on task performance; however, current research in this area has typically focused on tasks that are either entirely static during an interruption (i.e., the operator is returned to the same task with the environment in the same state) or entirely dynamic (i.e., the state of the task and the state of the environment may change during the interruption). Tasks faced by pilots can be of either type, and can also be a combination of the two (e.g., a pilot can return to the same task state but face a very different set of environmental conditions). For this reason, it is important to examine interruption recovery techniques used for both types of tasks and attempt to establish a method that will work for the aviation domain. For dynamic tasks in particular, this method will need to support pilots in understanding whether the environmental or task conditions have changed, and in what way.

In a series of studies that used a static task to examine the effect of providing users with a warning before the onset of an interruption, Altmann and Trafton (Altmann and Trafton, 2004, Trafton et al., 2005, Trafton et al., 2003) found that the availability of primary task information during the warning period improved recovery performance after the interruption. Additionally, they showed that when resuming the task, the presentation of primary task cues (such as the cursor position or last action performed before the interruption occurred) improved recovery performance. While this method is simple to implement and does provide some benefit for static tasks, it is limited in its application to dynamic tasks in that it does not provide any information about what occurred in the primary task while the user was interrupted.

Other interruption recovery assistance research has focused on addressing dynamic tasks by providing specific tools to address this limitation, including change logs (Smallman and St. John, 2003), instant replay tools (St. John et al., 2005), and integrated change log/instant replay tools (Scott et al., 2006, Scott et al., 2008). For example, Smallman and St. John (2003) showed that providing a dynamically updating log of events, sorted by importance, to participants performing an aircraft threat-classification task dramatically improved their performance after an interruption. In a series of studies examining interruptions in dynamic mission control tasks, Scott et al. (2006, 2008) showed that an instant replay tool or an integrated change log/instant replay tool could provide some benefit to interruption recovery performance. While these methods have all shown some success in improving interruption recovery performance, tools of this type tend to require dedicated display space, which is difficult to provide in a cockpit application, and generally require the user to interrupt their primary task to recover from a previous interruption.

2.4 Single Display Groupware (SDG)

The SDG model was introduced by Stewart et al. (1999) to describe “computer programs that enable co-present users to collaborate via a shared computer with a single shared display and simultaneous input devices.” These co-located multi-user computing systems differ from single user systems and “conventional” groupware because they have multiple, independent input channels (generally, one for each user) and a single output channel (shared between all users), as shown in Figure 2-3.

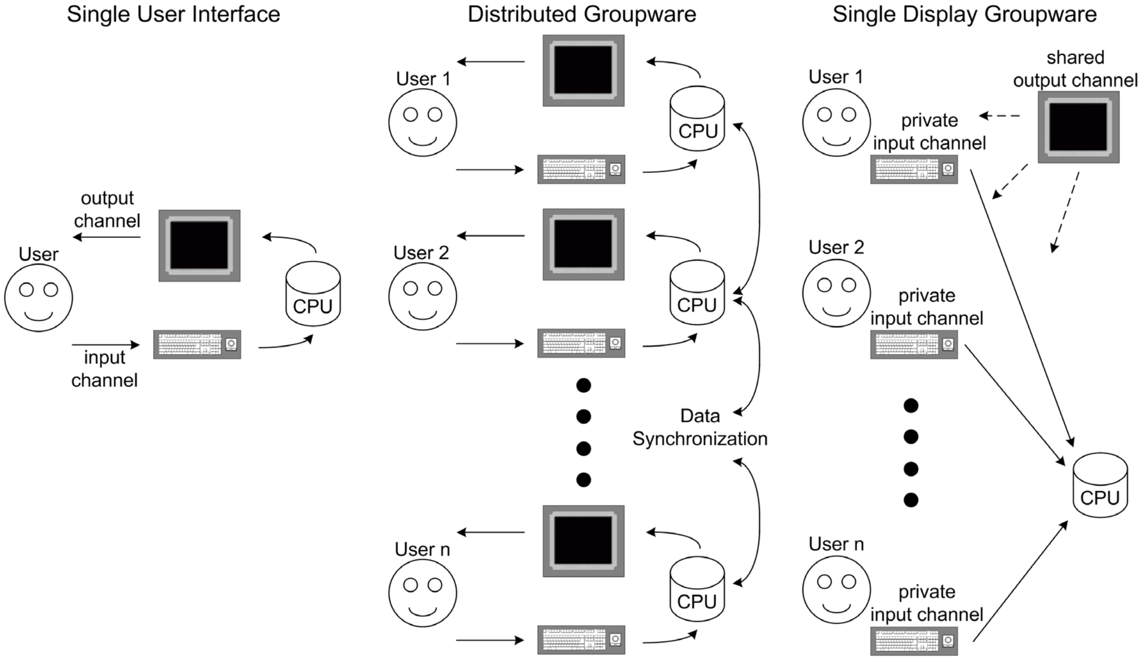


Figure 2-3 – Interaction models (adapted from Figures 1, 2, and 3 of (Stewart et al., 1999))

There are a number of potential advantages and disadvantages associated with SDG systems, and one of the potential disadvantages that has received research attention is the potential for conflicts to occur between users (defined as ‘interaction conflicts’ and discussed in more depth in Section 3.1, below). Research examining interaction conflicts in SDG systems often assumes that conflicts will occur and moves on to developing software techniques (Tse et al., 2004, Zanella and Greenberg, 2001) or general strategies (Morris et al., 2004) to mitigate them. One software technique proposed by Zanella and Greenberg (2001) used translucent pop-up menus to allow one user to make a menu selection while the other user continued working behind the pop-up menu. Morris et al. (2004) proposed a number of general conflict mitigation strategies for SDG systems; one example of such a strategy is the “no selections” policy, under which an input to the system would only be accepted if no other user has an active selection that would be affected by the input. Though this strategy and widget-driven approach can be effective in developing generalized methods for mitigating conflicts, when designing an interface for a specific task it would likely be more effective to examine the task to identify potential conflicts, determine their source and impact, and tailor the interface for optimal performance. Tse et al. (2004) used this type of approach to show that conflicts may not be an issue in SDG systems for tasks that lend themselves well to spatial separation.

Research on SDG systems has more recently expanded to include co-located collaborative systems that comprise one or more shared displays (such as tiled wall displays) that are simultaneously used by multiple people (Brignull et al., 2004), similar to the shared display in the cockpit architecture shown in Figure 1-1. While these systems do not strictly adhere to the definition of SDG, it is expected that lessons learned in SDG research should be applicable to the design of shared displays in more complex environments.

Even when considering the expansion of SDG research to include systems with multiple displays, there are a number of limitations in the work to this point that should be considered. First, research in this area has almost exclusively considered single function systems in which there is no need to change the overall content located on the shared display. These tasks also tend to be fairly simplistic; for example, Zanella and Greenberg (2001) used very basic pointing and selection tasks and Tse et al. (2004) used a simple tracing task. Additionally, and perhaps in part due to the simplicity of the tasks used, typical SDG research has primarily investigated situations in which the users are working on the same task, trying to accomplish a common goal.

2.5 Human Automation Collaboration

As mentioned previously in Section 2.1, while automation plays a significant role in modern aircraft, the actual interactions between aircraft crew members and automated systems are quite limited. This is true in most systems that combine human operators with automation; the widely used Sheridan-Verplank (SV) scale of levels of automation (Sheridan and Verplank, 1978) is essentially concerned with the division of workload and authority between human and automation and does not address any potential collaboration between the two. Parasuraman and Riley's (1997) paper "Humans and Automation: Use, Misuse, Disuse, Abuse" discusses the different ways in which the implementation of automation can be problematic:

i) misuse, which refers to situations in which operators use automation even though it is performing poorly (frequently as a result of overtrust or monitoring failures);

ii) disuse, which refers to situations in which operators do not use automation even though it has the capability to improve performance (often as a result of undertrust); and

iii) abuse, which refers to situations where automation is designed without appropriately considering the needs of the operator or the potential effects of automating the task.

In a review of "Humans and Automation", Lee (2008) makes the specific point that as automation continues to develop and become more complex, and operators come to rely more heavily on it, these issues of misuse, disuse, and abuse will become more and more problematic unless designers begin considering automation and human operators as elements of an overall collaborative system instead of as independent actors.

Some recent research in automated systems has begun to explore the possibilities of collaborative interactions between humans and automation. For example, Bruni et al. (2007) have developed a framework (the Human Automation Collaboration Taxonomy, or HACT) for shared decision making that allows more flexibility in defining the roles of the human and the automated system. However, most computer systems are still limited by the fact that their only awareness of a human operator is the direct inputs of that operator (an issue that is discussed further in Section 3.1).

Chapter 3

Domain Analysis

While the literature described above in Chapter 2 provides a useful starting point when considering cockpit design, very little published work has attempted to form an understanding of the potential challenges associated with new architecture cockpits. This chapter describes the process used to address this limitation, including the development of a framework for understanding interaction conflicts (one of the challenges associated with new architecture cockpits), and the analysis process that was used to develop a more detailed understanding of the aviation domain and the other potential design challenges associated with new architecture cockpits. The final product of this analysis process was a set of design challenges to be addressed in the design of collaborative cockpit interfaces (detailed in Section 3.3).

3.1 Interaction Conflict Framework

The background literature from the human-computer interaction (HCI) research community, in particular the HCI work focused on SDG systems discussed in Section 2.4 above, provided a basic understanding of interaction conflicts and potential ways of mitigating them. However, it was limited by focusing primarily on single function systems in which there is no need to change the overall content located on the shared display, and in which the users are working together on a relatively simple task, trying to accomplish a common goal. The new architecture style cockpit (as shown in Figure 1-1) is a multi-function system in which the shared display content can be changed depending on the state of the task and the environment, the task itself (aviation) is very complex, and the users are frequently working on separate tasks with different individual goals. To address this research limitation and develop a way of structuring investigations of more complex task domains using shared screens, I created a framework for understanding the different types of interaction conflicts that may occur.

When trying to understand computer use, it is important to consider both “active” and “passive” use. Design for active use is concerned with the design of effective mechanisms for direct interactions with or inputs to the system, while design for passive use is concerned with the design of effective information visualizations to help with decision making or data comprehension; these design elements are intended for the visual channel only and do not require user interaction. Both active (input) and passive (visual) computer use are relevant for a discussion of interaction conflicts in a shared display situation, and both are considered in the interaction conflict framework (illustrated in Figure 3-1).

		User 1 Interaction	
		Input	**Visual**
User 2 Interaction	**Input**	Input-Input Conflict	Input-Visual Conflict
	Visual	Input-Visual Conflict	Visual-Visual Conflict

Figure 3-1 – Interaction Conflict Framework

This framework contains three main categories of potential conflicts, detailed below:

i) Input-input (i.e. active-active) conflicts occur when the users attempt to issue mutually exclusive inputs to the system. The most obvious type of potential conflict in SDG-type systems, this type of conflict can occur in any system that has multiple inputs, including several existing cockpit systems such as flight controls (duplicate controls for the pilot and co-pilot that operate the same control surfaces) and traditional flight management systems (separate control display units that are linked in software). Because input-input conflicts can be easily detected by the system, conflict mitigation can be accomplished in a relatively simple manner using technological strategies. These strategies can range from very inflexible (lock the screen so that only one user can perform input at a time) to very flexible (allow all inputs on a "last-in wins" basis). For an aviation application, the very flexible solution has potential safety concerns and the very inflexible solution removes much of the benefit of the collaborative system, so it is likely that the ideal solution lies somewhere in between.

ii) Input-visual (i.e. active-passive) conflicts occur when one user attempts to issue an input that affects the display of the other user's desired output. Input-visual conflicts are more difficult to detect and address than input-input. This difficulty is due to the fact that only one user is directly interacting with the system and computerized systems are not typically aware of users' visual focus, which means that the system software does not know when a conflict of this type is occurring. This difficulty in implementing technological solutions means that input-visual conflict mitigation is much more likely to be left to procedural strategies or to social protocols. However, if a computer system is provided with awareness of users' visual focus (as in the design proposed in this thesis), technology-based solutions for input-visual conflicts become possible.

iii) Visual-visual (i.e. passive-passive) conflicts occur when the users desire information that requires mutually exclusive outputs. Because these conflicts do not involve any input to the system, they are the most difficult of the three conflict types to identify and address. Similarly to input-visual conflicts, it is difficult to develop technological conflict mitigation strategies for visual-visual conflicts unless the computer system is provided with awareness of users' visual focus.

Each potential interaction conflict in a shared display system architecture fits into one of these three categories, and the need to identify and assess potential conflicts of all three types should be considered when examining a task domain. This framework can also provide a focus when designing conflict mitigation strategies, as it is important to recognize that different conflict types may require different mitigation strategies.

Another important point to consider when studying interaction conflicts in the context of a shared display cockpit system is that most modern two-pilot cockpits are intended to be operated such that that one pilot is always "heads-up" (actively flying/monitoring the state of the aircraft) so that interaction conflicts should never occur in-flight. However, anecdotal evidence from pilots, as discussed by Hutchins & Klausen (1996), indicates that occurrences of "two heads in the cockpit" can readily occur in modern, automated aircraft. Thus, interaction conflicts are possible in the advanced cockpit and should be considered in the design of the shared cockpit display.

3.2 Domain Operations Analysis

3.2.1 Research Approach

A multi-stage analysis approach was used in developing an improved understanding of the task domain and the challenges associated with the emerging cockpit architecture. The first two stages in the approach focused on understanding the cockpit environment and the functional requirements for a modern cockpit system, while the third stage (conducted partially in parallel with the first two stages) aimed to confirm the validity of the first two analysis stages and provide an operational context for the results.

The first step in the domain investigation was the creation of a form of Operational Sequence Diagrams (OSDs) for representative scenarios of cockpit interaction. Operational sequence diagrams "are graphic representations of operator or user tasks, as they relate sequentially to both equipment and other operators" (Chapanis, 1996). The symbology used in the creation of OSDs consists of a variety of individual symbols to represent different types of tasks, as shown in Figure 3-2. These symbols are linked together to form a chronological map of the tasks (and links between the tasks) carried out by each operator or piece of equipment in the system being analyzed.

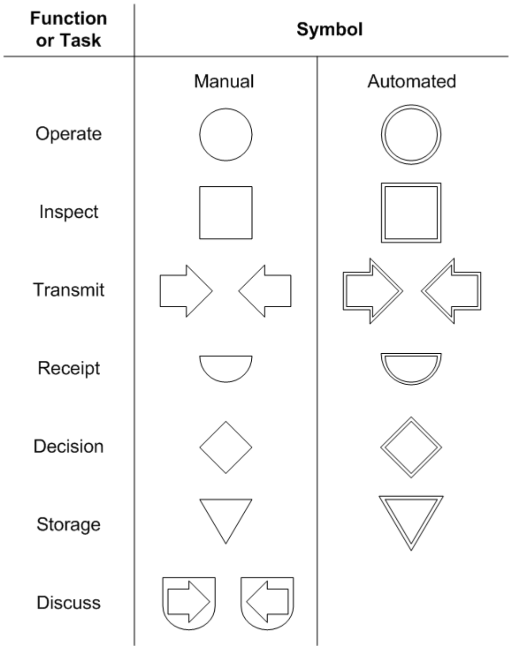


Figure 3-2 – OSD symbols

Since it is difficult to examine the richness of collaboration (for example, ongoing coordination of interrelated activities) using the low-level task approach that is typical to operational sequence modeling, a modified approach was used to instead consider coordination and collaboration in the performance of higher-level functions. For example, instead of considering each low-level task needed to complete a checklist (e.g. check landing gear position, check flap position, check altimeters), the functional approach would simply list the higher-level function (perform after take-off checklist). As the typical low-level OSD approach specifies all tasks in exacting detail, including those related to collaboration, it neglects the fact that collaborative activities such as discussions are flexible and can be performed in different ways depending on the personalities of the human operators involved. The higher-level functional approach ensured that this richness of collaboration was not lost through an overly analytical approach, yet still took advantage of the strength of OSDs to identify elements of operator work where collaboration is important. Using this approach also meant that the analysis could be performed at a level that is relatively independent of specific technology solutions, which should help to ensure that the results will generalize across low-level differences in cockpit technology.

The basis of operational sequence modeling is an appropriate operational scenario, and the scenario selected for this analysis was a change of approach during a flight. An approach consists of a set of information that determines how an aircraft will approach an airport and land at a specific runway, including details such as what equipment is needed on the aircraft, what speeds and altitudes must be flown, and the minimum visibility needed to land safely. A change of approach during a flight could occur for a variety of reasons, including (but not limited to) changing weather conditions or equipment failures on the aircraft or at the airport. The in-flight approach change scenario was selected to provide a series of tasks and decisions that were sufficiently challenging to stress the collaboration between the two pilots without making the analysis excessively complex, allowing the exploration of key information sources and interactions between humans within the cockpit, humans and automation, and humans outside the cockpit.

The high-level operational sequence modeling approach was used to create an OSD for this scenario (Appendix A), and this model was then used as the basis for a more detailed analysis of the action/information requirements of a limited set of tasks. These tasks, such as evaluating and discussing options for a new approach, were selected based on two main factors:

i) their relevance to the emerging cockpit architecture (i.e., if the task were performed in a cockpit based on the emerging architecture, would the shared display be used)

ii) the extent to which they were a venue for collaborative interactions between the two pilots (i.e., how much the OSD explicitly showed collaboration involved in the task).

Each of the functional requirements selected from the OSD was broken down into action requirements (the actions that must be carried out to accomplish the function) and information requirements (the pieces of information needed to carry out the actions). The full list of selected functional requirements and their associated action and information requirements can be found in Appendix B.

To supplement the above analysis activities, a third stage of analysis was performed in which several informal interviews were conducted with subject-matter experts. The interviews were conducted in parallel to, and helped inform, the functional and action/information requirements analysis. Interviews were conducted with five pilots (including three former Canadian Forces test pilots and two current civilian airline pilots, with a minimum of 3,000 flight hours) and one additional cockpit design expert. Results were used to refine the understanding of the tasks, information requirements, and typical actions within the example scenario.

The information gathered in these activities clarified my understanding of the collaborative design challenges that may arise in the advanced cockpit, as discussed below in Section 3.2.2.

The operational sequence modeling approach was valuable and useful in helping to identify design challenges and implications, but it is not intended to be a singular, comprehensive analysis. The amount and scope of information available from an OSD is directly related to how representative the scenario is of the operational environment, and in the aviation domain there is a large amount of variation in the potential usage scenarios and in the particular organization of events within a given scenario. For example, considering just the approach phase of flight, potential scenarios could include a simple textbook approach, an approach change, a missed approach followed by either a second attempt or a change of approach, or a number of other possibilities. However, even considering this limitation, the operational sequence modeling approach did serve as an initial step towards identifying several important design challenges related to the shared displays in an advanced cockpit.

3.2.2 Results

The results of the operational sequence modeling analysis identified several important design considerations and challenges, summarized below.

There is a need for team situation awareness. Follow-up interviews with subject matter experts (as described in Section 3.2.1 above) indicated that beyond the instances of collaboration that can be explicitly identified, pilots' work always includes an effort to maintain a common situational awareness picture. This refers to the fact that each pilot not only needs to be aware of the state of their aircraft and the environment (conventional SA (Endsley, 2000)), but also of the actions and the awareness of their teammate. This shared situation awareness, a common understanding that collaborators are "on the same page", is important for any team situation, but particularly relevant in time-critical or life-critical environments like aviation. Attempting to support this type of awareness in the proposed cockpit architecture became a significant part of the focus for this thesis.

Significant collaboration occurs between pilots, and between pilots and other actors. The limited scenario investigated using the OSD approach showed multiple instances where the pilots must communicate decision options, provide instructions, confirm and cross-check the other's actions, or otherwise engage in collaborative actions and activities. For example, the need to evaluate and discuss potential options for the alternative approach requires both information exchange both between the pilots, and between the pilots and other actors such as air traffic controllers. Examination of these instances of collaborative activity provided a focus for determining the pilots' action and information requirements.

Information used in collaborative activities can be historical and dynamic. Also identified in the follow-up interviews with pilots was the fact that outside of the single scenario considered in the OSD analysis, information requirements in the cockpit change over time, and information gathered during one activity can be used in others, including to support later collaboration decisions and actions. For example, information obtained during standard radio updates with a company dispatcher can help pilots to decide on routing adjustments or approach changes. These discussions illustrated the need to provide a cockpit design that promotes awareness of current and historical events. Such design support would be particularly important to assist pilots in quickly regaining awareness of their previous task status and of the updated system state when resuming a task following a task interruption.

Pilots operate in a multi-task environment. The analysis of the scenario showed that both pilots had at least one and often two tasks ongoing related to the change of approach in addition to the continuous task of flying and monitoring the progress of the aircraft. For example, Figure 3-3 shows a situation in which the First Officer is continuously monitoring the Captain's control of the aircraft while also carrying out a discussion about an approach change and identifying alternative options for the approach. In follow up interviews pilots indicated that it can be a significant challenge simply to keep up with all of the individual tasks required to safely complete a flight. This prevalence of multiple tasks competing for a pilot's limited attention resources highlights the need for the cockpit design to mitigate the costs of task switching, and to facilitate the pilot's ability to quickly resume a task after being interrupted by another ongoing task.

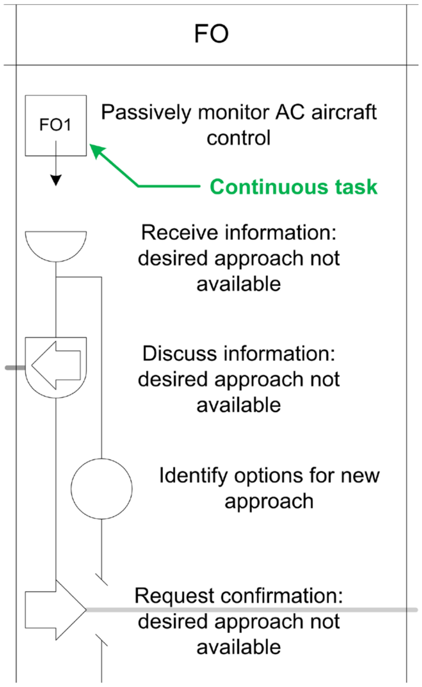


Figure 3-3 – Example of multi-task environment

A wide variety of information is required and a large amount of this information is shared. Tracing the information requirements needed to support the collaborative activities involved in selecting a new approach in the flight plan showed the variety and amount of information that must be shared and commonly understood by both pilots. The variety of information required is demonstrated by the sheer number of qualitatively different information requirements that appear during this scenario; even a single function from the OSD can require many different types of information. For example, when identifying options for a new approach, pilots need access to approach plates for the destination airport, current weather information, airport traffic information, and potentially several other situation-dependent items. Considering that each pilot may be performing more than one of these functions at a time, the amount of information that may be needed at a given time could be significant. The large amount of shared information appears in several different steps in the OSD, underlining the need to ensure that both pilots can access the information that they require without coming into conflict.

Automation plays a limited role in cockpit collaboration. The OSD analysis and pilot interviews also demonstrated that while automation plays a critical role in the control and navigation of modern aircraft, the interaction between cockpit computer systems and the crews operating them is typically limited to relatively simple control inputs by the crew and feedback displays from the computers. For example, a typical interaction with the system might consist of the crew entering a piece of data (such as a new approach) into the system, with the system then displaying that data to confirm it was entered. As mentioned previously in Section 2.5, this limitation represents an opportunity for improvement in the design of automated systems, and this thesis proposes a cockpit interface design that attempts to increase the level of collaborative activity between the flight crew and the automated cockpit systems.

3.3 Collaborative Cockpit Interface Design Challenges

Based on the literature review detailed in Chapter 2 and the domain analysis described above in Chapter 3, three significant challenges were selected as the most important candidates to be addressed in cockpits using the new architecture shown in Figure 1-1. These three challenges (supporting awareness, assisting interruption recovery, and mitigating interaction conflicts) are summarized in the following three sections. Section 3.3.4 then summarizes how the limitations of current automated systems contribute to these challenges.

3.3.1 Support Awareness

The importance of situation awareness in the aviation domain was described in Section 2.2, which also explained how the concept of SA is limited by its inherent focus on individual performance. However, group-focused awareness research from the CSCW domain was demonstrated to have potential for application to cockpit environments if some limitations are addressed. Additionally, the discussion of distributed cognition in Section 2.1 explained that supporting situation and team awareness in complex collaborative tasks can have a positive effect on team performance. The results of the domain analysis described in Section 3.2.2 showed that in addition to individual situation awareness, team awareness is important in multi-pilot cockpits. In summary, the design of new architecture cockpits should attempt to support both individual and team situation awareness.

3.3.2 Assist Interruption Recovery

The results of domain analysis (Section 3.2.2) indicated that the dynamic information and multi-task environment in the cockpit means that interruptions are likely and could significantly impact performance if effective recovery is not supported. Additionally, the interruption recovery literature described in Section 2.3 showed that current interruption recovery support concepts for both static and dynamic task environments have limitations that prevent them from being feasible and effective in the complex domain of aviation. For a new architecture cockpit, it will be important to create an interface design that is capable of assisting interruption recovery performance on static, dynamic, and combined tasks.

3.3.3 Mitigate Interaction Conflicts

Using single display groupware research as a context, the idea that interaction conflicts are a potential problem in systems with shared displays was presented in Section 2.4. This idea was expanded in Section 3.1 with a discussion of the different types of potential interaction conflict and some general methods for mitigating them. It was also suggested that in spite of standard “pilot flying/pilot monitoring” procedure in aviation, instances of “two heads in the cockpit” do occur, meaning that interaction conflicts could be a problem in cockpits with shared displays. Given this potential problem, the design of an interface for these new architecture cockpits, and particularly the shared displays, should assist the pilots in avoiding conflict situations and should mitigate the effects of these conflicts if they do occur.

3.3.4 Human Automation Collaboration

The background literature dealing with automation, described in Section 2.5, showed that while significant research has been conducted examining the interaction between humans and automation, very little of this research considered these interactions in the context of collaboration. This literature, along with the interaction conflict framework (Section 3.1) and the results of the domain analysis (Section 3.2.2), also showed that most current automated systems are limited by their lack of awareness of many human operator actions. Creating computer systems that are more aware of human operator(s) should allow them to begin taking a more active role in the collaborative process.

Chapter 4

Interface Design

Based on the results of the analysis process described in Chapter 2, three key challenges that need to be addressed to improve performance in cockpits using shared displays are: supporting individual and team awareness, assisting pilot interruption recovery, and mitigating potential interaction conflicts. This chapter presents a design concept that has the potential to address all of these challenges in shared display cockpit architectures.

4.1 Design

Existing implementations of designs using the emerging cockpit architecture (Figure 1-1) have a number of limitations related to the key challenges identified in Section 3.3, which can be briefly summarized as follows:

- 'peripheral' type awareness information that existed in older cockpits has been lost;
- support for interruption recovery is limited due to the lack of cues for pilots about what they were working on and what has changed while they were interrupted; and
- the system architecture creates the potential for interaction conflicts.

These limitations can potentially be addressed by augmenting the cockpit interface with visualizations of the history of each operator's interactions with the interface (using a concept similar to the "computational wear" shared awareness design approach discussed in Section 2.2). To further improve the utility of this interface augmentation, a gaze tracking system could be added to provide additional data to the system of the pilots' use of the cockpit displays. This concept has the potential to address the previously stated limitations by:

- re-creating lost peripheral awareness information by highlighting areas of the shared display that are being viewed or used by each pilot;
- assisting interruption recovery by (a) providing cues for pilots about what they were working on before they were interrupted and what has changed since, and (b) allowing the system to detect when a pilot has been interrupted; and
- mitigating interaction conflicts by (a) clearly indicating where each pilot is working, helping them to avoid conflicts, and (b) allowing the system to detect when input-visual conflicts are occurring.

Having identified the idea of augmenting the cockpit interface with operator usage history data (both input and visual) as a way of addressing the design challenges of the emerging cockpit architecture, it was necessary to confirm that gaze tracking was a feasible technology to use in the cockpit (discussion in Section 4.1.1) and develop a method for displaying the usage history data in the cockpit interface (described in Section 4.1.2).

4.1.1 Gaze Tracking

The concept of using gaze tracking as a component of a computer interface is not new; many researchers have experimented with different ways of using a gaze tracking system to augment or replace a cursor control device (Borah, 1995, Jacob, 1990, Kumar et al., 2007). However, much of this research has been hampered by the accuracy of gaze tracking systems which, even under ideal conditions, are limited to 0.5-1° of visual angle. At an eye-to-monitor viewing distance of 50 cm (the standard distance specified in MIL-STD-1472F (1999) for electronic displays), this allows for an on-screen gaze position tracking accuracy of 0.45-0.9 cm, while typical windows toolbar buttons can be as small as 0.3 cm. Some research has attempted to find ways of mitigating this accuracy problem (Kammerer et al., 2008, Kumar et al., 2007), but the limited success of these systems indicates that gaze tracking has limited real-world application as a general computer interface component. However, aviation computer interfaces are generally custom designed and are already required to use larger components than home computer systems; for example, MIL-STD-1472F (1999) dictates that "Aircraft display characters and symbols that must be read in flight shall subtend not less than 7 mrad (24 min) of visual angle" (almost 0.5°). While this means that individual characters and symbols may be too small for a gaze tracking system to identify, components such as buttons that are made up of multiple characters should be, and could be designed to be, sufficiently large.

When proposing the use of a gaze tracking system, it is also important to consider the usage environment and choose a system with an appropriate form factor. Current gaze tracking systems commonly use either head-mounted or remote-mounted sensors to track point-of-gaze. Each style of system has specific advantages and disadvantages, but the required accuracy of 0.5-1° of visual angle can be obtained with both styles. For the cockpit environment, a remote-mounted system would be preferable primarily because the system sensors could be incorporated into the cockpit instrument panel, eliminating the need for pilots to put on and adjust headgear before flight.

4.1.2 Design Rationale

Once it was determined that the idea of using a gaze tracking system in the cockpit could be feasible, a concept for the display of usage history information was developed to include both input and visual information. Based on the proximity compatibility principle (Wickens and Hollands, 2000, p. 97), it was decided that the usage history data would need to be integrated into existing cockpit display components. This would allow the data to be displayed in the context where it was relevant, reducing the cognitive load required to understand the information. Even without this cognitive benefit, displaying the data in context would be advantageous because this format can help reduce the amount of space required for data display, which is extremely desirable in an aviation application. The decision to use a contextual format meant that it was necessary to develop a data display method that could be generalized enough for application to any interface component and that could be used to indicate both input and visual usage history data.

Another major design decision was to develop a treatment that could take advantage of pre-attentive processing (Wickens and Hollands, 2000, p. 87) to further reduce the cognitive load for pilots using the interface. This decision dictated that the design needed to be composed of visual representations of data with characteristics that allow them be pre-attentively processed. To this end, Carpendale's (2003) review of Bertin's (1983) concept of 'visual variables' not only identifies different characteristics of visual representations that are pre-attentively processed (though Carpendale does not specifically refer to the concept of pre-attentive processing), but also discusses the different types of information that can be encoded using each visual variable. The visual variables concept proposes that there are basic 'marks' (points, lines, areas, surfaces, and volumes) that can be encoded with information using visual variables, according to the following five characteristics that determine what types of information can be encoded:

i) Selective: Can a change in this variable make a mark distinct from other marks of the same type?

ii) Associative: Can marks be sorted into groups based on this variable?

iii) Quantitative: Can changes in this variable be used to encode numerical data?

iv) Order: Are changes in this variable naturally perceived as having an order? (i.e., is one value of the variable naturally read as more or less than another value?)

v) Length: How many changes in this variable can be easily distinguished?

Table 4-1 summarizes the answers to the above five questions for each of the seven visual variables. A checkmark in the respective column indicates that the answer is “yes”, an X indicates “no”, and a tilde indicates “somewhat”.

Table 4-1 – Characteristics of Visual Variables (adapted from Tables 1-10 of (Carpendale, 2003))

Characteristic **Visual Variable**	**Selective**	**Associative**	**Quantitative**	**Order**	**Length**
Position changes in location	✓	✓	✓	✓	Theoretically infinite, practically limited by display resolution
Size change in length or area	✓	✓	~	✓	Theoretically infinite, practically limited to ~20
Shape infinite number of shapes	~	~	✗	✗	Theoretically infinite
Value changes from light to dark	✓	✓	✗	✓	Theoretically infinite, practically limited to ~10
Colour changes in hue at a given value	✓	✓	✗	✗	Theoretically infinite, practically limited to ~10
Orientation changes in alignment	✓	✓	✗	✗	Theoretically infinite, in practice should be limited to 4
Pattern repetitive use of shape changes	~	~	✗	✗	Theoretically infinite
Grain varying granularity	✓	✓	✗	✗	Theoretically infinite, practically limited to ~5
Texture a characteristic of the material	✓	✓	✗	✗	Theoretically infinite

To create an initial prototype for the interface design, it was necessary to identify all of the different items of usage history information that could potentially be added to the interface and the characteristics that would be needed to fully encode each item. The results of this process are summarized in Table 4-2.

Table 4-2 – Characteristics of Usage History Information

Usage History Information	Characteristics Needed
User identification	Selective, associative
Recency of use	Order, quantitative
Frequency of use	Order, quantitative
Total duration of use	Order, quantitative

Based on these results, visual variables were selected for use in encoding each item of usage history information. Examining Table 4-2, it can be seen that encoding user identification required a visual variable that was both selective and associative, to allow identification of the user and their interaction patterns. Possible visual variables that meet these criteria include position, size, value, colour, orientation, grain, and texture. Position and orientation were eliminated as possibilities because they were already defined by the interface components; out of the remaining choices, colour was selected as the one that provided the best method of showing a clear distinction between the users. Referring again to Table 4-2, it can be seen that recency, frequency, and total duration of use all required visual variables capable of encoding order and quantitative data. Having already eliminated position as a choice, size and value were the only two remaining variables that could be used to encode order, with neither one being particularly useful for encoding quantitative data. Based on this, it was decided that recency of use was the most important piece of data and that it should be encoded singularly, while the frequency and total duration of use could be combined to create a single composite piece of information. Finally, to choose which visual variable (size or value) to use for each piece of information, the 'length' of the two choices was examined; due to the limited space available on the interface, size would practically need to be limited to two or three distinct levels while the full range of value could be used. Given the decision that recency of use was more important than the combined frequency/total duration measure, value was selected to encode recency and size was selected for the combined measure of frequency/total duration.

After deciding which items of usage history information would be included and what visual variables would be used to encode them, it was necessary to choose a basic display treatment that could be encoded with the desired information using the selected visual variables. Based on an examination of the interface components on which the treatment would need to be applied (including windows, buttons, menus, and several other components), a simple border was selected as the basic treatment. While this basic treatment was effective as an initial design for carrying out a preliminary evaluation of the potential of the underlying concept, there are other treatment alternatives that should be compared in future work to determine which is most effective.

To reduce the potential for clutter in the interface, it was decided that visual interaction information (sourced from the gaze tracking data) would be provided only at a general window level, and input interaction information would be provided at an individual "widget" level (e.g. buttons, map symbols, etc). An initial prototype was developed using this treatment concept and refined through a user-centered design process using feedback obtained from colleagues, interface design experts, and pilots. The final prototype design concept is described in Section 4.1.3, below.

4.1.3 Final Prototype Design

After carrying out several iterations of prototyping and informal user feedback collection, a final prototype design was created based on a 'generic' style cockpit interface of the type that might be seen in a cockpit similar to Figure 1-1. The border treatments and the visual variables applied to them are shown applied to an example interface component in Table 4-3, with a description of how the treatments work dynamically to show the desired information.

Table 4-3 – Proposed interface treatments for providing cockpit interaction awareness

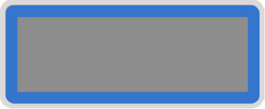
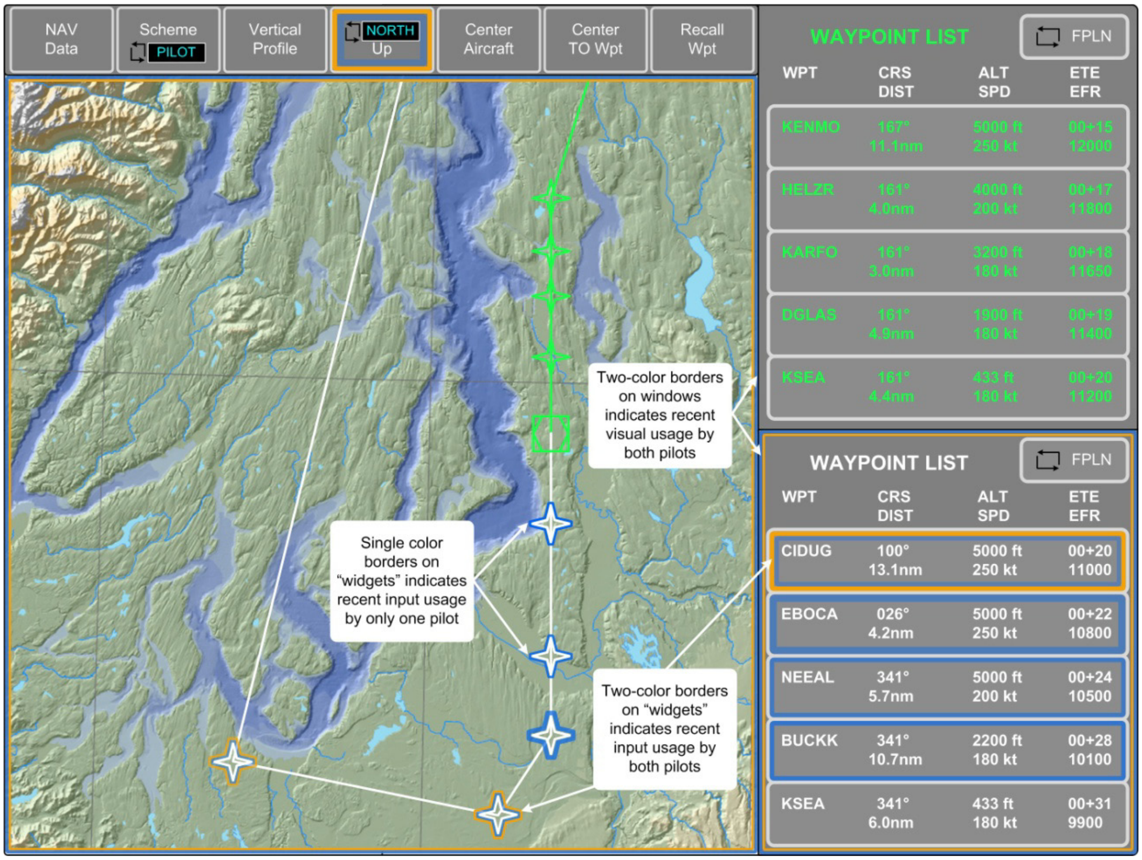
Treatment	**Meaning**	**Example**
Basic border and Colour	Basic treatment to identify user and show usage	
Value	Border fades over time to show recency of use; fade rates are variable depending on type of use (input or visual) and context	
Thickness	Thicker border is used to indicate greater importance (based on frequency and total duration of use)	
Relative position	Relative position of two borders indicates which user's interaction was more recent (outer border)	

Figure 4-1 shows an example of what this concept looked like when applied to a mock-up of an interactive flight planning map display (map image adapted from work by Finlayson (2005)).



WAYPOINT LIST

WPT	CRS DIST	ALT SPD	ETE EFR
KENMO	167° 11.1nm	5000 ft 250 kt	00+15 12000
HELZR	161° 4.0nm	4000 ft 200 kt	00+17 11800
KARFO	161° 3.0nm	3200 ft 180 kt	00+18 11650
DGLAS	161° 4.9nm	1900 ft 180 kt	00+19 11400
KSEA	161° 4.4nm	433 ft 180 kt	00+20 11200

WAYPOINT LIST

WPT	CRS DIST	ALT SPD	ETE EFR
CIDUG	100° 13.1nm	5000 ft 250 kt	00+20 11000
EBOCA	026° 4.2nm	5000 ft 250 kt	00+22 10800
NEEAL	341° 5.7nm	5000 ft 200 kt	00+24 10500
BUCKK	341° 10.7nm	2200 ft 180 kt	00+28 10100
KSEA	341° 6.0nm	433 ft 180 kt	00+31 9900

Two-color borders on windows indicates recent visual usage by both pilots

Single color borders on "widgets" indicates recent input usage by only one pilot

Two-color borders on "widgets" indicates recent input usage by both pilots

Figure 4-1 – Example showing interface treatments on an interactive flight planning map

4.2 Implementation

As a way of developing an understanding of the design decisions that would need to be made in implementing this concept, a video prototype of the design was created based on an existing cockpit interface. This video prototype also provided an effective visual aid to the process of clarifying the design of the dynamic features of the prototype, and served as a basis for the experimental study used to evaluate the design (described in Chapter 5). The decision to use a non-interactive video prototype instead of an interactive cockpit simulator involved several factors:

- Time. The amount of time that would have been required to create a fully interactive cockpit using two gaze tracking systems would have pushed the project well beyond the scope of this research.

- Expense. The cost of two gaze tracking systems capable of the accuracy and viewing area necessary was prohibitive for an initial implementation and a proof of concept evaluation.
- Concept maturity. The concept presented here, in the context of a cockpit interface design, is still in the very early stages of development and needs to show promise at a basic level before the cost of a full prototype is warranted. The video prototype provided a relatively low cost basis for the evaluation process which aimed to show the usefulness of the overall usage history information concept without necessarily proving that the proposed treatment was the ideal design.

The design of the cockpit interface and the limitations of the video editing software (detailed below) meant that a few modifications to the final prototype design were needed in order to create the implemented design described in Section 4.2.1.

Cockpit Interface Limitations. The design of the cockpit interface that was used as a basis for the prototype implementation had two main limitations that required changes to the prototype design. First, the interface had already been designed with specific colours associated with the pilot and co-pilot; this meant that to support a logical interface design, the augmentation was limited to the existing colour scheme (orange for the pilot and blue for the co-pilot). Second, the interface had not originally been designed to support the border treatments proposed, and a number of components were either too small or too close together to practically support individual borders. However, all of the components that presented this problem were elements of larger functional groups on the interface, so the only design change needed was to adjust the level of detail provided by the input interaction borders. The final design of these borders is described in Section 4.2.1.

Video Editing Software Limitations. The way the borders were implemented using the video editing software also presented two significant limitations to the proposed design. The first limitation was that the implementation of the fading treatment required a significant amount of time and had to be completely repeated in order to change the fade durations, making it extremely time consuming to use an iterative design process to determine appropriate fade durations. The second limitation was that implementing dynamic changes to the relative position or thickness of the borders would have required significantly more time (approximately double for changes in relative position). The added time required would be difficult to determine for thickness as further analysis would be required to determine which borders would be ‘thick’. For the purposes of this study, it was determined that:

i) the fade durations would be established using a brief iterative design process using a 2 minute segment of the scenario videos, and because of this would also be independent of context, with one duration for all visual borders and one for all input borders;

ii) the relative opacity (fading) of the borders should be sufficient to indicate which user had most recently interacted with a component, and the relative position concept could therefore be removed from the design; and

iii) the extra time required to develop and implement the thickness concept brought it outside the scope of this research, and it could be left for study in future work.

4.2.1 Implemented Design

The final design was implemented on full resolution screen capture videos from a prototype cockpit (discussed further in Section 5.2.2) that used a layout similar to the cockpit architecture in Figure 1-1. In this prototype cockpit, the two individual displays (left and right) were identical primary flight displays (PFDs), the upper shared screen was an interactive navigation display, and the lower shared screen was a combined flight information and synoptic display. These screens, and the implementation of the borders on each, are described in the following three sections: Section 4.2.1.1 for both primary flight displays (PFDs), Section 4.2.1.2 for the interactive navigation display, and

Section 4.2.1.3 for the combined flight information/synoptic display. Finally, Section 4.2.1.4 contains an example sequence of images showing how the borders behave over time. The fade durations (100% opacity to zero) for the input and visual borders were set at 60 s and 15 s respectively.

4.2.1.1 Primary Flight Displays

The two primary flight displays (one for each pilot) in the prototype cockpit were identical, each showing a standard set of integrated primary flight instruments on the upper half of the screen with the lower half divided into thirds showing the horizontal situation indicator, the radio controls, and the crew alerting system (which was not implemented in the prototype). An example of this display is shown in Figure 4-2.

Each PFD was augmented with six different usage history borders, with one visual history border each around the primary flight instrument window, the horizontal situation indicator window, and the radio control window, and one input history border around each radio control. No border was implemented on the crew alerting system window as it was not used in the prototype. Figure 4-3 shows what the co-pilot's interface would look like with all six borders at full opacity; the pilot's interface would look similar except with the visual borders in orange instead of blue. The input history borders around the radio controls were always shown (on both PFDs) in the colour of whichever pilot performed the input action; this is an example of a way in which the system could be a more active collaborator and 'push' awareness information from one side of the cockpit to the other. However, in the evaluation scenario (described in Section 5.2.1), all input actions were performed by the co-pilot, and as a result the radio input borders were always blue.

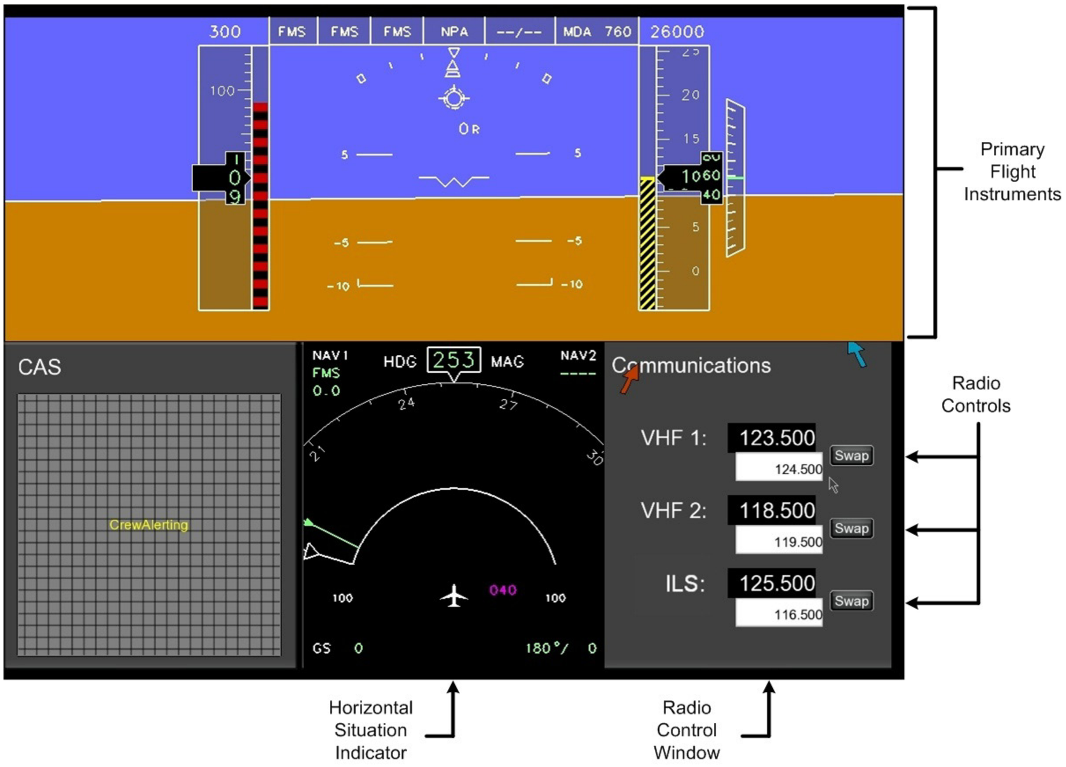


Figure 4-2 – Original Primary Flight Display

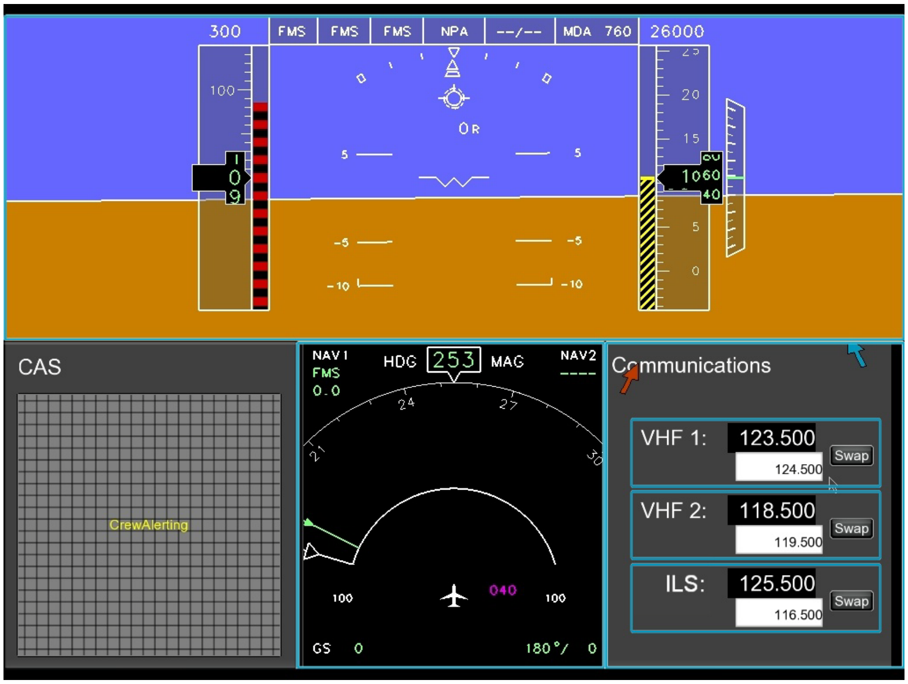
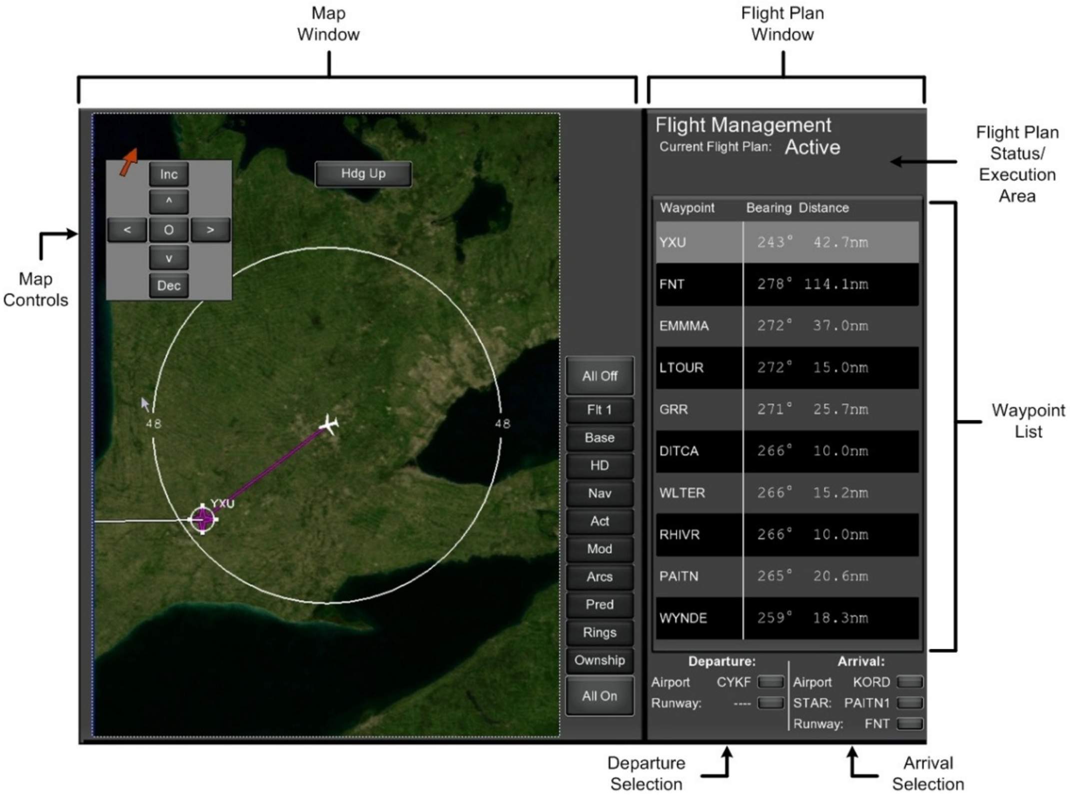


Figure 4-3 – Primary Flight Display, showing co-pilot treatments

4.2.1.2 Interactive Navigation Display

The interactive navigation display, shown in Figure 4-4, was divided into two ‘windows’, with two thirds of the screen showing a satellite map image, including a graphical representation of the flight plan, and one third showing the flight plan, including the departure, waypoint list, and arrival. The map display could also be re-oriented and zoomed in/out to allow pilots to review their current position and planned route.

The navigation display was augmented with a total of nine borders: one visual border for each pilot on both the map and flight plan windows, and one input border on each of the map controls, flight plan status/execution area, waypoint list, departure selection, and arrival selection. In a full implementation, two input borders would be needed for each component; for the purposes of this thesis, only one was needed because all input in the evaluation scenario was performed by the co-pilot. Figure 4-5 shows an image of the navigation display with all borders at full opacity.



Flight Management

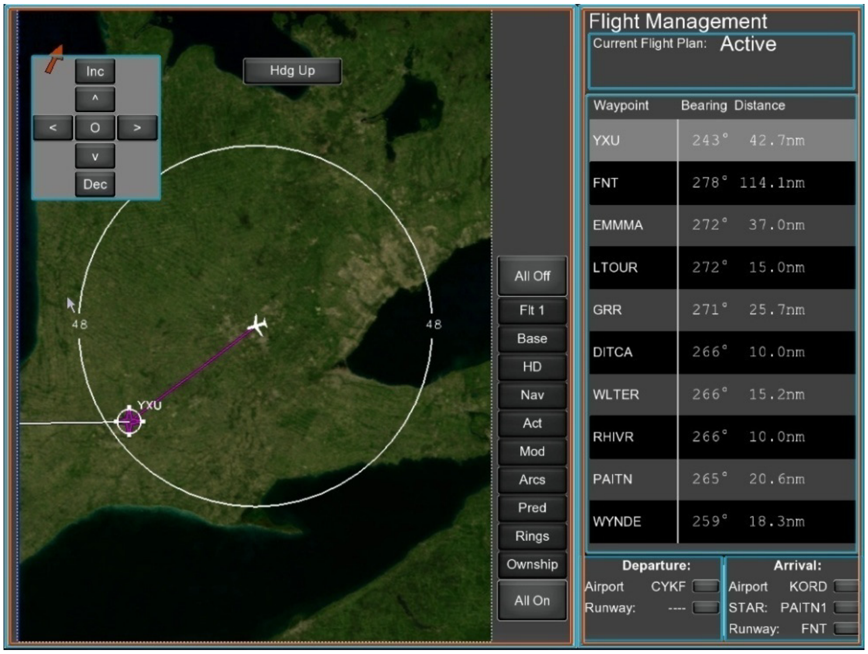
Current Flight Plan: Active

Waypoint	Bearing	Distance
YXU	243°	42.7nm
FNT	278°	114.1nm
EMMMA	272°	37.0nm
LTOUR	272°	15.0nm
GRR	271°	25.7nm
DITCA	266°	10.0nm
WLTER	266°	15.2nm
RHIVR	266°	10.0nm
PAITN	265°	20.6nm
WYNDE	259°	18.3nm

Departure:
Airport CYKF
Runway: ----

Arrival:
Airport KORD
STAR: PAITN1
Runway: FNT

Figure 4-4 – Original Navigation Display



Flight Management

Current Flight Plan: Active

Waypoint	Bearing	Distance
YXU	243°	42.7nm
FNT	278°	114.1nm
EMMMA	272°	37.0nm
LTOUR	272°	15.0nm
GRR	271°	25.7nm
DITCA	266°	10.0nm
WLTER	266°	15.2nm
RHIVR	266°	10.0nm
PAITN	265°	20.6nm
WYNDE	259°	18.3nm

Departure:
Airport CYKF
Runway: ----

Arrival:
Airport KORD
STAR: PAITN1
Runway: FNT

Figure 4-5 – Navigation Display, showing treatments

4.2.1.3 Flight Information/Synoptic Display

The second (lower) shared screen in the cockpit prototype consisted of five individual ‘window’ areas (as shown in Figure 4-6) that displayed weight and center of gravity information, takeoff data, landing data, a fuel system synoptic page, and an electronic checklist area. The electronic checklist system was not implemented in the prototype system, so the checklist window was static throughout the scenario.

The flight information/synoptic display was augmented with a total of eight visual borders (shown at full opacity in Figure 4-7), with one border for each pilot on each window with the exception of the checklist window. No borders were implemented on the checklist window as it was not functional in the prototype system. The flight information/synoptic display did not have any components with input functionality, so no input borders were needed.

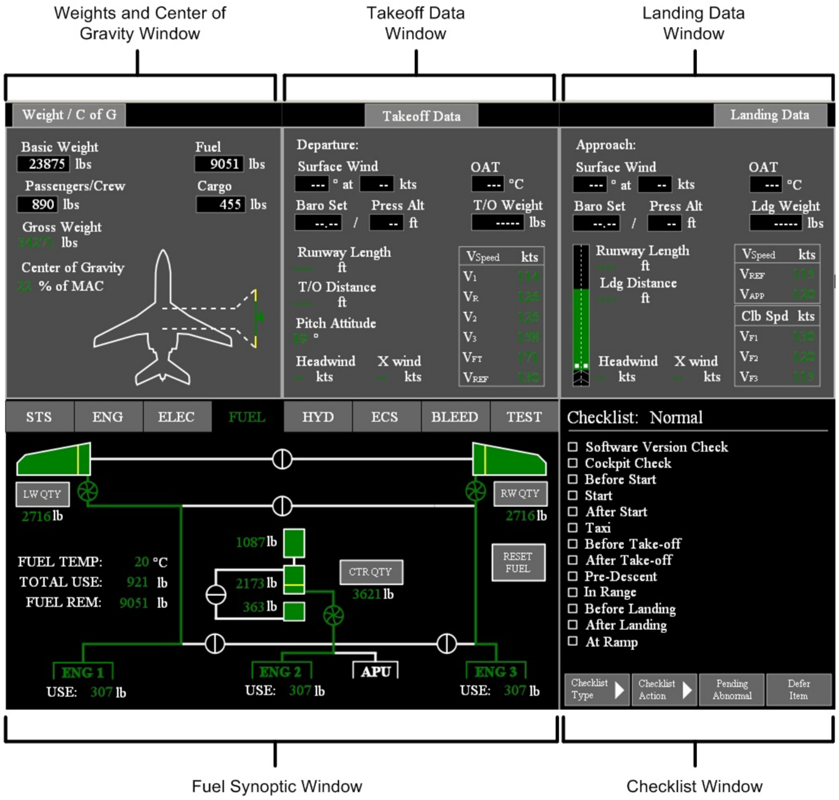


Figure 4-6 – Original Flight Information/Synoptic Display

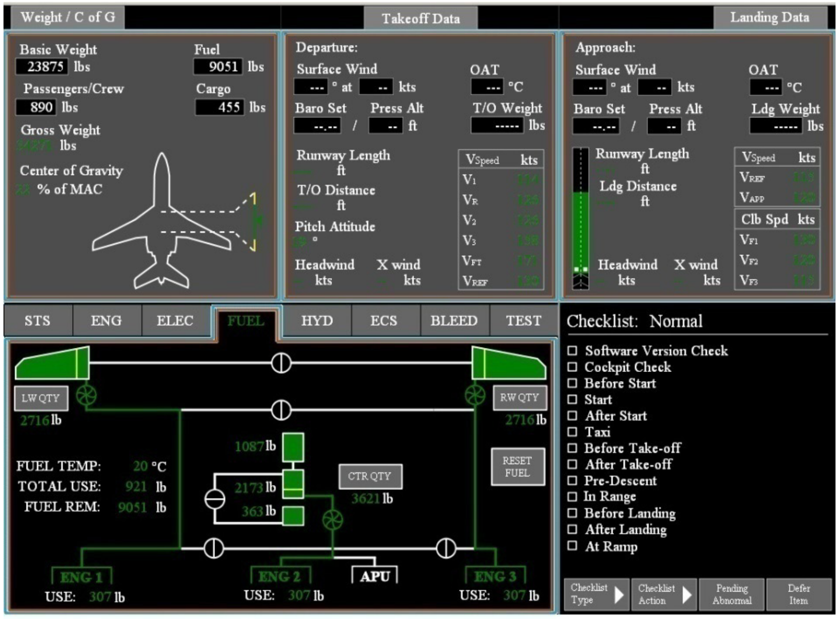
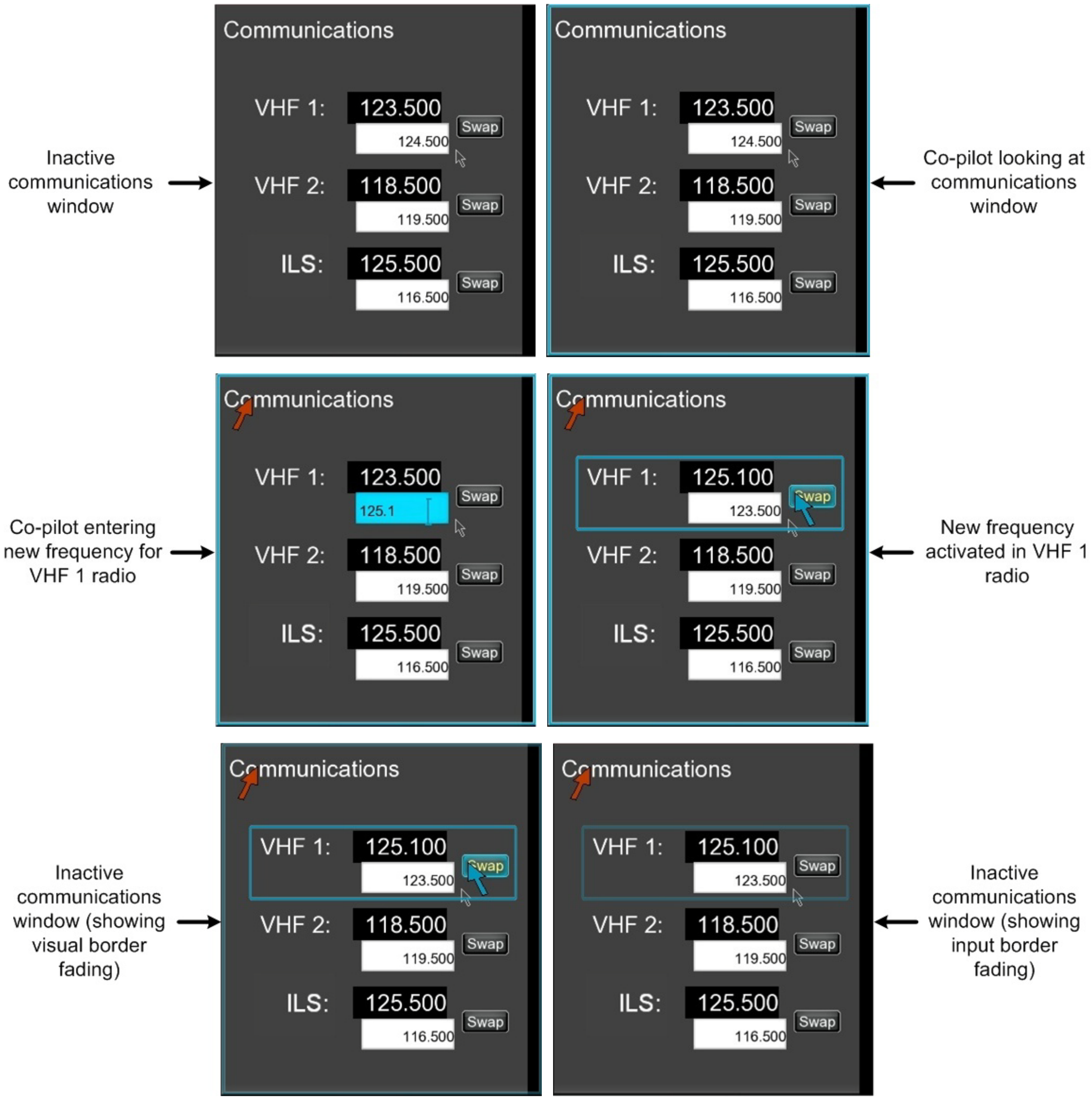


Figure 4-7 – Flight Information/Synoptic Display, showing treatments

4.2.1.4 Example Sequence

To demonstrate the functionality of the borders, the series of images in Figure 4-8 shows the communications window of the co-pilot's PFD as he enters a new frequency into the VHF 1 radio.


Inactive communications window

Co-pilot looking at communications window

Co-pilot entering new frequency for VHF 1 radio

New frequency activated in VHF 1 radio

Inactive communications window (showing visual border fading)

Inactive communications window (showing input border fading)

Figure 4-8 – Communications window input sequence example

Chapter 5

Evaluation Methodology

This chapter describes the novel two-stage evaluation process used to test the utility of the interface design proposed in Chapter 4. Section 5.1 describes the evaluation concept and discusses the hypotheses to be tested, and Sections 5.2 and 5.3 describe the first and second phases of the evaluation process.

5.1 Introduction

In order to evaluate the interface design described in Chapter 4, a two-stage evaluation was necessary. In the first stage, described in detail in Section 5.2, gaze tracking, flight, and communications data were recorded of two pilots flying a representative scenario in a simulated cockpit prototype. In the second stage, described in Section 5.3, the pilots' interaction data was integrated with the screen recordings to create a video based prototype interface (described in Section 4.2.1). This interface was then tested with experienced pilots to evaluate the effectiveness of the augmented interface design. Both stages of the evaluation process were reviewed and approved by the University of Waterloo Office of Research Ethics, and the clearance emails are attached in Appendix C.

The aim of the evaluation process was to answer three main research questions. The first two questions were derived from the original research question from Chapter 1 and the background research described in Chapter 2:

Does the proposed augmented cockpit interface effectively communicate usage history information?

Can this information promote pilot awareness, improve interruption recovery, and mitigate interaction conflicts?

Both of these questions were developed into a set of hypotheses to be tested:

(H1a) It is expected that pilots will be able to detect and understand the usage history information provided by the interface treatment.

(H1b) It is expected that pilots will be able to understand the difference between the treatments indicating visual and input usage.

(H2a) It is expected that the interface treatment will improve pilot awareness of the actions of their crew members.

(H2b) It is expected that the interface treatment will assist pilot interruption recovery.

(H2c) It is expected that the interface treatment will help mitigate interaction conflicts in a cockpit using shared displays.

Because the limitations discussed in Section 4.2 prevented the implementation of an interactive prototype, it was not possible to test the effect of the augmented interface on interaction conflicts; therefore, hypothesis H2c was discarded. Similarly, proving or disproving hypotheses H2a and H2b in the above form would be unlikely with the given design prototype; therefore, they were recast in a form that could be tested in the second stage of the evaluation process:

(H2d) It is expected that the interface treatment will improve participant awareness of the actions of the pilots in the scenario video.

(H2e) It is expected that the interface treatment will assist participant interruption recovery.

In addition, the use of a novel, non-interactive prototype in the evaluation provided the opportunity to examine whether this type of approach could provide useful insights into the effectiveness of an interface design meant for an interactive implementation. It was expected that the results of the evaluation would indicate, at minimum, whether the design concept has the potential to be useful in promoting team awareness and assisting interruption recovery and whether it would be worthwhile, in future, to develop an interactive prototype for further testing.

5.2 Evaluation Phase 1 – Experimental Platform

The first stage of the evaluation process involved developing a representative flight scenario and recording the cockpit displays, input and visual interaction data, and simulated radio and intercom conversation of the three pilots recruited to participate. The three pilots used in the data collection process were recruited as volunteers from the pool of pilots working at the company where the first stage of the evaluation was conducted. Pilots were selected from the pool of volunteers based on their level of experience and on how recently they had been actively flying:

- The pilot-in-command (referred to hereafter as PIC) had over 4000 flying hours (including 1700 hours as a pilot-in-command of both military and civilian multi-engine fixed-wing aircraft) and was then flying as Captain on a Cessna C750 business jet.
- The first officer (co-pilot) (referred to hereafter as FO) was no longer an active pilot, but was a qualified test pilot with over 3800 hours of flight experience when he retired from flying.
- The third pilot (referred to hereafter as ATC) was recruited to simulate the air traffic controllers that would communicate with the pilots throughout the scenario. ATC had also retired from flying, but was a qualified test pilot with over 3600 hours of flight experience at the time of his retirement.

The flight scenario used in this process is detailed in Section 5.2.1, and the cockpit prototype and data recording equipment are described in Section 5.2.2.

5.2.1 Scenario

Three main criteria were used to generate the flight scenario used in the evaluation process; they dictated that the scenario needed to be:

i) sufficiently challenging to stress the collaboration between the two pilots without generating an unreasonably high workload;

ii) long enough to allow time for interesting events to occur, but short enough to allow the participants in the second phase of the study to receive training about the interface, watch the scenario video, and complete their post-scenario data collection in a reasonable amount of time; and

iii) set in a location that would allow the participants in the second phase of the study to have some familiarity with the surroundings.

Using these three criteria, a basic scenario was developed that included an initial instrument flight rules[3] (IFR) flight plan from Waterloo, Ontario (CYKF) to Chicago, Illinois (KORD) and an in-flight re-route to Windsor, Ontario (CYQG). Constraining the pilots to IFR flight and requiring an in-flight re-route provided an unusual event that would stress cockpit collaboration, while ending the scenario when the crew were established in descent to CYQG meant that it lasted a reasonable 33 minutes. The selected airports provided the desired familiarity for local pilots because CYKF is the local airport, CYQG is close by (250 km direct flight), and KORD is commonly known among pilots in North America. With the help of a fourth experienced pilot (at the time, actively flying as a First Officer on Boeing 777 aircraft and with over 6000 total hours as pilot-in-command of both military and civilian aircraft), this basic scenario was developed into a detailed script that was followed during

---

[3] Instrument flight rules are used when flying in low visibility conditions, and require (a) the use of specialized navigation equipment, (b) more detailed and accurate flight plans, and (c) more communication with air traffic controllers.

the data collection process. This detailed scenario script can be found in Appendix D, while the information/consent letter the pilots were required to sign before participating is listed in Appendix E and the IFR flight plan, checklists, and charts provided to them at the beginning of the scenario can be found in Appendix F.

5.2.2 Cockpit Prototype and Data Recording Equipment

The cockpit used in the data collection process (shown in Figure 5-1, below) was an early-stage developmental mock-up of an interface based on the architecture shown in Figure 1-1. It used one keyboard and one trackball-type cursor control device per pilot for interaction with the four 17 in monitors (1024x768 resolution) used for the individual and shared displays, and included a single 24 in widescreen monitor (1920x1200 resolution) for the 'out the window' view. The prototype was run using two software suites; a commercially available flight simulator (X-Plane) to run the flight model and 'out the window' view, and a set of proprietary software (developed by the company that owned the simulator) to run the cockpit displays, flight management, and autopilot.

Figure 5-1 – Prototype Cockpit (cursor control devices out of view at bottom)

The recording equipment used to capture the data from this prototype cockpit included:

- a high-definition digital camcorder (Canon Vixia HG20) with a wide-angle lens (Raynox HD-5000 Pro, 0.5x magnification) allowing a field of view similar to that shown in Figure 5-1;
- three lapel microphones connected to the camcorder (one microphone for each pilot and one for the simulated air traffic controller);
- screen capturing software for all cockpit displays (Beepa® FRAPS for the PFDs and navigation display and TechSmith Camtasia Studio for the flight information/synoptic display and 'out the window' view); and
- a single gaze tracking system, described in Section 5.2.2.1, below.

Due to cost limitations, the only gaze tracking system available for use in the data collection process was the single system owned by the company that developed the cockpit prototype; thus, gaze data was only gathered for one of the two pilots. Because the co-pilot performed all of the input actions in the flight scenario, it was decided that it would be more useful for him to wear the gaze tracking system. However, because the proposed interface designs (described in Section 4.2.1) required point-of-gaze information for both pilots, it was necessary to develop a method to obtain these data for the pilot as well. Point-of-gaze data were approximated by using the video recorded on the digital camcorder (which indicated the screen the pilot was looking at), by interviewing the pilot after the scenario (which helped to determine his normal visual scan patterns), and by watching the screen capture videos (which provided the context of the situation that helped determine what the pilot was most likely to be looking at). Further analysis of the advantages and limitations of using this method (as compared to using a gaze tracking system) is presented in Section 7.2.2.

5.2.2.1 Gaze Tracking Equipment

The gaze tracking equipment used was an Applied Science Laboratories[4] (ASL) Eye-Trac 6000 system mounted on a custom headband, as shown in Figure 5-2. This system uses head mounted infrared optics and an Ascension Technology[5] "Flock of Birds" magnetic head tracker to record eye-gaze position across a user-defined set of planes (in this case, the co-pilot's PFD and the two shared screens). To ensure a sufficient level of accuracy when using the magnetic head tracking system, all mounting components for the transmitter and receiver were built using wood, plastic, or other non-magnetic materials.

Figure 5-2 – ASL Eye-Trac 6000 mounted on a custom headband

Careful configuration and calibration of this system in the prototype cockpit environment ensured that the ASL analysis software could be used to generate a time-stamped log of the co-pilot's eye-fixations, including information about which window he was looking at during each fixation.

5.3 Evaluation Phase 2 – Video Prototype and Experimental Study

The next stage in the evaluation process involved integrating the video and interaction data collected in phase 1 to create video-based interface prototypes of the proposed interface augmentation

---

[4] www.a-s-l.com

[5] www.ascension-tech.com

(described in Sections 5.3.1 and 5.3.2) and using these interface prototypes in a study involving experienced pilots to evaluate the utility of the interface design (described in Sections 5.3.3 through 5.3.7).

5.3.1 Cockpit Interface Video Prototype Development

Using the data recorded in the first phase of the evaluation process, two sets of videos were created: one "control" set with no interface treatments applied, and one "treatment" set augmented with the interface treatment as described in Section 4.2.1. These videos were produced using Adobe Premiere Pro CS4 by creating images of the borders (as shown in Figure 5-3) and superimposing them over the correct positions in the videos (as shown in Section 4.2.1). The Adobe Premiere 'opacity' control was then used to create the fading effect for each border, based on the visual use data (gaze tracking for the co-pilot and approximate for the pilot) and input data (which was obtained by watching the co-pilot's input actions in the screen capture videos).

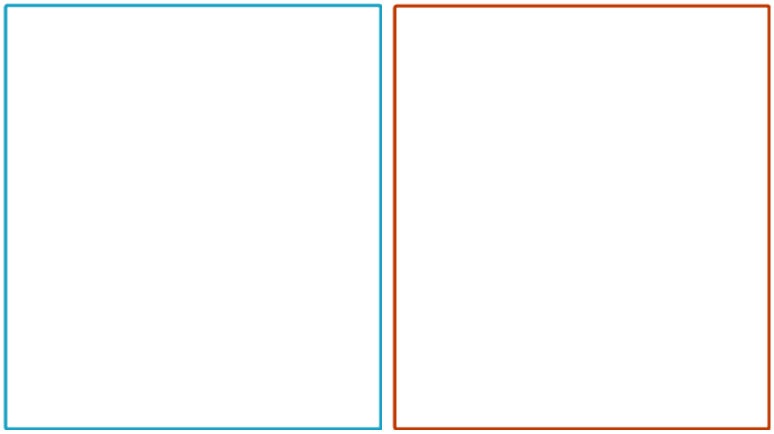


Figure 5-3 – Examples of treatment border images

5.3.2 Experimental Setup

After creating the necessary videos, an experimental display setup was assembled with a form similar to the prototype cockpit shown in Figure 5-1, except with no input devices. This experimental display setup used four 19 in monitors (1280x1024 resolution) for the cockpit displays and one 24 in widescreen monitor (1920x1200 resolution) for the ‘out the window’ view, as shown in Figure 5-4. Audio of the simulated radio conversations was played using standard desktop speakers (not shown in Figure 5-4).

Figure 5-4 – Video Prototype Display Setup

Playback of the videos in the video cockpit prototype was synchronized using a small piece of video playing software (Wallace, 2009) that was able to create multiple instances of Windows Media Player and play, stop, or pause all of them at once using a master control panel. The cockpit display videos were fit to the larger resolution monitors using Windows Media Player by scaling them to the width of the display and leaving blank space at the top and bottom.

5.3.3 Participants

In order to evaluate the interface design using the video-based prototypes, an experimental study was conducted with experienced pilots. Experienced pilots (with a minimum of 15 hours of flight experience) were recruited to help reduce the time required for participant training. The recruitment criteria were not gender specific; however, all eleven pilots that responded and participated in the study were male. The age range of the participants was 19-51 years with a median of 21 and a mean of 28.5. The flight experience of the participants ranged from 80.2-6500 flight hours with a median of 160 and a mean of 756.8. Participants were compensated for their time at a rate of $10 per hour up to a maximum of $20 for two hours.

5.3.4 Experimental Tasks

The participants were given two main tasks to perform while watching the scenario videos:

i) take the role of an evaluator by paying attention to the events of the scenario and the actions of the flight crew and, at the end of the scenario, rating their performance both individually and as a group; and

ii) fill out the scenario log sheet (attached in Appendix G) as completely as possible.

The evaluation task was selected because it provided a continuous cognitive task for the participants; instead of simply watching and listening for cues about log sheet information, they needed to pay attention to the scenario events and integrate them into an overall understanding of the scenario and the performance of the flight crew.

Filling out the scenario log sheet involved recording information about checklists, radio frequency changes, communications with air traffic control, autopilot settings, and flight crew errors. Additional space was provided to allow participants to record any ‘other events’ that they deemed noteworthy. They were also asked to include a time with each recorded piece of data, based on a stopwatch-style timer that displayed the minutes and seconds elapsed since the beginning of the scenario. Asking the participants to fill out this log sheet served two purposes:

i) The completed log sheets became a source of data that provided some insight into how well the participants followed and understood the events of the scenario and the actions of the flight crew (i.e., a measure of situation and team awareness).

ii) Because participants could not watch the displays and record information on the log sheet at the same time, filling out the sheet also served as an interruption to the primary task of watching the scenario videos (i.e., it enabled an examination of the effects of the interface design on interruption recovery).

The major limitation associated with using the log sheets for the purpose of measuring awareness was that they were a self-reported source of data. This meant that the reliability of the log sheet information depended on how well the participants cooperated with the data recording process (i.e., the participants may have perceived and understood more information than they recorded on the scenario log sheet).

5.3.5 Experimental Design

The experiment was a between-subjects design with one independent variable: the interface type (a two level fixed factor). While a within-subjects design likely would have generated more qualitative evaluation data from participants, it would have required the development of a different scenario for each treatment condition, as well as greater amounts of time with each evaluation participant so that they could watch two scenario videos. It also would have introduced the potential for learning effects that could affect participant performance across the two treatment conditions, increasing the complexity of the quantitative evaluation. Using the between-subjects design, participants were randomly assigned to either the control or treatment condition, with a total of five participants assigned to the control condition and six to the treatment condition.

5.3.5.1 Dependent Variables

Several dependent variables (detailed below) were used in this experiment to test for differences between the two participant groups.

Flight crew performance ratings. As part of the post-scenario questionnaire (described in Section 5.3.6), participants were asked to rate the performance of the pilots individually and as a team. This set of questions was included primarily to validate the idea that the participants were acting as flight evaluators, but also provided an opportunity to examine whether the interface treatment had any effect on the perceived performance of the crew.

Scenario log sheet confidence score. The post-scenario questionnaire also included a question that asked participants to rate their confidence in the completeness of their scenario log sheet; this question was included for three reasons:

i) It provided an opportunity for participants to indicate their own level of satisfaction with their recorded notes.

ii) If the participant confidence rating was low, it provided a lead-in for a potential interview question asking what parts of the log sheet the participants felt were incomplete and why.

iii) It enabled testing of whether the interface treatment had an effect on participants' perception of their own level of awareness of the scenario events and actions of the flight crew.

Scenario description confidence score. As part of the interview process (detailed in Section 5.3.6), participants were asked to give a brief description of the scenario events and then rate their confidence in that description on a 7 point scale. This question provided a method of testing whether the interface augmentation had an effect on participants' self-reported perception of their overall understanding of the scenario (i.e., how well the participants thought they understood the scenario).

Scenario log sheet scores. To generate completeness scores for the scenario log sheet, it was necessary to develop a 'master' log sheet (attached in Appendix G) that included all events that could have been recorded for each category to be scored. The categories and their scoring criteria were:

- Checklists, which were considered to have been recorded correctly if participants had a checklist name resembling the correct name and a time ± one minute from the actual time;
- Radio frequency changes, which had one score for recording the correct frequency and one for the correct radio (the recorded times of the events were not considered in the scoring of the radio frequency changes);
- ATC clearances, which were considered to have been recorded correctly if participants had a reasonably accurate description of the clearance and a time ± one minute from the actual time;
- Autopilot setting changes, which had one score for recording a change (irrespective of the time), one for the correct speed setting, and one for the correct altitude setting (the speed and altitude settings were based on the speed and altitude bugs on the PFD); and
- Flight crew errors[6], which were considered to have been recorded correctly if participants had a description of an error that did occur in the scenario with a time ± one minute from the actual time.

---

[6] The flight crew errors present in the scenario were not introduced intentionally, and in most cases resulted from either a lack of realism in the experimental platform or the inability to easily correct errors as a result of using a video prototype. For example, one of the flight crew errors occurred when the co-pilot entered frequency 135.285 in his radio instead of 135.825; in reality, 135.285 is an invalid frequency and the radio would not have accepted it.

5.3.6 Procedure

Before the evaluation began, participants were asked to sign an informed consent form and fill out a background questionnaire (both found in Appendix G). Participants were then seated in front of the prototype display setup (Figure 5-4) and given a verbal explanation of the interface that included:

- a description of the overall concept of operation (a two pilot cockpit using cursor control devices and keyboards, with individual and shared displays);
- details about which colours were used to represent each pilot (orange for left seat, blue for right);
- detailed descriptions of each of the cockpit displays;
- for participants in the treatment condition, an explanation of the interface treatment, including the differences between the input and visual history borders.

Next, the participants were instructed in how to perform the experimental tasks during the scenario (described in Section 5.3.4) and provided with the same flight scenario materials (attached in Appendix F) that were used by the pilots in the scenario. After allowing the participants a few minutes to review these materials, the scenario videos were started. At the conclusion of the scenario, participants were asked to fill out a brief questionnaire (attached in Appendix G) that included three questions asking them to rate the crew's performance and one question asking them to rate their confidence that their scenario log sheet captured all the relevant scenario information. They were asked to answer each question by circling a value on a provided 7 point scale. After completing the post-scenario questionnaire, participants were interviewed using a semi-structured process that attempted to elicit additional qualitative evaluation details by discussing three general topics: the post-scenario questionnaire, the information on the scenario log sheet, and the cockpit interface.

Participants in the treatment condition were asked an additional set of questions dealing specifically with the interface augmentation. The interview question guide is attached in Appendix G.

5.3.7 Data Collection

The experimental data was collected using the scenario log sheet (described in Section 5.3.4), the post-scenario questionnaire (described in Section 5.3.6), and the semi-structured interview (described in Section 5.3.6). Audio and video were also recorded during all evaluation sessions, beginning with the training procedures and ending at the conclusion of the interview. The recording equipment consisted of a high definition digital camcorder (Sony HDR-SR11) connected to two lapel microphones (one for the participant and one for the experimenter). The experimenter also observed the participants during the scenario process and made notes of any interesting events or participant behaviour to be reviewed later using the video recordings. The results of the evaluation process are listed in Chapter 6 and discussed in Chapter 7.

Chapter 6

Results

This chapter presents the results of the evaluation process described in Chapter 5. The quantitative results obtained from the measurement of the dependent variables (described in Section 5.3.5.1) are discussed in Section 6.1, while the qualitative results obtained from all other data sources are discussed in Section 6.2.

6.1 Quantitative Results

In order to determine the impact of display treatment on the quantitative dependent measures, unpaired two-sample t-tests were conducted on the flight crew performance, pilot performance, and co-pilot performance ratings, the scenario log sheet confidence score and scenario description confidence score, and the eight scenario log sheet completeness scores (all detailed in Section 5.3.5.1). The mean and standard deviation of each rating and score for the two participant groups are listed in Table 6-1, below. No statistically significant differences were found between the treatment groups (complete details of the statistical analysis can be found in Appendix H). However, an interesting behavioural phenomenon was observed during the study, related to the scenario log sheet data: pilots with high levels of experience (in the case of this study, one pilot test participant with over 8000 flight hours and one evaluation participant (P4) with 6500) tended to record less information on the log sheet despite having a clear understanding of the scenario events (demonstrated by their responses and discussions during the post-scenario interviews).

Table 6-1 – Means and Standard Deviations for Quantitative Results

Rating or Score	Range	Control Condition		Treatment Condition	
		Mean	Std Dev.	Mean	Std Dev.
Overall crew performance rating	1-7	5.200	0.447	5.500	0.548
Pilot performance rating	1-7	5.200	0.447	5.333	0.816
Co-Pilot performance rating	1-7	5.000	0.707	5.167	0.753
Log sheet confidence rating	1-7	4.300	1.204	5.000	0.894
Scenario description confidence rating	1-7	5.500	0.577	6.200	0.837
Checklist score	/4	3.000	1.225	2.000	2.191
Radio frequency score	/13	9.600	1.673	9.333	1.751
Correct radio score	/13	9.200	1.924	8.000	3.578
ATC clearance score	/8	7.400	0.894	7.500	0.837
Autopilot setting score	/5	4.400	0.894	3.167	2.483
Autopilot altitude setting score	/5	3.200	1.924	3.167	2.483
Autopilot speed setting score	/5	3.400	2.074	2.833	2.317
Flight crew error score	/9	0.800	0.837	0.500	0.837

6.2 Qualitative Results

The qualitative analysis results are described below, with results obtained from observations of the participants performing the experimental task described in Section 6.2.1 and results obtained from the participant interviews described in Section 6.2.2.

6.2.1 Observation of Participants

During the evaluation process, the participants were observed while carrying out their scenario tasks, and notes were made of any interesting events that related to participant awareness or interruption recovery. In particular, two noteworthy events were observed:

- One participant (P4) in the treatment condition, while listening to the pilot's takeoff briefing and following along using the provided flight scenario information, failed to notice the co-pilot making a radio frequency change. When the takeoff briefing was completed, the participant looked up at the displays, noticed that the frequency had been changed, and recorded it on his log sheet.

- One participant (P6) in the control condition missed the first of two consecutive radio frequency changes while examining the flight plan information on the navigation screen. He finished examining the flight plan information in time to notice the second radio frequency change as it was occurring.

6.2.2 Semi-Structured Interviews

Several of the questions used in conducting the semi-structured interview process facilitated some interesting discussions with participants; these questions and the content of the discussions are summarized below. The participant responses and discussions presented here were selected either because they were related to awareness or interruption recovery, or because they provided some insight into the interface design or the evaluation process. Chapter 7 contains a further discussion of the responses and their relation to the design challenges.

Which information did you find easy to keep track of?

Two participants (one in each condition) had interesting responses to this question, both related to radio frequency changes. The participant in the treatment condition mentioned specifically that he found the radio frequency changes easy to track in part because of how the input borders on the radio controls allowed him to re-trace the steps of the pilots. The participant in the control condition said he found the radio frequency changes easy to track overall, but explained that it was harder without being able to see an arm reaching out to the instrument panel.

Were there any particular times or sequences of time during the scenario that you found challenging to track? If so, when and why?

Ten of the eleven participants answered yes to this question, and all of them specified either a high-activity sequence (e.g. takeoff, in-flight re-plan) or simply any sequence where multiple events occurred close together. One treatment condition participant that specified the events up to and including takeoff as challenging also mentioned specifically that during the takeoff briefing, he did not realize where the V speed[7] information was coming from until he noticed the visual usage borders on the takeoff data window of the flight information/synoptic display.

Did you find the interface helpful in tracking what the pilots were doing?

This question was answered affirmatively by all participants in both conditions. Three participants in the treatment condition specifically mentioned the borders; one said he "found [them] really helpful" and another said they were "extremely helpful" and that without them he "almost would have been lost". The third participant, however, mentioned that he felt they could have been more salient as he "had to almost blur [his] vision a little to see [them] popping up".

What could have been added to the interface to make these tasks easier?

Most answers to this question involved adding functionality that would be present on normal aircraft but was not implemented in the cockpit prototype used to record the scenarios (e.g., engine instrumentation or traffic alert and collision avoidance system (TCAS) readouts). However, one participant in the control condition said that he did not think anything needed to be added for pilots, but "from the point of view of just watching it [without the pilots] sitting here, maybe there could have been more cues about who was doing what".

Were you easily able to understand the difference between the "window" (visual) type borders and the "component" (input) type borders?

---

[7] V speeds are reference speeds specific to an aircraft type, labeled as V_x (where x can be a letter, number, or set of letters and numbers). For example, V_1 is the maximum speed at which pilots can abort a takeoff and still be able to stop safely without leaving the runway.

Of the six participants in the treatment condition, five said they were easily able to understand the difference between the visual and input borders (including one who noted that the input borders usually imply the existence of a visual border), and the sixth participant said he probably could have if he had paid more attention to the borders.

Which type of border did you find more useful?

Of the six participants in the treatment condition, two said they found the visual borders more useful, with one noting specifically that they drew his attention to where the pilots' attention was. Three of the four remaining participants indicated that they found the input borders more useful. The final participant who, interestingly, had the most flight experience of all the participants in the evaluation, said he did not pay much attention to the borders.

Did you find the level of detail in the "component" type borders appropriate? Would you have liked more or less?

The five participants in the treatment condition who said they used the borders all said that the level of detail used for the input borders was appropriate. One participant explained his opinion by stating "I guess the border just kind of alerts you to watch what the mouse is doing, so then you're able to see it – it just kind of brings your attention there."

Did your usage or understanding of the borders change over the course of the scenario?

Of the five participants who made use of the borders, one said his usage and understanding did not change during the scenario, one said his understanding was consistent but his usage was higher in the early parts of the scenario, and the remaining three said their usage and understanding increased as the scenario progressed. One participant in particular explained that he "didn't really understand at first, but really got used to it and knew what was happening more with [the borders] than just with my ear and listening to the [pilots] say something." He also explained that the appearance of the borders helped to confirm that an action mentioned by the pilots over the intercom (like a radio frequency change) had actually been completed.

Do you have any other questions or comments about anything (scenario, process, etc)?

Two participants had additional comments about the cockpit prototype that were relevant to the evaluation process. The participant (P4) with the most flight experience (6500 hours) mentioned that the prototype setup was similar to an instrument procedures trainer, which is a simplified cockpit mock-up used to help pilots learn a cockpit's layout, flows, and procedures before moving into a fully representative (and expensive) simulator. Another participant (P5) in the treatment condition noticed a difference in the activity of the visual borders between the pilot and co-pilot; he pointed out that the pilot's gaze didn't seem to "dart around" as much as the co-pilot's (recall that the pilot's gaze was not recorded during phase 1 of the evaluation, but instead was estimated in order to create the video-based prototype for phase 2).

Chapter 7

Discussion

This chapter begins by revisiting the hypotheses presented in Chapter 5 and discussing them in terms of the results described in Chapter 6, and continues with a further discussion explaining the insights that these results provide in terms of both the interface design and the evaluation methodology that are not specifically related to the hypotheses. Finally, recommendations are made for improving both the interface design and the evaluation methodology.

7.1 Hypotheses

As previously discussed in Chapter 5, the research hypotheses for this evaluation process were:

(H1a) It is expected that pilots will be able to detect and understand the usage history information provided by the interface treatment.

(H1b) It is expected that pilots will be able to understand the difference between the treatments indicating visual and input usage.

(H2d) It is expected that the interface treatment will improve participant awareness of the actions of the pilots in the scenario video.

(H2e) It is expected that the interface treatment will assist participant interruption recovery.

Each of these hypotheses and the results relevant to them are discussed in one of the following two sections, with hypotheses H1a and H1b in Section 7.1.1 and H2d and H2e in Section 7.1.2.

7.1.1 Perceptibility and Attention Effects of Interface Augmentation

The results of the semi-structured interviews provide strong evidence that the participants were able to both detect and understand the information provided by the interface treatment (supporting hypotheses H1a and H1b). In fact, many of the participants in the treatment condition mentioned that the appearance of borders on the interface tended to draw their attention, suggesting that the visual augmentation was quite easily detectable. This result is not surprising, as research shows that visual attention can be attracted by stimuli that appear suddenly, particularly if these stimuli appear in peripheral areas (Wickens and Hollands, 2000, p. 75). While this result does indicate that the design was at least somewhat successful at providing peripheral awareness information, capturing pilot attention in a cockpit (whether intentionally or unintentionally) can have important safety implications. Therefore, the effects of the interface treatment on pilot attention allocation would need to be further studied in an interactive context before the treatment is considered for implementation for live cockpit operations.

The participants' understanding of the information provided by the interface treatment is demonstrated in a variety of participant comments indicating the different ways in which they used the information. For example, one participant reported that he understood that the input borders could be used as a source of confirmation that an action that the pilots had talked about had actually been completed on the interface; this demonstrates that the participant understood the input borders. Another participant demonstrated his understanding of the visual borders by commenting that he used them as a way of identifying areas of the interface where the pilots were focusing their attention.

The participants were also generally affirmative in their responses to the question that asked if they were able to understand the difference between the input and visual borders. While these responses do not necessarily confirm the participants' understanding of the difference (because they were self-evaluations of understanding), other participant comments explaining the ways in which they used the different borders demonstrated that the difference was well understood. For example, one participant commented that when input borders appeared, it usually implied the existence of a visual border as well (because the pilot would generally be looking at his input).

7.1.2 Utility of Interface Augmentation

The general utility of the interface augmentation to the participants in the evaluation is evident in the comments of participants from both the treatment and control conditions. For example, several participants in the treatment condition specifically mentioned the value of the borders in the discussions about how well the interface supported their understanding of what the pilots in the scenario were doing, including one who said without the borders he “almost would have been lost”. Additionally, when asked what could have been added to the interface to have made his tasks easier, one of the participants in the control condition suggested that it may have helped to have had more cues about what each of the pilots were doing, which is one of the types of information that the borders seemed to provide very effectively.

In terms of the specific hypothesis H2d (that the interface treatment will improve participant awareness of the actions of the pilots), there was ample evidence that this support was needed and at least some evidence suggesting the interface could provide it. One example supporting the need for awareness is provided by the participant (P6) in the control condition who missed a radio frequency change while viewing the navigation display; if this participant had been in the treatment condition, it is possible (perhaps even likely, given the number of comments about the borders drawing attention) that the appearance of the input border would have provided a sufficient peripheral awareness cue for him to notice the frequency change. Similarly, another control condition participant described the difficulty of tracking radio frequency changes without being able to see an arm reaching out to the instrument panel; this is exactly the sort of peripheral awareness information that the interface augmentation attempted to provide.

One previously mentioned example of the treatment condition participant (P4) who missed a radio frequency change while looking at a chart, but was able to notice and record the change when he looked back up at the display, suggests that the interface treatment can support awareness as it helped the participant gain awareness of the radio frequency change. This example also provides some support for hypothesis H2e (that the interface treatment will assist participant interruption recovery), as the participant was able to recover from the interruption and record the frequency change. Based on the previously discussed comments about the utility of the augmentation, the border was almost certainly a contributing factor in this participant's recovery from the interruption.

7.2 Further Insights

Apart from the results that specifically addressed the hypotheses (discussed in Section 7.1 above), a number of results provided some unexpected insights into the interface design (described in Section 7.2.1) or the evaluation process (described in Section 7.2.2).

7.2.1 Design Insights

The first insight into the interface design that became evident throughout the evaluation process was that the participants seemed to use the augmentation more as an indication of what was currently happening than as a way of keeping track of what had happened in the recent past. There are several possible explanations for this effect:

- as a result of the attention capturing effect of the borders appearing, combined with the relatively fast eye movements of the pilots, the participants' attention may have become focused primarily on the actively appearing visual borders;
- the relatively limited design process for the fade rates and the fact that those rates were not adjusted based on context (as discussed in Section 4.2) may mean that the fade duration of the input borders was too brief for some situations;
- the participants may not have attempted to use the borders as a source of history information due to the nature of the experimental task, which required them to record events as they occurred and not retrospectively.

It is likely that the participants' use of the interface augmentation as an indication of current events resulted from some combination of these three effects (and potentially others).

A second design insight was provided by the participants' generally positive responses to the question asking whether the level of detail used for the input borders was appropriate. Although the participants did not experience alternative interface designs, this result does suggest that the input borders worked reasonably well the way they were implemented, and could be a good starting point for further investigations of this interface design solution.

Another interesting effect that was observed during the evaluation process relates to the experience level of the participants. Participants with high levels of flight experience did not seem to make use of the interface augmentation at the same level as those with less experience. In fact, the participant with the most flight experience stated that he paid little attention to the borders, while another highly experienced pilot who served as a pre-evaluation test participant mentioned that he found the augmentation had a tendency to distract him from his normal cockpit scan pattern. The feedback from pilots with lower levels of experience (15-300 flight hours), on the other hand, was almost universally positive. It is likely that this effect is in part a result of the force of habit in the more experienced pilots; because they are so strongly trained on current systems, it is more difficult for them to accept the new interface. This could likely be addressed by using a longer training time for more experienced pilots, allowing them to adjust from their conventional approach and use the new system effectively. This experience effect also suggests that the interface augmentation may be particularly effective as a training tool; an idea that is also supported by the previously mentioned tendency of participants to use the interface augmentation to follow current events. Creating augmented video using point-of-gaze data from expert pilots could help in demonstrating effective cockpit scan patterns to pilots in training. In fact, the comment from participant P4 that the prototype cockpit was similar to an instrument procedures trainer suggests that the experimental platform used in this research could be turned into a training tool with relatively little modification.

7.2.2 Evaluation Process Insights

The level of experience of the participants was also the source of an important insight about the evaluation process, as the more highly experienced pilots had a tendency to record much less information on the scenario log sheet in spite of having a clear understanding of the scenario events (demonstrated by their responses and discussions during the post-scenario interviews). This result was a manifestation of the limitations of using a self-reported method of data collection (as discussed in Section 5.3.4) and could be explained by a number of factors. One possible explanation is that experienced pilots may have decided not to record the information on the sheet because they had a better understanding of the scenario context and could remember details if needed. It is also possible that because that one of the experimental tasks given to the participants was to act as flight evaluators, the experienced pilots may have been paying less attention to the log sheet so that they could focus more on watching and evaluating the actions of the flight crew, while the participants with less experience were not as confident in their ability to evaluate the flight crew and instead focused on the log sheet task.

A second valuable insight gleaned from the evaluation process is related to the decision to create approximate data for the pilot's point-of-gaze (as discussed in Section 5.2.2). It was initially unclear whether this approximate data might present a significant limitation to the evaluation process if participants were easily able to notice a difference in activity between the pilot and co-pilot visual borders. If this difference was severe enough, the participants might distrust the interface treatment and begin to ignore it. However, of the six participants in the treatment condition, only one (P5) mentioned having noticed a difference between the two, and it was limited to an observation that the co-pilot's eye movements seemed to "dart around" more than the pilot's. Experiences gained from conducting this evaluation indicate that obtaining real gaze data for use in the prototype did have some benefits in that it made the process of prototyping the visual borders somewhat faster and ensured that they were accurate representations of the co-pilot's eye activity, but these benefits came at a cost in terms of the time required to set up and calibrate the gaze tracker and record and analyze the point-of-gaze data. It is suspected that using only the approximate method would have greatly accelerated the first phase of the evaluation process with little cost to the realism of the prototype.

Another point that became clear during the evaluation process was that the cockpit prototype used as a base for the interface augmentations could have benefit from more time spent in the first phase of the evaluation process. As a result of the limited, developmental implementation, the training process for second phase participants required nearly as much time to be spent explaining parts of the interface that were not functional as was spent explaining parts that were. Additionally, in response to the question asking what could have been added to the interface, most of the participants commented about items that would have been present in a more fully implemented prototype. Depending on the complexity of software development needed, further implementation of the system software for such complex task domains (if possible) may help to reduce the total time needed to complete the entire evaluation process. In particular, it would have definite benefits in increasing the realism of the scenario for experienced domain participants.

In summary, the non-interactive process provided a significant amount of qualitative insight into some of the potential benefits and drawbacks of the design. In addition, it involved significantly less time, cost, and computer expertise than would have been required for the development of a fully interactive prototype.

7.3 Recommendations

7.3.1 Design Recommendations

Based on the design process used to create the augmented interface concept and results of the evaluation process, the following concepts were developed as potential design improvements:

- Based on the fact that the treatments tended to capture attention as a result of their sudden appearance, the idea of implementing a 'fade-in' for the visual treatments should be considered as a way of reducing their tendency to overwhelm the input treatments.
- When developing a treatment to display input history information on an interface, showing this information based on functional groups (as opposed to individual components) is likely an effective way of reducing interface clutter while providing most of the potential benefits.
- As a way of increasing the potential for the input treatments to be a useful source of history information, their fade duration should be increased. Further investigation is needed to determine the appropriate increase.

7.3.2 Evaluation Methodology Recommendations

In addition to the design improvements suggested above, the evaluation process used in this thesis could also be improved in several ways:

- The cockpit prototype used as a base for the treatments should be more fully developed to increase the realism of the evaluation scenario. In particular, implementation of elements such as the autopilot mode control panel and flight mode annunciators, the engine instruments, and the traffic collision avoidance system display would bring the prototype closer to simulating an operational cockpit.

- Participants with higher levels of experience may benefit from longer training times, which would allow them to develop a better understanding of how to use the information available in the new interface as an effective part of their conventional scan patterns.
- The data collection process used for participants watching the scenario could include a more direct evaluation of awareness, perhaps using a variant of one of the situation awareness evaluation techniques mentioned in Section 2.2.
- This data collection process could also include a more direct evaluation of participant interruption recovery performance by creating a situation in the scenario where an event (such as a radio frequency change) occurs while the participants are interrupted (i.e., recording another item on the log sheet). This would also help in creating a situation in which the input treatments would see use as a source of history data instead of an indicator of current events.
- For evaluations of this type using non-interactive prototypes of systems that make use of point-of-gaze data, preparation time could be reduced by creating approximate gaze tracking data instead of recording actual data.

Chapter 8

Conclusions

The recent emergence of a cockpit architecture that uses both individual and shared displays, with input accomplished through cursor control devices and keyboards, has allowed for the creation of cockpit interface designs that have a number of notable advantages over older style cockpits. These improvements, however, come with a price in the form of design challenges that need to be addressed in order for new cockpits to take full advantage of the flexibility offered by the shared display architecture. This research was motivated by a desire to address some of these important design challenges, including the loss of peripheral awareness information, the need to support interruption recovery, and the potential for interaction conflicts.

The augmented cockpit interface design described in this thesis used visualizations of operator usage history in an effort to address these challenges and improve collaboration and performance in the cockpit. An evaluation of this concept, using a non-interactive prototype, provided sufficient evidence of the potential of this concept to support crewmember awareness and interruption recovery to merit further study, particularly as a possible tool for pilot training.

8.1 Research Objectives and Findings

The objectives of this research were to examine current practices for effective collaboration in cockpit environments, design an improved cockpit interface for use in modern (new architecture) cockpits, and evaluate the utility of the improved interface for addressing the identified design challenges. The first research objective was addressed by reviewing cockpit and collaborative design literature (detailed in Chapter 2) and conducting an operational sequence modeling analysis of a representative flight scenario (described in Chapter 3). This process identified three major challenges for the design of modern cockpits: supporting awareness, assisting interruption recovery, and mitigating interaction conflicts.

The second research objective was addressed in Chapter 4, which took advantage of existing cognitive research and visual design concepts to propose ways in which these challenges could be addressed by augmenting the cockpit with visualizations of operator interaction history. Namely, the concept proposed to provide pilots with peripheral awareness of each other's actions, to provide pilots returning from an interruption with information about where they were previously working and what happened while they were interrupted, and to allow pilots to easily see where their crew member is working and use this information to avoid interaction conflicts.

Finally, the third research objective was addressed in Chapters 5-7 by evaluating the proposed design concept using a two-stage process that culminated with a human participant study of the non-interactive video prototype. This evaluation process was an initial step in determining whether the proposed interface design addressed the three design challenges in a cockpit context. The evaluation results provided important insights into the design, including an indication that the design solution has sufficient potential to merit further consideration and evaluation using a fully interactive prototype.

8.2 Recommendations and Future Work

Based on the results of this thesis, a number of recommendations can be made for potential future work related to the proposed design concept. First, given that the proposed design concept showed potential to be useful in addressing the identified design challenges, further research should be done to determine the most effective form for the display treatments. This should include an examination of the basic treatment (i.e., what basic treatments could be used instead of borders, and which treatment is most effective?), a comparison of different fade durations (including context-based dynamic durations), and testing of the whether the concept of using thickness to indicate importance can be implemented and is useful.

Future studies of the design concept should also include the development of an interactive prototype to address the limitations of the non-interactive evaluation method (i.e., to examine the utility of the concept for live cockpit operations, and to further investigate the potential issues of distractions raised by the evaluation in this thesis).

Based on the results indicating that the interface treatment may show the most benefit for pilots with lower levels of experience, the idea of developing and evaluating the interface treatment as part of a focused training tool should also be considered as an option for future work.

Finally, the concept of using gaze tracking systems in the cockpit should be studied further to include other potential applications. Examples of such applications could include accident and incident investigation (i.e., adding point-of-gaze as part of the data set for flight data recorders so that investigators could establish where pilots were looking during an accident or incident) or “smart” cockpits (e.g., if the cockpit knows where the pilots are looking, can it begin to guess their intentions and make suggestions about possible control actions?).

References

U.S. Department of Defense, 1999. Department of Defense Design Criteria Standard - Human Engineering, MIL-STD-1472F.

Altmann, E. M. and Trafton, J. G., 2004. Task Interruption: Resumption Lag and the Role of Cues. in *Proceedings of 26th Annual Conference of the Cognitive Science Society* (Chicago, Illinois, USA), Lawrence Erlbaum, Mahwah, New Jersey, USA.

Bailey, B. P., Konstan, J. A., and Carlis, J. V., 2001. The Effects of Interruptions on Task Performance, Annoyance, and Anxiety in the User Interface. in *Proceedings of INTERACT 2001: 8th IFIP International Conference on Human-Computer Interaction* (Tokyo, Japan), IOS Press, Amsterdam, The Netherlands.

Bertin, J., 1983. *Semiology of graphics: diagrams, networks, maps*. original publication 1967, translated by Berg, W. J. University of Wisconsin Press, Madison, Wisconsin, USA.

Borah, J., 1995. Investigation of Eye and Head Controlled Cursor Positioning Techniques *AL/CF-SR-1995-0018*, Armstrong Laboratory, Crew Systems Directorate, Air Force Material Command, Wright-Patterson Air Force Base, Ohio, USA.

Brignull, H., Izadi, S., Fitzpatrick, G., Rogers, Y., and Rodden, T., 2004. The introduction of a shared interactive surface into a communal space. in *Proceedings of CSCW 2004: ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA), ACM, New York, New York, USA.

Bruni, S., Marquez, J. J., Brzezinski, A., Nehme, C., and Boussemart, Y., 2007. Introducing a Human-Automation Collaboration Taxonomy (HACT) in Command and Control Decision-Support Systems. in *Proceedings of the 12th International Command and Control Research and Technology Symposium* (Newport, Rhode Island, USA).

Cadiz, J., Venolia, G., Jancke, G., and Gupta, A., 2002. Designing and Deploying an Information Awareness Interface *MSR-TR-2002-87*, Microsoft Research, Redmond, Washington, USA.

Carpendale, S., 2003. Considering Visual Variables as a Basis for Information Visualisation. Department of Computer Science, University of Calgary, Calgary, Alberta, Canada.

Chapanis, A., 1996. *Human Factors in Systems Engineering*. John Wiley & Sons, Inc., New York, New York, USA.

Cooke, N. J., Stout, R. J., and Salas, E., 2001. A knowledge elicitation approach to the measurement of team situation awareness. in McNeese, M., Endsley, M. R. and Salas, E. eds. *New Trends in Cooperative Activities: System Dynamics in Complex Settings*, Human Factors & Ergonomics Society, Santa Monica, California, USA, 114-139.

Dietz, P. and Leigh, D., 2001. DiamondTouch: a multi-user touch technology. in *Proceedings of the 14th annual ACM symposium on User interface software and technology* (Orlando, Florida), ACM, New York, New York, USA.

Dismukes, R. K., Young, G., and Sumwalt, R., 1998. Cockpit interruptions and distractions: Effective management requires a careful balancing act. *ASRS Directline* (10). Available at http://asrs.arc.nasa.gov/publications/directline/dl10_distract.htm.

Endsley, M. R., 1987. SAGAT: A methodology for the measurement of situation awareness. NOR DOC 87-83, Northrop Corporation, Hawthorne, California, USA.

Endsley, M. R., 1988. Design and evaluation for situation awareness enhancement. in *Proceedings of the Human Factors Society 32nd Annual Meeting*, Human Factors Society, Santa Monica, California, USA.

Endsley, M. R., 1999. Situation awareness in aviation systems. in Garland, D. J., Wise, J. A. and Hopkin, V. D. eds. *Human factors in aviation systems*, Lawrence Erlbaum, Hillsdale, New Jersey, USA, 257-276.

Endsley, M. R., 2000. Theoretical Underpinnings of Situation Awareness: A Critical Review. in Endsley, M. R. and Garland, D. J. eds. *Situation Awareness Analysis and Measurement*, Lawrence Erlbaum Associates, Mahwah, New Jersey, USA.

Finlayson, D. P., 2005. Combined bathymetry and topography of the Puget Lowland, Washington State, University of Washington School of Oceanography.

Foushee, H. C., Lauber, J. K., Baetge, M. M., and Acomb, D. B., 1986. Crew factors in flight operations: Part 3. The operational significance of exposure to short-haul air transport operations (Technical Memorandum 88322). NASA Ames Research Center, Moffett Field, California, USA.

Grudin, J., 1994. Groupware and social dynamics: eight challenges for developers. *Communications of the ACM*, *37* (1). 92-105.

Gutwin, C. and Greenberg, S., 1998. Design for Individuals, Design for Groups: Tradeoffs between Power and Workspace Awareness. in *Proceedings of CSCW '98: ACM Conference on Computer-Supported Cooperative Work* (Seattle, Washington, USA), ACM, New York, New York, USA.

Han, J. Y., 2005. Low-cost multi-touch sensing through frustrated total internal reflection. in *Proceedings of the 18th annual ACM symposium on User interface software and technology* (Seattle, WA, USA), ACM, New York, New York, USA.

Hill, J. and Gutwin, C., 2004. The MAUI Toolkit: Groupware Widgets for Group Awareness. *Computer Supported Cooperative Work*, *13*. 539-571.

Hill, W. C., Hollan, J. D., Wroblewski, D., and McCandless, T., 1992. Edit Wear and Read Wear. in *Proceedings of CHI 92: Conference on Human Factors in Computer Systems* (Monterey, California, USA), ACM, New York, New York, USA.

Hutchins, E., 1990. The technology of team navigation. in Galegher, J., Kraut, R. E. and Egido, C. eds. *Intellectual Teamwork: Social and technological foundations of cooperative work*, Lawrence Erlbaum Associates, Hillsdale, NJ, 191-220.

Hutchins, E., 1995. How a cockpit remembers its speeds. *Cognitive Science*, *19*. 265-288.

Hutchins, E. and Klausen, T., 1996. Distributed Cognition in an Airline Cockpit. in Engeström, Y. and Middleton, D. eds. *Cognition and communication at work.*, Cambridge University Press, New York, 15-34.

Jacob, R. J. K., 1990. What You Look At Is What You Get: Eye Movement-Based Interaction Techniques. in *Proceedings of CHI 90: Conference on Human Factors in Computer Systems* (Seattle, Washington, USA), ACM, New York, New York, USA.

Johanson, B., Hutchins, G., Winograd, T., and Stone, M., 2002. PointRight: experience with flexible input redirection in interactive workspaces. in *Proceedings of the 15th annual ACM symposium on User interface software and technology* (Paris, France), ACM, New York, New York, USA.

Kammerer, Y., Schleiter, K., and Beinhauer, W., 2008. Looking my Way through the Menu: The Impact of Menu Design and Multimodal Input on Gaze-based Menu Selection. in *Proceedings of ETRA 2008: Eye Tracking Research and Application Symposium* (Savannah, Georgia, USA), ACM, New York, New York, USA.

Kramer, L. J., Prinzel III, L. J., Bailey, R. E., Arthur III, J. J., and Parrish, R. V., 2004. Flight Test Evaluation of Synthetic Vision Concepts at a Terrain Challenged Airport, NASA/TP-2004-212997.

Kumar, M., Paepcke, A., and Winograd, T., 2007. EyePoint: Practical Pointing and Selection Using Gaze and Keyboard. in *Proceedings of CHI 2007: Conference on Human Factors in Computer Systems* (San Jose, California, USA), ACM, New York, New York, USA.

Lee, J. D., 2008. Review of a Pivotal Human Factors Article: "Humans and Automation: Use, Misuse, Disuse, Abuse". *Human Factors*, *50* (3). 404-410.

Loukopoulos, L. D., Dismukes, R. K., and Barshi, I., 2001. Cockpit Interruptions and Distractions: A Line Observation Study. in *Proceedings of the 11th International Symposium on Aviation Psychology* (Columbus, Ohio, USA), The Ohio State University, Columbus, Ohio, USA.

McFarlane, D. C., 1999. Coordinating the Interruption of People in Human-Computer Interaction. in *Proceedings of INTERACT 99: 7th IFIP International Conference on Human-Computer Interaction* (Edinburgh, Scotland, UK), IOS Press, Amsterdam, The Netherlands.

Morris, M. R., Ryall, K., Shen, C., Forlines, C., and Vernier, F., 2004. Beyond "social protocols": multi-user coordination policies for co-located groupware. in *Proceedings of CSCW 2004: ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA), ACM, New York, New York, USA.

Nomura, S., Hutchins, E., and Holder, B. E., 2006. The uses of paper in commercial airline flight operations. in *Proceedings of CSCW 2006: ACM Conference on Computer Supported Cooperative Work* (Banff, Alberta, Canada), ACM, New York, New York, USA.

NTSB, 1988. Aircraft Accident Report--Northwest Airlines, Inc., McDonnell Douglas DC-9-82, N312RC, Detroit Metropolitan Wayne County Airport, Romulus, Michigan, August 16,1987 *NTSB/AAR-88/05*, National Transportation Safety Board.

Parasuraman, R. and Riley, V., 1997. Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*. 250-253.

Perry, M., 2003. Distributed Cognition. in Carroll, J. M. ed. *HCI Models, Theories, and Frameworks: Toward an Interdisciplinary Science*, Morgan Kaufmann, San Francisco, 193-223.

Rekimoto, J., 1997. Pick-and-drop: a direct manipulation technique for multiple computer environments. in *Proceedings of the 10th annual ACM symposium on User interface software and technology* (Banff, Alberta, Canada), ACM, New York, New York, USA.

Salas, E., Prince, C., Baker, D. P., and Shrestha, L., 1995. Situation Awareness in Team Performance: Implications for Measurement and Training. *Human Factors*, *37* (1). 123-136.

Scott, S. D., Mercier, S., Cummings, M. L., and Wang, E., 2006. Assisting Interruption Recovery in Supervisory Control of Multiple UAVs. in *Proceedings of the Human Factors and Ergonomic Society 50th Annual Meeting* (San Francisco, California, USA), Human Factors and Ergonomic Society, Santa Monica, California, USA.

Scott, S. D., Wan, J., Sasangohar, F., and Cummings, M. L., 2008. Mitigating Supervisory-level Interruptions in Mission Control Operations. in *Proceedings of 2nd International Conference on Applied Human Factors and Ergonomics* (Las Vegas, Nevada, USA), (CD-Proceedings), Las Vegas, Nevada, USA.

Sheridan, T. B. and Verplank, W., 1978. Human and Computer Control of Undersea Teleoperators. Man Machine Systems Laboratory, Department of Mechanical Engineering, MIT, Cambridge, Massachusetts, USA.

Smallman, H. S. and St. John, M., 2003. CHEX (Change History EXplicit): New HCI Concepts for Change Awareness. in *Proceedings of the Human Factors and Ergonomics Society 47th Annual Meeting*, Human Factors and Ergonomics Society, Santa Monica, California, USA.

St. John, M., Smallman, H. S., and Manes, D. I., 2005. Recovery from interruptions to a dynamic monitoring task: The beguiling utility of instant replay. in *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, Human Factors and Ergonomics Society, Santa Monica, California, USA.

Stewart, J., Bederson, B., and Druin, A., 1999. Single display groupware: a model for co-present collaboration. in *Proceedings of CHI 99: Conference on Human Factors in Computer Systems* (Pittsburgh, Pennsylvania, USA), ACM, New York, New York, USA.

Streitz, N. A., Jörg, G. l., Holmer, T., Konomi, S. i., Christian, M.-T., Reischl, W., Rexroth, P., Seitz, P., and Steinmetz, R., 1999. i-LAND: an interactive landscape for creativity and innovation. in *Proceedings of the SIGCHI conference on Human factors in computing systems: the CHI is the limit* (Pittsburgh, Pennsylvania, United States), ACM, New York, New York, USA.

Taylor, R. M., 1990. Situational awareness rating technique (SART): The development of a tool for aircrew systems design. in *Situational Awareness in Aerospace Operations* (AGARD-CP-478),NATO-AGARD, Neuilly Sur Seine, France.

Toups, Z. O. and Kerne, A., 2007. Implicit coordination in firefighting practice: design implications for teaching fire emergency responders. in *Proceedings of CHI 2007: Conference on Human Factors in Computer Systems* (San Jose, California, USA), ACM, New York, New York, USA.

Trafton, J. G., Altmann, E. M., and Brock, D. P., 2005. Huh, what was I doing? How people use environmental cues after an interruption. in *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, Human Factors and Ergonomics Society, Santa Monica, California, USA.

Trafton, J. G., Altmann, E. M., Brock, D. P., and Mintz, F. E., 2003. Preparing to resume an interrupted task: effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human-Computer Studies*, *58*. 583-603.

Tse, E., Histon, J., Scott, S. D., and Greenberg, S., 2004. Avoiding interference: how people use spatial separation and partitioning in SDG workspaces. in *Proceedings of CSCW 2004: ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA), ACM, New York, New York, USA.

Vidulich, M. A. and Hughes, E. R., 1991. Testing a subjective metric of situation awareness. in *Proceedings of the Human Factors and Ergonomics Society 35th Annual Meeting*, Human Factors and Ergonomics Society, Santa Monica, California, USA.

Wallace, J., 2009. Multiplayer Video Software, University of Waterloo, Waterloo, Ontario, Canada.

Wickens, C. D. and Hollands, J. G., 2000. *Engineering Psychology and Human Performance, 3rd Edition*. Prentice Hall, Upper Saddle River, New Jersey, USA.

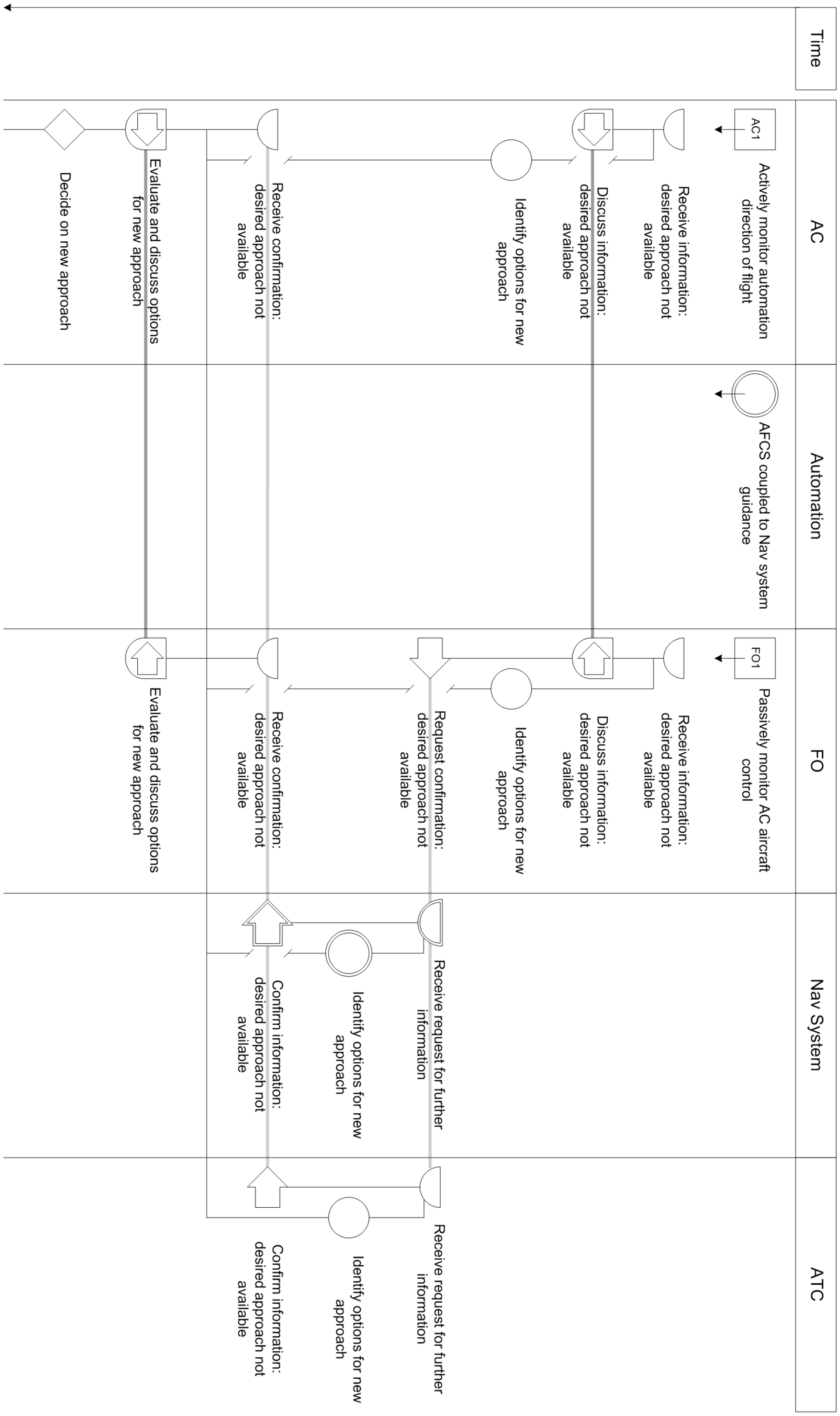
Wiener, E. L., Kanki, B. G., and Helmreich, R. L., 1993. *Cockpit resource management*. Academic Press, San Diego, California, USA.

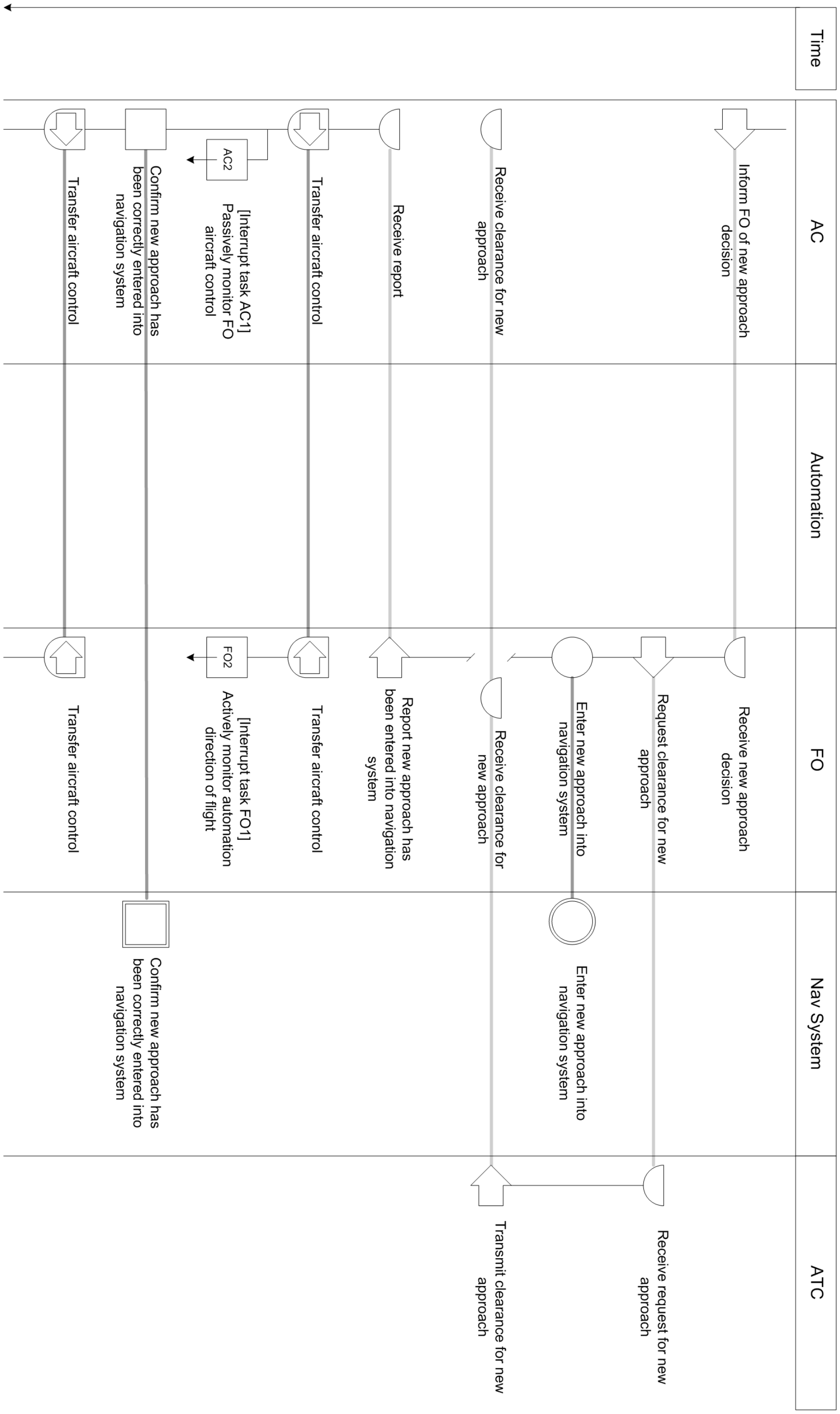
Yankelovich, N., Walker, W., Roberts, P., Wessler, M., Kaplan, J., and Provino, J., 2004. Meeting central: making distributed meetings more effective. in *Proceedings of CSCW 2004: ACM Conference on Computer Supported Cooperative Work* (Chicago, Illinois, USA), ACM, New York, NY, USA.

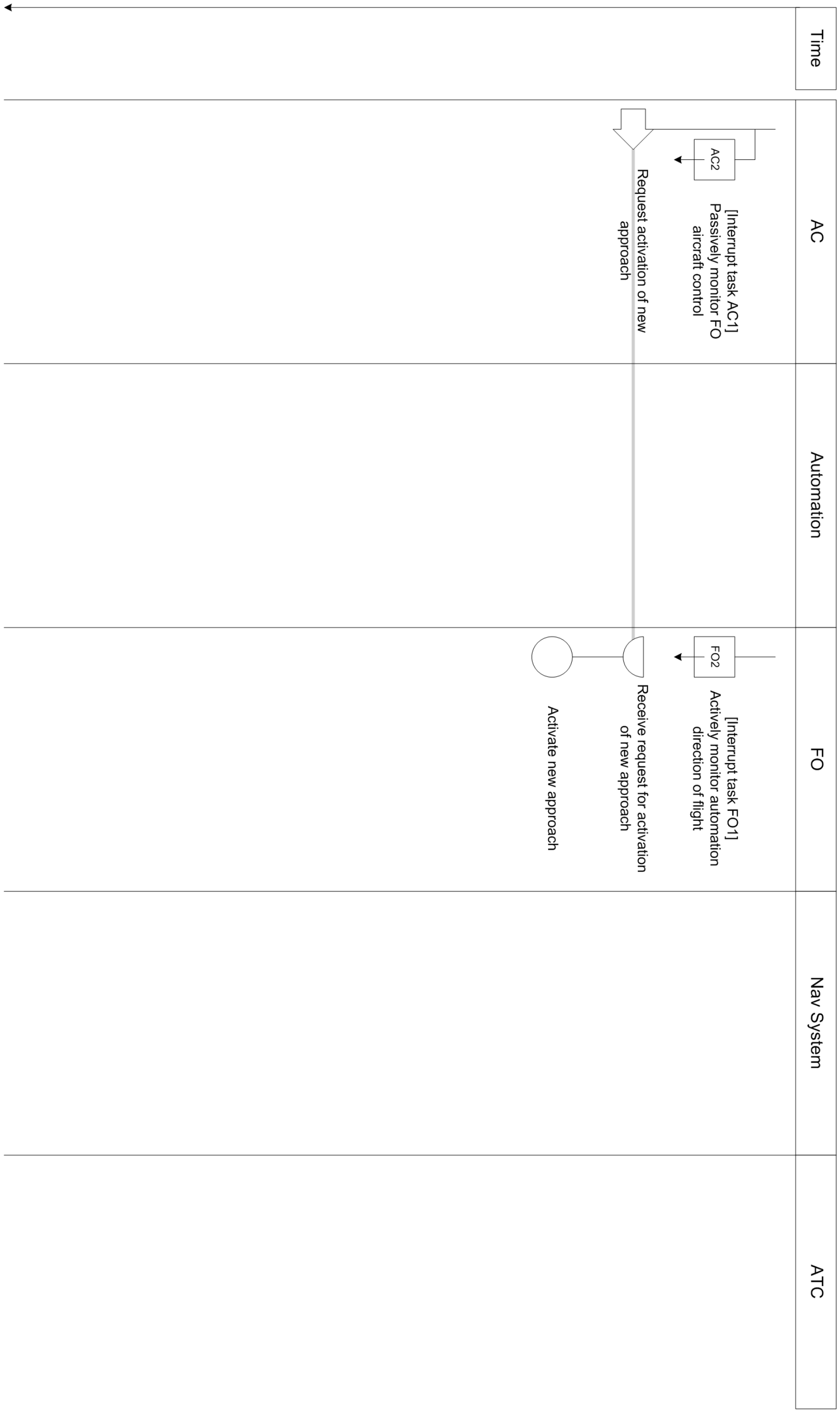
Zanella, A. and Greenberg, S., 2001. Avoiding interference through translucent interface components in single display groupware. in *Proceedings of CHI 2001: Conference on Human Factors in Computer Systems* (Seattle, Washington, USA), ACM, New York, New York, USA.

Appendix A

Operational Sequence Diagram




Time
AC
Automation
FO
Nav System
ATC
Inform FO of new approach decision
Receive new approach decision
Request clearance for new approach
Receive request for new approach
Enter new approach into navigation system
Enter new approach into navigation system
Transmit clearance for new approach
Receive clearance for new approach
Receive clearance for new approach
Report new approach has been entered into navigation system
Receive report
Transfer aircraft control
Transfer aircraft control
AC2
[Interrupt task AC1] Passively monitor FO aircraft control
FO2
[Interrupt task FO1] Actively monitor automation direction of flight
Confirm new approach has been correctly entered into navigation system
Confirm new approach has been correctly entered into navigation system
Transfer aircraft control
Transfer aircraft control


Time
AC
Automation
FO
Nav System
ATC
AC2
[Interrupt task AC1]
Passively monitor FO
aircraft control
Request activation of new
approach
FO2
[Interrupt task FO1]
Actively monitor automation
direction of flight
Receive request for activation
of new approach
Activate new approach

Appendix B

Action/Information Requirements

Functional Requirements	Action Requirements	Information Requirements
Identify options for new approach	Review airport approaches	Airport map
		List of approaches/approach plates
	Identify approach constraints	Weather – ceiling, visibility, wind
		Traffic
		Equipment problems (air/ground)
		Reason for approach change
		Aircraft capabilities
	Identify approaches that meet constraints	All above information
Evaluate and discuss options for new approach	Identify advantages and disadvantages of potential new approaches	All above information
		Original approach
	Decide on relative importance of various advantages/disadvantages	Pilot/Co-Pilot strengths/weaknesses
		Aircraft flight characteristics
Enter new approach into navigation system	Enter new STAR (if necessary)	Original planned route (waypoints)
		New STAR chart
		New approach plate
	Enter new runway (if necessary)	New approach plate
	Enter new approach	New approach plate
	Enter new transition	Original planned route (waypoints)
		New approach plate
		New STAR chart (if applicable)

Functional Requirements	**Action Requirements**	**Information Requirements**
Transfer aircraft control	Ensure pilot and co-pilot both understand the state of the aircraft and automation	Automation status (on/off)
		Automation mode
		Current flight plan
		Aircraft attitude
		Aircraft altitude
		Aircraft speed
	Transfer control of aircraft	Pilot readiness for control transfer
		Co-Pilot readiness for control transfer
Confirm that new approach has been correctly entered into navigation system	Enter new STAR (if necessary)	Original planned route (waypoints)
		New STAR chart
		New approach plate
		New approach information as entered into navigation system
	Enter new runway (if necessary)	New approach plate
		New approach information as entered into navigation system
	Enter new approach	New approach plate
		New approach information as entered into navigation system
	Enter new transition	Original planned route (waypoints)
		New approach plate
		New STAR chart (if applicable)
		New approach information as entered into navigation system

Appendix C

Ethics Clearance Emails

Dear Researcher:

The recommended revisions/additional information requested in the ethics review of your ORE application:

Title: Design of Collaborative Systems for Modern Cockpits

ORE #: 15534

Collaborator: Jonathan M. Histon

Faculty Supervisor: Stacey D. Scott

Student Investigator: Paul McKay

have been reviewed and are considered acceptable. As a result, your application now has received full ethics clearance.

A signed copy of the Notification of Full Ethics Clearance will be sent to the Principal Investigator or Faculty Supervisor in the case of student research.

****

Note 1: This clearance is valid for four years from the date shown on the certificate and a new application must be submitted for on-going projects continuing beyond four years.

Note 2: This project must be conducted according to the application description and revised materials for which ethics clearance have been granted. All subsequent modifications to the protocol must receive prior ethics clearance through our office and must not begin until notification has been received.

Note 3: Researchers must submit a Progress Report on Continuing Human Research Projects (ORE Form 105) annually for all ongoing research projects. In addition, researchers must submit a Form 105 at the conclusion of the project if it continues for less than a year.

Note 4: Any events related to the procedures used that adversely affect participants must be reported immediately to the ORE using ORE Form 106.

Best wishes for success with this study.

Susanne Santi, M. Math.,

Senior Manager

Office of Research Ethics

NH 1027

Dear Researcher:

The recommended revisions/additional information requested in the ethics review of your ORE application:

Title: Design of Collaborative Systems for Modern Cockpits (Phase 2)

ORE #: 15715

Collaborator: Jonathan M. Histon (jhiston@uwaterloo.ca)

Faculty Supervisor: Stacey D. Scott (s9scott@uwaterloo.ca)

Student Investigator: Paul McKay (pdmckay@uwaterloo.ca)

have been reviewed and are considered acceptable. As a result, your application now has received full ethics clearance.

A signed copy of the Notification of Full Ethics Clearance will be sent to the Principal Investigator or Faculty Supervisor in the case of student research.

****

Note 1: This clearance is valid for four years from the date shown on the certificate and a new application must be submitted for on-going projects continuing beyond four years.

Note 2: This project must be conducted according to the application description and revised materials for which ethics clearance have been granted. All subsequent modifications to the protocol must receive prior ethics clearance through our office and must not begin until notification has been received.

Note 3: Researchers must submit a Progress Report on Continuing Human Research Projects (ORE Form 105) annually for all ongoing research projects. In addition, researchers must submit a Form 105 at the conclusion of the project if it continues for less than a year.

Note 4: Any events related to the procedures used that adversely affect participants must be reported immediately to the ORE using ORE Form 106.

Best wishes for success with this study.

Susanne Santi, M. Math.,

Senior Manager

Office of Research Ethics

NH 1027

Appendix D

Flight Scenario Details

Cockpit Interaction Awareness Display Operational Testing Scenario

Summary

The initial flight plan will be from Kitchener/Waterloo Int'l to Chicago O'Hare Int'l. The plan will consist of takeoff, a direct flight/climb to London VOR, flight along the J547 airway to the Flint VOR, taking the Flint transition into the PAITN arrival for ILS RWY 27L.

Between KW and London, the crew will receive information that there has been a security event and they must land at Windsor. They will re-plan the flight to Windsor, using the London transition into the PICES arrival for ILS RWY 25 and landing.

Phase-of-Flight Approach

This scenario can be summarized using "Phase-of-flight" divisions as follows:

Preflight – Prepare aircraft for required operational use given representative mission parameters.

Take-Off – Operate aircraft from take-off roll to commencement of "CLIMB" phase-of-flight including aircraft reconfigurations.

Climb (CLB) – Operate aircraft up to CRUISE phase-of-flight, to include both lateral and vertical flight path modifications.

Cruise (CRZ) – Operate aircraft up to DESCENT phase-of-flight to include flight plan modification.

Descent (DES) – Scenario ends when aircraft is established in descent and passes waypoint "AXXIS".

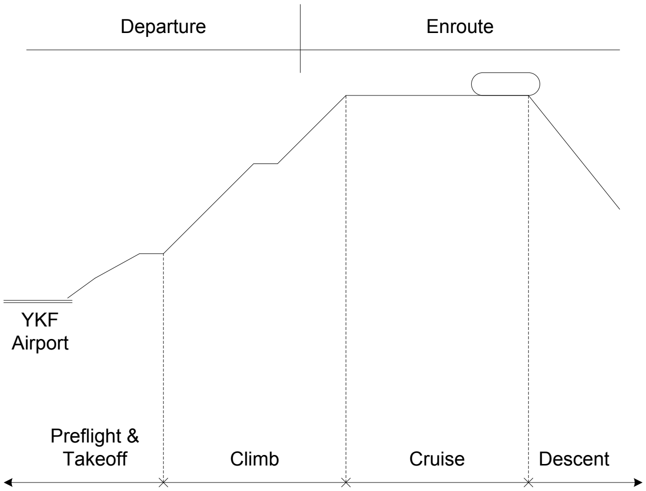


Figure C-1 – Operational Scenario for the Testing of a Cockpit Interaction Awareness Display

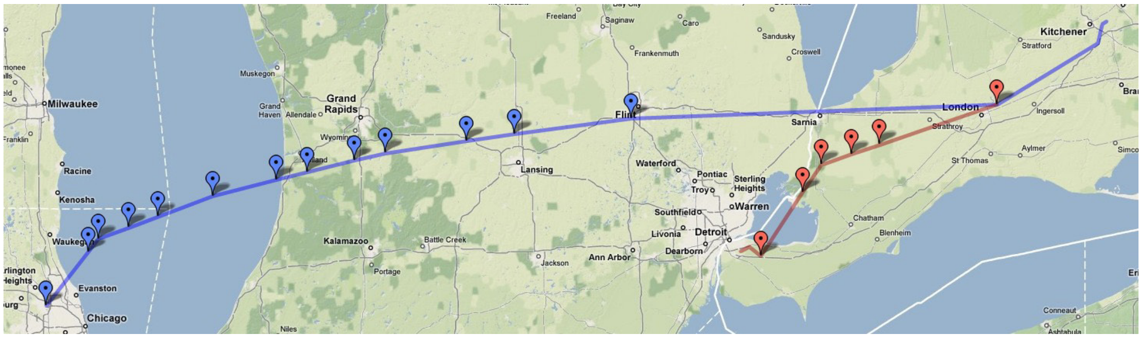


Figure C-2 – Map of Operational Scenario (initial route in blue, final in red)

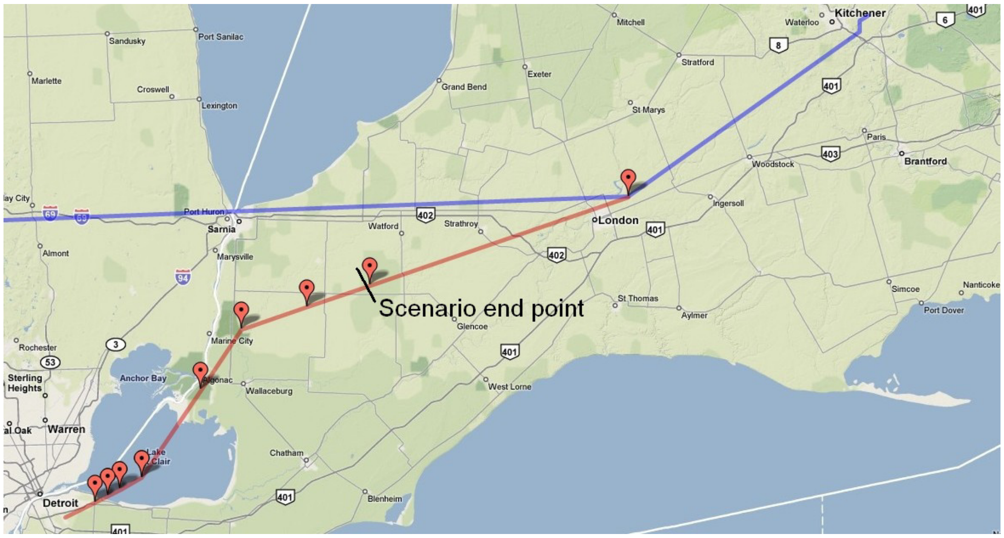


Figure C-3 – Close-up Map of Actual Flight Scenario

PREFLIGHT

Step	ATC Action?	Task Sequence	Info Needed
1		Seat aircrew and explain current aircraft configuration.	Aircraft state
2		Provide copy of Flight Plan, aircraft checklists and Departure / Arrival charts.	
3		Crew to review Flight Plan	IFR flight plan
4		PIC directs FO to obtain "ATIS" information on Freq 125.1	Radio control operation
5	📢	ATC proxy transmits "ATIS" information: "Waterloo ATIS information Alpha, 2000 feet overcast; visibility four miles in haze; temperature 10 dew point 10 degrees Celsius; wind 273 @ 23 knots; and Altimeter 2998. Departing runway 26; Landing ILS Runway 26; Advise Toronto ATC you have received information Alpha"	
6		Aircrew copies "ATIS" information.	
7		PIC prompts FO to contact ground for departure clearance	
8		FO selects freq 121.8 and requests: "ground, flight Charlie Golf Foxtrot X-ray, standing by airways Chicago, flight level 260"	
9	📢	ATC proxy returns: "Charlie Golf Foxtrot X-ray is cleared Chicago airport depart runway 26 waterloo two, flight plan route to maintain FL 260 squawking 0777."	
10		Aircrew copies clearance and PIC prompts FO to select required transponder code.	
11		FO initializes interactive flight plan map and inserts route (PIC monitors with flight plan in hand).	IFR flight plan, charts

TAKEOFF

Step	ATC Action?	Task Sequence	Info Needed
12		Assume taxi completed and Before Takeoff Checks done.	
13		PIC instructs FO to obtain T/O clearance on frequency 126.0	Radio controls
14	📢	ATC proxy provides clearance: "Charlie Golf Foxtrot X-ray is cleared for takeoff Waterloo two departure, maintain 4000 feet, contact Toronto departure frequency 128.275 when airborne."	
15		Aircrew engages appropriate automatic modes and advances throttles	
16		Once airborne and "positive climb", PIC initiates "Gear up" sequence	
17		PIC engages auto-pilot when aircraft greater than 400 feet AGL	
18		PIC directs FO to contact Departure 128.275	
19	📢	ATC proxy, once contacted, provides clearance: "Charlie Golf Foxtrot X-ray, continue waterloo two until 4000 feet then proceed on course flight level 260"	
20		PIC inputs auto-flight system to fly departure	
21		Above "Acceleration Altitude", PIC calls for: "Flaps Zero" above "F" speed	
22		Above "Acceleration Altitude", PIC calls for: "Slats Retract" above "S" speed	
23		At "O" speed, PIC calls for: "After Take-off check"	

CLB

Step	ATC Action?	Task Sequence	Info Needed
24	📢	At TBD time, ATC proxy directs: "Charlie Golf Foxtrot X-ray, for traffic level off at 10,000 feet"	
25		PIC sets 10,000 in altitude selector.	
26	📢	At TBD time, ATC proxy directs: " Charlie Golf Foxtrot X-ray, traffic is no threat, now cleared FL 260, contact Toronto enroute frequency 135.825"	
27		PIC directs FO to proceed on course and climb to FL 260 cruise altitude, switch radio to frequency 135.825.	
28		Monitor level off at cruise altitude. (May be before or after routing change)	

CRZ

Step	ATC Action?	Task Sequence	Info Needed
29	📢	Approx 2 minutes after step 27, ATC proxy notifies “Charlie Golf Foxtrot X-ray, due to a security event you are being re-routed to land in Windsor, anticipate Pices one arrival to ILS runway 25, contact Detroit approach control 126.85”	Windsor charts
30		PIC to advise FO to contact Detroit approach control and request hold at London to permit cockpit preparation including receipt of ATIS information.	
31		FO calls ATC requesting hold	
32	📢	ATC proxy directs: "Charlie Golf Foxtrot X-ray, hold at London FL 260 inbound track 260"	
33		Once aircrew are on their way to hold, PIC directs FO to obtain destination ATIS information on frequency 134.5	

CRZ

Step	ATC Action?	Task Sequence	Info Needed
34	📢	ATC proxy advises: "Windsor ATIS information Charlie, 800 feet overcast; visibility two miles in haze; temperature 11 dew point 10 degrees Celsius; wind 258 @ 18 knots; and Altimeter 2995. Departing runway 25; Landing ILS Runway 25;. Advise Detroit approach control you have received information Charlie"	
35		PIC and FO copy ATIS information. PIC advises FO to make descent and arrival preparations.	
36		PIC directs FO to enter new approach into flight plan	
37		FO inserts route into interactive flight plan map	
38		PIC briefs the approach.	
39		PIC directs FO to conduct pre-descent check, then advise ATC ready for approach.	Pre-descent checklist
40		FO advises ATC ready for approach	
41	📢	ATC proxy directs: "Charlie Golf Foxtrot X-ray, cleared to WINZZ 4000 feet via PICES one arrival, report crossing DROME."	
42		Crew exits hold and engages new approach.	
43		Exercise ends when crew is established in descent and crosses AXXIS.	

Appendix E
Initial Phase Information/Consent Letter

INFORMATION LETTER

Title of Project: ***Design of Collaborative Systems for Modern Cockpits***

Student Investigator: ***Paul McKay, 519-888-4567 ext 36813, pdmckay@uwaterloo.ca***

Faculty Supervisor: ***Stacey Scott, 519-888-4567 ext 32236, s9scott@uwaterloo.ca***

Summary of the Project:
This project is the first part of a two-part research study aimed at improving the design of modern collaborative cockpit systems involving shared displays. Our proposed interface improvement involves the display of interaction history data as a part of the cockpit interface. In order to create a prototype system that includes this information, it is necessary to collect interaction data that can be shown as part of a recorded scenario. The researchers hope to obtain this interaction data through observation of experienced pilots performing representative flight scenarios. The information gathered by the combined, two-part study will be used to develop design suggestions for cockpit interfaces that support effective individual and collaborative work.

Procedure:
Your participation in this study is voluntary. Participation involves performing simulated flight scenarios in a prototype cockpit environment (located in the CMC Human Factors Lab), potentially while wearing a head-mounted gaze tracking system. Because the study involves flight scenarios, all participants should be experienced pilots. The gaze tracking system consists of an adjustable headband that is used to mount an infrared camera and mirror for eye tracking and a magnetic sensor for head tracking. The total assembly weighs approximately 1 lb. You will complete these flight scenarios with a partner. A description of each activity follows.

You will receive an introduction to the study, the prototype cockpit, and the gaze tracking system and then you will be asked to:

- Perform the gaze tracking system calibration procedure (this applies only for the participant who is assigned to wear the gaze tracking system).
- Complete a brief training flight scenario to familiarize yourself with the prototype cockpit.
- Complete a full flight scenario that will be recorded.
- Review the recording with the researcher and indicate the areas of the interface that you were viewing during the scenario.

This session will take approximately 90 minutes.

During the session a researcher will observe and take notes regarding your interactions with the cockpit systems, as well as your interactions with your partner in the sessions. Your computer-based interactions (in the form of a screen capture video) and point-of-gaze data (if applicable) will also be captured and stored in a computer log file. A video recording of the full flight scenario will also be made, and any task materials produced during the session will remain with the researcher. You may decline to respond to questions if you wish. You may withdraw your participation at any time without penalty.

Confidentiality and Data Security:
All information you provide is considered completely confidential. Your name will not appear in any publication resulting from this study; however, with your permission anonymous quotations may be used. In these cases participants will be referred to as Aircraft Captain (AC) or First Officer (FO) or collectively as a flight crew (Crew A, B,…). Data collected during this study will be retained indefinitely in locked cabinets or on password protected desktop computers in a secure location. Electronic data will not include personal identifying information such as names.
You will be explicitly asked for consent for the use of photo/video/audio data, captured from the scenario recording, for the purpose of reporting the study's findings. If consent is granted, these data will be used only for the purposes associated with teaching, scientific presentations, publications, and/or sharing with other researchers and you will not be identified by name.

Risks and Benefits:
There are no known or anticipated risks from participation in this study. There are no direct benefits to you from participation. However, the results of this research may contribute to the knowledge base of Human Systems Engineering research and help in developing improved cockpit interfaces.

Research Ethics Clearance:
I would like to assure you that this study has been reviewed and received ethics clearance through the Office of Research Ethics at the University of Waterloo. However, the final decision about participation is yours. Should you have comments or concerns resulting from your participation in this study, please contact Dr. Susan Sykes in the Office of Research Ethics at 519-888-4567, Ext. 36005.

Thank you for your assistance in this project.

CONSENT FORM

Project: ***Design of Collaborative Systems for Modern Cockpits***

I have read the information presented in the information letter about a study being conducted ***by Paul McKay*** of the Department of ***Systems Design Engineering***, under the supervision of Professor ***Stacey Scott.*** I have had the opportunity to ask any questions related to this study, to receive satisfactory answers to my questions, and any additional details I wanted.

I am aware that I may allow video, audio, and/or point-of-gaze recordings of my full flight scenario to be used in further research (shown to participants in a follow-on study) with the understanding that I will not be identified by name.

Sometimes a certain image and/or segment of video or audio recording clearly shows a particular feature or detail that would be helpful in teaching or when presenting the study results at a scientific presentation or in a publication.

I am aware that I may allow video and/or digital images which show my computer interactions to be used in teaching, scientific presentations, publications, and/or sharing with other researchers with the understanding that I will not be identified by name. I am aware that I may allow excerpts from the conversational data collected for this study to be included in teaching, scientific presentations and/or publications, with the understanding that any quotations will be anonymous.

I am aware that I may withdraw my consent for any of the above statements or withdraw my study participation at any time without penalty by advising the researcher.

This project has been reviewed by, and received ethics clearance through, the Office of Research Ethics at the University of Waterloo. I was informed that if I have any comments or concerns resulting from my participation in this study, I may contact the Director, Office of Research Ethics at (519) 888-4567 ext. 36005.

	Please Circle One	**Please initial Your Choice**
With full knowledge of all foregoing, I agree, of my own free will to participate in this study.	YES NO	________
I agree to use the gaze tracking system and allow my point-of-gaze to be recorded.	YES NO	________
I agree to be videotaped.	YES NO	________

I agree to let my conversation during the study be directly quoted, YES NO ________
anonymously, in presentations of research results.

I agree to let data from the full flight scenario recording be used YES NO ________
In the development of a prototype for further research.

I agree to let data from the full flight scenario recording be used YES NO ________
for presentations of the research results.

Participant Name: ____
(Please print)

Participant Signature: _____________________________________

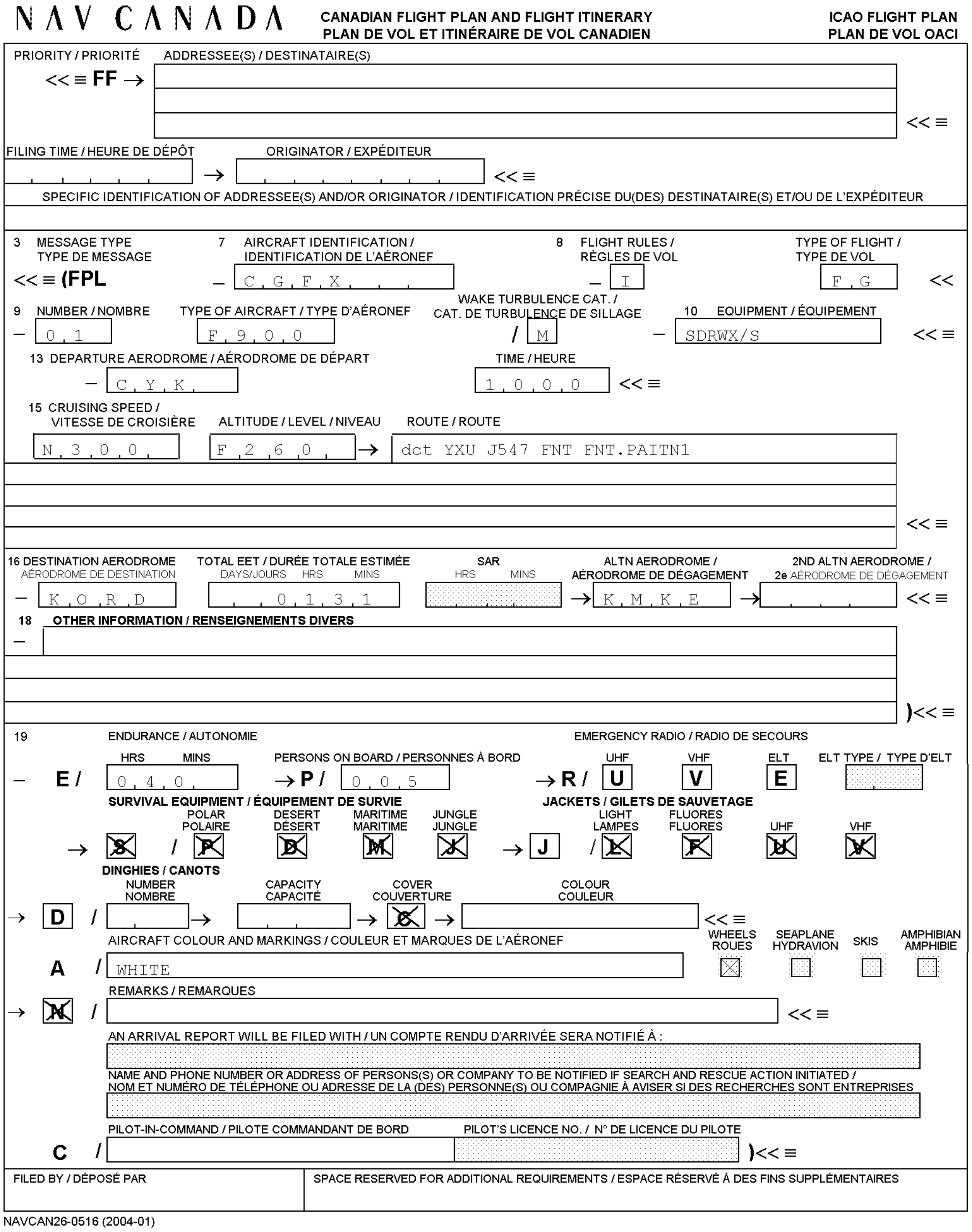
Witness Name: ____
(Please print)

Witness Signature: _______________________________________

Date: ________________________________

Appendix F

Flight Scenario Materials



NAV CANADA

CANADIAN FLIGHT PLAN AND FLIGHT ITINERARY
PLAN DE VOL ET ITINÉRAIRE DE VOL CANADIEN

ICAO FLIGHT PLAN
PLAN DE VOL OACI

PRIORITY / PRIORITÉ: << ≡ FF →
ADDRESSEE(S) / DESTINATAIRE(S): << ≡

FILING TIME / HEURE DE DÉPÔT: →
ORIGINATOR / EXPÉDITEUR: << ≡

SPECIFIC IDENTIFICATION OF ADDRESSEE(S) AND/OR ORIGINATOR / IDENTIFICATION PRÉCISE DU(DES) DESTINATAIRE(S) ET/OU DE L'EXPÉDITEUR

3 MESSAGE TYPE / TYPE DE MESSAGE: << ≡ (FPL
7 AIRCRAFT IDENTIFICATION / IDENTIFICATION DE L'AÉRONEF: – C G F X
8 FLIGHT RULES / RÈGLES DE VOL: – I
TYPE OF FLIGHT / TYPE DE VOL: F G <<

9 NUMBER / NOMBRE: – 0 1
TYPE OF AIRCRAFT / TYPE D'AÉRONEF: F 9 0 0
WAKE TURBULENCE CAT. / CAT. DE TURBULENCE DE SILLAGE: / M
10 EQUIPMENT / ÉQUIPEMENT: – SDRWX/S << ≡

13 DEPARTURE AERODROME / AÉRODROME DE DÉPART: – C Y K
TIME / HEURE: 1 0 0 0 << ≡

15 CRUISING SPEED / VITESSE DE CROISIÈRE: N 3 0 0
ALTITUDE / LEVEL / NIVEAU: F 2 6 0 →
ROUTE / ROUTE: dct YXU J547 FNT FNT.PAITN1 << ≡

16 DESTINATION AERODROME / AÉRODROME DE DESTINATION: – K O R D
TOTAL EET / DURÉE TOTALE ESTIMÉE (DAYS/JOURS HRS MINS): 0 1 3 1
SAR (HRS MINS):
ALTN AERODROME / AÉRODROME DE DÉGAGEMENT: → K M K E
2ND ALTN AERODROME / 2e AÉRODROME DE DÉGAGEMENT: → << ≡

18 OTHER INFORMATION / RENSEIGNEMENTS DIVERS: –)<< ≡

19 ENDURANCE / AUTONOMIE (HRS MINS): – E / 0 4 0
PERSONS ON BOARD / PERSONNES À BORD: → P / 0 0 5
EMERGENCY RADIO / RADIO DE SECOURS: → R / UHF U VHF V ELT E ELT TYPE / TYPE D'ELT

SURVIVAL EQUIPMENT / ÉQUIPEMENT DE SURVIE: → ~~S~~ / POLAR POLAIRE ~~P~~ DESERT DÉSERT ~~D~~ MARITIME MARITIME ~~M~~ JUNGLE JUNGLE ~~J~~

JACKETS / GILETS DE SAUVETAGE: → J / LIGHT LAMPES ~~L~~ FLUORES FLUORES ~~F~~ UHF ~~U~~ VHF ~~V~~

DINGHIES / CANOTS: → D / NUMBER NOMBRE → CAPACITY CAPACITÉ → COVER COUVERTURE ~~C~~ → COLOUR COULEUR << ≡

AIRCRAFT COLOUR AND MARKINGS / COULEUR ET MARQUES DE L'AÉRONEF: A / WHITE
WHEELS ROUES ☒ SEAPLANE HYDRAVION ☐ SKIS ☐ AMPHIBIAN AMPHIBIE ☐

REMARKS / REMARQUES: → ~~N~~ / << ≡

AN ARRIVAL REPORT WILL BE FILED WITH / UN COMPTE RENDU D'ARRIVÉE SERA NOTIFIÉ À :

NAME AND PHONE NUMBER OR ADDRESS OF PERSONS(S) OR COMPANY TO BE NOTIFIED IF SEARCH AND RESCUE ACTION INITIATED / NOM ET NUMÉRO DE TÉLÉPHONE OU ADRESSE DE LA (DES) PERSONNE(S) OU COMPAGNIE À AVISER SI DES RECHERCHES SONT ENTREPRISES

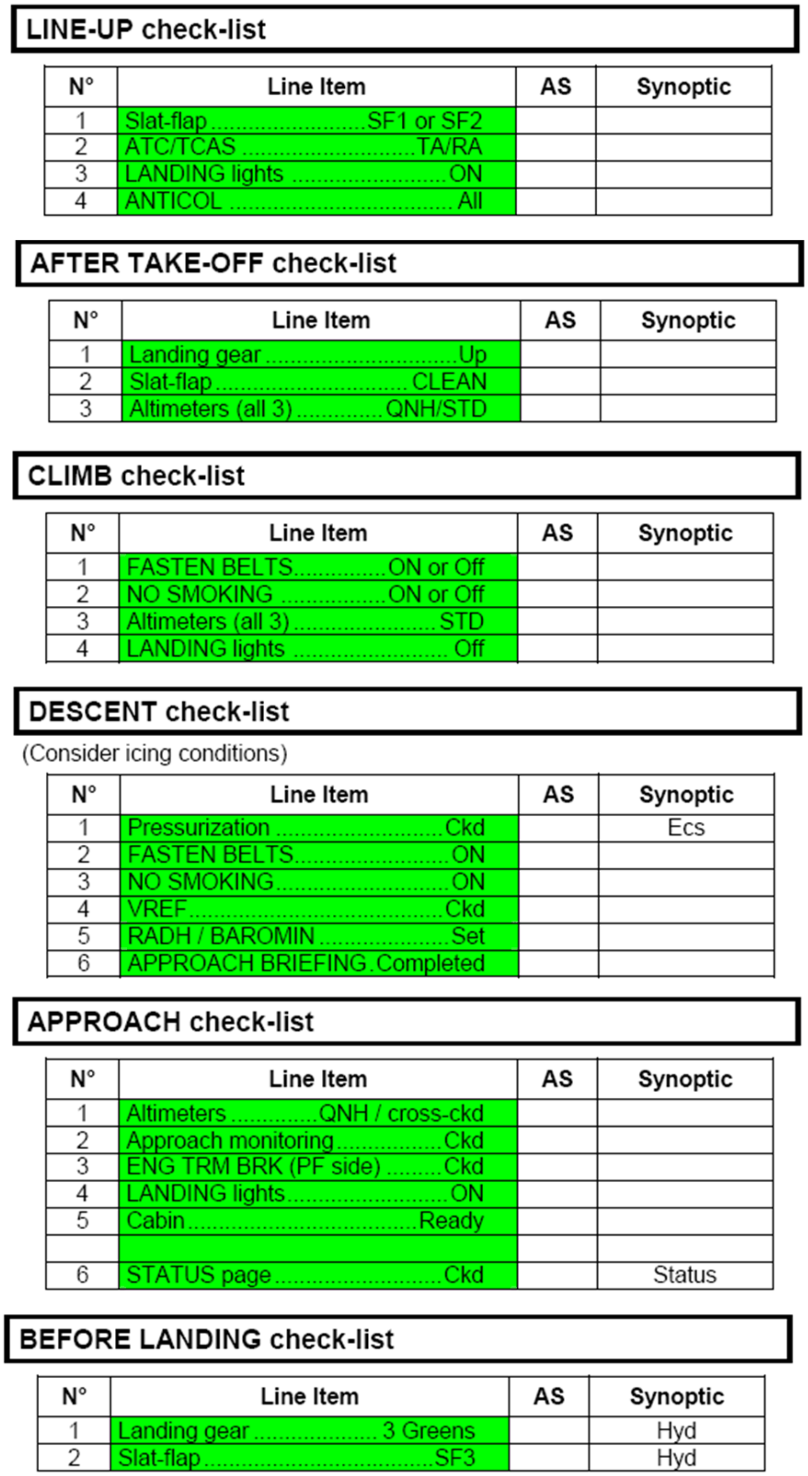
PILOT-IN-COMMAND / PILOTE COMMANDANT DE BORD: C /
PILOT'S LICENCE NO. / N° DE LICENCE DU PILOTE:)<< ≡

FILED BY / DÉPOSÉ PAR

SPACE RESERVED FOR ADDITIONAL REQUIREMENTS / ESPACE RÉSERVÉ À DES FINS SUPPLÉMENTAIRES

NAVCAN26-0516 (2004-01)

Figure E-1 – Scenario IFR Flight Plan



LINE-UP check-list

N°	Line Item	AS	Synoptic
1	Slat-flap ....SF1 or SF2		
2	ATC/TCAS ....TA/RA		
3	LANDING lights ....ON		
4	ANTICOL ....All		

AFTER TAKE-OFF check-list

N°	Line Item	AS	Synoptic
1	Landing gear ....Up		
2	Slat-flap .... CLEAN		
3	Altimeters (all 3).............QNH/STD		

CLIMB check-list

N°	Line Item	AS	Synoptic
1	FASTEN BELTS...............ON or Off		
2	NO SMOKING ....ON or Off		
3	Altimeters (all 3) ....STD		
4	LANDING lights .... Off		

DESCENT check-list

(Consider icing conditions)

N°	Line Item	AS	Synoptic
1	Pressurization ....Ckd		Ecs
2	FASTEN BELTS..........................ON		
3	NO SMOKING.............................ON		
4	VREF...Ckd		
5	RADH / BAROMIN ....Set		
6	APPROACH BRIEFING .Completed		

APPROACH check-list

N°	Line Item	AS	Synoptic
1	Altimeters ....QNH / cross-ckd		
2	Approach monitoring.................Ckd		
3	ENG TRM BRK (PF side) ....Ckd		
4	LANDING lights..........................ON		
5	Cabin....................................Ready		
6	STATUS page ....Ckd		Status

BEFORE LANDING check-list

N°	Line Item	AS	Synoptic
1	Landing gear .... 3 Greens		Hyd
2	Slat-flap....................................SF3		Hyd

Figure E-2 – Scenario Aircraft Checklists

Figure E-3 – Scenario High Level Chart, original route shown in yellow

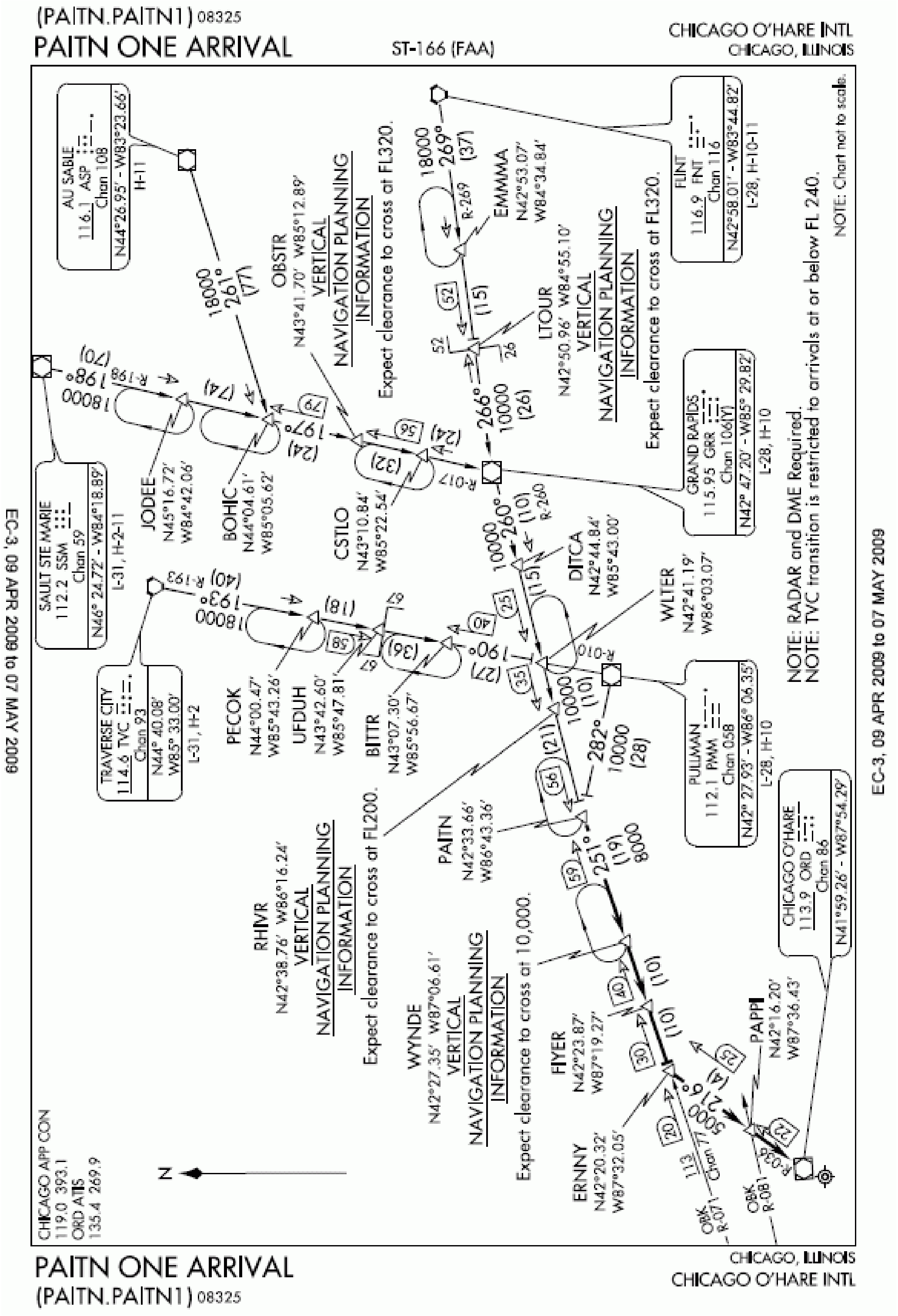


Figure E-4 – Chicago (KORD) PAITN1 STAR Chart, Page 1

(PAITN.PAITN1) 08325
PAITN ONE ARRIVAL
ST-166 (FAA)
CHICAGO O'HARE INTL
CHICAGO, ILLINOIS

EC-3, 09 APR 2009 to 07 MAY 2009

ARRIVAL DESCRIPTION

AU SABLE TRANSITION (ASP.PAITN1): From over ASP VOR/DME via ASP R-261 to BOHIC then via GRR R-017 to GRR VOR/DME then via GRR R-260 to PAITN. Thence....

FLINT TRANSITION (FNT.PAITN1): From over FNT VORTAC via FNT R-269 to GRR VOR/DME then via GRR R-260 to PAITN. Thence....

GRAND RAPIDS TRANSITION (GRR.PAITN1): From over GRR VOR/DME via GRR R-260 to PAITN. Thence....

PULLMAN TRANSITION (PMM.PAITN1): From over PMM VOR/DME via PMM R-282 to PAITN. Thence....

SAULT STE MARIE TRANSITION (SSM.PAITN1): From over SSM VOR/DME via SSM R-198 to BOHIC then via GRR R-17 to GRR VOR/DME then via GRR R-260 to PAITN. Thence....

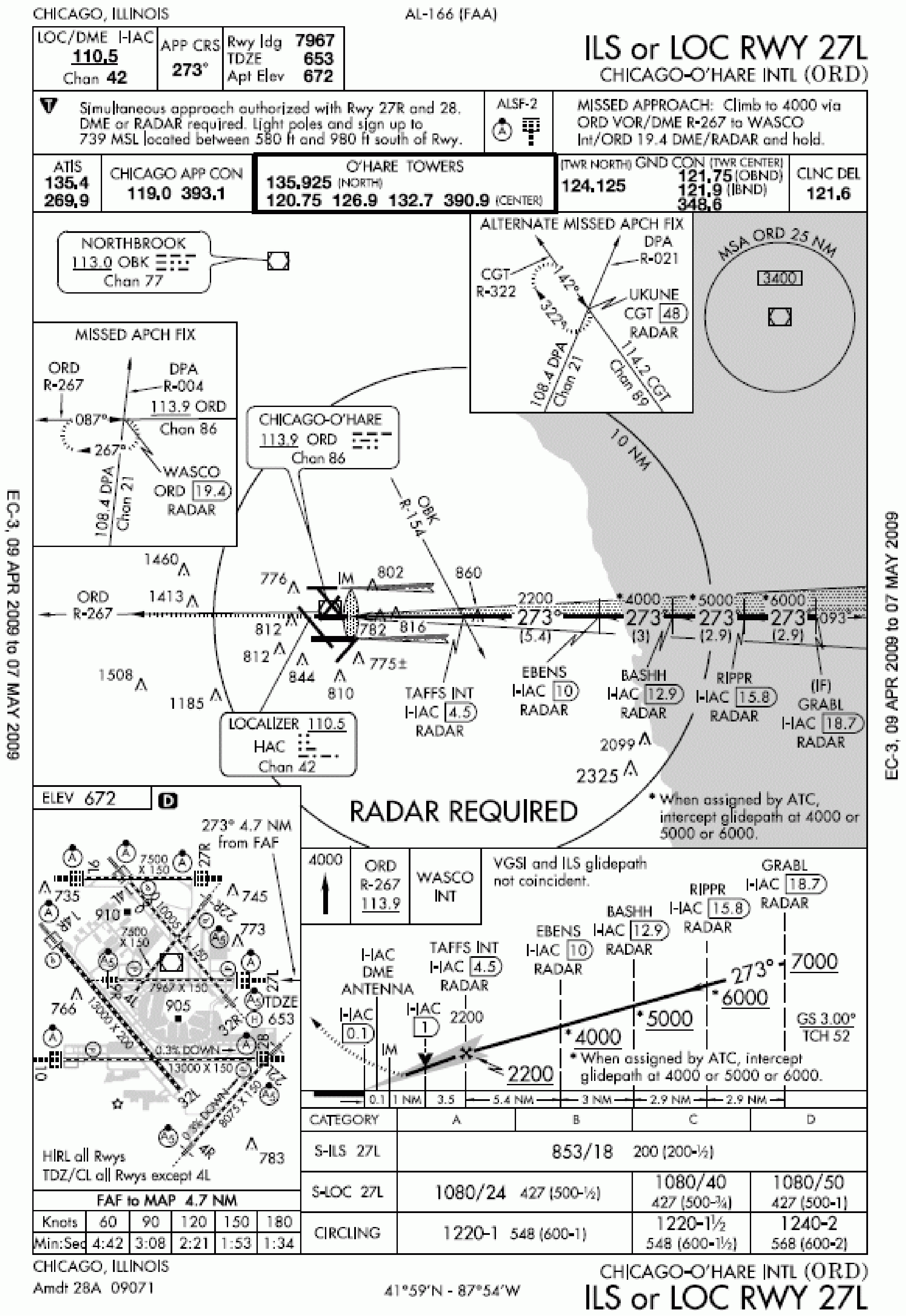
TRAVERSE CITY TRANSITION (TVC.PAITN1): From over TVC VORTAC via TVC R-193 to BITTR/TVC R-190 to WLTER/TVC 122 DME then via GRR R-260 to PAITN. Thence....

....From over PAITN via OBK VOR/DME R-071 to WYNDE, then via OBK VOR/DME R-071 to FIYER, then via OBK VOR/DME R-071 to ERNNY, then via ORD VOR/DME R-036 to PAPPI, then via ORD VOR/DME R-036 to ORD VOR/DME. Expect radar vectors to final approach course.

EC-3, 09 APR 2009 to 07 MAY 2009

PAITN ONE ARRIVAL
(PAITN.PAITN1) 08325
CHICAGO, ILLINOIS
CHICAGO O'HARE INTL

Figure E-5 – Chicago (KORD) PAITN1 STAR Chart, Page 2



CHICAGO, ILLINOIS AL-166 (FAA)

LOC/DME I-HAC 110.5 Chan 42	APP CRS 273°	Rwy ldg 7967 TDZE 653 Apt Elev 672

ILS or LOC RWY 27L

CHICAGO-O'HARE INTL (ORD)

Simultaneous approach authorized with Rwy 27R and 28. DME or RADAR required. Light poles and sign up to 739 MSL located between 580 ft and 980 ft south of Rwy.

ALSF-2

MISSED APPROACH: Climb to 4000 via ORD VOR/DME R-267 to WASCO Int/ORD 19.4 DME/RADAR and hold.

ATIS	CHICAGO APP CON	O'HARE TOWERS	(TWR NORTH) GND CON (TWR CENTER)	CLNC DEL
135.4 269.9	119.0 393.1	135.925 (NORTH) 120.75 126.9 132.7 390.9 (CENTER)	124.125 121.75 (OBND) 121.9 (IBND) 348.6	121.6

CATEGORY	A	B	C	D
S-ILS 27L	853/18 200 (200-½)			
S-LOC 27L	1080/24 427 (500-½)		1080/40 427 (500-¾)	1080/50 427 (500-1)
CIRCLING	1220-1 548 (600-1)		1220-1½ 548 (600-1½)	1240-2 568 (600-2)

FAF to MAP 4.7 NM

Knots	60	90	120	150	180
Min:Sec	4:42	3:08	2:21	1:53	1:34

CHICAGO, ILLINOIS

Amdt 28A 09071

41°59'N - 87°54'W

CHICAGO-O'HARE INTL (ORD)

ILS or LOC RWY 27L

Figure E-6 – Chicago (KORD) Approach Plate, ILS Runway 27L

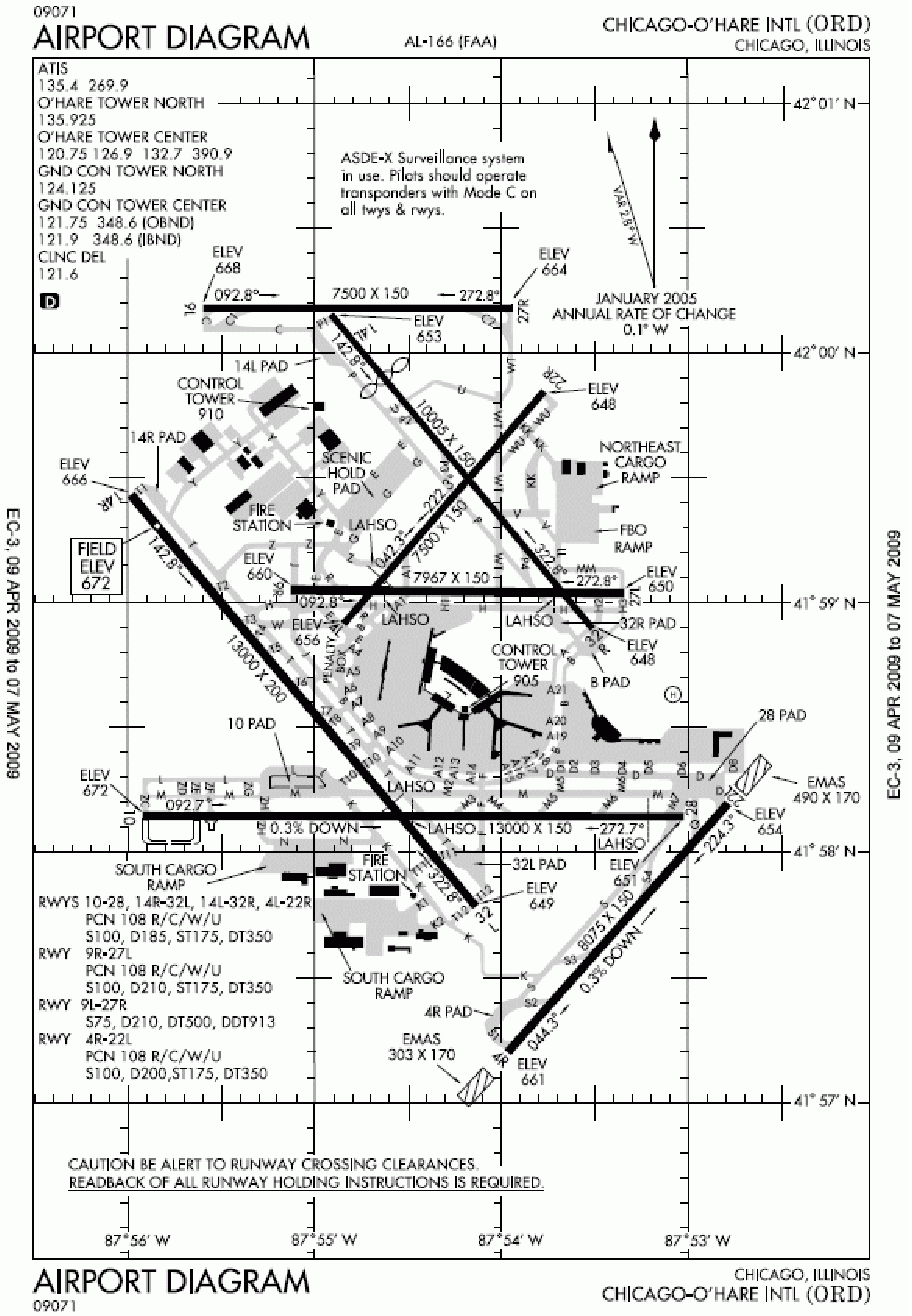
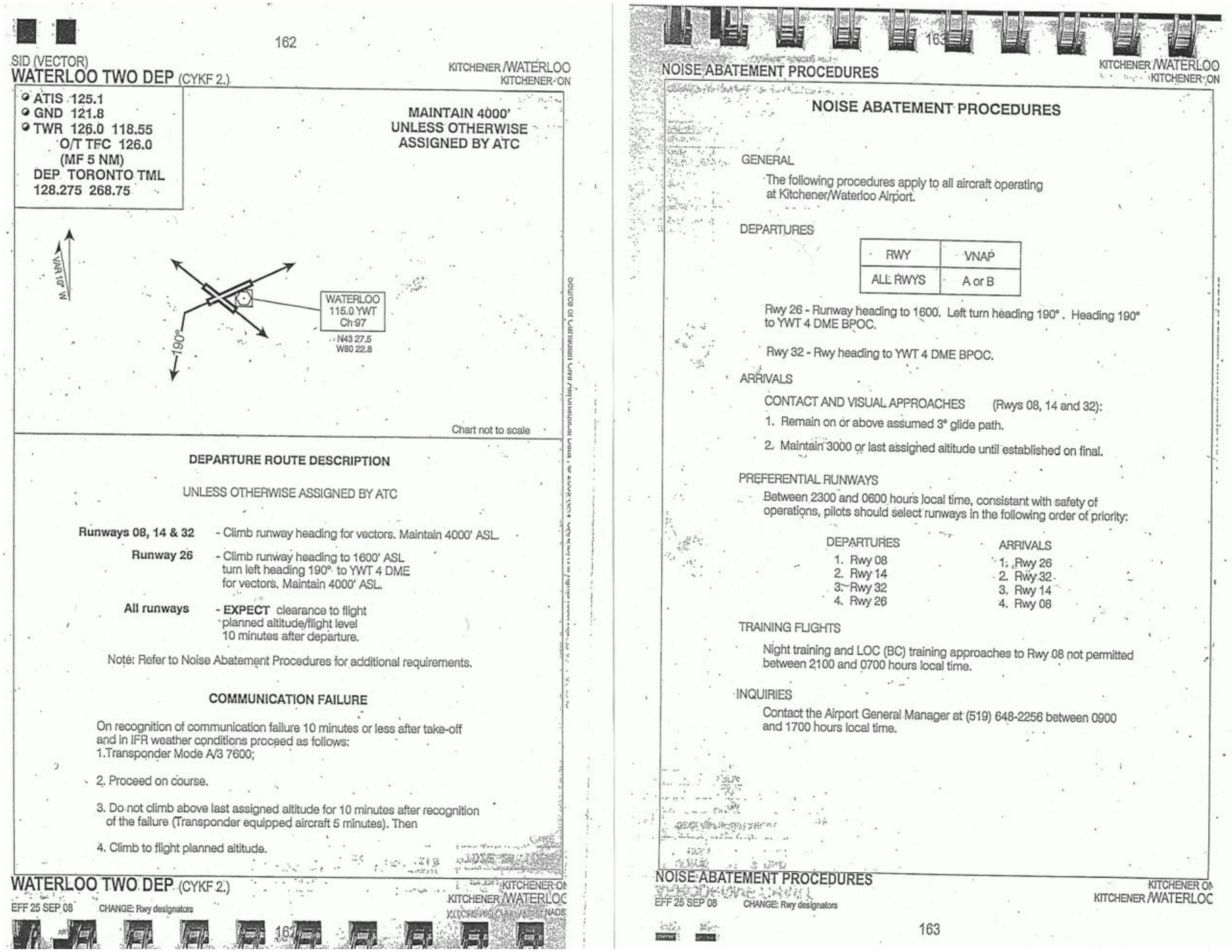


Figure E-7 – Chicago (KORD) Airport Diagram



162

SID (VECTOR)

WATERLOO TWO DEP (CYKF 2.)

KITCHENER /WATERLOO
KITCHENER-ON

ATIS 125.1
GND 121.8
TWR 126.0 118.55
O/T TFC 126.0
(MF 5 NM)
DEP TORONTO TML
128.275 268.75

MAINTAIN 4000'
UNLESS OTHERWISE
ASSIGNED BY ATC

DEPARTURE ROUTE DESCRIPTION

UNLESS OTHERWISE ASSIGNED BY ATC

Runways 08, 14 & 32	- Climb runway heading for vectors. Maintain 4000' ASL.
Runway 26	- Climb runway heading to 1600' ASL turn left heading 190° to YWT 4 DME for vectors. Maintain 4000' ASL.
All runways	- **EXPECT** clearance to flight planned altitude/flight level 10 minutes after departure.

Note: Refer to Noise Abatement Procedures for additional requirements.

COMMUNICATION FAILURE

On recognition of communication failure 10 minutes or less after take-off and in IFR weather conditions proceed as follows:
1.Transponder Mode A/3 7600;

2. Proceed on course.

3. Do not climb above last assigned altitude for 10 minutes after recognition of the failure (Transponder equipped aircraft 5 minutes). Then

4. Climb to flight planned altitude.

WATERLOO TWO DEP (CYKF 2.)

KITCHENER-ON
KITCHENER /WATERLOO

EFF 25 SEP 08 CHANGE: Rwy designators

162

163

NOISE ABATEMENT PROCEDURES

KITCHENER /WATERLOO
KITCHENER-ON

NOISE ABATEMENT PROCEDURES

GENERAL

The following procedures apply to all aircraft operating at Kitchener/Waterloo Airport.

DEPARTURES

RWY	VNAP
ALL RWYS	A or B

Rwy 26 - Runway heading to 1600. Left turn heading 190°. Heading 190° to YWT 4 DME BPOC.

Rwy 32 - Rwy heading to YWT 4 DME BPOC.

ARRIVALS

CONTACT AND VISUAL APPROACHES (Rwys 08, 14 and 32):

1. Remain on or above assumed 3° glide path.
2. Maintain 3000 or last assigned altitude until established on final.

PREFERENTIAL RUNWAYS

Between 2300 and 0600 hours local time, consistant with safety of operations, pilots should select runways in the following order of priority:

DEPARTURES	ARRIVALS
1. Rwy 08	1. Rwy 26
2. Rwy 14	2. Rwy 32
3. Rwy 32	3. Rwy 14
4. Rwy 26	4. Rwy 08

TRAINING FLIGHTS

Night training and LOC (BC) training approaches to Rwy 08 not permitted between 2100 and 0700 hours local time.

INQUIRIES

Contact the Airport General Manager at (519) 648-2256 between 0900 and 1700 hours local time.

NOISE ABATEMENT PROCEDURES

KITCHENER ON
KITCHENER /WATERLOO

EFF 25 SEP 08 CHANGE: Rwy designators

163

Figure E-8 – Waterloo (CYKF) Two Departure Chart, Pages 1 and 2

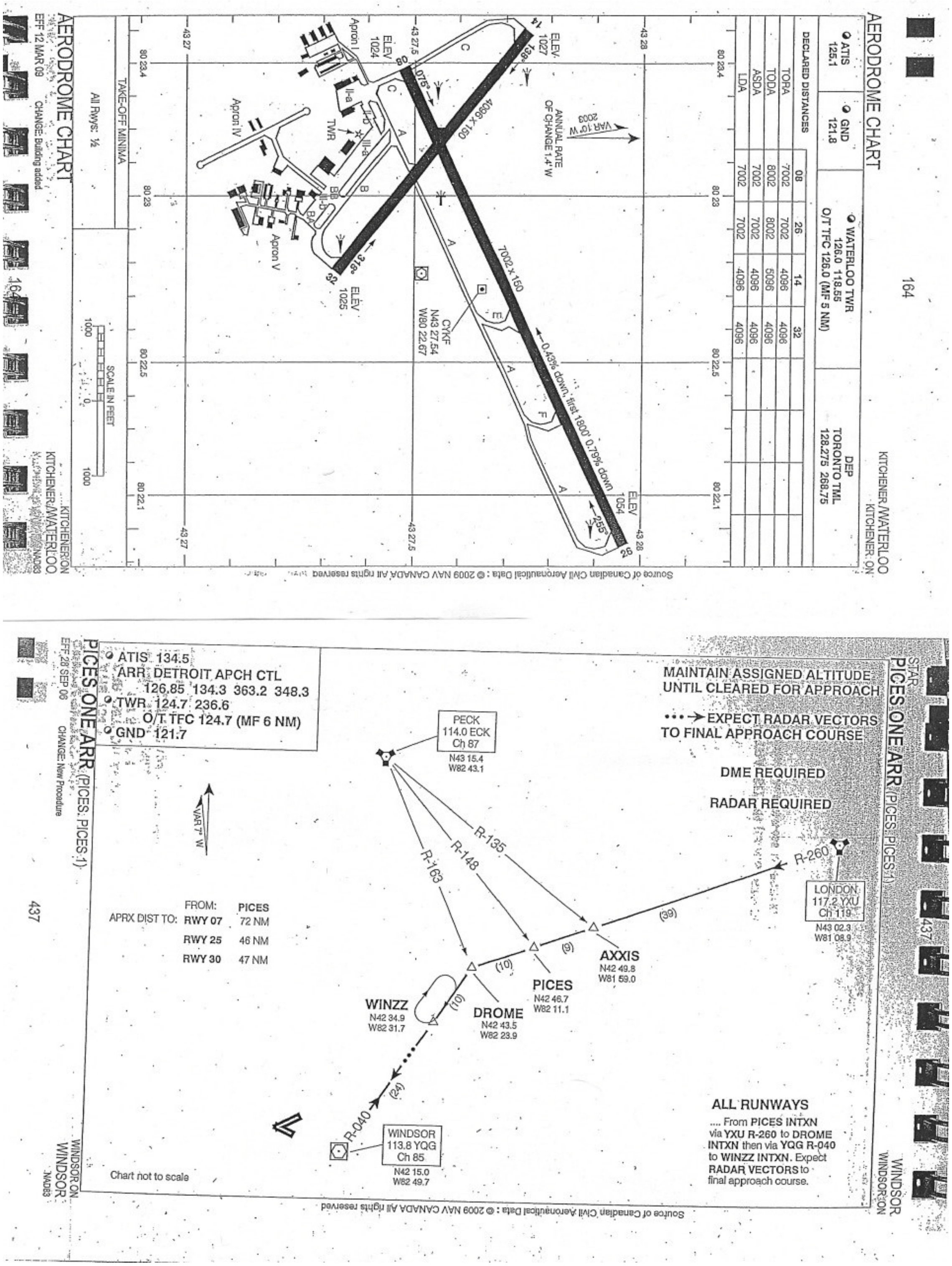


Figure E-9 – Waterloo (CYKF) Airport Diagram and Windsor (CYQG) PICES1 STAR Chart

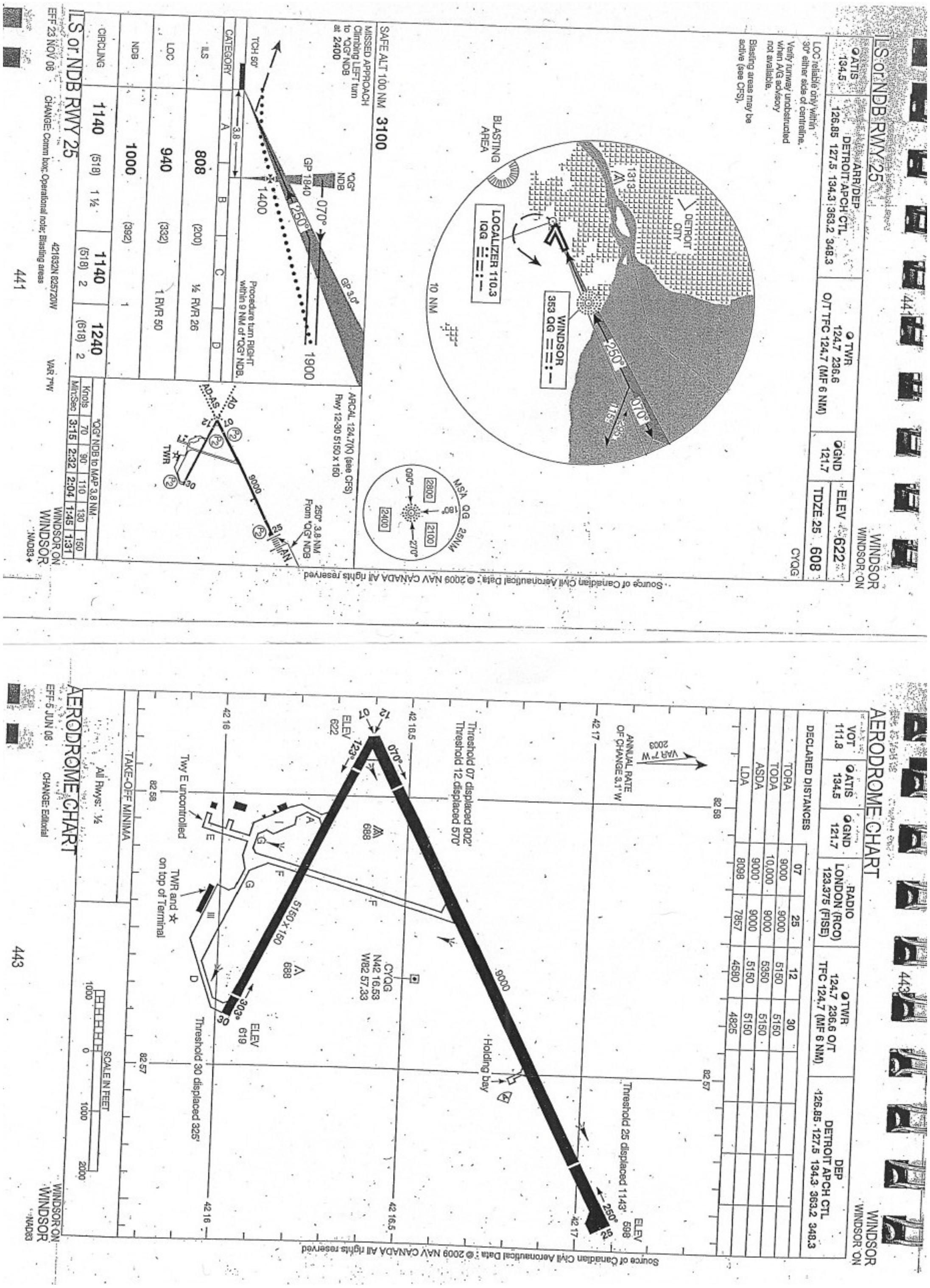


Figure E-10 – Windsor (CYQG) Approach Plate ILS Runway 25 and Airport Diagram

Appendix G
Video Prototype Evaluation Materials

INFORMATION LETTER

Title of Project: ***Design of Collaborative Systems for Modern Cockpits***

Student Investigator: ***Paul McKay, 519-888-4567 ext 36813, pdmckay@uwaterloo.ca***

Faculty Supervisor: ***Stacey Scott, 519-888-4567 ext 32236, s9scott@uwaterloo.ca***

Summary of the Project:
This project is the second part of a two-part research study aimed at improving the design of modern collaborative cockpit systems involving shared displays. The information gathered by the combined, two-part study will be used to develop design suggestions for cockpit interfaces that support effective individual and collaborative work.

Procedure:
Your participation in this study is voluntary. Participation involves viewing a video of a flight scenario on a prototype cockpit display setup and recording notes about the progress of the flight, including any interesting events or potential safety concerns. Because the study involves flight scenarios, all participants must have previous piloting experience, with a minimum of 15 flight hours. Participants also must have normal or corrected-to-normal vision. You will view the flight scenario individually, taking the role of a flight evaluator. A description of each activity follows.

You will be asked to fill out a brief background questionnaire (including a colour-blindness test) and will receive an introduction to the study and the prototype cockpit. You will then be asked to:

- Watch a full flight scenario and record notes about the status and progress of the flight.
- Complete a short questionnaire and a brief interview.

This session will take no more than two hours.

During the session a researcher will observe and take notes regarding your interactions with the cockpit systems. With your permission, a video recording of the full flight evaluation scenario will also be made, and any task materials produced during the session will remain with the researcher. You may decline to respond to questions if you wish. You may withdraw your participation at any time without penalty.

Confidentiality and Data Security:
All information you provide is considered completely confidential. Your name will not appear in any publication resulting from this study; however, with your permission anonymous quotations may be used. In these cases, participants will be referred to as Participant 1, 2, ... (or P1, P2, ...). Data collected during this study will be retained indefinitely in locked cabinets or on password protected desktop computers in a secure location. Electronic data will not include personal identifying information such as names.

You will be explicitly asked for consent for the use of photo/video/audio data, captured from the scenario recording, for the purpose of reporting the study's findings. If consent is granted, these data will be used only for the purposes associated with teaching, scientific presentations, publications, and/or sharing with other researchers and you will not be identified by name.

Remuneration for Your Participation:
You will receive remuneration for your participation in this study, for a total of $20 if you complete the session. If you choose to withdraw your participation from the study prior to study completion, you will be remunerated at a rate of $10 per hour of participation.

Risks and Benefits:
There are no known or anticipated risks from participation in this study. There are no direct benefits to you from participation. However, the results of this research may contribute to the knowledge base of Human Systems Engineering research and help in developing improved cockpit interfaces.

Research Ethics Clearance:
I would like to assure you that this study has been reviewed and received ethics clearance through the Office of Research Ethics at the University of Waterloo. However, the final decision about participation is yours. Should you have comments or concerns resulting from your participation in this study, please contact Dr. Susan Sykes in the Office of Research Ethics at 519-888-4567, Ext. 36005.

Thank you for your assistance in this project.

CONSENT FORM

Project: ***Design of Collaborative Systems for Modern Cockpits***

I have read the information presented in the information letter about a study being conducted by Paul McKay of the Department of Systems Design Engineering, under the supervision of Professor Stacey Scott. I have had the opportunity to ask any questions related to this study, to receive satisfactory answers to my questions, and any additional details I wanted.

Sometimes a certain image and/or segment of video or audio recording clearly shows a particular feature or detail that would be helpful in teaching or when presenting the study results at a scientific presentation or in a publication.

I am aware that I may allow video and/or digital images which show my computer interactions to be used in teaching, scientific presentations, publications, and/or sharing with other researchers with the understanding that I will not be identified by name. I am aware that I may allow excerpts from the conversational data collected for this study to be included in teaching, scientific presentations and/or publications, with the understanding that any quotations will be anonymous.

I am aware that I may withdraw my consent for any of the above statements or withdraw my study participation at any time without penalty by advising the researcher.

This project has been reviewed by, and received ethics clearance through, the Office of Research Ethics at the University of Waterloo. I was informed that if I have any comments or concerns resulting from my participation in this study, I may contact the Director, Office of Research Ethics at (519) 888-4567 ext. 36005.

	Please Circle One	**Please initial Your Choice**
With full knowledge of all foregoing, I agree, of my own free will to participate in this study.	YES NO	________
I agree to be videotaped.	YES NO	________
I agree to let any verbal comments made during the study be directly quoted, anonymously, in presentations of research results.	YES NO	________
I agree to let video, audio, or written data from the study be used for reports or presentations of the research results.	YES NO	________

Participant Name: ____
(Please print)

Participant Signature: _____________________________________

Witness Name: ____
(Please print)

Witness Signature: _______________________________________

Date: _______________________________

Background Questionnaire

Participant: P__ Condition: ___ Date: ________________________

Please fill out this questionnaire as accurately as possible. None of the information will be personally linked to you in any way. Please do not write your name anywhere on the questionnaire.

1. What is your sex?

Female Male

2. What is your age?

3. What is your occupation?

If student, what degree/program are you in?

4. How many flight hours have you completed?

5. What aircraft type(s) do you have experience flying (as pilot-in-command or co-pilot)?

6. Do you have experience using flight simulators of any kind?

If so, what flight simulators have you used, and how much would you estimate you use them? (e.g. 1 hour per week, 1 hour per month)

7. On the laptop screen in front of you, you'll see a short colour blindness-test, please write down the six numbers on your screen in order in the respective box on this page. (you might need to scroll down in the web browser)

1.	2.
3.	4.
5.	6.

Flight Scenario Log Sheet

Checklists

Time	Checklist	Time

Radio Frequency Changes

Time	Radio			New frequency
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	
	VHF1	VHF2	ILS	

ATC Communications/Clearances

Time	Communication/Clearance	Time

Autopilot Setting Changes

Time	On/Off	Speed target	Altitude target

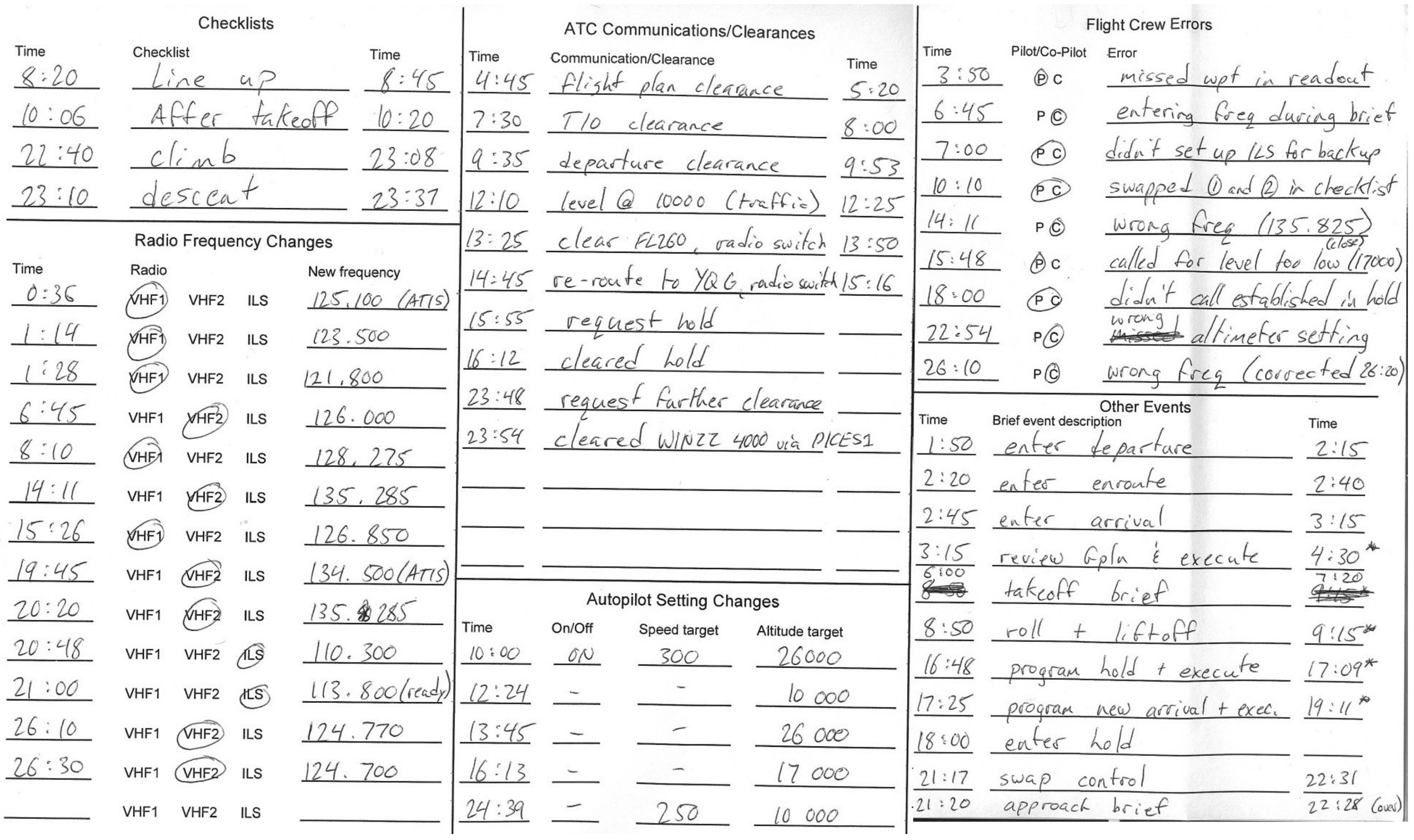
Flight Crew Errors

Time	Pilot/Co-Pilot	Error
	P C	
	P C	
	P C	
	P C	
	P C	
	P C	
	P C	
	P C	
	P C	

Other Events

Time	Brief event description	Time

Figure F-1 – Flight Scenario Log Sheet



Checklists

Time	Checklist	Time
8:20	Line up	8:45
10:06	After takeoff	10:20
22:40	climb	23:08
23:10	descent	23:37

Radio Frequency Changes

Time	Radio	New frequency
0:36	[✓ VHF1] VHF2 ILS	125.100 (ATIS)
1:14	[✓ VHF1] VHF2 ILS	123.500
1:28	[✓ VHF1] VHF2 ILS	121.800
6:45	VHF1 [✓ VHF2] ILS	126.000
8:10	[✓ VHF1] VHF2 ILS	128.275
14:11	VHF1 [✓ VHF2] ILS	135.285
15:26	[✓ VHF1] VHF2 ILS	126.850
19:45	VHF1 [✓ VHF2] ILS	134.500 (ATIS)
20:20	VHF1 [✓ VHF2] ILS	135.285
20:48	VHF1 VHF2 [✓ ILS]	110.300
21:00	VHF1 VHF2 [✓ ILS]	113.800 (ready)
26:10	VHF1 [✓ VHF2] ILS	124.770
26:30	VHF1 [✓ VHF2] ILS	124.700
	VHF1 VHF2 ILS	

ATC Communications/Clearances

Time	Communication/Clearance	Time
4:45	flight plan clearance	5:20
7:30	T/O clearance	8:00
9:35	departure clearance	9:53
12:10	level @ 10000 (traffic)	12:25
13:25	clear FL260, radio switch	13:50
14:45	re-route to YQG, radio switch	15:16
15:55	request hold	
16:12	cleared hold	
23:48	request further clearance	
23:54	cleared WINZZ 4000 via PICES1	

Autopilot Setting Changes

Time	On/Off	Speed target	Altitude target
10:00	ON	300	26000
12:24	-	-	10 000
13:45	-	-	26 000
16:13	-	-	17 000
24:39	-	250	10 000

Flight Crew Errors

Time	Pilot/Co-Pilot	Error
3:50	[✓ P] C	missed wpt in readout
6:45	P [✓ C]	entering freq during brief
7:00	[✓ P C]	didn't set up ILS for backup
10:10	[✓ P C]	swapped ① and ② in checklist
14:11	P [✓ C]	wrong freq (135.825) (close)
15:48	[✓ P] C	called for level too low (17000)
18:00	[✓ P C]	didn't call established in hold
22:54	P [✓ C]	~~missed~~ wrong altimeter setting
26:10	P [✓ C]	wrong freq (corrected 26:20)

Other Events

Time	Brief event description	Time
1:50	enter departure	2:15
2:20	enter enroute	2:40
2:45	enter arrival	3:15
3:15	review Fpln & execute	4:30*
~~8:50~~ 6:00	takeoff brief	~~9:15*~~ 7:20
8:50	roll + liftoff	9:15*
16:48	program hold + execute	17:09*
17:25	program new arrival + exec.	19:11*
18:00	enter hold	
21:17	swap control	22:31
21:20	approach brief	22:28 (over)

Figure F-2 – Master Flight Scenario Log Sheet

Flight Scenario Questionnaire

Participant: P__ Condition: ___ Date: ________________________

Rate your answers to the questions below on the provided 7-point scale.

Overall, how would you rate the performance of the flight crew?

Very Poor Excellent

1 2 3 4 5 6 7

How would you rate the performance of the pilot-in-command?

Very Poor Excellent

1 2 3 4 5 6 7

How would you rate the performance of the co-pilot?

Very Poor Excellent

1 2 3 4 5 6 7

How confident are you that your flight scenario notes captured all of the scenario information?

Not Confident Very Confident

1 2 3 4 5 6 7

Interview Questions

Questions for all participants (control and augmented display):

- (If confidence rating on last question of questionnaire is 4 or less) – What part of your flight scenario notes do you feel is missing information?
- Can you give a brief description of what happened in the scenario?
- On a scale of 1 to 7, how confident are you that your description is accurate?
- Can you briefly explain your performance ratings of the flight crew?
- Which information did you find easy to keep track of?
- Which information was more difficult?
- Were there any particular times or sequences of time during the scenario that you found challenging to track? If so, when and why?

NOTE: Participants will be shown a static picture of the interface while answering questions below.

- Did you find the interface helpful in tracking the status and progress of the flight?
- Did you find the interface helpful in tracking what the pilots were doing?
- If so, what aspects of the interface did you find useful in carrying out these tasks?
- What could have been added to the interface to make these tasks easier?

Extra questions for participants using augmented display:

- Were you easily able to understand the difference between the "window" type borders and the "component" type borders?
- Which type of border did you find more useful?
- Did you find the level of detail in the "component" type borders appropriate? Would you have liked more or less?
- Did your usage or understanding of the borders change over the course of the scenario?

Final question for all participants:

- Do you have any other questions or comments about anything (scenario, process, etc)?

Appendix H

Statistical Analysis

Unpaired t-test for Crew Rating
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	-.300	9	-.980	.3527

Group Info for Crew Rating
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	5.200	.200	.447	.200
Treatment	6	5.500	.300	.548	.224

Figure G-1 – Statistical analysis of overall crew performance rating

Unpaired t-test for Pilot Rating
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	-.133	9	-.325	.7527

Group Info for Pilot Rating
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	5.200	.200	.447	.200
Treatment	6	5.333	.667	.816	.333

Figure G-2 – Statistical analysis of pilot performance rating

Unpaired t-test for Co-Pilot Rating
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	-.167	9	-.376	.7159

Group Info for Co-Pilot Rating
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	5.000	.500	.707	.316
Treatment	6	5.167	.567	.753	.307

Figure G-3 – Statistical analysis of co-pilot performance rating

Unpaired t-test for Log Sheet Conf.
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	-.700	9	-1.108	.2967

Group Info for Log Sheet Conf.
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	4.300	1.450	1.204	.539
Treatment	6	5.000	.800	.894	.365

Figure G-4 – Statistical analysis of scenario log sheet confidence rating

Unpaired t-test for Secnario Descr. Conf.
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	-.700	7	-1.416	.1996

Group Info for Secnario Descr. Conf.
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	4	5.500	.333	.577	.289
Treatment	5	6.200	.700	.837	.374

Figure G-5 – Statistical analysis of scenario description confidence ratings

Unpaired t-test for Checklist
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	1.000	9	.905	.3893

Group Info for Checklist
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	3.000	1.500	1.225	.548
Treatment	6	2.000	4.800	2.191	.894

Figure G-6 – Statistical analysis of checklist scores

Unpaired t-test for Radio Frequency
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	.267	9	.256	.8033

Group Info for Radio Frequency
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	9.600	2.800	1.673	.748
Treatment	6	9.333	3.067	1.751	.715

Figure G-7 – Statistical analysis of radio frequency scores

Unpaired t-test for Correct Radio
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	1.200	9	.670	.5198

Group Info for Correct Radio
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	9.200	3.700	1.924	.860
Treatment	6	8.000	12.800	3.578	1.461

Figure G-8 – Statistical analysis of correct radio scores

Unpaired t-test for ATC
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	-.100	9	-.191	.8525

Group Info for ATC
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	7.400	.800	.894	.400
Treatment	6	7.500	.700	.837	.342

Figure G-9 – Statistical analysis of ATC clearance scores

Unpaired t-test for Autopilot
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	1.233	9	1.047	.3222

Group Info for Autopilot
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	4.400	.800	.894	.400
Treatment	6	3.167	6.167	2.483	1.014

Figure G-10 – Statistical analysis of autopilot scores

Unpaired t-test for Correct Alt
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	.033	9	.024	.9810

Group Info for Correct Alt
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	3.200	3.700	1.924	.860
Treatment	6	3.167	6.167	2.483	1.014

Figure G-11 – Statistical analysis of autopilot altitude setting scores

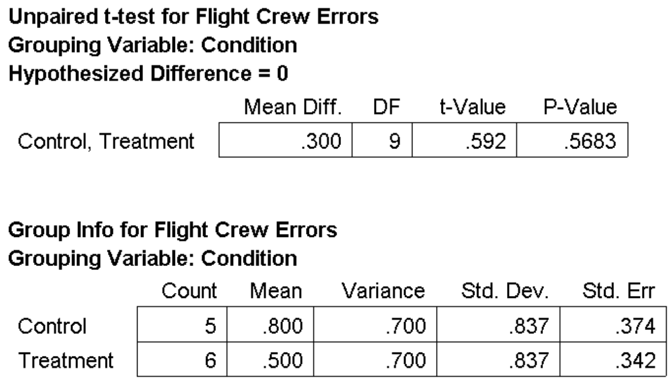
Unpaired t-test for Correct Speed
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	.567	9	.423	.6822

Group Info for Correct Speed
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	3.400	4.300	2.074	.927
Treatment	6	2.833	5.367	2.317	.946

Figure G-12 – Statistical analysis of autopilot speed setting scores



Unpaired t-test for Flight Crew Errors
Grouping Variable: Condition
Hypothesized Difference = 0

	Mean Diff.	DF	t-Value	P-Value
Control, Treatment	.300	9	.592	.5683

Group Info for Flight Crew Errors
Grouping Variable: Condition

	Count	Mean	Variance	Std. Dev.	Std. Err
Control	5	.800	.700	.837	.374
Treatment	6	.500	.700	.837	.342

Figure G-13 – Statistical analysis of flight crew error scores