NAVAL
POSTGRADUATE
SCHOOL

MONTEREY, CALIFORNIA

THESIS

**A HUMAN FACTORS EVALUATION OF THE
SPACE SHUTTLE COCKPIT AVIONICS UPGRADE**

by

Michael Westenhaver

September 2012

Thesis Advisor:	John Schmidt
Thesis Co-Advisor:	Daniel Bursch
Second Reader:	Christian Smith

Approved for public release; distribution is unlimited

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE			*Form Approved OMB No. 0704-0188*
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instruction, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188) Washington DC 20503.			
1. AGENCY USE ONLY *(Leave blank)*	**2. REPORT DATE** September 2012	**3. REPORT TYPE AND DATES COVERED** Master's Thesis	
4. TITLE AND SUBTITLE A Human Factors Evaluation of the Space Shuttle Cockpit Avionics Upgrade		**5. FUNDING NUMBERS**	
6. AUTHOR(S) Michael S. Westenhaver			
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943-5000		**8. PERFORMING ORGANIZATION REPORT NUMBER**	
9. SPONSORING /MONITORING AGENCY NAME(S) AND ADDRESS(ES) N/A		**10. SPONSORING/MONITORING AGENCY REPORT NUMBER**	
11. SUPPLEMENTARY NOTES The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. Government. IRB Protocol number ___N/A____.			
12a. DISTRIBUTION / AVAILABILITY STATEMENT Approved for public release; distribution is unlimited		**12b. DISTRIBUTION CODE** A	
13. ABSTRACT (maximum 200 words) During the late 90s, NASA retrofitted the Space Shuttle fleet with a "glass cockpit." The new displays replicated legacy formats developed in the 70s, and did not leverage 20 years of display technology and human factors advances. To address this shortcoming the Cockpit Avionics Upgrade (CAU) was initiated to reduce mental workload (MW), increase situational awareness (SA), and enhance performance. Despite the CAU demonstrating improvements in MW, SA, and performance, it was cancelled. Consequently, recorded astronaut data from using the baseline and CAU cockpit configurations was never tied back to cockpit design. This study assesses the CAU design employing human factors principles, evaluates baseline and CAU simulation data, and traces MW and SA differences back to CAU design modifications. Significant improvements were found in all measures and across all conditions. These improvements were found to be greater for ascent scenarios than for entry. From the findings, recommendations for the design and evaluation of future spacecraft cockpits are made.			
14. SUBJECT TERMS Human Factors, Mental Workload, Situation Awareness, Space Shuttle, NASA, Display Design			**15. NUMBER OF PAGES** 119
			16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	**18. SECURITY CLASSIFICATION OF THIS PAGE** Unclassified	**19. SECURITY CLASSIFICATION OF ABSTRACT** Unclassified	**20. LIMITATION OF ABSTRACT** UU

THIS PAGE INTENTIONALLY LEFT BLANK

A HUMAN FACTORS EVALUATION OF THE
SPACE SHUTTLE COCKPIT AVIONICS UPGRADE

Michael S. Westenhaver
Lieutenant, United States Navy
B.S., Illinois Institute of Technology, 2004

Submitted in partial fulfillment of the
requirements for the degrees of

MASTER OF SCIENCE IN SPACE SYSTEMS OPERATIONS

and

MASTER OF SCIENCE IN HUMAN SYSTEMS INTEGRATION

from the

NAVAL POSTGRADUATE SCHOOL
September 2012

Author: Michael Westenhaver

Approved by: John K. Schmidt
               Thesis Advisor

               Daniel Bursch
               Thesis Co-Advisor

               Christian Smith
               Second Reader

               Robert Dell
               Chair, Department of Operations Research

               Rudolph Panholzer
               Chair, Space Systems Academic Group

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

During the late 90s, NASA retrofitted the Space Shuttle fleet with a "glass cockpit." The new displays replicated legacy formats developed in the 70s, and did not leverage 20 years of display technology and human factors advances. To address this shortcoming the Cockpit Avionics Upgrade (CAU) was initiated to reduce mental workload (MW), increase situational awareness (SA), and enhance performance. Despite the CAU demonstrating improvements in MW, SA, and performance, it was cancelled. Consequently, recorded astronaut data from using the baseline and CAU cockpit configurations was never tied back to cockpit design. This study assesses the CAU design employing human factors principles, evaluates baseline and CAU simulation data, and traces MW and SA differences back to CAU design modifications. Significant improvements were found in all measures and across all conditions. These improvements were found to be greater for ascent scenarios than for entry. From the findings, recommendations for the design and evaluation of future spacecraft cockpits are made.

THIS PAGE INTENTIONALLY LEFT BLANK

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF ACRONYMS AND ABBREVIATIONS

ADI	Attitude Director Indicator
AMI	Alpha/Mach Indicator
APU/HYD	Auxiliary Power Unit/Hydraulic System
ATO	Abort to Orbit
AVVI	Altitude/Vertical Velocity Indicator
BFS	Backup Flight System
BWRS	Bedford Workload Rating Scale
CAU	Cockpit Avionics Upgrade
CDR	Commander
CRT	Cathode Ray Tube
DOD	Department of Defense
DPS	Data Processing System
DST	Dynamic Skills Trainer
ECLSS	Environmental Control/Life Support System
EPS	Electrical Power System
FDF	Flight Data File
GNC	Guidance, Navigation, and Control
GPC	General Purpose Computer
HFE	Human Factors Engineering
HSI	Horizontal Situation Indicator
H-SIT	Horizontal Situation Display
HTA	Hierarchical Task Analysis
IDP	Integrated Display Processor
LCD	Liquid Crystal Display
MCC	Mission Control Center
MDU	Multifunction Display Unit
MEDS	Multifunction Electronic Display System
MFD	Multifunction Display
MPS	Main Propulsion System
MS2	Mission Specialist Two

MW	Mental Workload
NASA	National Air and Space Administration
OMS	Orbital Maneuvering System
OI	Orbiter Increment
PASS	Primary Avionics System Software
PLT	Pilot
RCS	Reaction Control System
RTLS	Return to Launch Site
SA	Situation Awareness
SAGAT	Situation Awareness Global Assessment Technique
SCOM	Shuttle Crew Operations Manual
SM	System Management
SUM	Summary
TAL	Transoceanic Abort Landing
TLX	Task Load Index
TRAJ	Trajectory Display

EXECUTIVE SUMMARY

When the Space Shuttle was originally developed in the 1970s, the cockpit was based on the latest technology at the time including fly by wire controls and computer displays. As that technology became obsolete, NASA retrofitted the Space Shuttle fleet with a modern "glass cockpit." At first the new displays replicated legacy formats, and did not leverage nearly 30 years of advances in display technology and human factors design. To address this shortcoming, the Cockpit Avionics Upgrade (CAU) project was initiated to update display formats to reduce mental workload (MW), increase situational awareness (SA), and enhance performance. Despite demonstrating improvement in terms of MW, SA, and performance over the baseline design, the CAU project was cancelled in 2004 due to budgetary constraints. Consequently, astronaut performance data using the baseline and CAU configurations was not completely analyzed and tied back to the adopted design modifications. The purpose of this study was to critique the CAU design employing human factors principles, evaluate baseline and CAU MW and SA data, and trace back specific design modifications that contributed to improvements in MW and SA.

A literature review was conducted to identify how human factors principles can be used to effectively organize and present information in a design concept for a cockpit. It included consideration of relevant methods and criteria to objectively critique a design concept, including task analysis, display design principles, design eye position, and display-control compatibility. This was followed by an examination of the theories and methods used in design assessment, including measures of mental workload and situation awareness. The information reviewed includes scientific journals, textbooks, NASA documents, technical standards, and related prior work.

Based on the literature review, a process was established to critically evaluate a cockpit design change and assessing its effectiveness. The evaluation began with a task analysis. This was followed by a task-oriented evaluation of

the layout of the displays from the operator perspective, the formatting of information on each display, the hardware attributes, and the display-control relationship. The design critique noted that the changes in the use of color, the consolidation of information, and the use of pictorial representations of systems are relevant to all crew positions and in all phases of flight. The critique also identified the predictive aiding features of the ascent horizontal situation display as the likely cause of the variation between ascent and entry scenarios.

To assess the effectiveness of the CAU design changes over the baseline design, MW and SA data were analyzed. For every MW and SA measure, there was a significant improvement favoring the CAU design vs. the baseline for all crew positions and in all phases of flight. The analysis also revealed that these improvements were significantly greater for ascent operations than for entry.

It is recommended that future manned spacecraft cockpits incorporate the design principles demonstrated in the CAU, including the use of a consistent color scheme, task-centric consolidation of information, and the use of pictorial representations of complex systems. It is further suggested that predictive aiding displays should be developed wherever possible for use during all dynamic phases of flight. Finally, it is recommended that the evaluation of these future systems should include a mechanism for gathering operator feedback on the relative usefulness of individual design principles.

ACKNOWLEDGMENTS

First and foremost, I would like to thank my beautiful wife, Kari. I am eternally grateful for your patience, love and understanding throughout this difficult journey. I would also like to sincerely thank my advisor, CAPT John Schmidt, whose mentorship and guidance have extended far beyond this thesis. Thanks to Dr. Kip Smith for providing clarity where it was needed most. To CAPT Dan Bursch, thanks for the technical insight and perspective. To Johnny O'Connor, your timely contribution was essential. To Dr. McCann, this wouldn't have been possible without your continual support. Lastly, I would like to recognize the late CAPT Alan Poindexter, without whose enthusiastic support this thesis would not have happened.

THIS PAGE INTENTIONALLY LEFT BLANK

I. INTRODUCTION

A. BACKGROUND

With the Space Shuttle program ending in 2011, the United States no longer possesses the capability to launch astronauts into space. Many follow on programs were proposed over the years, but none have been realized to date (Frank, 2010). The Orion Multi-Purpose Crew Vehicle is currently being developed as NASA's next generation manned spacecraft, but it is not expected to be operational until at least 2021 (Stanfield, 2012). This gap will be the longest the United States has gone without possessing its own launch capability since the 1961 dawn of manned spaceflight. It is inevitable that manned spaceflight will resume in the United States, but until then it is imperative that we maintain our technological expertise, and continue to build on the lessons learned in past programs.

When the Space Shuttle was originally developed in the 1970s, the cockpit was based on the latest technology available at the time. The avionics suite incorporated technologies such as fly by wire controls and computer displays, which were considered quite advanced at the time (McCandless, 2004). The two decades that followed saw the development of glass cockpit technology in the commercial sector, driven by the ever-increasing complexity of air transport operations. Despite the fact that NASA was very much involved in this effort, the Space Shuttle cockpit remained unchanged for this period (Tranthien, 1995).

While the shuttle's original cockpit equipment was extremely reliable, it was nonetheless costly to maintain over time (McCandless, 2004). In 2000, the original electro-mechanical cockpit instrumentation was replaced with the Multifunction Electronic Display System (MEDS) glass cockpit (McCandless et al., 2005). The MEDS cockpit upgraded all of the hardware to modern liquid crystal displays, but in many areas it retained the legacy layouts of information (Marchant, Eastin, & Ferguson, 2001). While this approach solved the

1

maintenance and supportability issues, it did not resolve any of the legacy human factors concerns, and did not fully leverage the advances that had been made in glass cockpit technology. The system still displayed graphics and text in monochrome, still required multiple key-presses to read system and subsystem information, and still presented information in closely spaced rows of digital numbers, making it difficult to locate off-normal values (McCandless et al., 2005).

Figure 1.   Shuttle Glass Cockpit (From NASA, 2000)

To address these remaining shortcomings, the Cockpit Avionics Upgrade (CAU) project was then initiated in 1999 (Marchant et al., 2001). CAU was intended to update the display formats shown on the cockpit displays in order to reduce mental workload (MW), enhance situational awareness (SA), and increase performance, thereby improving overall system safety (McCandless, 2004). Based on human factors research, the program produced a design for enhanced display formats and keysets. The new formats were intended to be task-oriented and tailored to each phase of flight (McCandless, 2004).

Despite demonstrating considerable improvement over the baseline MEDS cockpit in workload, SA, and performance, CAU was cancelled in 2004,

and no orbiter vehicle ever flew with the enhanced cockpit (McCandless, 2004). To evaluate the effectiveness of the CAU, thousands of data points on human performance were collected as part of this project, but since the cancellation, those data have never been fully analyzed (McCandless, 2004).

B. PURPOSE

As part of the CAU evaluation, the data were analyzed only so far as to determine that significant improvement did exist in the CAU cockpit in MW and SA. To date a complete analysis has not been completed to assess why astronaut performance improved, what design changes contributed to these improvements, or what crew activities were most affected. Furthermore, the breadth of data collected allows the opportunity to examine which measures are most useful in predicting overall crew performance. Therefore, there are three primary objectives of this thesis: 1) complete a thorough assessment of the CAU evaluation data in order to gain understanding of why an improvement exists, 2) develop design recommendations for future manned space systems, and 3) develop recommendations for methods of evaluating astronaut performance in these future systems.

C. PROBLEM STATEMENT

The CAU project was intended to leverage over 20 years of advances that had been made in human factors since the Shuttle was originally designed in the 1970s. Despite demonstrating considerable improvements in MW and SA, CAU was cancelled in 2004 and the data that were collected during the program was never been fully analyzed. With the retirement of the Space Shuttle, the impetus to revisit the data has waned as well.

It is inevitable that the United States eventually will return to manned spaceflight in an American spacecraft. When this happens, it is essential that existing lessons learned in safety, situational awareness, and mental workload all be leveraged. Through the CAU program, a great deal of data have already

been collected in all of these areas. It is imperative that this data be translated into actionable knowledge that can be applied to future manned spacecraft.

In examining the data for the CAU evaluation study the following research questions are proposed:

1) Did the CAU design changes lead to improvements in crew mental workload and/or situation awareness?

2) If mental workload and/or situation awareness improvements exist, do they vary by crew station or phase of flight?

3) Can crew mental workload and/or situation awareness changes be traced back to specific CAU design modifications?

D. HUMAN SYSTEM INTEGRATION DOMAIN APPLICABILITY

NASA describes Human System Integration as "an umbrella term for several areas of 'human factors' research that include human performance, technology design, and human-computer interaction" (NASA, 2012). The Department of Defense (DOD) further identifies seven domains of Human Systems Integration: human factors engineering (HFE), personnel, habitability, manpower, training, safety and occupational health, and survivability (DOD, 2008). This thesis relates to HFE, training, and system safety domains.

HFE is defined as "the comprehensive integration of human characteristics into system definition, design, development, and evaluation to optimize the performance of human-machine combinations" (Booher, 2003). This study involves the application of HFE principles in evaluating a set of display designs of the Space Shuttle. The lessons learned from this evaluation can be applied to future activities in the HFE domain.

Training is defined as the "the requisite knowledge, skills, and abilities needed by the available personnel to operate and maintain systems under operational conditions" (Booher, 2003). The design modifications evaluated in

this study alter the task requirements, and therefore affect the knowledge and skills required to operate the system. Consequently, the training domain is affected.

System Safety is defined as "the inherent ability of the system to be used, operated, and maintained without accidental injury to personnel" (Booher, 2003). The primary objective of the CAU program was to improve overall system safety by improving the quality of the human-computer interface (McCandless, 2004). Therefore, this evaluation of the CAU program relates to the system safety domain.

E.   ORGANIZATION

This thesis is organized into five chapters. Chapter I provides an overview of the Space Shuttle Cockpit Avionics Upgrade program, including rationale and objectives. Chapter II provides a review of scientific and technical literature regarding Space Shuttle systems, cockpit evaluation, and flight crew performance assessment. Chapter III describes the methods by which the CAU modifications were evaluated. Chapter IV presents the results of the evaluation. Chapter V offers conclusions and recommendations based on the findings.

THIS PAGE INTENTIONALLY LEFT BLANK

II. LITERATURE REVIEW

A. OVERVIEW

The intent of this chapter is to provide the context necessary for understanding and evaluating the Shuttle CAU program. It begins with a historical and technical review of shuttle cockpit avionics systems. Next there is a review of relevant methods and criteria by which the design upgrades can be objectively critiqued. Finally, there is a review of the theories and methods by which the effectiveness of the upgrades can be assessed experimentally. The information reviewed includes scientific journals, textbooks, NASA documents, technical standards, and related prior work.

The literature review process began with a review of prior work in the CAU program. This included materials obtained directly from the NASA Ames Research Center, as well as those that were found by searching Google Scholar for the terms "Shuttle" and "CAU." To provide the necessary conceptual foundations, the search was extended to include such topics as mental workload, situation awareness, and information processing. Additional electronic resources searched included The Naval Postgraduate School Library, Defense Technical Information Center, and NASA Technical Report Server. Available textbooks were also reviewed for these topics. To gain a practical perspective on cockpit evaluation methods, additional resources were obtained from the U.S. Naval Test Pilot School.

B. SHUTTLE COCKPIT SYSTEMS

The Space Shuttle was one of the first major applications of fly-by-wire control technology (Marchant et al., 2001). The onboard Data Processing System (DPS) controls almost all shuttle operations from ascent to landing, including flight control, system management (SM), and guidance navigation and control

(GNC). Most tasks during ascent and entry are fully automated, with the crew monitoring for various takeover criteria (NASA, 2002).

At the heart of the DPS system were five IBM APC-101S general purpose computers (GPCs) (NASA, 2002). Four of these GPCs ran the shuttle's Primary Avionics System Software (PASS), while the fifth ran the Backup Flight System (BFS). PASS was designed to manage all shuttle operations from launch to landing. BFS was much simpler, intended to provide only the capabilities needed to safely launch and recover the shuttle. Each software suite was developed separately to minimize the possibility that a programming defect in one would affect the other (NASA, 2002). Both software suites performed critical functions, and crewmembers often were required to utilize both of them throughout normal flight operations (Holland & Vanderark, 1993; NASA, 2002).

DPS software was divided by phase of flight into Operational Sequences (OPS), and further subdivided into major modes (see Figure 2). Transitions between OPS were initiated by the crew, but transitions between MMs could be done automatically (NASA, 2002). Each MM included an associated display page that presented the crew with information concerning that portion of the mission phase (NASA, 2002).
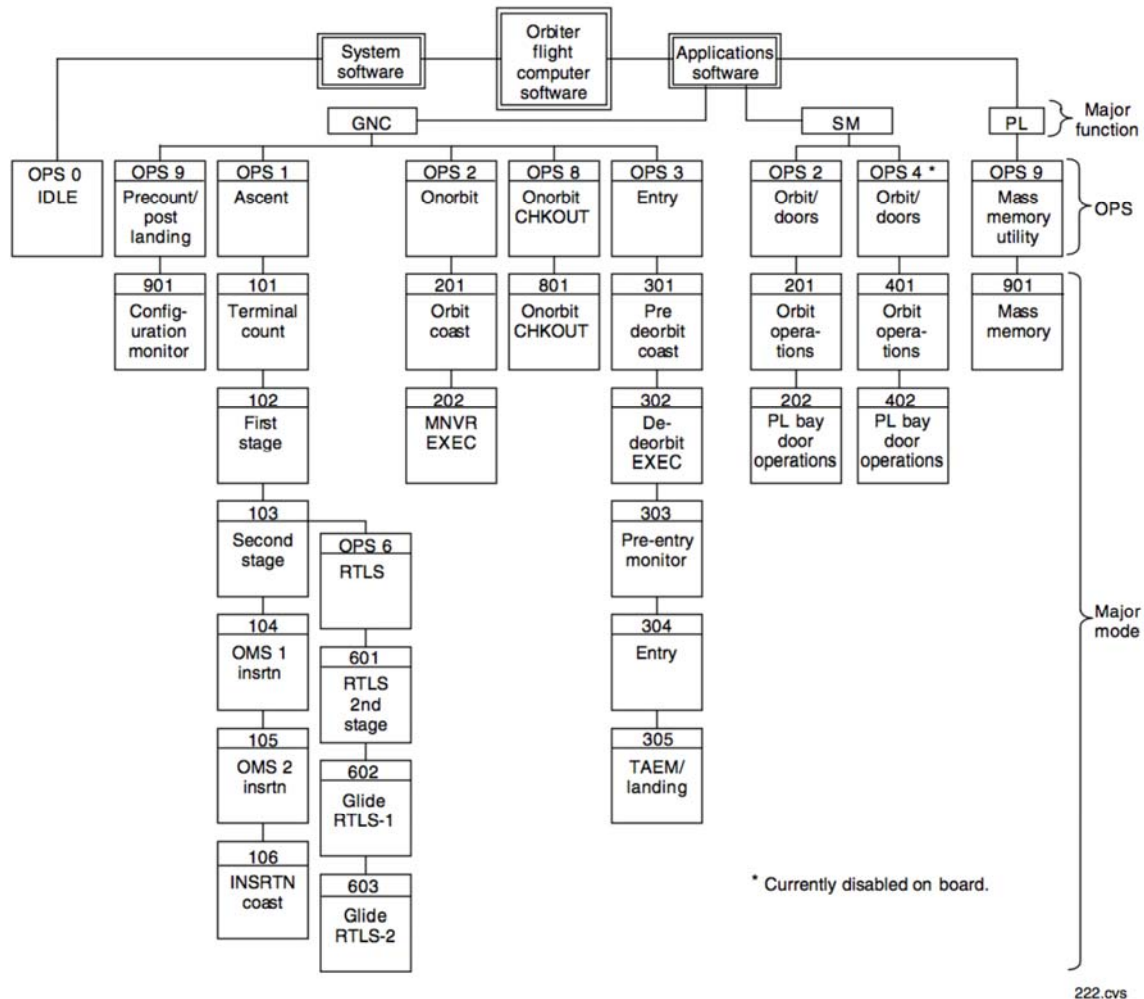
Figure 2.   Shuttle Flight Computer Software (NASA, 2002)

In order to discuss the shuttle's glass cockpit in context, it is helpful to begin by reviewing the layout of the original electro-mechanical cockpit (see Figure 3). In this legacy arrangement, electrically driven mechanical displays were provided for monitoring flight performance and critical systems (Tranthien, 1995). This included an Attitude Director Indicator (ADI), Horizontal Situation Indicator (HSI), Alpha/Mach Indicator (AMI), Altitude/Vertical Velocity Indicator (AVVI), and various moving-tape instruments for monitoring critical systems such as Main Propulsion (MPS), Orbital Maneuvering (OMS), and the Auxiliary Power Unit driven Hydraulic system (APU/HYD). Three multi-function cathode ray tube

(CRT) displays and two corresponding keypads were provided to interface with the DPS. Display pages unique to each MM were presented on these three displays. Due to the degree of computer control aboard the shuttle, the bulk of operating parameters were monitored by the crew solely through this DPS interface (Marchant et al., 2001). Each CRT Display/ Keyboard pair could be switched to connect to any of the five GPCs (NASA, 2002).
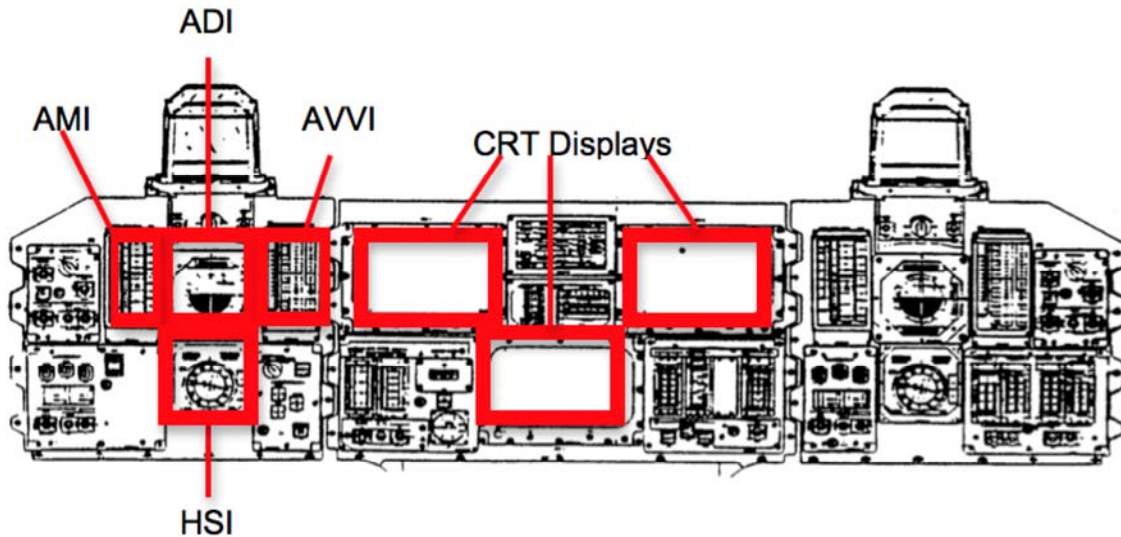
Figure 3.   Legacy Electromechanical Cockpit Layout (From Tranthien, 1995)

The original (MEDS) glass cockpit upgrade had nine identical Multifunction Display Units (MDUs) installed in the forward panel of the shuttle cockpit (see Figure 4) (NASA, 2002). Two MDUs were installed directly in front of the commander and pilot stations, respectively labeled CDR1, CDR2, PLT1, and PLT2. The remaining five MDUs, labeled MFD1, MFD2, CRT1, CRT2, and CRT3 are located between the two crew stations (NASA, 2002).
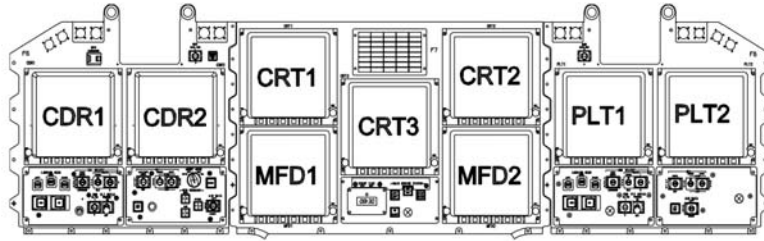
Figure 4.   MEDS Cockpit Display Layout (From NASA, 2002)

The nine displays are driven by four onboard Integrated Display Processors (IDPs), which receive information from five data busses and four analog-to-digital converters (NASA, 2002).   This information is processed for display and sent to the MDUs via a 1553B data bus (see Figure 5).   The CDR, PLT, and MFD displays are each connected to a primary and secondary IDP. The CRT displays are each only connected to a single IDP.   The system is configured to ensure that at least one CDR, one PLT, and one CRT display remain available in the event of a dual IDP failure.   The five GPCs, along with the PASS and BFS software, were not affected by the glass cockpit installation (Tranthien, 1995; NASA, 2002).
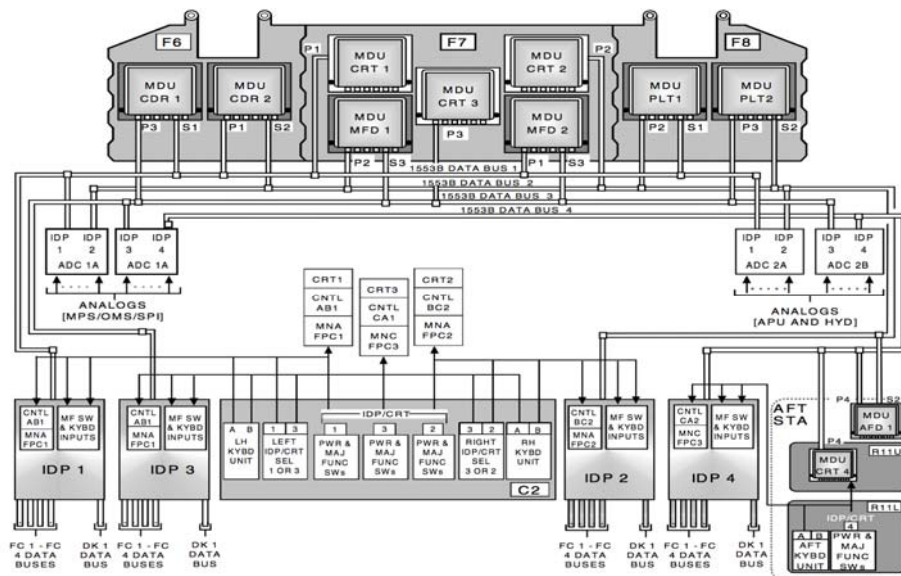
Figure 5.   MEDS System Overview (From NASA, 2002)

The Shuttle Crew Operating Manual (SCOM) divides the set of available display formats in the baseline MEDS configuration into flight, subsystem, and DPS formats (NASA, 2002). Flight formats display graphical representations of the legacy electromechanical flight instruments (ADI, HSI, AMI, AVVI, etc.). Subsystem displays include APU/HYD and MPS/OMS subsystem vertical tapes, as well as the Surface Position Indicator (SPI). The DPS display formats (which varied according to the current MM) were carried over directly from the legacy CRT display formats. In general, the CDR and PLT MDUs are used to show the flight instruments, the MFD MDUs are used to show the subsystem displays and the CRT MDUs are used to show DPS displays (see Figure 6).
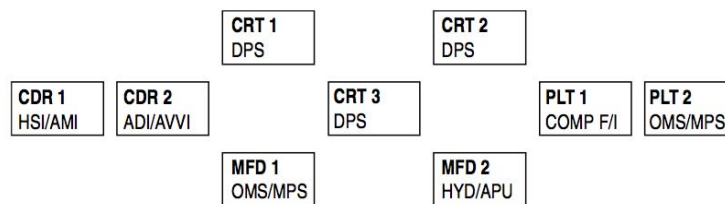
Figure 6.   Typical Baseline Ascent Display Configuration (From NASA, 2002)

The CAU called for minimal hardware upgrades, however the display formats, however, differed greatly from the baseline MEDS (Hayashi et al., 2005). CAU formats were divided into Flight, System, and Fault categories (Patrick & Mastracchio, 2002). The Flight category included primary flight display (PFD), trajectory (TRAJ), and horizontal situation (H-SIT) formats, each tailored to the current MM. The System category included specific pages for ECLSS, DPS, Navigation, Control, RCS, OMS/MPS, APU/Hyd, and EPS subsystems. Fault pages provided overall failure annunciation and diagnostic information, and were intended to be the focal point for critical systems monitoring (Reisman, 2002). Figure 7 presents a typical arrangement of CAU display formats.
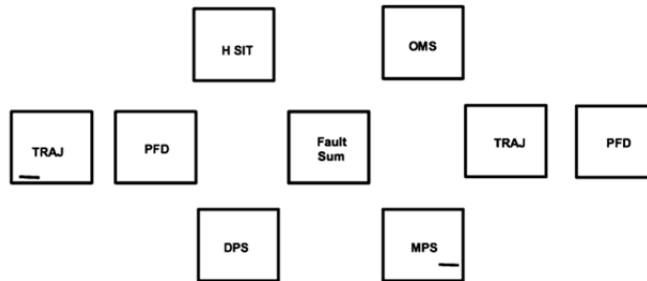
Figure 7.   Typical CAU Display Configuration for Ascent (From Reisman, 2002)

C.   DESIGN CRITERIA

The U.S. Naval Test Pilot School Flight Test Manual provides a methodology and recommended guidelines for conducting a cockpit evaluation. A central element of this evaluation method is a focus on the mission and tasks to be performed by the flight crew (Masters et al., 2005). The list of cockpit elements to be evaluated should include the layout of displays and controls as seen from the design eye position, the hardware attributes of the various components, the formatting of display content, and the interrelationships between controls and displays (Masters et al., 2005). To enable a thorough cockpit evaluation using this method, it is helpful to review the relevant aspects of task analysis, cockpit layout, display formatting, hardware attributes, and the display-control relationship.

1.   Task Analysis

Task analysis is defined as the study of the actions and cognitive processes required of an operator to achieve a system goal (Kirwan & Ainsworth, 1992). There are two complementary task analysis procedures that will be employed in this evaluation. The first is hierarchical task analysis (HTA), which is a method used to define an activity by breaking it down into its respective components (Kirwan & Ainsworth, 1992). The second method, task decomposition, is a way of expanding the information from the task description into a series of detailed statements which are of analytical interest (Kirwan &

Ainsworth, 1992). For this evaluation, the goal of the HTA was to identify the display-related crewmember tasks during ascent and entry operations. The goal of the task decomposition was to identify information requirements, challenges, and opportunities for error for each identified task.

Kirwan and Ainsworth (1992) identified several potential means of gathering information for a task analysis, including observation, interviews with subject matter experts, and examination of documentation. Potential document sources can include operating manuals and emergency procedures (Kirwan & Ainsworth, 1992). Several sources are available for gathering information on Space Shuttle flight crew tasks. Most notably, the SCOM provides an overview of all normal procedures for all phases of flight, and the crew pocket checklist provides details on contingency and emergency procedures (NASA, 2002). Additionally, Holland and Vanderark (1993) used operating manuals, astronaut interviews, and observations in a Space Shuttle simulator to conduct a detailed task analysis of flight crew activities during entry and landing phases of flight.

2.   Design Eye and Display Placement

The Design Eye Position is defined in MIL-STD-1333B as the intended viewpoint of the average operator (see Figure 8). It is a reference point from which all crew station dimensions are related and referenced (DoD, 1987). The placement of displays within a cockpit is determined with this point in mind. For military systems, this placement is standardized in MIL-STD-1333B and MIL-STD-1472G. NASA-STD-3001 serves a similar function for manned spacecraft.
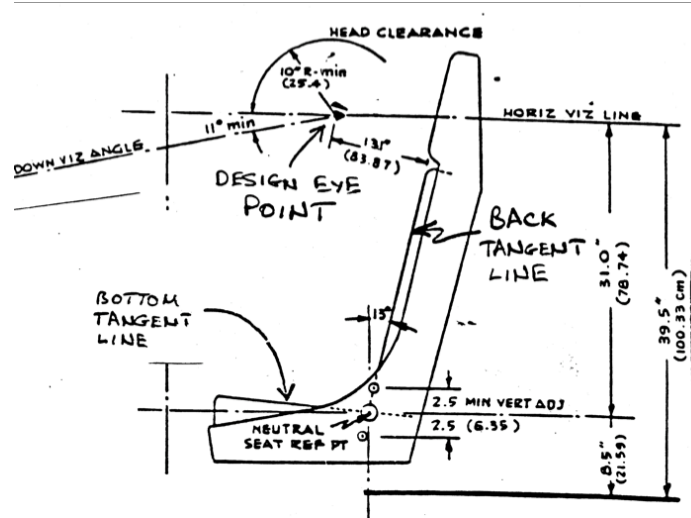
Figure 8.   Design Eye Position (From Woodson, 1981)

The NASA standard provides three relevant guidelines for the placement of displays: (1) The most important and most frequently used displays shall be provided privileged positions in the crew's viewing zone in order to ensure quick processing and reaction; (2) Displays and controls shall be visible and be within the functional reach envelope of the crew under all conditions (e.g., suited, seated, restrained, and unrestrained); and (3) Displays and controls shall be located and designed so that they may be used to the required degree of accuracy by the crew in normal operating positions (NASA, 2011).

Huchingson (1981) asserted that the most important and frequently used instruments should be grouped together within a 30-degree cone of vision centered about the primary line of sight. Huchingson further identified a maximum visual field for display placement as +/- 35 degrees horizontally, and +40/-20 degrees vertically about the primary line of sight. These limits are based on eye movement alone, without head movement (Huchingson, 1981).

15
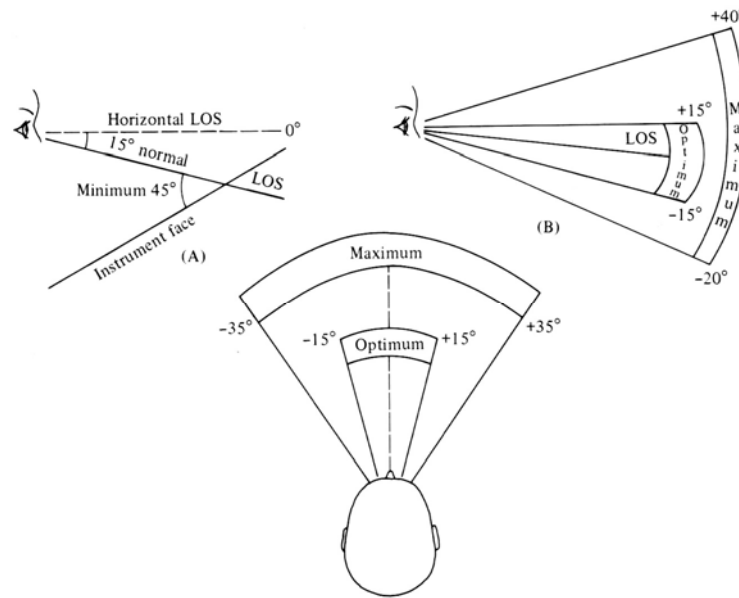
Figure 9.   Optimum and Maximum Visual Fields (From Huchingson, 1981)

3.   Display Design

In order to describe a cockpit design change from a human factors perspective, it is helpful to first identify a universal framework that can be used to evaluate the quality of a cockpit display. One such widely accepted framework is the thirteen principles of display design (Wickens, Lee, Liu, & Gordon-Becker, 2004).   This list is organized into four distinct categories: (1) perceptual principles, (2) mental model principles, (3) principles based on attention, and (4) memory principles.

a.   Perception

Perceptual principles influence the way an operator initially perceives the information being displayed. These principles emphasize presenting information clearly and unambiguously so as to promote timely perception and avoid confusion (Wickens et al., 2004). Perception can be further aided by creating context, or by using familiar representations, such as fonts and icons (Wickens & Carswell, 2006). Considerations for these principles are summarized in Table 1.

16

Principle	Considerations
Make displays legible	Consider contrast, visual angle, illumination, noise, etc.
Avoid absolute judgment limits	Avoid requiring a user to judge more than 5-7 levels of a single variable (color, size, etc).
Top-down processing	People interpret signals based on what they expect to perceive, and therefore off-normal conditions should be emphasized.
Redundancy gain	Correct interpretation is more likely when a signal is expressed in more than one place, especially when an alternate form is used.
Use discriminable elements	Two signals that look alike are likely to be confused.

Table 1.   Perceptual Principles (From Wickens et al., 2004)

b.   Mental Models

Mental model principles relate to the operator's expectations, or "mental model" of the system being displayed (Wickens et al., 2004). Mental models can be defined as a mechanism to generate descriptions of system purpose and form, explanations of system functioning, and predictions of future system states (Rouse & Morris, 1985). When a display is perceived, operators tend to interpret its appearance and movement in terms of this mental model (Wilson & Rutherford, 1989).   Mental Model principles are summarized in Table 2.

Principle	Considerations
Principle of pictorial realism	The display should look like the variable that it represents.
Principle of the moving part	Moving elements on a display should match the user's mental model of the system being represented.

Table 2.   Mental Model Principles (From Wickens et al., 2004)

c.   Attention

Attentional principles are important for displays that have multiple elements (Wickens et al., 2004). Three components of attention are needed to process these types of displays: (1) selective attention (or attention allocation) for choosing the necessary information source, (2) focused attention for avoiding distraction from neighboring sources, and (3) divided attention for allowing parallel processing from two sources as needed (Wickens et al., 2004). Attention allocation can be either a knowledge drive top-down process, or a bottom-up process which is driven by the salience of the cue (Yantis, 1993). In either case it can be inhibited when excessive effort is required to shift attention(Wickens et al., 2004). The attentional principles summarized in Table 3 provide ways to capitalize on human strengths or mitigate human weaknesses with respect to attention.

Principle	Considerations
Minimizing information access cost	The amount of effort needed to move between required pieces of information (menus on a display screen, checklist pages, etc) should be minimized.
Proximity compatibility principle	Two related pieces of information that must be integrated (such as a graph and its legend) should be displayed in close proximity to each other. Close proximity need not be exclusively in terms of space. It can also be achieved by using a common color or pattern.
Principle of multiple resources	Divide information across resources, such as visual and auditory, to facilitate concurrent processing.

Table 3.   Principles Based on Attention (From Wickens et al., 2004)

d.   Memory

Memory principles deal with the limited capacities of both long term and short term memory (Wickens et al., 2004). Short term memory is heavily limited in the number of "chunks" of information that can be retained at any one time (Cowan, 2010). Long term memory limitations include both forgetting important information, as well as persisting in following outdated or incorrect information (Wickens et al., 2004). These principles, presented in Table 4, emphasize overcoming these limitations.

Principle	Considerations
Replace memory with visual information	Do not require that all important information be retained solely in working memory or retrieved from long term memory.
Principle of predictive aiding	Predicting future states is a cognitively demanding task.
Principle of consistency	Displays should be designed in a way that is consistent with what the user is already familiar with.

Table 4.   Memory Principles (From Wickens et al., 2004)

4.   Hardware Attributes

The NASA Space Flight Human-System Standard for Human Factors, Habitability, and Environmental Health (NASA-STD-3001 Volume 2) defines a set of technical requirements by which a spacecraft's display hardware can be evaluated (see Table 5). Despite the fact that the Space Shuttle is not a military system, MIL STD 1472G can be used for evaluation as well. MIL-STD-1472 is widely regarded as an authoritative source of good human factors design practices (Woodson, 1981).

Metric	Minimum	Maximum	Notes
Ambient contrast ratio	10		Includes ambient illumination
Viewing angle	-45 deg	+45 deg	4-point viewing angle (left, right, up, down), contrast and color gamut criteria met
Spatial resolution	32 pixels/deg		
Frame rate	60 Hz		
Number of colors	4096		

Table 5.   Required Visual Display Parameters (From NASA, 2011)

While evaluating against the NASA standard provides a sense of the overall quality of a display, an additional level of analysis is needed to determine whether or not existing hardware is suitable for displaying the proposed display formats.   To this end, NASA conducted a study to evaluate the color and luminance properties of the existing MDU hardware (McCandless, 2003).   Using a colorimeter and a luminance meter, measurements were taken of each required color from crew design eye positions under various lighting conditions. Two relevant key findings came from this study: 1) Viewing angle has negligible effect on all proposed CAU colors except orange, which can vary in appearance from yellow to red depending on angle; and 2) The display can not produce true black, but instead appears dark blue when black is commanded.

5.   Display Control

Fitts and Seeger (1953) identified the importance of the relationship between a control and the display for the entity being controlled in the principle of stimulus-response compatibility.   Good compatibility, which is associated with timelier and/or more accurate response to stimuli, is achieved when the control is

located in close proximity to the display of the entity being controlled (Wickens et al., 2004). Stimulus-response compatibility is also affected by the coding (i.e., verbal vs. spatial) of the control and the associated display (Wickens, Vidulich, & Sandry-Garza, 1984). Design requirements based on these principles are articulated in NASA STD 3001; which states that displays and controls shall be grouped according to purpose or function, and that the relationship between the display and control shall be logical and explicit (NASA, 2011).

MIL STD 1472 provides additional guidelines for the physical properties of display control hardware. The standard specifies that a physical keyboard shall be used for entering any alpha-numeric data, and that positive feedback shall be provided on screen during such data-entry tasks (DoD, 2012). It further requires that individual keys in a vehicle-mounted keypad should be no smaller than 0.75 inches in width, and be separated by 0.5 inches (DoD, 2012). Finally, the standard requires that all keys and pushbuttons (regardless of location) shall provide tactile feedback (DoD, 2012).

D. ASSESSING DESIGN EFFECTIVENESS

MW and SA are both considered to be important factors in the design of aviation systems. (Selcon, Taylor, & Koritsas, 1991). Tsang and Vidulich (2006) noted that these two concepts are highly interrelated, and are affected by many of the same factors. Cockpit design improvement efforts are often focused on these common factors, with an overall objective of reducing mental workload and improving SA (e.g., Carmody-Bubb & Maybury, 1998; Weinstein & Wickens, 1992)

Two goals of the CAU program were to reduce workload and enhance situation awareness. Therefore, prior to assessing the effectiveness of the program, it is necessary to review these two concepts. Both MW and SA are reviewed in terms of underlying theory, measurement technique, and practical

application. Measurements are assessed in terms of validity, sensitivity, selectivity, diagnosticity, intrusiveness, and reliability (see Table 6).

Validity	The extent to which a dependent measure actually assesses what it is intended to measure
Sensitivity	the degree to which a measure distinguishes between differing conditions or states
Selectivity	the degree to which a measure is sensitive only to changes in the construct of interest
Diagnosticity	the degree to which a measure not only identifies changes but identifies the cause of any variation
Intrusiveness	the degree to which a measure interferes with the primary task
Reliability	the degree to which a measure is consistent

Table 6.   Measurement Criteria (After Marchant et al., 2001;
Uhlarik & Comerford, 2002)

1.   Mental Workload

a.   Theory and Issues

The most basic way to express the MW concept is in terms of spare capacity. The simplest and most intuitive way to express spare capacity is in terms of the ratio of time required to complete a set of tasks and the time available (TR/TA) (Wickens et al., 2004). While this single dimension can not fully capture MW's complexity and multidimensionality, it has nonetheless been shown to be an effective approximation in some situations (Hendy, Liao, & Milgram, 1997). In addition to roughly predicting MW, the TR/TA ratio also should predict the point at which excess workload begins to degrade performance. Whenever time required exceeds time available (TR/TA > 1.0) a

22

person can be said to be in overload. Conversely, when time required is less than time available (TR/TA < 1.0) a person can be said to have spare capacity (Wickens et al., 2004). In situations where spare capacity exists, the amount of spare capacity can be used as a MW measure (Lysaght et al., 1989; Roscoe & Ellis, 1990).

While time makes for a good approximation in many situations, it is very limited as an expression of overall workload. For example, tasks that are very practiced can often be carried out with little conscious thought, regardless of how much time is required to actually complete them (Logan, 1985). These kinds of issues can be accounted for by viewing workload not in terms of limitations in time, but in terms of a limited capacity central processor (Moray, 1967). This view is foundational to so-called 'capacity' or 'resource' theories, which postulate that limited attentional resources are available for conducting tasks, and that more difficult tasks leave fewer resources available for conducting concurrent tasks. The available resource supply fluctuates based on the individual's level of arousal, and performance degradations occur when resource demands exceed resource supply (Kahneman, 1973).

As a result of studies conducted in the 1970s, evidence emerged that task performance could not be attributed to a single resource pool, and that the presence of multiple, separate limited resource pools appeared to better account for performance (Wickens, 2002a). The presence of multiple pools related to specific abilities is a foundational principle of multiple-resource theory. Wickens's multiple-resource model proposes four dimensions to account for variance in multitasking performance: (1) processing stages, (2) perceptual modalities, (3) visual channels, and (4) processing codes. In each dimension, two separate pools of resources are postulated to exist. The dimension of stages refers to information processing stages, including perception, cognition, and response. Wickens (2002) identified that the resources used for perception and cognition appear to be separate and distinct from those resources used for responding. The dimension of modalities refers to the distinction between auditory and visual

23

modalities. Wickens identified that the resources used for each of these two modalities appears to be separate and distinct. The dimension of channels refers to a distinction between focal and ambient visual processing, that appear to use separate resources. Finally, the dimension of codes refers to the distinction between spatial and verbal processing. Wickens found that in any cognitive stage, spatial and verbal processes appear to depend on separate resources. Across each of these dimensions, much work has been done suggesting that when controls and displays do not all require the same resource pool, there is less competition for limited resources, and less overall workload as a result (Lysaght et al., 1989).

b. *Measurement*

The techniques commonly used to empirically assess MW include task performance measures, physiological measures, and subjective measures (Lysaght et al., 1989). Muckler & Seven (1992) noted that the "distinction between 'objective' and 'subjective' measurement is neither meaningful nor useful in human performance studies." Casali and Wierwille (1984) found that with highly trained participants (such as aircraft pilots) subjective rating scales are highly sensitive instruments for measuring MW. Therefore, of the three methods listed above, subjective measures are of greatest interest to this study.

Roscoe and Ellis (1990) showed that pilots easily adapt to subjectively expressing MW through the concept of spare capacity, and developed the Bedford Workload Rating Scale (BWRS) based on this concept. The BWRS, which was adapted from the Cooper Harper rating scale for aircraft handling, uses a simple decision tree (see Figure 10) to arrive at a subjective assessment of workload on a scale from 1 (Workload insufficient) to 10 (Pilot unable to apply sufficient effort). While easy to use, the authors recognized that there are significant drawbacks to this method. One of the most glaring is the fact that it requires active subject participation, and therefore cannot be used to measure workload during a task without a significant disruption. Another important limitation is that normalization is required as BWRS results tend to be highly

individualized to each pilot. The BWRS has been extensively used in cockpit workload evaluation studies, and is well understood by flight crews (Lysaght et al., 1989). It has been found to correlate well with other MW measures, such as heart rate, and is most effective when used during or shortly after the flight task in question (Roscoe & Ellis, 1990). The BWRS is also somewhat limited in its complexity and sensitivity.
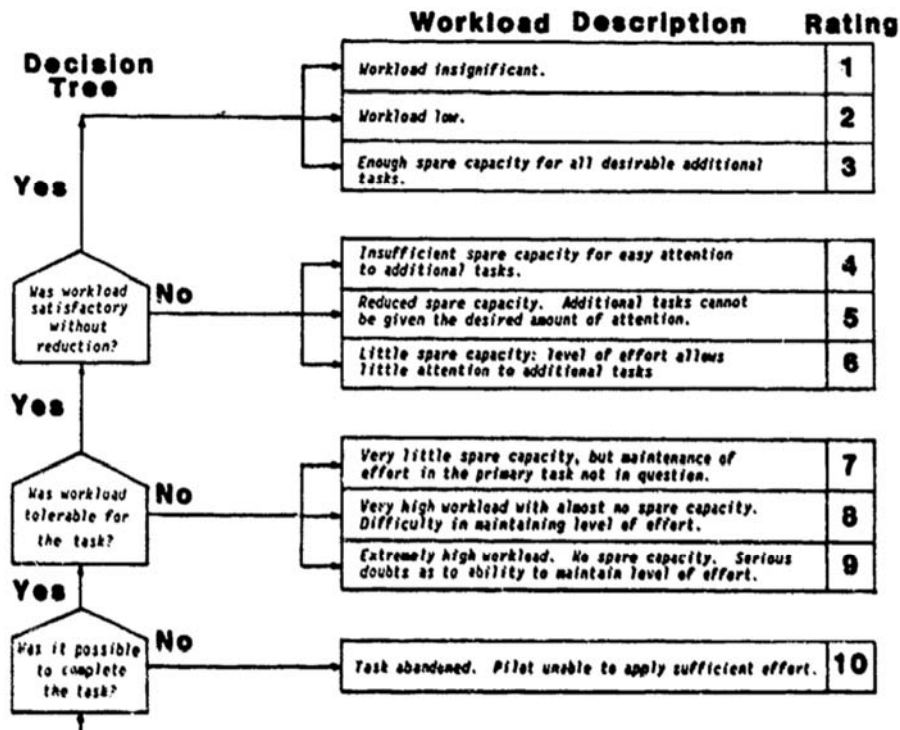
Figure 10.   Bedford Workload Rating Scale (From Roscoe & Ellis, 1990)

Because workload is a multidimensional concept, various researchers have argued that subjective ratings should consist of more than just a single scale (Wickens et al., 2004). In developing the NASA Task Load Index (TLX), Hart & Staveland (1988) proposed using six dimensions to assess workload: mental demand, physical demand, temporal demand, performance, effort, and frustration. Subjective ratings are then obtained separately for each of these dimensions. Each dimension is weighted according to how much it contributed to overall workload, and an aggregate score is computed by summing the weighted

ratings from each scale. While TLX provides more dimensionality than BWRS, it suffers from some of the same limitations. It cannot be used to measure workload in the middle of a task without disrupting that task, and it requires normalization for each participant. Hart (2006), in reviewing the use of TLX over a 20 year period found that it had been referenced in over 500 studies, primarily in aviation settings. A literature review conducted by Cain (2007) indicated that TLX performs comparably to other multidimensional workload instruments, and is much more sensitive than one-dimensional scales such as the BWRS. A similar review conducted by Lysaght (Lysaght et al., 1989) found TLX to be a valid, reliable, and sensitive technique for workload assessment.

1. Mental demand: How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

Low   High

2. Physical demand: How much physical activity was required (e.g., pushing, pulling, turning, controlling activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

Low   High

3. Temporal demand: How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

Low   High

4. Performance: How successful do you think you were in accomplishing the goals of the task set by the researchers (or yourself)? How satisfied were you with your performance in accomplishing these goals?

Good   Poor

5. Effort: How hard did you have to work (mentally and physically) to accomplish your level of performance?

Low   High

6. Frustration: How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

Low   High

Figure 11.   NASA TLX Components (From Beutter et al., 2006)

2. Situation Awareness

a. Theory and Issues

Endsley (1999) defined 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." This is the most frequently cited definition to be found in the scientific literature (Uhlarik & Comerford, 2002). Encapsulated within Endsley's definition are three distinct levels of SA. Level 1 SA involves perceiving critical elements in the environment. Level 2 SA involves understanding these critical elements and relating them to the aircrew's goals. Level 3 SA involves projection these elements into the near future and understanding what will happen. Wickens (2002b) further identified a clear distinction between spatial awareness and system awareness, with distinct requirements for each type. Spatial awareness requires knowledge of elements such as attitude, heading, velocity, vertical velocity, aircraft capabilities, and projected flight path; whereas System awareness requires awareness of such elements as system status, settings, and the impact of system configurations and malfunctions on overall system performance and flight safety (Wickens, 2002b)

As a process, SA is closely linked with perception and working memory, and is therefore subject to limits in attentional capacity (Tsang & Vidulich, 2006). Consequently, high workload tasks that consume attentional resources can lead to poor SA. Therefore, the way in which a cockpit presents information to the operator has a huge impact on crew SA (Tsang & Vidulich, 2006). Presenting too much information at one time can lead to excessive workload, which can be as much of a hindrance to good SA as presenting too little information (Uhlarik & Comerford, 2002). In general, Endsley (1999) notes, "The more complex the systems are to operate, the greater the increase the mental workload required to achieve a given level of SA. When that demand exceeds human capabilities, SA will suffer."

b. Measurement

In their review of the scientific literature pertaining to MW measurement, Uhlarik & Comerford (2002) identified three categories of SA measurement: Explicit measures, implicit measures, and subjective measures. Explicit measures are those that simply ask the subject to report information about the recent state of an aircraft; implicit measures are those that infer a level of SA from other measures of task performance; and subjective measures are based solely on opinion.

Explicit measures can provide specific detailed information about the operator's concept of a situation that can then be objectively evaluated against reality (Endsley, 1995). This method typically takes the form of questionnaires that can be administered following the completion of a task, during the course of its execution, or while the task is paused (DoD, 2012; Uhlarik & Comerford, 2002; Wickens et al., 1984; etc.). Each approach has its benefits and drawbacks. Post-task questionnaires allow the subject ample time to respond to lengthy questionnaires, but are limited by the level of detail that can be remembered throughout an entire task (Endsley, 1995). Questions asked during the execution of a task overcome the memory limitation, but may impose additional workload that can alter performance. The intrusiveness of the measure can be reduced when the task can be frozen, as by pausing a simulator and blanking the displays (Tsang & Vidulich, 2006).

One such tool that employs a "freeze technique" is the Situation Awareness Global Assessment Tool (SAGAT) (Endsley, 1995). This tool is based on developing a global list of SA requirements for a given system and task, and then asking the subject questions based on this list at random intervals throughout the simulation. SAGAT is generally considered to offer high sensitivity to changes in situation awareness, but it can be intrusive when the simulator is resumed (Tsang & Vidulich, 2006).

Implicit measures are objective, non-intrusive, and relatively easy to collect in a simulator, but they are limited in their selectivity and sensitivity (Endsley, 1995). These measures can be collected globally, or for just a specific sub-task of interest. Global measures are not particularly reliable for measuring SA due to the multitude of cognitive factors (many unrelated to SA) that can affect overall performance (Uhlarik & Comerford, 2002). Imbedded task measures offer improved selectivity over global measures, however the results may be misleading because changes in a single component of SA can greatly impact other components of SA (Endsley, 1995). For this reason, imbedded task measures are rarely used (Uhlarik & Comerford, 2002).

Subjective measures of SA are inexpensive and easy to use, but have many significant limitations. Self-rated subjective measures of SA are inherently limited by the fact that an operator (by definition) does not know whether his own SA is incomplete or inaccurate (Uhlarik & Comerford, 2002). Furthermore, subjective self-ratings tend to be confounded by the effects of performance and workload (Endsley, 1995). Conversely, observer-rated measures of SA are inherently limited in what the observer can know about an operator's concept of a situation. An observer can reliably detect overt errors related to SA, but little more (Uhlarik & Comerford, 2002).

3.   Using Multiple Measures

There is no consensus in the literature on a single preferred instrument for measuring mental workload or situation awareness. Selcon, Taylor, and Koritsas (1991) acknowledged that measures of workload and SA have a lot in common, but each captures certain aspects of performance that the other one does not. Rubio, Diaz, Martin and Puente (2004) evaluated several such measures of workload for intrusiveness, sensitivity, validity, diagnosticity, and acceptance, and found that no single measure outperformed the others in every category. Tsang and Vidulich (2006) noted several situations that could lead to dissociation of measures of SA and workload, a condition they say can be informative, "if one is

cognizant of the idiosyncratic properties of the different measures." As a result of this diversity of opinion, multiple measures of workload and SA are often used in aviation human performance studies. The choice of which measures to use is situation dependent, and requires the experimenter to have good working knowledge of the strengths and limitations of each of the available tools (Lysaght et al., 1989; Muckler & Seven, 1992).

E.   SUMMARY

The design of the Space Shuttle's data processing system provides the flight crew with a large volume of information gathered from a variety of onboard and external sources. The chief objective of this literature review has been to survey the human factors principles that relate to the way in which this information is organized and presented in a cockpit design. Together, these principles support a method of critically evaluating a cockpit design change and assessing its effectiveness.

The critical evaluation must begin with a thorough task analysis in order to gain an understanding of what is required of the design. With this in mind, the evaluation must consider the layout of the displays from the operator perspective, the formatting of information on each display, the hardware attributes, and the display-control relationship. Guidelines have been established for each of these characteristics.

In order to assess the effectiveness of design changes, measures of flight crew mental workload and situation awareness are needed. Several such measures can be found in the literature, each with its own advantages and disadvantages. Multiple measures are therefore needed to ensure a thorough assessment. Careful analysis of MW and SA data can show what improvements exist and where. Tracing these results back to the critical design evaluation can offer insight into why. The use of multiple measures of MW and SA presents an opportunity to examine which measures work best in this setting.

III. METHODS

A. OVERVIEW

The evaluation begins with a critique of the CAU program design modifications, and then proceeds to an assessment of the effectiveness of these changes. The design critique is centered around a task analysis of crew operations in both ascent and entry phases of flight. It examines the modifications in terms of cockpit layout, design principles, hardware attributes, and display control. Paired astronaut crews flying identical simulated scenarios in the baseline and upgraded CAU configurations were then assessed using MW and SA measures. These results were then analyzed for statistical significance.

B. DESIGN CRITIQUE

The design changes between the baseline and CAU cockpit display formats were evaluated according to established human factors principles. Central to this evaluation was a task analysis of flight crew activities during ascent and entry phases of flight. This task analysis was used to define a set of information display requirements. From this, the cockpit display upgrades were evaluated in terms of cockpit layout, display design, hardware attributes, and display control. The evaluations were conducted using a wide variety of documents that describe Space Shuttle cockpit hardware and software characteristics (e.g., Jenkins, 2010; Thomsen & Hancock, 1994; Tranthien, 1995).

1. Task Analysis

Two complementary task analysis techniques were used to evaluate the design changes. First, a basic hierarchical task analysis was used to identify the display-related crewmember tasks during ascent and entry operations. Once display-related tasks were identified, the HTA was not developed further. Based on this HTA, a task decomposition was then used to identify information

requirements and display support in each cockpit for each identified task. The results of this stage of the evaluation were used to guide subsequent stages.

2. Design Eye and Display Placement

The astronauts' field of view must be taken into consideration for this critique. Although no changes were made in the physical arrangement of display hardware, changes made to the display formats can affect the availability of key pieces of information. This portion of the critique therefore began by finding the design eye position for each crew station. Next, optimum and maximum visual fields were used to determine what displays are most readily visible to each crew station. The optimum and maximum fields were based on the limits established by Huchingson (1981), defined in the previous chapter. Viewing angle limits of the display hardware were considered as well. Finally, by using typical layouts of display formats for both the baseline and CAU cockpits, lists were developed for each crew station detailing what information is most readily visible.

3. Display Formats

Every display format in the shuttle cockpit was examined and modified as part of the CAU program. In addition, the organizational structure was modified such that there is not a one-to-one mapping of old display formats to new ones. Display formats were variously added, deleted, consolidated, and rearranged in order to meet the design objectives.

For this critique, the display formats were grouped according to task. Task categories include: (1) Monitoring flight parameters (altitude, speed, heading, etc.); (2) Systems Management; and (3) Monitoring trajectory. Display formats within each of these categories share certain characteristics that allowed them to be evaluated collectively. Within each category, differences were highlighted between the CAU and baseline displays. These differences are then critiqued according to the thirteen principles of display design.

4. Display Hardware

Although no changes have been made to the shuttle's MDU hardware, there are nonetheless some hardware attributes worth examining in this evaluation. The display hardware was examined first for its general suitability within the cockpit, and then for its suitability in displaying the CAU display formats. The first part of the critique was conducted by checklist method, wherein display hardware specifications are evaluated against NASA and DoD standards. The second part of the critique was done by comparing MDU color capabilities with the requirements of the CAU formats.

5. Display Control

Although no major changes were made to the layout of the keypads and bezel keys used to control the displays, changes to the display formats can significantly affect the display-control relationship. This portion of the critique examined that relationship. The hardware characteristics of the controls were first evaluated according to NASA and DoD standards to examine their suitability for data entry and display configuration tasks. Next, the relationships between the control hardware and the information displays were evaluated for both the baseline and CAU designs. The evaluation focused on principles of stimulus-response compatibility, and examined both proximity and coding.

C. DESIGN ASSESSMENT

The testing process was designed around a set of minimum success criteria, which provided the standard by which the updated displays would be evaluated (McCandless, 2004). Achieving the minimum success criteria would indicate that the program was successful, whereas failure to meet the criteria would have resulted in either a redesign of the displays, or cancellation of the entire project. The goals for the minimum success criteria were baselined by the Office of Spaceflight at NASA Headquarters and confirmed during four

independent reviews (McCandless, 2004). The minimum success criteria were divided into three categories, presented in Table 7.

Category	Definition	Goals
Situation Awareness	Situation awareness is defined as a crew member's understanding of his or her dynamic flight environment with respect to the mission.	a) Trajectory, horizontal situation, alternate landing sites, abort determination and monitoring: 100% improvement b) DPS, EPS, MPS, OMS monitoring and failure recognition: 50% improvement c) Other system monitoring and failure recognition: 20% improvement
Mental Workload	Workload is defined as the mental effort necessary to perform a task.	Reduce workload, as measured by Bedford and NASA TLX
Performance	Performance is a measure of how well the crew accomplished the appropriate tasks in the cockpit.	a) Reduction in unidentified malfunctions and recognition time b) Reduction in errors c) Reduction in keystrokes

Table 7.   Minimum Success Criteria (After McCandless, 2004)

1.   Participants

The evaluation was conducted with six crews of three astronauts each. Each crew included a flown pilot or commander in the commander seat, an unflown pilot in the pilot seat, and an unflown mission specialist in the mission specialist 2 (MS2) seat. These three positions are the primary "stake-holders" in the CAU upgrade, which makes them the appropriate subjects for the evaluation. The goal with this crew make-up was to test the displays with relatively new astronauts who would have the least amount of ingrained experience and repetition with the Baseline displays, yet who also had some experience with shuttle operations.

Training in the baseline configuration was not required because of the level of familiarity the participants had already acquired through several years of experience and training. Training in the CAU configuration was conducted over the course of 3-4 weeks. It consisted of 20 hours of classroom training, 5 hours of Dynamic Skills Trainer (DST) training, and 10 hours of familiarization sessions in the Shuttle Mission Simulator.

2.   Instruments

a.   Subjective Workload

The workload measures taken for the evaluation, the BWRS (see Table 9) and the NASA TLX (see Table 10), are subjective assessments. Both methods were discussed in the previous chapter. Workload measures were collected from each crewmember at the completion of each run.

Rating	Description
1	Workload insignificant.
2	Workload low.
3	Enough spare capacity for all desirable additional tasks.
4	Insufficient spare capacity for easy attention to additional tasks.
5	Reduced spare capacity. Additional tasks cannot be given the desired amount of attention
6	Little spare capacity. Level of effort allows little attention to additional tasks.
7	Very little spare capacity, but maintenance of effort in the primary tasks is not in question
8	Very high workload with almost no spare capacity. Difficulty in maintaining level of effort.
9	Extremely high workload. No spare capacity. Serious doubts as to ability to maintain level of effort.
10	Task abandoned. Unable to apply sufficient effort.

Table 8.   Bedford Workload Rating Scale (After McCandless, 2004)

Component

1) **Mental demand** (from low to high on a scale of 1 to 10): How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?

2) **Physical demand** (from low to high on a scale of 1 to 10): How much physical activity was required (e.g., pushing, pulling, turning, controlling activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?

3) **Temporal demand** (from low to high on a scale of 1 to 10): How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?

4) **Performance** (from good to poor on a scale of 1 to 10): How successful do you think you were in accomplishing the goals of the task set by the researchers (or yourself)? How satisfied were you with your performance in accomplishing these goals?

5) **Effort** (from low to high on a scale of 1 to 10): How hard did you have to work (mentally and physically) to accomplish your level of performance?

6) **Frustration** (from low to high on a scale of 1 to 10): How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?

Table 9.   NASA TLX Rating Components (After McCandless, 2004)

b.   Situation Awareness

The SA data can be divided into objective and subjective categories. Objective questions were those that had a definitive correct or incorrect answer. These questions were based on the SAGAT discussed in the previous chapter. Objective SA questions related to trajectory management, critical system status, and non-critical system status. An example of an objective question is:

What was your trajectory energy state at the beginning of the run (circle one)?

High Low Nominal Don't Know

Subjective ratings were based on crewmembers' opinions, and had no correct or incorrect answer. Subjective SA questions related to trajectory management, critical system status, and non-critical system status. An example of a subjective question is:

Rate your situational awareness of the fuel cell problem as provided by the displays.

1 2 3 4 5 6 7 8 9 10

The subjective rating scale used for rating situation awareness is presented in Table 8:

Rating Description
Insufficient SA
1   Completely unaware of situation
2   Mostly unaware of situation and totally unable to monitor/follow-up
3   Somewhat aware of situation and mostly unable to monitor/follow-up
Reduced SA
4   Somewhat aware of situation and somewhat able to monitor/follow-up
5   Mostly aware of situation and somewhat able to monitor/follow-up
Adequate SA
6   Mostly aware of situation and mostly able to monitor/follow-up
7   Completely aware of situation and somewhat able to monitor/follow-up
Excellent SA
8   Completely aware of situation and mostly able to monitor/follow-up
9   Completely aware of situation and completely able to monitor/follow-up
Perfect SA
10   Completely aware of situation, completely able to monitor/follow-up, and aware of next worst failure

Table 10.   Description of Situation Awareness Scale (After McCandless, 2004)

3. Observers

One to two observers were in the cockpit for every run, all were either astronauts or members of the NASA evaluation team. Additional observers were in the instructor station. Cockpit observers, as trained evaluators, provided a unique perspective on non-verbal communication. These observers provided subjective data by estimating crew situation awareness and workload. Observers in the instructor station were used to monitor crew actions and record objective data on errors, unrecognized malfunctions, and recognition times.

4. Apparatus

a. Simulator

The evaluation was conducted in the Shuttle Mission Simulator (SMS) at NASA JSC in Houston, Texas. This full motion simulator replicates all aspects of the shuttle cockpit and provides high-fidelity simulation of shuttle ascent and entry operations. For this evaluation, the simulator was configured with two classes of displays: Baseline, and CAU. Baseline displays are the Orbiter Increment-29 (OI-29) displays in the MEDS cockpit. CAU displays are the OI-41 redesigned displays.

The simulator was used to record the number of key presses on the DPS keypads. These data were divided into key presses related to display navigation, and those use for other functions. The number of display edge keys pressed and switches thrown was also recorded.

b. Flight Data File

The evaluation only included ascent and entry flight phases. The five FDF procedure books associated with ascent and entry were updated to account for the technical changes imposed by the CAU system. Additional FDF

modifications were avoided to eliminate the potential of misinterpreting FDF improvements as cockpit improvements. The Baseline testing evaluations used the FDF from shuttle flight STS-112.

5. Procedure

The evaluation process was centered around the execution of scripted formal scenarios simulated with both the baseline and CAU cockpit displays. These scenarios included normal ascents, aborts (including RTLS, TAL, and ATO), normal entry (under both PASS and BFS guidance), and off normal entry.

The purpose of the Baseline testing was to evaluate the existing Baseline shuttle cockpit to provide a basis of comparison for the CAU cockpit. After an interlude of 8 to 11 months, the CAU displays were evaluated under the same conditions as the Baseline testing. The interlude was intended to reduce the chances that the crews would remember the details of the scripted scenarios during the testing on the CAU displays. Baseline testing was run over October - December, 2003, whereas CAU testing was run over August - September, 2004.

For each cockpit, the crews participated in three data collection sessions, each lasting about 2 hours. Each session included eight scripted runs, which were identical in both baseline and CAU evaluation. Session 1 was comprised of entry runs and sessions 2 and 3 were comprised of ascent runs. The runs consisted of short (approximately 10 minute) time slices of a given flight phase. The start and end point for each script varied based on the required test objectives. Each run contained several simulated malfunctions, which were designed to test the effect of the modifications to the displays. Throughout each scenario, observers recorded various objective measures of crew performance. At the end of each script, the simulator was frozen and the questionnaires were administered to the crew.

D. DATA COLLECTION AND ANALYSIS

Table 11 provides a summary of the data collected in each run. Data for each measure were collected in both the ascent and entry phases of operations, and in both the baseline and CAU design configurations. Individual measures were collected for each of the six commanders, pilots, and mission specialists.

Source	Category	Subcategory	Measures
Individual (by crew position)	MW	Subjective	BWRS NASA TLX
	SA	Subjective	Trajectory Awareness Critical System Awareness Non-critical System Awareness
		Objective	Trajectory Status Questions Critical System Status Non-critical System Status
Observer	MW	Objective	BWRS
	SA	Subjective	Trajectory Awareness Critical System Awareness Non-critical System Awareness
	Performance	Objective	Number of Errors Malfunction Recognize Time
Simulator	Performance	Objective	Navigation Key Presses Non-navigation Key Presses Switches Thrown MDU Edge Key Presses

Table 11.   List of MW, SA, & Performance Measures Collected

The data were examined with a series of statistical tests intended to address the research questions of this thesis. The first research question asked if the CAU design changes led to improvements in MW and SA. It was then asked whether or not these changes were greater for one specific crew position or phase of operations versus the others. These two questions required a

statistical evaluation of how MW and SA were affected by differences in design, crew position, and phase of operations. To this end, a series of Analysis of Variance (ANOVA) were conducted.

Because of data set limitations, not every measure was included in the ANOVA. Some measures were under-sampled for certain conditions, and some errors were made in recording data. Ultimately, only one measure of MW, and one type of SA measure were used. The decision of which measures to use was based on available sample sizes, and variance within each measure. For MW, BWRS results were used. For each of the components of SA (Trajectory, Critical Systems, and Non-Critical Systems), subjective measures were used.

An initial round of ANOVA tests were conducted that included a nested, within-subjects design to examine the significance of individual differences. The results of these tests indicated that the effect of subjects was insignificant. Having demonstrated that individual differences are minimal, the design was reduced to a purely between-factors ANOVA. Although this approach bends the assumption of strict independence of observations, this deviation is appropriate given the homogeneity of the sample population.

Ultimately, the ANOVA tests for MW and SA included three fixed factors: design (Baseline vs CAU), phase of operations (Ascent vs Entry), and position (CDR vs PLT vs MS2). For each of these tests, results were considered statistically significant if analysis rejected the null hypothesis of no difference between levels of a factor or interaction between factors at an alpha level of .05. Where significant effects were found, post-hoc analysis was conducted using Tukey's Honestly Significant Difference (HSD) test.

THIS PAGE INTENTIONALLY LEFT BLANK

IV.   RESULTS

A.   DESIGN CRITIQUE

The organization of information in the baseline MEDS cockpit is constrained by the inherent limitations of the legacy cockpit from which it is functionally derived. Information was divided according to its source, which could include analog instruments, PASS, and BFS software. As a result, information about one system or even one flight parameter is sometimes scattered across two or more display formats, not all of which can be viewed at any one time. Extensive display navigation is sometimes needed to gather required system and subsystem information. Furthermore, some tasks (such as identifying available landing sites based on energy state) require that the crew consult printed Flight Data Files (FDF), which contain the tables needed to manually perform the required calculation.

A design objective of the CAU program was to resolve these problems by consolidating information from multiple sources onto a single display to create display formats that are more task-oriented. These formats include single-source system summary pages (such as MPS Sum, APU/HYD Sum, etc), as well as single-source pages for monitoring trajectory performance (such as Ascent TRAJ and Entry TRAJ).

1.   Task Analysis

The results of the hierarchical task analysis for ascent and entry display-related tasks are presented in Figures 12 and 13, respectively. In general, tasks can be divided into GNC and SM functions that are performed concurrently throughout Ascent and Entry phases. Each phase also includes an additional set of tasks associated with responding to off-nominal conditions. Each bottom-level task presented represents a requirement for an information display.   These requirements are further decomposed in the sections that follow.
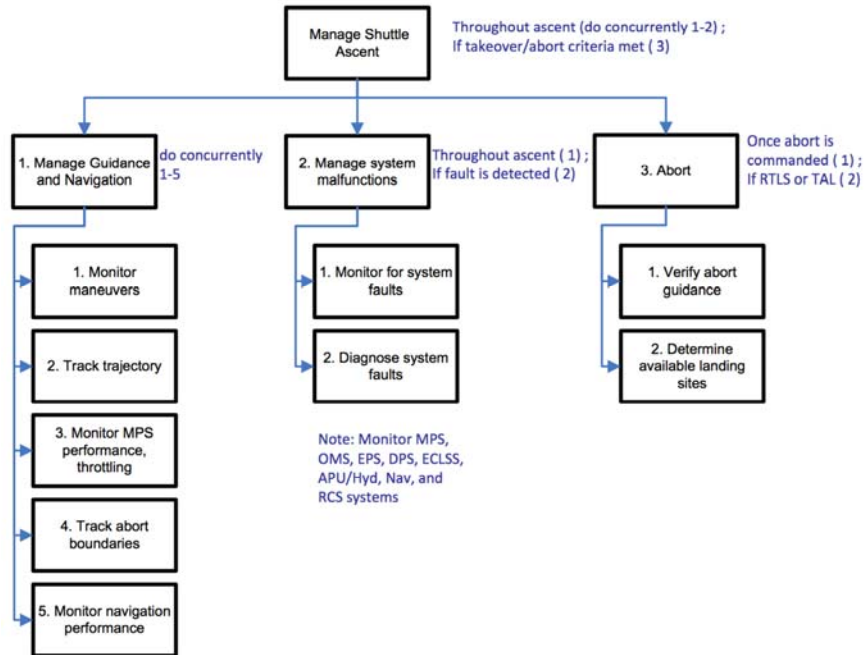
Figure 12.   Display Related Tasks During Ascent and Abort
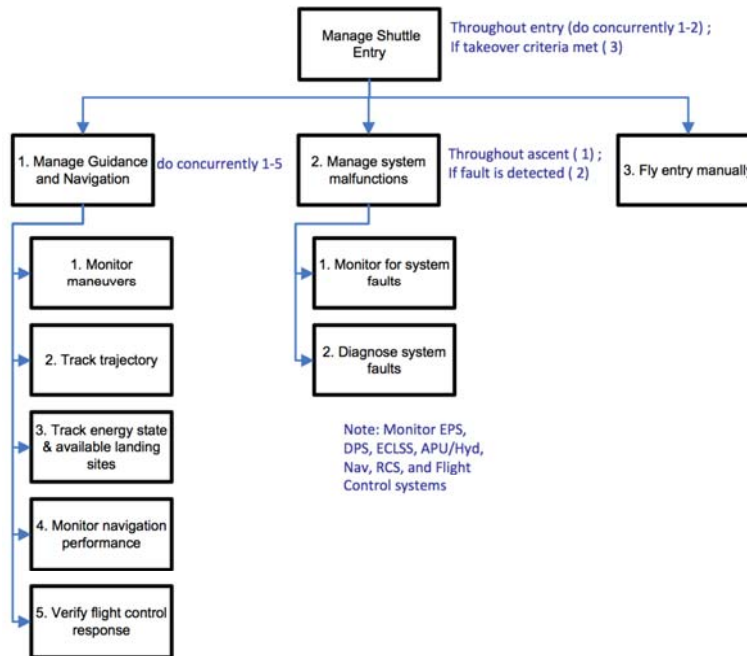
Figure 13.   Display Related Tasks During Entry

The manner in which these information requirements are met differs greatly between the baseline and CAU cockpit configurations. These differences were examined through a task decomposition process in which specific information requirements for each task were identified, along with the associated displays within each cockpit design. These results are presented in Tables 12 and 13 for Ascent and Entry operations, respectively. Of note, a minimum of 10 necessary display formats were identified in the baseline design, compared to four in the CAU.

Operation	Information Requirements	Baseline Formats	CAU Formats
Monitor Maneuvers	Vehicle Attitude	ADI/AVVI and HSI/AMI; or Comp FI	PFD
Track Trajectory (Ascent)	DPS Trajectory data	BFS Ascent TRAJ and/or PASS Ascent TRAJ	Ascent TRAJ
Monitor MPS Performance, Throttling	Power setting, fuel remaining	OMS/MPS Sum	Ascent TRAJ
Track Abort Boundaries	DPS Trajectory data	BFS Ascent TRAJ	Ascent TRAJ
Monitor Navigation Performance	Δ between PASS and BFS guidance	BFS Ascent TRAJ and PASS Ascent TRAJ	Ascent TRAJ
Monitor for System Faults	Failure annunciation	Fault Sum	Fault Sum
Diagnose System Faults	Parameters for MPS, OMS, DPS, EPS, ECLSS, APU/Hyd, Nav, RCS systems	BFS Sys Sum, GNC Sys Sum, MPS/OMS sum, APU/Hyd Sum	CAU System-Specific Displays
Abort Guidance	DPS Trajectory data, Primary Flight Instruments	ADI/AVVI, HSI/AMI, BFS TRAJ (DPS), and PASS Ascent TRAJ	PFD, Ascent TRAJ
Determine Available Landing sites	Runways reachable based on current energy state	Relayed by radio or computed by energy state and cue cards	H Sit

Table 12.   Summary of Display Related Tasks During Ascent

Operation	Information Requirements	Baseline Formats	CAU Formats
Monitor Maneuvers	Vehicle Attitude	ADI/AVVI & HSI/AMI; or Comp FI	PFD
Track Trajectory (Entry)	DPS Trajectory data	BFS &/or PASS Entry Trajectory	Entry TRAJ
Track Energy state & available landing sites	Runways reachable based on current energy state	Relayed by radio or computed by energy state and cue cards	H Sit
Monitor Navigation Performance	Δ between PASS and BFS guidance	BFS Entry TRAJ & PASS Entry TRAJ	Entry TRAJ
Monitor for system faults	Failure annunciation	Fault Sum	Fault Sum
Diagnose system faults	Parameters for MPS, OMS, DPS, EPS, ECLSS, APU/Hyd, Nav, RCS systems	BFS Sys Sum, GNC Sys Sum, MPS/OMS sum, APU/Hyd Sum	CAU System-specific Displays
Fly Entry manually	Basic flight instruments, DPS computed trajectory	HSI/AMI & ADI/AVVI, BFS or PASS Entry TRAJ	PFD, Entry TRAJ

Table 13.   Summary of Display Related Tasks During Entry

Task responsibilities are different for each crew position. For redundancy, individual crew positions are assigned overlapping Primary and Secondary responsibilities. These requirements can be varied by individual crews, but a typical distribution is presented in Table 14.

Position	Primary	Secondary
CDR	Track Trajectory Track Abort Boundaries Monitor DPS, ECLSS systems	Monitor OMS, RCS, Flight Control Systems
PLT	Monitor MPS, EPS, APU/HYD, OMS, RCS, and Flight Control Systems	Track Trajectory
MS2	Overhead Panels Overhead Gauges	Track Abort Boundaries Consult Reference Data (cue cards)

Table 14.   Typical Flight Crew Duties During Ascent and Entry
(From NASA, 2002)

2.   Design Eye and Display Placement

No changes were made to the shuttle seating positions or display locations as a result of the CAU program. Therefore, DEPs and fields of view for both the baseline and the upgraded cockpits are identical. The DEPs of the Commander and Mission Specialist crew stations are presented from front, side, and overhead views in Figures 14, 15, and 16, respectively. The position of the pilot station mirrors that of the commander station.
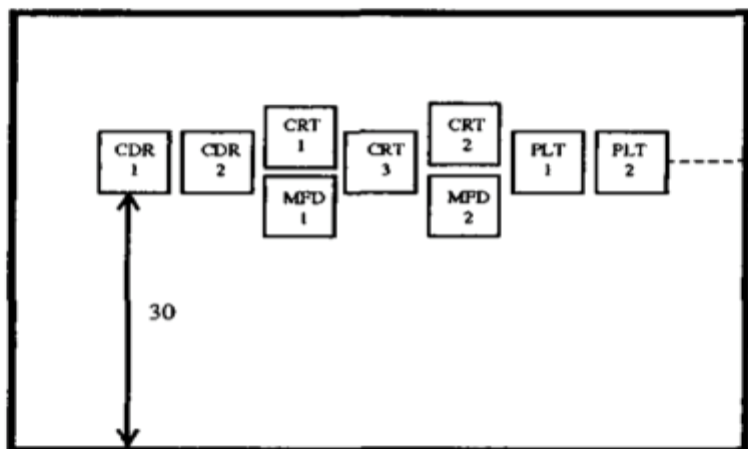
Note: Linear units are inches and angular units are degrees.

Figure 14.   Side View of Shuttle Cockpit Measurements
(From McCandless, 2003)

47
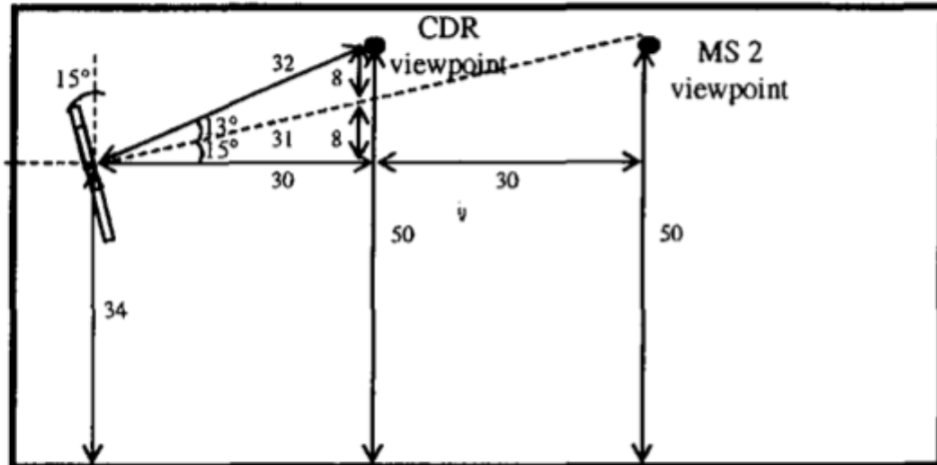
Note: Linear units are inches and angular units are degrees.

Figure 15.   Side View of Shuttle Cockpit Measurements
(From McCandless, 2003)
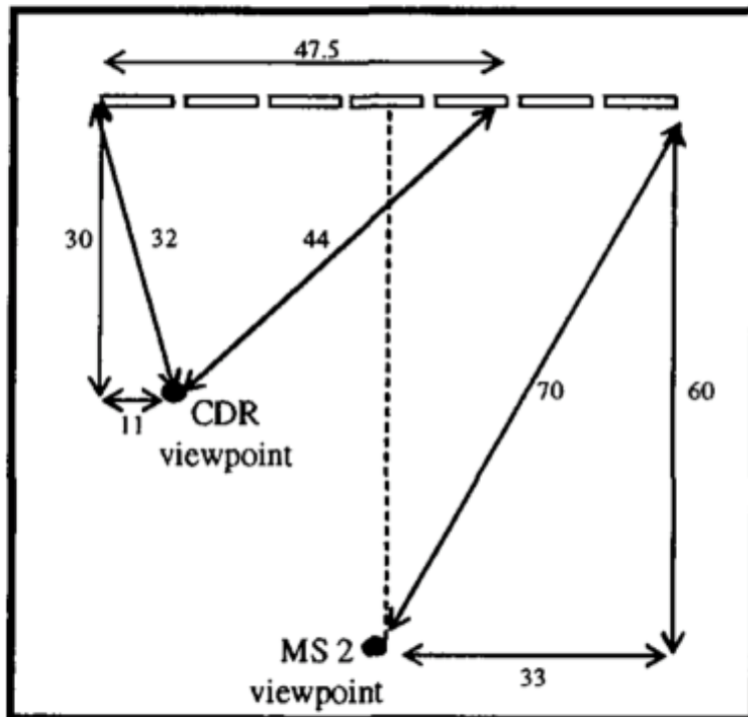
Note: Linear units are inches and angular units are degrees.

Figure 16.   Top View of Shuttle Cockpit Measurements
(From McCandless, 2003)

The optimum visual field for the commander's station was determined to be a 16-inch diameter circle centered on the top edge of the CDR2 display unit. This circle encompasses the entire CDR2 display, and the rightmost portions of the CDR1 display. The maximum field of view was determined to be a rectangle 42 inches wide, and 35 inches in height. This rectangle includes the remainder of the CDR1 display, as well as the complete CRT1 and MFD 1 units. Allowing for head movement, the commander is limited only by the viewing angle limits of the display hardware. Given these limits, the commander can see all MDUs with the exception of the two PLT units. The pilot's side mirrors the commander's side. The visibility of each MDU from the Commander's position is summarized in Figure 17.

Within optimum field of view
Within maximum field of view
Within hardware viewing angle limits
(Head movement required)
Not viewable
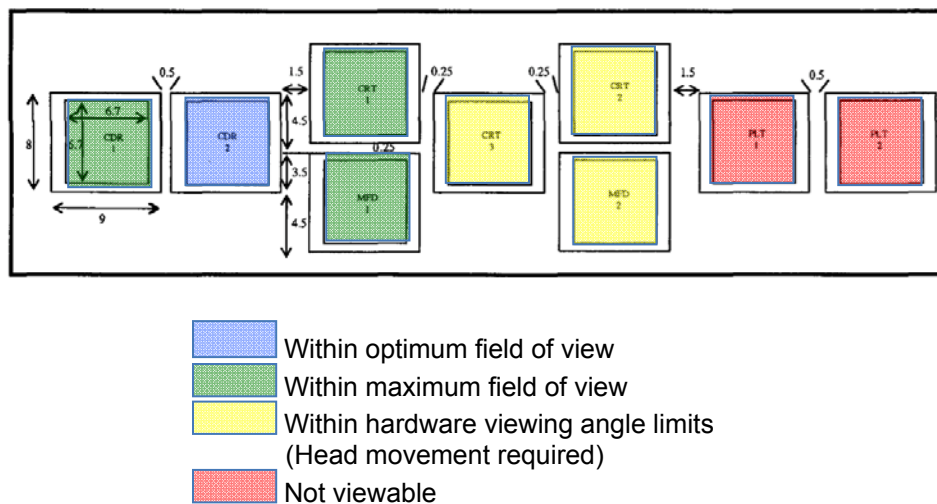
Figure 17.   Display Availability for Commander Position

The optimum visual field for the mission specialist station was determined to be a 32-inch diameter circle centered on the CRT3 display. This circle encompasses all five of the central MDUs. The maximum field of view was determined to include the entire forward instrument panel. The Mission Specialist is within hardware viewing angle limits for all MDUs, however the view

of the outboard-most units is partially obstructed by the commander and pilot seats. The visibility of each MDU from the Mission Specialist's position is summarized in Figure 18.

Within optimum field of view
Within maximum field of view
Within hardware viewing angle limits
(Head movement required)
Not viewable
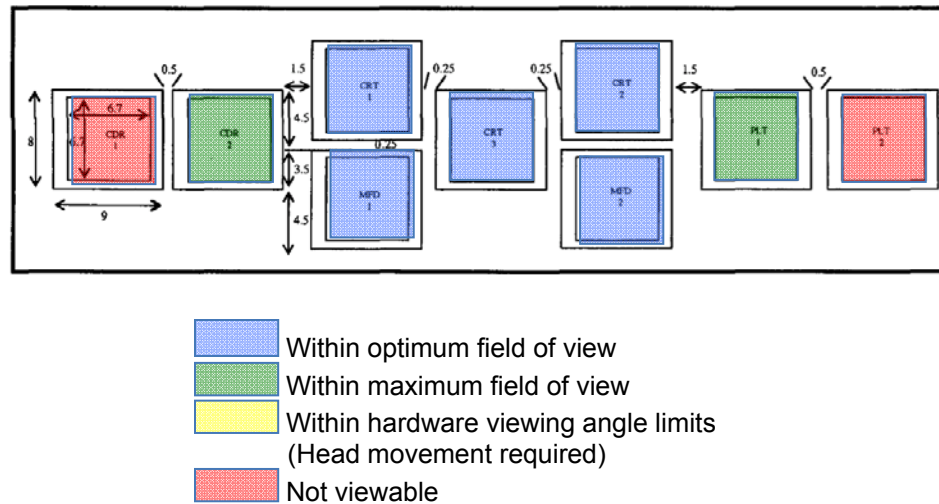
Figure 18.   Display Availability for MS2 Position

In the baseline configuration, the arrangement of display formats across each of the MDUs can be varied based on crew preference and phase of flight. A typical display layout for the baseline configuration is described in the Shuttle Crew Operations Manual (see Figure 19). Using this typical configuration, the relative accessibility of each display format was determined for each crew position (see Table 15).
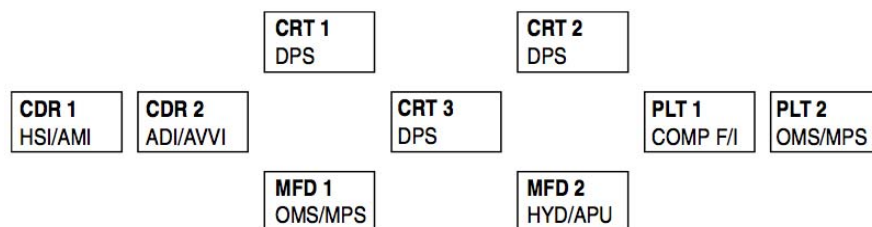
Figure 19.   Typical Baseline Display Configuration

Position	Displays within optimum FOV	Displays within maximum FOV	Displays visible with head movement
CDR	ADI/AVVI	HSI/AMI DPS (closest) OMS/MPS	HYD/APU DPS
PLT	COMP FI	OMS/MPS HYD/APU DPS (closest)	DPS
MS2	DPS OMS/MPS HYD/APU	ADI/AVVI COMP FI	

Table 15.   Display Format Availability in a Typical Baseline Configuration

As with the baseline design, the layout of display formats across each of the MDUs can be varied. A typical display layout for the CAU configuration is described in the CAU display dictionary (see Figure 20). Using this typical configuration, the relative accessibility of each display format was determined for each crew position (see Table 16).
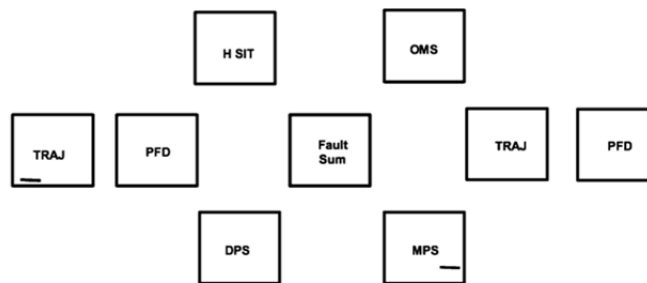
Figure 20.   Typical CAU Display Configuration

Position	Displays within optimum FOV	Displays within maximum FOV	Displays visible with head movement
CDR	PFD	Trajectory DPS Status Horizontal Situation	Fault Summary MPS OMS
PLT	Trajectory	PFD MPS OMS	Fault Summary DPS Status Horizontal Situation
MS2	Fault Summary DPS Horizontal Situation MPS OMS	PFD Trajectory	

Table 16.   Display Format Availability in a Typical CAU Configuration.

3.   Display Formats

Display format changes were evaluated according to the principles of display design (see Chapter II). Several dozen distinct display formats exist in each cockpit configuration, and there is no direct relationship between Baseline and CAU formats. Therefore, to evaluate design changes, displays were grouped according to function. Comparisons were made using representative example display formats from each cockpit design. Where applicable, design changes were described in terms of Perceptual Principles, Attentional Principles, Mental Model Principles, and Memory Principles.

a.   Primary Flight Instruments

Display formats were provided in both baseline and CAU cockpits that present graphical representations of legacy ADI, HSI, AMI, and AVVI instruments. These instruments provided information about basic flight parameters, such as attitude, speed, altitude, rate of climb, angle of attack, and heading. In the baseline cockpit, this information was provided by the HSI/AMI (see Figure 21) and ADI/AVVI (see Figure 22) display formats. Alternately, a composite ADI/HSI display (see Figure 23) was provided that combined these

instruments onto a single page. In the CAU cockpit, this information was provided solely by the Primary Flight Display (PFD) format (see Figure 24). The arrangement of these instruments in the CAU's PFD format was similar to the baseline cockpit's Composite ADI/HSI display, but some changes were made in the presentation of the information. The changes are discussed in the sections that follow.
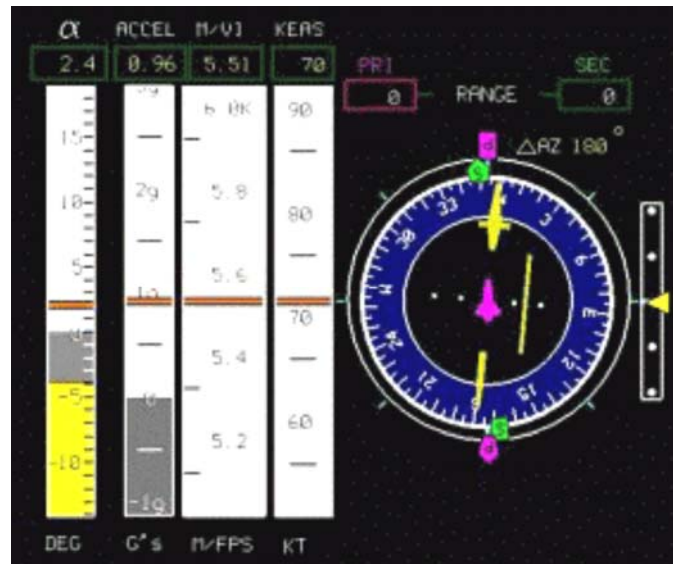
Figure 21.   Baseline HSI/AMI Display
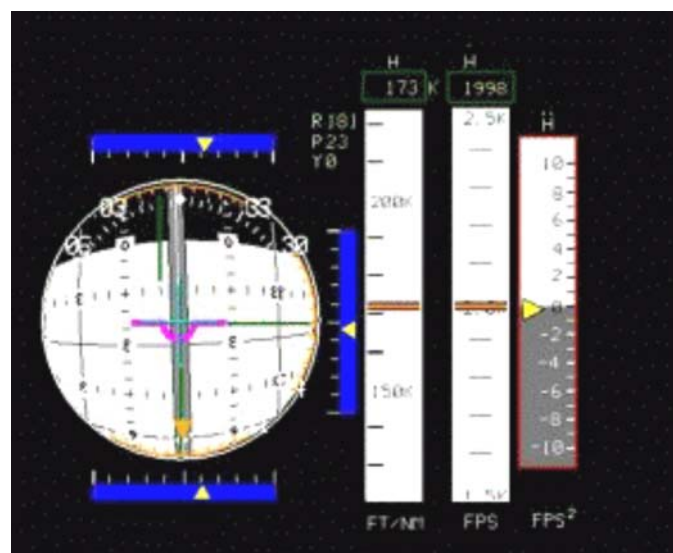
Figure 22.   Baseline ADI/AVVI Display

53

Figure 23.   Baseline Composite ADI/HSI Display
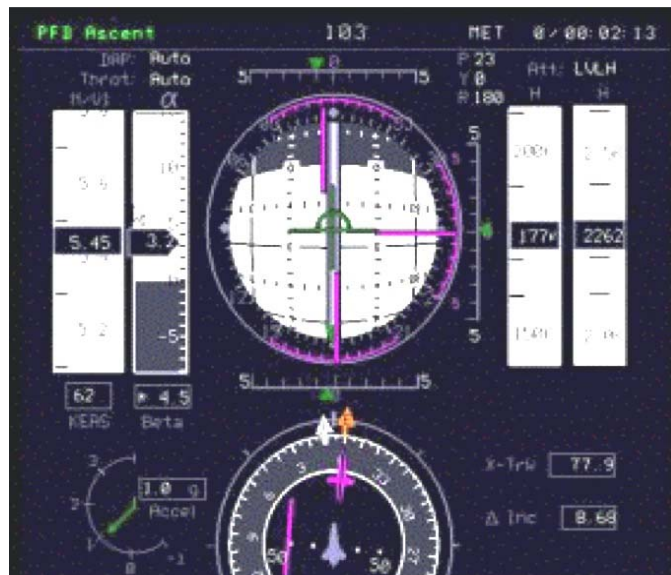
Figure 24.   CAU Primary Flight Display (Ascent Configuration)

Perceptual and Attentional Principles

The changes made in the use of color are immediately apparent when comparing the two displays. While the baseline display used several inconsistent colors to display its various elements, the CAU display presents a single, unified

color scheme. In the baseline display, separator lines and non-critical elements sometimes have as much contrast with the background as the messages themselves. For example, the boxes surrounding the numerical readouts are colored in green and magenta, and the backgrounds for the rate sliders around the ADI are blue. This unnecessary contrast decreases the relative salience of the important signals, thereby hindering perception. The CAU cockpit format presents these less critical elements in dark grey to reduce their contrast with the background, while presenting important signals in white, maximizing contrast.

A similar effect was achieved by minimizing the clutter in the bottom half of the display. The format of the accelerometer (g-meter), for example, was changed from a prominently colored moving tape display in the baseline cockpit to a less salient dial and pointer display in the CAU. Other display elements have simply been removed when not needed. This reduction in clutter makes it easier to maintain focused attention on the more critical elements.

The location of the digital readouts associated with each of the vertical tape instruments was changed from being above the tapes in the baseline cockpit, to being centered on the tapes in the CAU. This change exploits the proximity compatibility principle, as it makes it easier for the astronaut to mentally integrate the digital readout with the moving tape. In addition, this change brings the digital readouts in line with the centerline of the ADI, enabling a quick scan in a straight horizontal line. This helps to minimize the cost of moving selective attention from one instrument to the next.

Memory Principles

With the exception of the previously mentioned accelerometer, CAU instruments move and operate in the same way as the baseline instruments. The CAU displays therefore remain consistent with the experience of the existing user base, and do not require extensive re-learning to understand.

b. Guidance and Navigation

In general, there are two types of display formats used throughout both cockpits that provide guidance and navigation information: (1) Trajectory, and (2) Horizontal situation. Trajectory displays (sometimes called vertical situation displays) present a side view of the shuttle's trajectory, and horizontal situation displays present a top-down view. These formats are task specific, with different variants being used at each stage of ascent, entry, and abort. However, these formats had many common characteristics and were evaluated collectively.

In the baseline cockpit, this information was provided by a variety of MM-specific DPS formats. For ascent, only trajectory formats were provided (see Figure 25). For entry, both trajectory and horizontal situation formats were provided. In each case, separate formats were provided by PASS and BFS software. In the CAU cockpit, both trajectory and horizontal situation formats were provided for both ascent and entry (see Figures 26 and 27). Information from PASS and BFS software was integrated into a single page. The changes are discussed in the sections that follow.

Figure 25. Baseline BFS Ascent Trajectory

Figure 26.   CAU Ascent Trajectory Display

Perceptual Principles

The baseline formats used a simple green on black color scheme for all information. The CAU formats, in comparison, used multiple colors to distinguish between various classes of information (see Appendix B).   This change increases the salience and discriminability of the important signals, thereby improving perception.

Attentional Principles

The CAU displays consolidated information that, in the baseline cockpit, was spread across multiple displays. For example, the CAU ascent trajectory format (see Figure 25) combines the relevant portions of the baseline PASS and BFS ascent trajectory displays.   In addition, the CAU format incorporates performance information for the main propulsion system, which must be mentally integrated with trajectory information during ascent. In the baseline cockpit, MPS performance information was provided on a separate display format.   This change exploits the proximity compatibility principle, and reduces the division of attention needed to perform trajectory monitoring.

The CAU's ascent horizontal situation display (see Figure 27) has no direct antecedent in the baseline cockpit. This display format provides information about the horizontal flight path and abort options during ascent. This information was not provided on any display in the baseline cockpit. The determination of abort options was computed automatically at mission control center, and was relayed to the crew by voice communication when needed. In the event of communication failure, the determination had to be done manually by reference to primary flight instruments and a printed flight data file. Incorporating this information into a single display greatly reduces the information access cost associated with understanding abort options.

Figure 27.   CAU Ascent Horizontal Situation Display

c.   Systems & Fault Management

System summary display formats were developed for the CAU in order to consolidate status information about ECLSS, DPS, Navigation, Control, RCS, OMS, MPS, APU/HYD, and EPS subsystems. While each system's display is

unique, they all have many characteristics in common and were therefore evaluated collectively. The changes are discussed in the sections that follow.

Attentional Principles

In the baseline cockpit, subsystem information was spread across multiple displays, organized according to its source. For example, system parameters that were monitored by BFS software were found on the BFS system summary display format (see Figure 28). Some other system parameters were presented on dedicated display formats (e.g., MPS/OMS summary, see Figure 29). The CAU display formats were instead organized according to the subsystem(s) being represented. Information for a given subsystem was pulled from all available sources, and presented on a single display (see Figure 30). The CAU design changes, therefore, both exploited the proximity compatibility principle, and reduced the effort needed to shift selective attention between different information sources.

Figure 28.   Baseline BFS System Summary Display

Figure 29.   Baseline OMS/MPS Display

Figure 30.   CAU Main Propulsion Summary Display
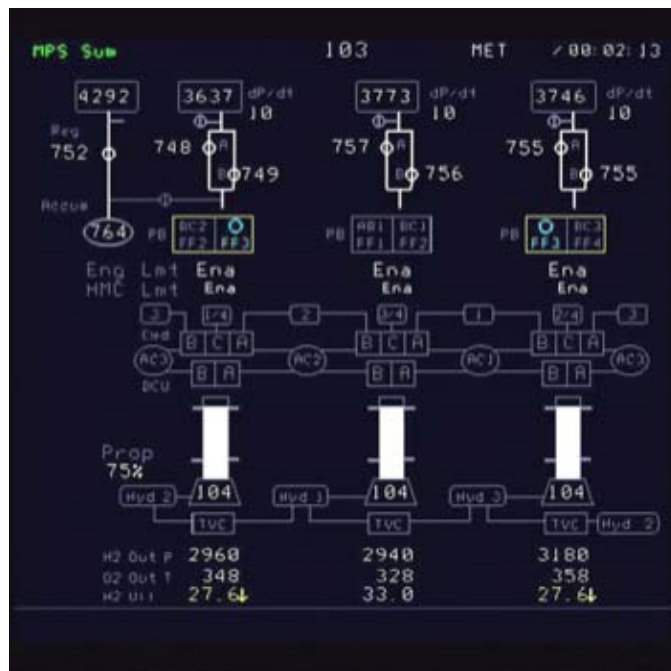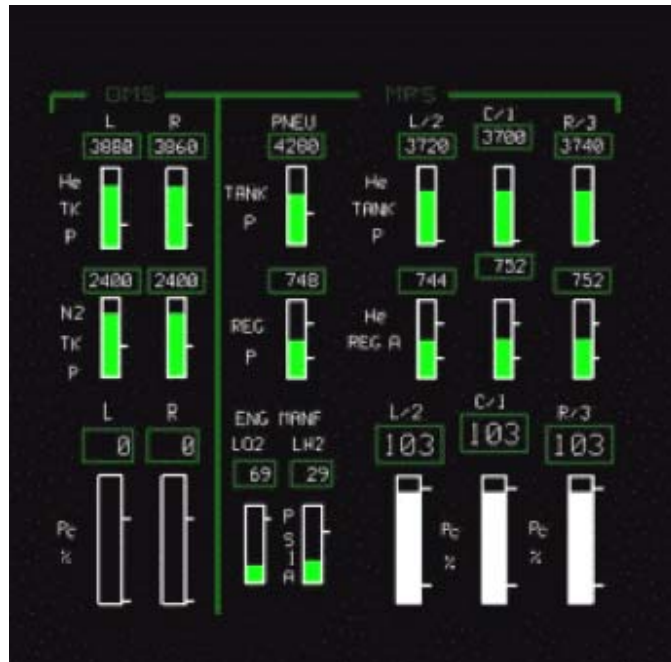
Perceptual Principles

In the baseline cockpit, most system information was presented in the form of green text over a black background. Numbers, labels, and dividing lines were all presented in the same color. The upgraded CAU formats instead exploited color to enable the crew to differentiate between the various classes of information, especially during off-nominal conditions (see Appendix B).

Mental Model & Memory Principles

Baseline and CAU cockpits differ greatly in the use of graphics. While the baseline display formats used matrices of numbers to convey information about a system, the CAU formats (where appropriate) incorporated this information into a rudimentary system schematic. The schematics depicted critical relationships between various system components. Icons were added to show the status of valves, and malfunctions are indicated by changing the color of the represented component. Instead of using a matrix of numbers, important values were placed on the schematic according to the component they relate to.

4.   Display Hardware

No changes were made to the shuttle display hardware as a result of the CAU program. Therefore, both the baseline and the upgraded cockpits have the same physical attributes. Each MDU consists of a 6.7 inch square full-color active matrix liquid crystal display (AMLCD), a brightness control knob, six programmable edge keys, and an on-off switch. All of the displays are readable in sunlight and can be viewed from +/- 60 degrees horizontally, and -10 deg/+45 degrees vertically. They have a pixel density of 172 ppi, a refresh rate of 60Hz, and a high ambient contrast ratio exceeding 6:1 (Thomsen & Hancock, 1994). All of these parameters are within current NASA standards with the exception of ambient contrast ratio, which is required to by 10:1. For comparison, current department of defense standards only require ambient contrast ratio of 5:1 for this type of display (DoD, 2012).

5. Display Control

In both the baseline and CAU configurations, display content is controlled through display edge-keys and keypads. No changes were made to the locations of these keys. The edge keys are used to navigate among available display formats. The keypads are used to provide data entry within certain display formats. In the baseline configuration, this data entry was only applicable to the three CRT display units when displaying DPS formats. In the CAU configuration, data can be entered into any display unit whenever an appropriate format is being displayed (McCandless, 2004). This change allows for a greater level of proximity compatibility when data entry is done on displays that are closer to the keypad (such as the MFD or CDR2 display), but it also opens the possibility for poor proximity compatibility when data entry is done on more distal displays.

NASA STD 3001 requires that there be a "logical and explicit" relationship between a control and its associated display. For display navigation, this requirement is met equally with the display edge keys in both the CAU and baseline cockpit. For data entry, this requirement is met by both the baseline and CAU configurations, but is best achieved in the CAU configuration when displays formats requiring data entry are selected on the MDUs that are closest to the keypad.

B. DESIGN ASSESSMENT

1. Mental Workload

The results of the ANOVA for MW are presented in Table 17. The three-way interaction was not found to be significant. Significant two-way interaction effects were found between Design and Phase of Operations, and between Design and Position. Both of these effects were found to be significant at the $p<.01$ level. The results also revealed a highly significant main effect for Design. This effect was far greater than that of any interaction, and is therefore worth

addressing. Plots of the data, Figures 31 and 32, reveal that the CAU design greatly outperforms the Baseline design in terms of MW.

Source of Variation	Sum of Squares	df	Mean Square	F	p
Design	1503.34	1	1503.34	643.79	0.000
Ops	49.80	1	49.80	21.30	0.000
Position	8.40	2	4.20	1.80	0.267
Design x Ops	23.04	1	23.04	9.86	0.002
Design x Position	29.54	2	14.77	5.05	0.002
Ops x Position	12.76	2	6.38	2.73	0.066
Design x Ops x Position	10.54	2	5.27	2.25	0.106
Error	1655.50	708	2.34		
Total	3448.99	719			

Table 17.   ANOVA for Mental Workload

To identify the details of Design/Operations interaction, the differences between Baseline and CAU designs were plotted for both Ascent and Entry operations (see Figure 31). Between Baseline and CAU designs, crews appeared to show a greater decrease in MW during ascent phases when compared to entry phases.

Post hoc comparisons were used to verify the significance of this interaction. The mean workload measure for the Baseline Ascent condition (M=7.0) was found to be significantly different from the Baseline Entry condition (M=6.1). However, the CAU ascent condition (M=3.7) did not significantly differ from the CAU entry condition (M=3.5). Both baseline conditions were significantly different from both CAU conditions. Taken together, these results indicate that the reduction in MW between baseline and CAU designs was significantly greater for Ascent operations as compared to Entry operations.
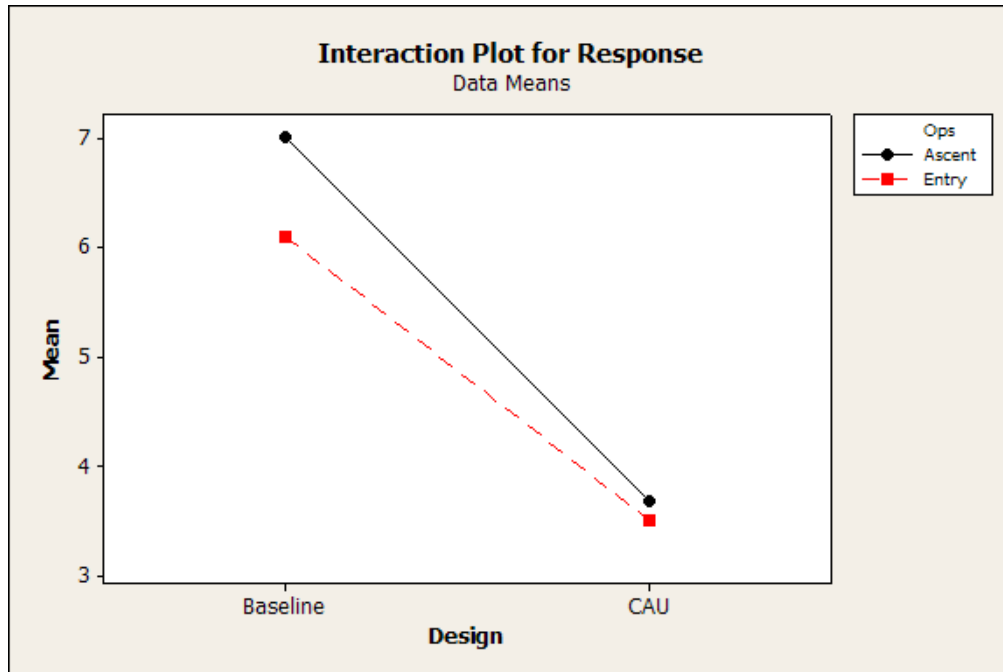
Figure 31.   Interaction between Design and Phase of Operations
for Mental Workload

To identify the details of Design/Operations interaction, the differences between Baseline and CAU designs were plotted for each of the crew positions (see Figure 32).   Between Baseline and CAU designs, Mission Specialists appeared to show a lesser reduction in MW than did Pilots and Commanders.

Post hoc comparisons were used to verify the significance of this interaction. The mean workload measure for the baseline Mission Specialist position (6.0) did not significantly differ from the baseline Pilot position (M=6.7). However, the CAU Mission Specialist position (M=4.0) was found to be significantly different from the CAU Pilot position (M=3.2). All baseline positions were significantly different from all CAU positions. Taken together, these results indicate that the reduction in mental workload between baseline and CAU designs was significantly less for the Mission Specialist position compared to Pilot position.
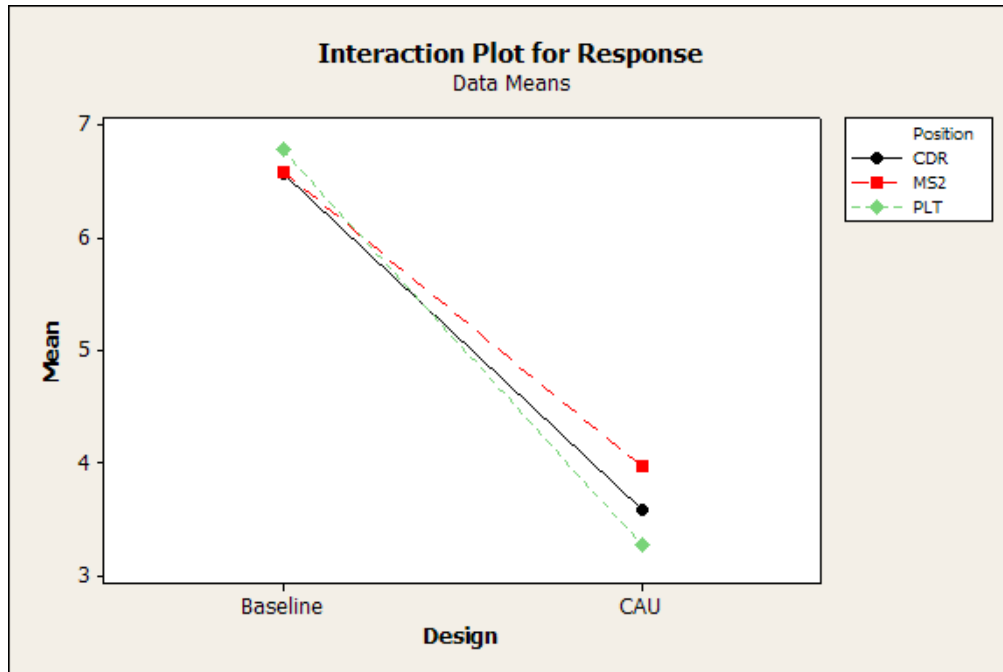
Figure 32.   Interaction between Design and Crew Position
for Mental Workload

2.   Situation Awareness of Trajectory

The results of the ANOVA for situation awareness of trajectory revealed a significant three-way interaction at the $p < .01$ level (see Table 18). To identify the details of this interaction, the data were divided by crew position, and the changes between Baseline and CAU designs were plotted for both Ascent and Entry operations (see Figures 33-35). The plots for CDR and MS2 positions both indicated a greater increase in SA for ascent phases when compared to entry phases. The plot for PLT position, by contrast, shows an equivalent increase in SA for entry and ascent phases.

Post hoc comparisons were used to verify the significance of this difference. CAU ascent PLT (M=8.6) was found to be significantly different from CAU entry PLT (M=7.8). The CAU ascent/entry pairs for PLT and MS2 positions were not found to be significantly different from each other. Baseline entry results for all three crew positions were found to be significantly different than the

corresponding baseline ascent results. Taken together, these results indicate that in terms of trajectory SA, pilots responded differently to the CAU upgrades than did commanders and mission specialists. Between baseline and CAU designs, commanders and mission specialists showed a greater increase in SA during ascent phases when compared to entry phases. Pilots, by contrast, improved at a more uniform rate across both phases of flight.

The results also revealed a highly significant main effect for Design. This effect was far greater than that of the interaction, and is therefore worth addressing. Figures 33, 34, and 35, clearly reveal that the CAU design greatly outperforms the Baseline design in terms of situation awareness of trajectory.

Source of Variation	Sum of Squares	df	Mean Square	F	p
Design	8046.8	1	8046.8	3282.95	0.000
Ops	209.27	1	209.27	85.38	0.000
Position	30.40	2	15.20	6.20	0.002
Design x Ops	40.57	1	40.57	16.55	0.000
Design x Position	13.84	2	6.92	2.82	0.060
Ops x Position	1.69	2	0.84	0.34	0.709
Design x Ops x Position	23.73	2	11.86	4.84	0.008
Error	3897.23	1590	2.45		
Total	14732.04	1601			

Table 18. ANOVA of Trajectory Situation Awareness
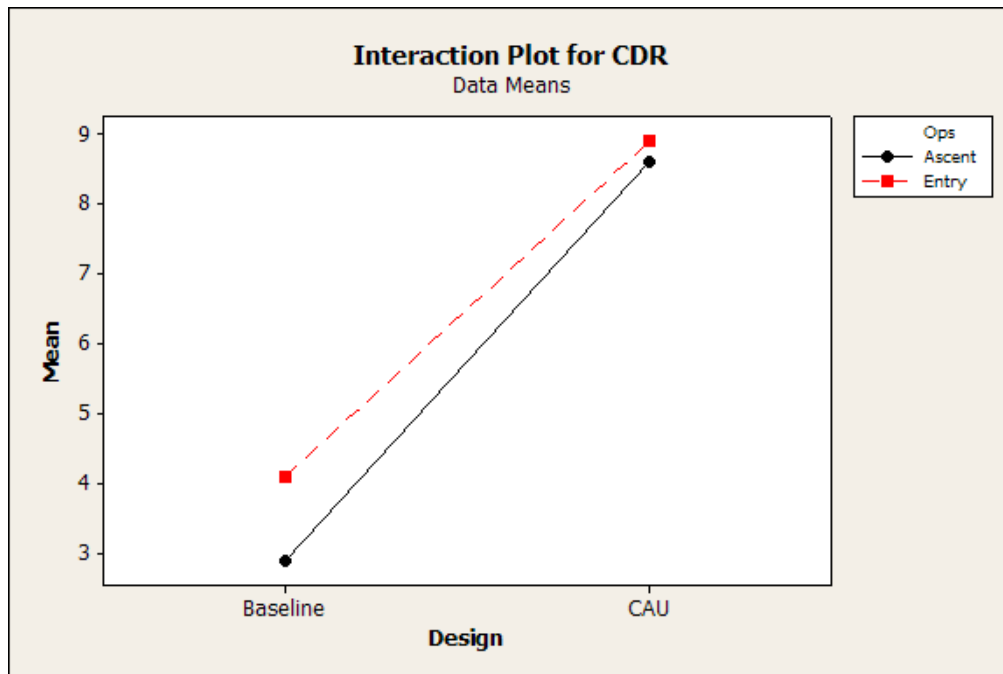
Figure 33.   Trajectory SA Interaction between Design and Phase of Operations for the Commander Crew Position
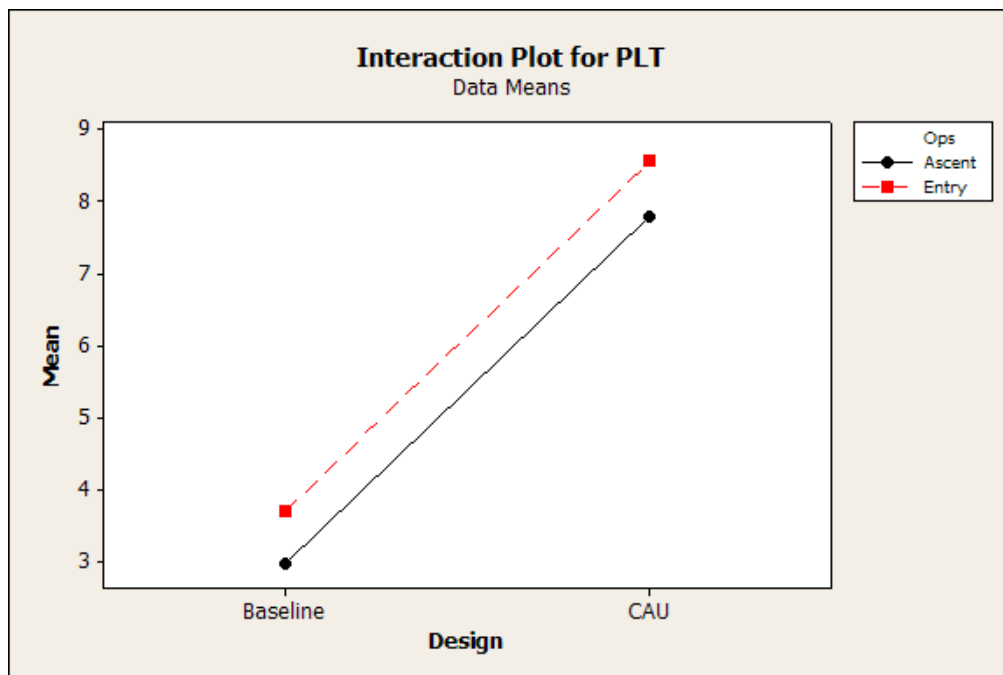
Figure 34.   Trajectory SA Interaction between Design and Phase of Operations at the Pilot Crew Position
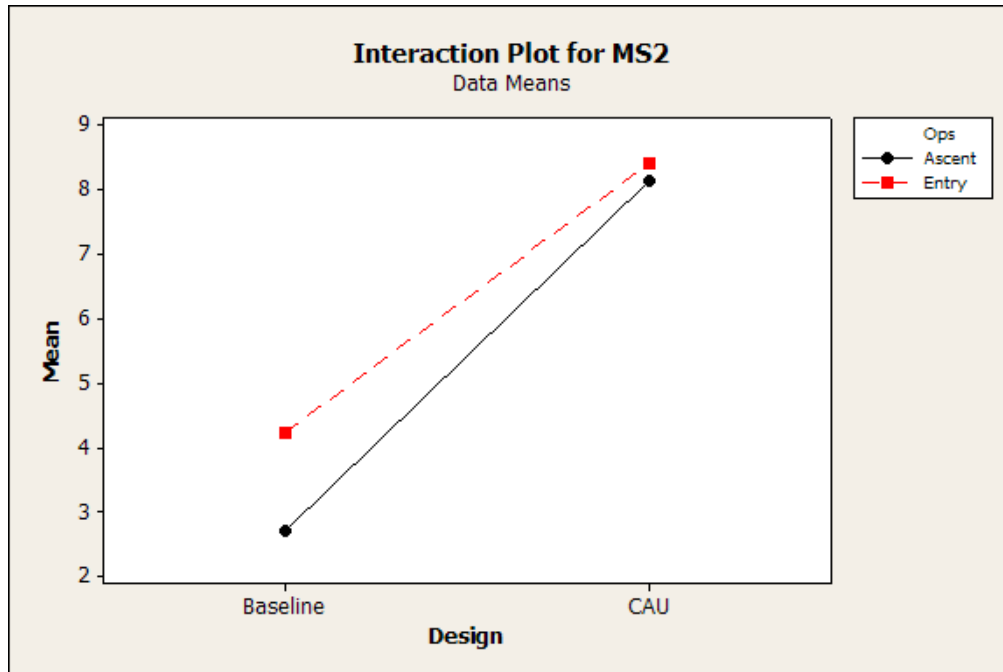
Figure 35. Trajectory SA Interaction between Design and Phase of Operations
at the Mission Specialist Crew Position

3. Situation Awareness of Critical Systems

The ANOVA for situation awareness of critical systems did not reveal a significant three-way interaction. However, a significant two-way interaction was found between Design and Operations at the $p<.01$ level (see Table 19). To identify the details of this interaction, the differences between baseline and CAU designs were plotted for both Ascent and Entry operations (see Figure 36). Between baseline and CAU designs, crews appeared to show a greater increase in critical system SA during ascent phases when compared to entry phases.

Post hoc comparisons were used to verify the significance of this difference. Baseline ascent SA was found to be significantly different from baseline entry SA. CAU ascent and entry SA were not found to be significantly different from each other. Both baseline conditions were found to be significantly different than both ascent conditions. Taken together, these results indicated that the design changes affected crew SA of critical systems differently in ascent

operations as compared to entry operations. Specifically, the changes yielded a greater increase in SA during Ascent operations than during Entry operations.

The results also revealed a highly significant main effect for Design. This effect was far greater than that of any interaction. As seen in Figure 36, the CAU design greatly outperforms the Baseline design in terms of situation awareness of critical systems.

Source of Variation	Sum of Squares	df	Mean Square	F	p
Design	10964.0	1	10964.0	5361.26	0.000
Ops	130.3	1	130.3	63.70	0.000
Position	48.4	2	24.2	11.84	0.000
Design x Ops	151.1	1	151.1	73.87	0.000
Design x Position	8.6	2	4.3	2.11	0.121
Ops x Position	8.1	2	4.0	1.98	0.139
Design x Ops x Position	8.7	2	4.3	2.13	0.119
Error	6306.9	3084	2.0		
Total	24262.1	3095			

Table 19.   ANOVA of Critical System Situation Awareness
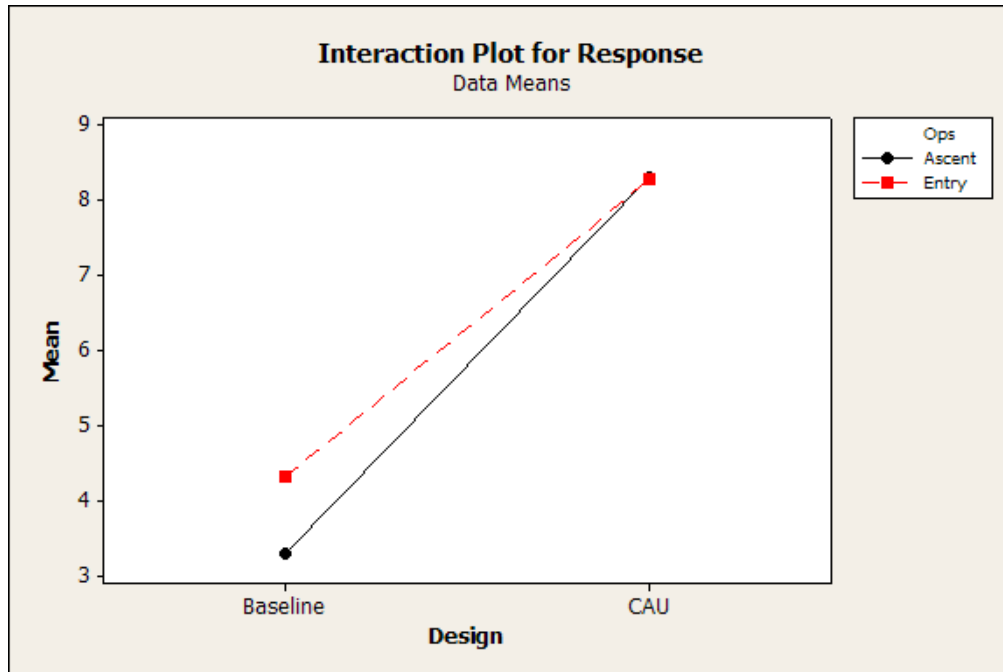
Figure 36.   Critical System SA Interaction between Design and
Phase of Operations

4.   Situation Awareness of Non-critical Systems

The ANOVA for situation awareness of non-critical systems did not reveal a significant three-way interaction.   However, a significant two-way interaction was found between Design and Operations at the $p<.01$ level (see Table 20). To identify the details of this interaction, the differences in SA between baseline and CAU designs were plotted for both Ascent and Entry operations (see Figure 37). Between baseline and CAU designs, crews appeared to show a greater increase in non-critical system SA during ascent phases when compared to entry phases.

Post hoc comparisons were used to verify the significance of this difference.   Baseline ascent SA was found to be significantly different from baseline entry SA. CAU ascent and entry SA were not found to be significantly different from each other. Both baseline conditions were found to be significantly different than both ascent conditions. Taken together, these results indicated that the design changes affected crew SA of non-critical systems differently in

ascent operations as compared to entry operations. Specifically, the changes yielded a greater increase in SA during Ascent operations than during Entry operations.

The results also revealed a highly significant main effect for Design. This effect was far greater than that of any interaction. Once again, the CAU design greatly outperforms the Baseline design in terms of situation awareness of non-critical systems (see Figure 37).

Source of Variation	Sum of Squares	df	Mean Square	F	p
Design	5480.17	1	5480.17	2586.71	0.000
Ops	35.63	1	35.63	16.82	0.000
Position	17.38	2	8.69	4.10	0.017
Design x Ops	18.33	1	18.33	8.65	0.003
Design x Position	11.30	2	5.65	2.67	0.070
Ops x Position	6.11	2	3.05	1.44	0.237
Design x Ops x Position	5.72	2	2.86	1.35	0.260
Error	2402.48	1134	2.12		
Total	8137.70	1145			

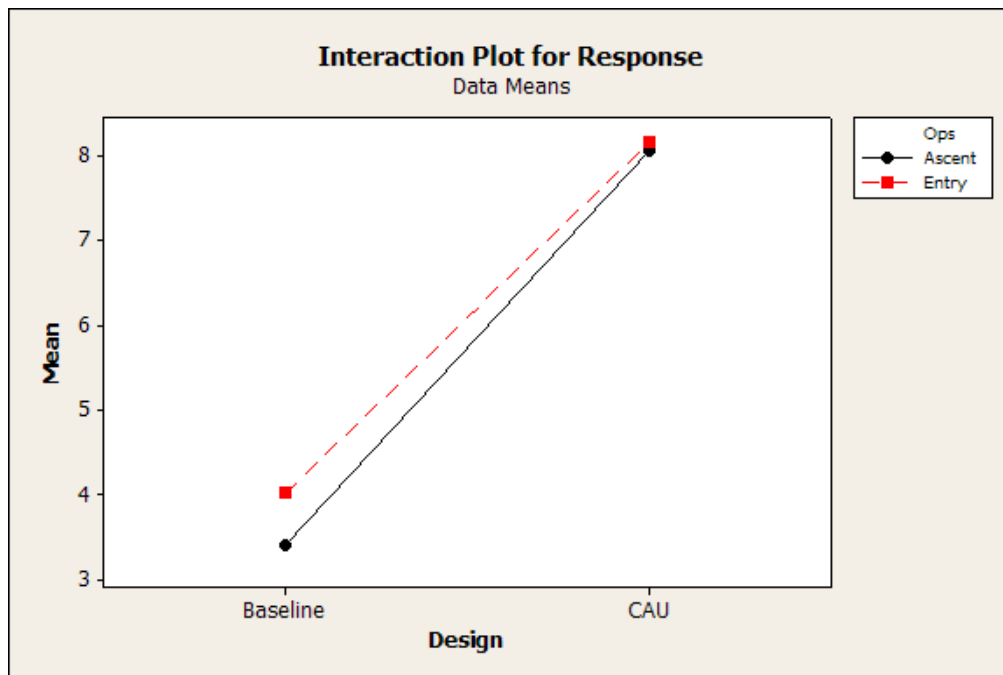Table 20.   ANOVA of Non-Critical Situation Awareness

Figure 37.   Non-Critical System SA Interaction between Design
and Phase of Operations

V. CONCLUSIONS AND RECOMMENDATIONS

A. SUMMARY

The primary purpose of the study was to complete a thorough assessment of the Shuttle CAU design improvements and to develop recommendations for the design and evaluation of future systems based on the results of this assessment. The evaluation began with a critique of the CAU design modifications in terms of task requirements, panel layout, design principles, hardware attributes, and display control. Finally, the effectiveness of these modifications was examined using measures of MW and SA. The statistical analysis focused on the measures with the highest quality data, which included Bedford Workload Rating Scale, and Subjective Situation Awareness.

For every performance measure examined, a significant improvement was found between the baseline and CAU designs for all crew positions and in all phases of flight. This was by far the strongest effect, with p-values on the order of 10^{-30} for every measure. The analysis also revealed that these improvements were significantly greater for ascent operations than for entry in all but one measure. That exception was found in the Trajectory Situation Awareness measure, wherein pilots improved at the same rate in both ascent and entry operations.

The Ascent H-SIT display is a development that is applicable exclusively during ascent scenarios in the CAU cockpit. This display is intended to improve ascent trajectory SA, particularly during highly dynamic (and attention-demanding) abort scenarios. It is typically visible from the Commander and Mission Specialist crew positions, but outside the maximum field of view for pilots. The design critique also noted several changes in design that apply throughout all conditions. Namely, the changes in the use of color, the consolidation of information, and the use of pictorial representations of systems are relevant to all crew positions and in all phases of flight.

B.   DISCUSSION

1.   Research Questions

Three research questions were proposed for examination in this thesis: (1) Did the CAU design changes lead to improvements in crew mental workload and/or situation awareness?; (2) If mental workload and/or situation awareness improvements exist, do they vary by crew station or phase of flight?; and (3) Can crew mental workload, and/or situation awareness changes be traced back to specific CAU design modifications?

The first research question was answered definitively. Highly significant improvements in MW and SA were found for every measure between the baseline and CAU designs. Furthermore, the design critique identified several factors in the CAU design changes, including use of color, consolidation of information, and use of pictorial system representations that, according to the literature, should have resulted in reduced MW and increased SA. It is therefore evident that the CAU design changes were behind the significant improvements that were identified.

Answers to the second research question were found in the analyses of variance. In particular, significantly greater improvements in MW and SA were found for ascent operations than for entry, and pilots were found to differ from the other crew positions in ascent trajectory SA improvements. These results definitively support that the improvements in MW and varied measurably by both crew station and phase of flight.

The third research question was answered by synthesizing the results of the statistical analysis with those of the design critique. Overall improvements between baseline and CAU designs can be traced to the use of color, the consolidation of imagery, and the pictorial representation of systems. However, the relative contributions of each of these factors cannot be determined.

74

The difference in improvement between ascent and entry operations, however, points to the ascent H-SIT display. From a display design perspective, the development of this format is the most significant change that affects only ascent operations. This determination is further supported by the observation that astronauts in the pilot position did not experience the same improvement as other astronauts in measures of trajectory SA. Since the CAU H-SIT is outside the normal field of view from the pilot's seat but within the normal field of view for commanders and mission specialists, it follows that pilots would not experience the same benefits as the other positions.

2.   Human Systems Integration Domain Applicability

The study identified a connection between human interface design changes, and operator MW and SA. This connection is directly relevant to the HFE domain. The lessons learned from this evaluation are broadly applicable to future HFE design and evaluation activities.

The design critique also identified changes in task requirements between the baseline and CAU designs. The connection between changing task demands and operator performance has implications in the training domain. This connection demands consideration of the effect of design on the knowledge, skills, and abilities that relate to the operation of the system.

The review of the scientific literature identified a clear connection between operator MW and SA, and system safety (e.g., Wickens, 2004; Tsang & Vidulich, 2006). Consequently, the findings of this thesis that relate to MW and SA also represent a contribution to the HSI domain of Safety. This design evaluation represents one way in which safety of future systems can be improved.

3.   Assessment Criteria

MW and SA are central concepts within aerospace human factors research (Wickens, 2002). Because of this, several measures of MW and SA were collected during the course of this evaluation. MW data were collected

using both BWRS and NASA-TLX instruments, and SA data were collected using both objective and subjective measures. Of these, BWRS and Subjective SA measures proved to be the most beneficial to this evaluation. For the SA measures, this decision was primarily based on the greater sampling of subjective data vs. objective data. For the MW measures, however, TLX measures showed less internal consistency than BWRS measures, and included possible indications of scale reversal by the participants. The BWRS, in comparison, appeared to be well understood by all participants.

C.   CONCLUSIONS

Collectively, the results indicate that the use of color, consolidation of information, and use of pictures dramatically improve MW and SA. These findings are consistent with the potential effects described in Wickens et al. (2004). However, as the experiment was not designed distinguish between the effects of these individual changes, no conclusion can be reached about their relative contributions. It is nonetheless clear that these design principles are highly effective when used together.

The evidence does support a more specific conclusion regarding the ascent H-SIT display. The difference in improvement between ascent and entry operations indicates that the predictive aiding provided by this display was particularly beneficial. The observation that astronauts who could not see the display experienced less of an improvement in trajectory SA than those that could further supports this conclusion.

D.   RECOMMENDATIONS

Future manned space systems would almost certainly benefit from incorporating the design principles demonstrated in the CAU. The use of a consistent color scheme, task-centric consolidation of information, and the use of pictorial representations of complex systems are clear, positive factors in reducing MW and increasing SA throughout all operations. Furthermore,

predictive aiding displays should be developed for use during all dynamic phases of flight. Because of the ubiquity of the underlying principles, these recommendations can be further generalized to manned aircraft as well.

In evaluating future systems, measurements of MW and SA should include the BWRS and Subjective SA. Both of these instruments appeared to be well understood by participants and provided data that are both meaningful and internally consistent. In addition, in future evaluations it would be of tremendous benefit to inquire of the participants what design features they felt had the greatest impact on overall improvement. Short of designing an experiment that examines design features in isolation, this would be the most appropriate way to distinguish specific design contributions to MW and SA.

THIS PAGE INTENTIONALLY LEFT BLANK

APPENDIX A.   EXAMPLES OF CAU BENEFITS

Note: The material in this Appendix has been condensed and adapted from *Evaluation of the Space Shuttle Cockpit Avionics Upgrade (CAU) Displays* by J.W. McCandless, Revision G, dated 2004, unpublished. Used with permission.

A.   ELECTRICAL BUS FAILURE

As an example illustrating the differences between the Baseline and CAU cockpits, this section discusses the process of diagnosing a failure of the Aft Power Controller (APC) #6, a component that provides power to a number of sensors, heaters, engine gimbals, valves, and other redundant systems. This example demonstrates changes that primarily relate to internal, system SA. This process is discussed for both Baseline and CAU displays.

1.   Baseline

When the APC6 failure occurs, the first indicator is a set of BFS driven Fault Messages generated by the Caution and Warning System. The crew member then views these fault messages on the "BFS FAULT" summary display (see Figure 38). The first of the three messages, "MPS LH2/LO2 ULL" indicates a problem with the MPS. The failure indicated is an effect of the loss of an MPS sensor. The crewmember must process several such indications of failures across multiple systems in order to infer the root cause of the problems.

```
1021/   /099    FAULT              5 000/00:01:55
                                 BFS 000/00:00:00
 CRT     FAULT           C/W  GPC       TIME
 ID
      MPS LH2/LO2 ULL      *    5    000/00:01:50
 SM2 APU SPD LO 2          *    5    000/00:01:50
 SM0 THRM APU                   5    000/00:01:50
```

Figure 38.   Baseline BFS Fault Summary Display

The crewmember has further insight to the MPS via the "BFS GNC SYS SUMM 1" display (Figure 39). As a function of the APC 6 failure, the "R MPS ULL P LH2" and "LO2" values show 'off-nominal' values of "12.0" with an accompanying down arrow for "LH2" and zero ("0") for "LO2." These numbers indicate how much pressure is available in the external propellant storage tanks to supply liquid hydrogen and liquid oxygen to each main engine. If the "LH2" pressure is too low, the engine will stop running from lack of fuel. With no other confirming cues, this signature looks like an impending loss of pressurization in the External Tank (ET).

The remaining fault messages, "THRM APU" and "APU SPD LO 2" refer to the APU system that provides hydraulic pressure to the shuttle. "THRM APU" indicates an off nominal temperature reading for one of the APUs, and "APU SPD LO 2" indicates a low turbine speed reading for APU 2. The "BFS SM SYS SUMM 2" display provides more specific information, showing that APU 2 has an off-scale low "SPEED" indicated by an "L," and APU 3 has an off scale low "GG BED" temperature, also indicated by an "L" (Figure 40).

```
1021/   /018    GNC SYS SUMM 1      5 000/00:02:00
                                  BFS 000/00:00:00


     SURF        POS    MDM    DPS    1 2 3 4
     L OB                      MDM FF
       IB                          FA
     R IB                          PL
       OB
      AIL
      RUD                      FCS CH 1 2 3 4
     SPD BRK
     BDY FLP
 MPS           L      C      R   NAV  1 2 3 4

 HE TK P     3280   3230   3240  IMU
   REG P A    784    768    768  TAC
         B    776    766    770  ADTA
    dP/dT      10     10     10  MPS PNEU HE P
 ULL P LH2    31.5   31.7   12.0↓      TK   4250
       LO2    21.1   21.0    0.0      REG    748
                                     ACUM   740
 GH2 OUT P   3400   3358   3425  MANF P LH2   30
 GO2 OUT T    420    418    427         LO2   60
```

Figure 39.   Baseline BFS GNC Sys Summ 1 Display

```
0001/   /079    SM SYS SUMM 2      5 000/00:02:30
                                  BFS 000/00:00:00
   CRYO TK   1     2     3     4     5    MANF1 MANF2
 H2 PRESS   208   208   206   206   206    208   207
 O2 PRESS   816   815   814   814   814    815   815
 HTR T1    -248  -248  -248  -248  -248
    T2     -248  -248  -248  -248  -248
   APU     1     2     3    HYD    1     2     3
 TEMP EGT  942   942   942  PRESS 3064  3064  3064
  B/U EGT  942   942   942  ACUM P 3080  3080  3080
  OIL  IN  250   250   250  RSVR T  116   153   142
      OUT  264   264   264
  GG BED   511H  511H    0L   QTY   72    74    71
  INJ     1271  1271  1271
 SPEED %    99     0L   101   W/B
 FUEL QTY   59    60    62  H2O QTY  78    73    78
  PMP LK P  14    14    14  BYP VLV BYP   BYP   BYP
 OIL OUT P  42    42    41
 FU TK VLV
       A T  63    65    62     THERM CNTL   1    28
       B T  63    65    62  H2O PUMP P    23    63
   AV BAY   1     2     3  FREON FLOW   2384  2384
 TEMP       97    97    83  EVAP OUT T     38    38
 FAN ΔP   3.45  3.46  3.39
```

Figure 40.   Baseline BFS SM System Summary 2 Display

Taken at face value, the signatures appear to imply the loss of 2 out of 3 APUs. However, in this instance the MPS and APU signatures are the result of instrumentation failures caused by the APC 6 failure. To determine the exact failure, the crewmember must utilize secondary cues across multiple displays to

81

confirm that the SSMEs and APUs are actually running normally, and conclude that these indications are instrumentation failures. The crewmember must then find the commonality between the sensors that failed simultaneously. Given the multitude of electrical buses and sensors on the vehicle, this "root cause determination" is a difficult task.

Inability to associate these signatures into a "root cause" of APC6 can have serious consequences. A crewmember failing to identify the proper failure means that they have low situation awareness of the state of the vehicle, and the correct procedure and recovery actions are not taken. This can lead to other equipment failures or problems later in the flight.

2.   CAU

Looking at an APC6 failure in CAU cockpit, we find that the Fault Messages annunciated for the failure are very similar. However, unlike the Baseline "BFS FAULT" summary display, the CAU "Fault Sum" display has a direct indication of a problem in the Electrical Power System (EPS) section (see Figure 41). The red "c" in the lower right of the section labeled "EPS" indicates a sub-bus failure on the Main C electrical bus. The "Fault Sum" also shows no indications of MPS or APU-related failures. The red "c" directs the crewmember to navigate to the EPS Summary ("EPS Sum") display.

Figure 41.   CAU Fault Sum Display

The "EPS Sum" display shows a layout of the shuttle's electrical bus system (see Figure 42). The EPS is divided into three Main buses (A, B, and C) that each have a set of sub-buses. APC6 is a sub-bus of Main C. The "EPS Sum" display colors the APC6 box and interior label red, indicating that the APC6 bus has failed. By directing the crews' attention to the appropriate location on the display via color-coding, the source of the failure is immediately apparent, eliminating the need to perform time-consuming and difficult cognitive operations to determine the root cause of the failure. Once the root cause is known, the crewmembers can reference their Flight Data File (FDF) to determine capabilities and redundancies lost throughout the cockpit and take the appropriate recovery/reconfiguration actions. These losses are also displayed to the crew on the appropriate displays, including the MPS and APU displays.
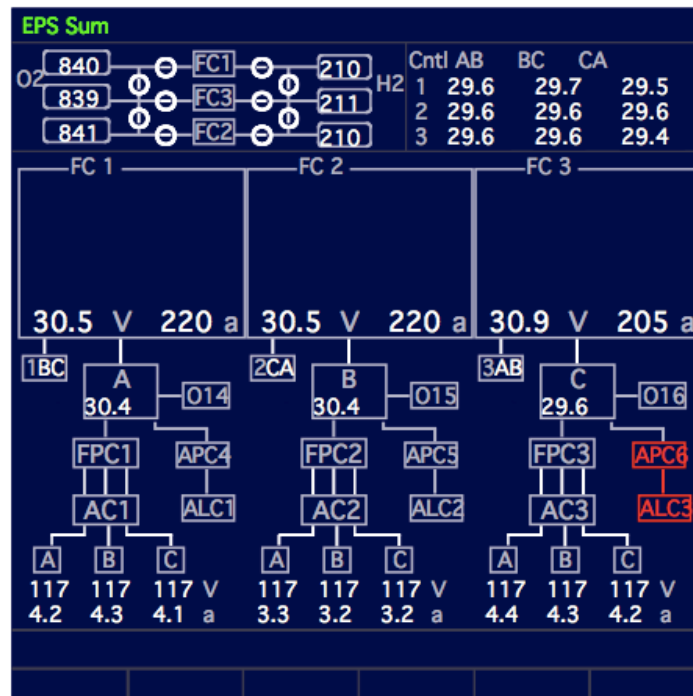
Figure 42.   CAU EPS Summary Display

If the crewmember looked at the "MPS Sum" display to see what the SSME impacts were to the APC6 failure, he or she would see that the Ullage pressure ("H2 Ull") value for the right engine has been replaced with the missing data symbol 'm' in cyan, indicating that the data is missing (see Figure 43). Since the APC6 failure removed power from this sensor, the CAU cockpit statuses the data as missing rather than showing invalid data as in the MEDS cockpit. Additionally, the CAU display shows that Helium valve A on the Right SSME is failed to the closed portion by coloring it red. This valve is one of two redundant valves on the R SSME that provides helium required for the engine to run. This failure indication is very valuable in letting the crewmember know not to close the "B" valve at any time in the future while the engine is running. Otherwise, the engine would be starved of helium and immediately shutdown. An engine shutting down due to a loss of helium is a potentially catastrophic condition. The "failed closed" status of the helium valve is not readily available in the MEDS cockpit.

Figure 43.   CAU MPS Sum Display

B.   TRAJECTORY

Another example illustrating the benefits of the CAU displays concept over the Baseline configuration focuses on the vehicle trajectory. This example demonstrates changes that primarily relate to spatial/temporal SA. This process is discussed for both Baseline and CAU displays.

1.   Baseline

The "H-SIT" display contains an overhead picture of the relationship of the shuttle and the runway (see Figure 44). In other words, it depicts what the shuttle and runway would look like if viewed from an altitude above the shuttle looking straight down. The shuttle symbol in the lower center of this section is in a fixed position on the display. The three dots in front of the shuttle are a prediction of where the shuttle will be in 20, 40, and 60 seconds based on parameters including speed and bank angle. The runway graphic (the circle with

the line in it) represents the Heading Alignment Cone (HAC), and the direction of the runway. The shuttle approaches the runway and makes a partial circle approach prior to lining up with the runway and landing. The runway graphic in this display example shows the runway off to the left of the nose of the shuttle, and the shuttle banking left (the 3 dots arc to the left) to approach it. Noticeably absent from this display is any indication of whether the shuttle has enough "energy" to make the runway. In other words, does the shuttle currently have enough speed and altitude to fly to the runway?

Figure 44.   Baseline Horizontal Situation Display

With Baseline displays, the only onboard approximation of this is to use range (R) and altitude (H) data from the ENTRY ALPHA cue card (see Figure 45). Using this cue card along with knowledge of the current altitude from the displays, the crew can compare their current range to the expected nominal range on the cue card. The crew notes their current (actual) altitude and range (indicated on the displays), then compares the range on the card for their given altitude with the actual range. For example, if the displays show that the space shuttle altitude is 227K feet, the crew would look on the ENTRY ALPHA cue card and determine that they should be 1508 nautical miles from the landing site. If

the crew's actual range (shown on the displays) does not match the range on the card, the vehicle is either high or low energy. To continue the preceding example, if the displays show the actual range is 1397 nautical miles, the crew would know they are in a high energy state. However, the crew cannot determine whether the selected site is still achievable. Additionally, the displays show the range to only the currently selected site (not alternate sites). If energy to the selected site is in question, the crewmember must select alternate sites, and then evaluate the range and altitude

ENTRY ALPHA

VR	αref	R	H	Href	Rref
25	40	4282	400		
24	HI 40 LO	2539	247	-48	L79
23	43 40 37	2114	239	-65	69
22	43 40 37	1774	234	-84	63
21	43 40 37	1508	227	-106	60
20	43 40 37	1292	221	-123	59
19	43 40 37	1122	215	-144	59
18	43 40 37	982	209	-165	60
17	43 40 37	865	203	-182	60
16	43 40 37	769	197	-190	61

Figure 45. Entry Alpha Cue Card

2. CAU

The CAU version of the Entry H-SIT contains a similar horizontal situation representation of that of the Baseline version of the display (see Figure 46). The shuttle symbol is placed at a fixed position on the display, and the magenta line connects the shuttle to the runway, which moves based on its horizontal relationship to the shuttle. In the figure provided, the runway is currently off to the left of the shuttle nose. This indicator also shows that the shuttle is banked to the left towards the site. This gives the crew the awareness that they are moving towards the site, not away from it.
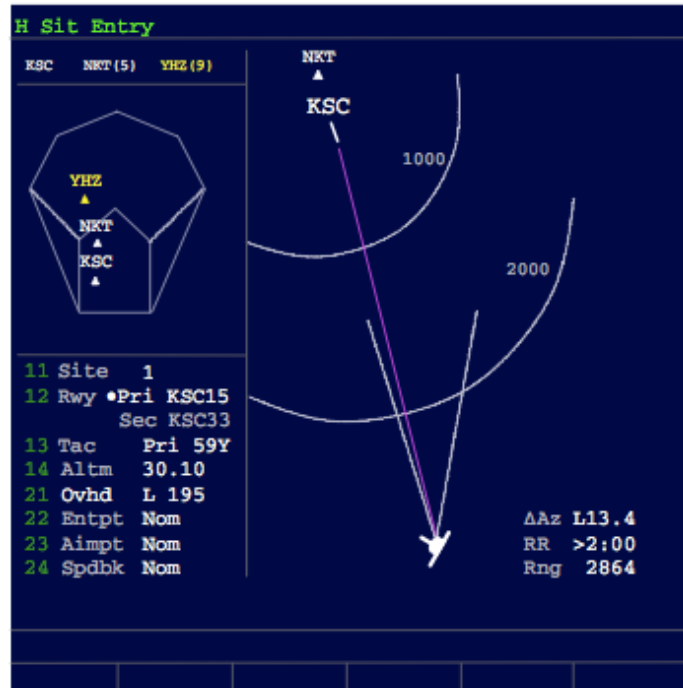
Figure 46.   CAU Entry Horizontal Situation Display

The CAU display also adds some key pieces of situation awareness that are not available on the Baseline displays. The display provides an indication on when the shuttle will command a "roll reversal." During the course of entry, the shuttle does not fly straight at the runway, but rather flies a series of left and right turns banking one way and then the other as it descends through the atmosphere. While the shuttle is banked left, its flight path turns to the left and the runway moves off to the right of the display. Eventually the shuttle has to reverse roll and bank to the right to avoid the runway moving too far off to the right to make a landing. The two gray lines extending from the shuttle graphic (they look like a big "V") indicate at what point the shuttle will command a roll reversal. In the example picture, the shuttle is banked to the left towards the site (the runway is moving to the right). When the runway moves far enough to the right, the magenta line will cross the right line of the "V" and the either the autopilot or the crewmember (flying manually) will reverse the bank angle and

88

turn back toward the runway. A countdown timer to the next roll reversal is located at the bottom right of the display. The display example indicates that there is more than two minutes to the next roll reversal. As the vehicle gets closer, the timer will begin counting down from two minutes and begin flashing when the roll reversal commences. Additionally, the display also provides digital readouts of the "ΔAz" (delta azimuth, called Del Az) which is the angle between the shuttle's current velocity vector and the vector to the landing site and range ("Rng"). These two pieces of data were available in Baseline conditions on another display. The range is also depicted in graphical form indicated by the "range rings." In the above figure, the plot shows the shuttle approximately 2900 nautical miles away from the site.

When the crew is flying manually, the CAU display also gives the crew a flashing alert when a roll reversal is approaching. When there is less than 10 seconds to a roll reversal, a "Roll in <10s" alert/message is displayed in this section. When a roll reversal needs to be commanded a "Reverse Roll" alert is displayed in its place. These alerts help keep the crew aware and help avoid missing or drastically delaying a roll reversal when flying manually. Not performing roll reversals promptly could potentially result in not achieving landing on a runway.

The top-left portion of this display also contains information on the shuttle's current energy state. This is a completely new set of information for the crew. This section is comprised mainly of an "energy footprint." This footprint displays information on what runways are achievable for landing based on the current energy state (in a basic sense, current velocity, altitude and distance from the site). The center region (looks like an upside down home plate) is the nominal energy region. Runways that appear in this section are achievable using nominal guidance commands. The regions outside of the nominal energy region are low energy regions. The sites located in these areas of the footprint are achievable only when the shuttle flies a profile designed to "stretch" its energy

out with special flying techniques that minimize the energy loss during the entry profile. Sites that are in the nominal footprint are colored white, which is the nominal color. Sites that are in the low energy footprint are colored yellow, which indicates that they are an off-nominal condition. Sites that are outside of the nominal and low energy regions are colored red. Red and yellow are the warning and caution colors, respectively. In the example above, both "KSC" and "NKT" runways are available in the nominal energy footprint, and "YHZ" (colored yellow) is available through the low energy techniques. The CAU display greatly simplifies this energy assessment process, allowing the crew to make quick and correct decisions on which runway to fly towards. The logic used in the energy evaluation is much more precise than the cue card method, and allows the crew to assess up to three sites simultaneously. The remaining portions of the Baseline display were moved to other CAU navigation displays to fit in with the task oriented design concept.

When the shuttle gets closer to the runway and enters the Terminal Area Energy Management (TAEM) phase of flight (below Mach 2.5 for a nominal entry), the CAU display tailors itself to flying in that regime. The energy state footprint is no longer needed, and the runway representation is swapped to show the HAC. Thus, the display provides the most important information when it is needed.

APPENDIX B.   CAU COLOR STANDARDS

The upgraded CAU display formats use color to enable the crew to differentiate between the various classes of information, especially during off-nominal conditions. Colors were primarily chosen based on usability principles, but were subject to hardware limitations. The color scheme used throughout the CAU display formats is presented in Table 21.

Color	Associated Elements
Dark Blue	Background
Dark Grey	Separator lines, non-critical elements
Light Blue/ Grey	Display labels
White	Nominal information
Magenta	Commanded messages which are critical for crew to read
Light Green	Display title, navigation elements, highlighting
Red	warning
Yellow	caution
Orange	Disagreement between primary and backup software
Cyan	Data unavailable

Table 21.   CAU Color Standards (After McCandless, 2004)

THIS PAGE INTENTIONALLY LEFT BLANK

LIST OF REFERENCES

Bentley, D. (2001). *CAU ascent traj display dictionary* (Revision C.). Houston: NASA Space Shuttle Cockpit Council.

Booher, H. R. (2003). Introduction: Human Systems Integration. In H. R. Booher (Ed.), *Handbook of Human Systems Integration* (pp. 1–30). Hoboken, NJ: John Wiley & Sons.

Cain, B. (2007). *A review of the mental workload literature* (No. RTO-TR-HFM-121-Part-II). Toronto: NATO RTO.

Carmody-Bubb, M. A., & Maybury, D. A. (1998). *Evaluation of prototype display of enemy Launch Acceptability Region (LAR) on the F/A-18 HUD*. In M. A. Carmody-Bubb (Ed.), (pp. 9–16). Presented at the Situational Awareness in the Tactical Air Environment, Piney Point, Maryland, 1998.

Casali, J. G., & Wierwille, W. W. (1984). On the measurement of pilot perceptual workload: a comparison of assessment techniques addressing sensitivity and intrusion issues. *Ergonomics*, *27*(10), 1033–1050.

Cowan, N. (2010). The Magical Mystery Four: How Is working memory capacity limited, and why? *Current Directions in Psychological Science, 19*(1), 51–57. doi:10.1177/0963721409359277

Department of Defense. (1987). *Aircrew station geometry for military aircraft (MIL-STD-1333B)*. Washington, DC: Government Printing Office.

Department of Defense. (2008). *Operation of the Defense Acquisition System (DoDI 5000.02)*. Washington, DC: Government Printing Office.

Department of Defense. (2012). *Department of Defense design criteria Standard: human engineering (MIL-STD-1472G)*. Washington, DC: Government Printing Office.

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *37*(1), 65–84. doi:10.1518/001872095779049499

Endsley, M. R. (1999). Situation awareness in aviation systems. *Handbook of Aviation Human Factors*, 257–276.

Fitts, P. M., & Seeger, C. M. (1953). SR compatibility: spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology*, *46*(3), 199.

Frank, A. (2010, February 2). The Vacant Sky: The end of the manned space program? *NPR*. Retrieved August 24, 2012, from http://www.npr.org/blogs/13.7/2010/02/the_vacant_sky_the_end_of_the.html

Hart, S. G. (2006). NASA-task load index (NASA-TLX); 20 years later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *50*(9), 904–908.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. *Human Mental Workload*, *1*, 139–183.

Hayashi, M., Huemer, V. A., Renema, F., Elkins, S., McCandless, J. W., & McCann, R. S. (2005). Effects of the space shuttle cockpit avionics upgrade on crewmember performance and situation awareness. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *49*(1), 54–58.

Hendy, K. C., Liao, J., & Milgram, P. (1997). Combining time and intensity effects in assessing operator information-processing load. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(1), 30–47.

Holland, A. W., & Vanderark, S. T. (1993). *Task Analysis of Shuttle Entry and Landing Activities*. Washington, DC: NASA.

Huchingson, R. D. (1981). *New horizons for human factors in design*. New York: McGraw Hill.

Jenkins, D. R. (2010). *Space Shuttle: The history of the national space transportation system* (3rd ed.). Author.

Kahneman, D. (1973). *Attention and effort*. Englewood Cliffs, NJ: Prentice-Hall.

Kirwan, B., & Ainsworth, L. K. (1992). *A guide to task analysis*. London: Taylor & Francis.

Logan, G. D. (1985). Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *39*(2), 367.

Lysaght, R. J., Hill, S. G., Dick, A. O., Plamondon, B. D., Linton, P. M., Wierwille, W. W., Zakland, A. L., et al. (1989). *Operator workload* (No. 851) (p. 263). U.S. Army Research Institute for the Behavioral and Social Sciences.

Marchant, C., Eastin, D., & Ferguson, R. (2001). Upgrading the US Space Shuttle fleet with a new smart cockpit. *Digital Avionics Systems, 2001. DASC. The 20th Conference*, *2*, 8B5–1–8B5–10 vol. 2.

Masters, G., Gordon, V., Lenahan, T., Landmann, M., O'Connor, J., Young, R., & Denihan, S. (2005). *US Naval test pilot school flight test manual (USNTPS-FTM-NO. 109).* (R. Cassol, J. Carty, S. Josselyn, D. Park, & T. Schumacher, Eds.). Patuxent River: Naval Air Warfare Center.

McCandless, J. W. (2003). *Color and luminance analysis of the space shuttle Multifunction Display Units (MDUs)* (No. NASA/CR-2003-212258) (pp. 1–30). Moffett Field: NASA Ames Research Center.

McCandless, J. W. (2004). *Evaluation of the space shuttle cockpit avionics upgrade (CAU) displays* (Rev. G.). Moffet Field: NASA Ames Research Center.

McCandless, J. W., McCann, R. S., Berumen, K. W., Gauvain, S. S., Palmer, V. J., Stahl, W. D., & Hamilton, A. S. (2005). Evaluation of the space shuttle cockpit avionics upgrade (CAU) displays. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *49*(1), 10–14.

Moray, N. (1967). Where is capacity limited? A survey and a model. *Acta Psychologica*, *27*, 84–92.

Muckler, F. A., & Seven, S. A. (1992). Selecting performance measures:" Objective" versus" subjective" measurement. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *34*(4), 441–455.

NASA. (2000). JSC2000-E-10522 [Photograph]. *STS-101 Shuttle Mission Imagery*. NASA. Retrieved August 13, 2012, from http://spaceflight.nasa.gov/gallery/images/shuttle/sts-101/html/jsc2000e10522.html

NASA. (2002). *Shuttle crew operations manual (OI-29)* (No. SFOC-FL0884) (Rev. B.).

NASA. (2011). *NASA SPACE FLIGHT HUMAN-SYSTEM STANDARD* (No. NASA-STD-3001) (Vol. 2). Washington, DC.

NASA. (2012, June 25). Human factors 101. (P. So, Ed.)*NASA Ames HSI division*. Retrieved August 13, 2012, from http://human-factors.arc.nasa.gov/awards_pubs/hf101.php

Reisman, G. (2002). *CAU fault sum display dictionary* (Revision D). Houston: NASA Space Shuttle Cockpit Council.

Roscoe, A. H., & Ellis, G. A. (1990). *A Subjective rating scale for assessing pilot workload in flight* (No. TR90019) (p. 16). Farnborough: Royal Aerospace Establishment.

Rouse, W. B., & Morris, N. (1985). *On looking into the black box: Prospects and limits in the search for mental models* (DTIC AD-A159080). Atlanta, GA: Georgia Institute of Technology, Center for Human-Machine Systems Research.

Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *Applied Psychology*, *53*(1), 61–86.

Selcon, S. J., Taylor, R. M., & Koritsas, E. (1991). Workload or situational awareness?: TLX vs. SART for aerospace systems design evaluation. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *35*(2), 62–66.

Stanfield, J. (2012, July 5). Orion's first test flight offers space launch system a first look at hardware operation, integration. Retrieved August 13, 2012, from http://www.nasa.gov/exploration/systems/sls/eft1_ksc.html

Thomsen, S. V., & Hancock, W. R. (1994). High-performance, AMLCD-based "smart" display for the Space Shuttle glass cockpit. *Digital Avionics Systems Conference, 1994. 13th DASC., AIAA/IEEE*, 281–288.

Tranthien, B. (1995). *The space shuttle Orbite's advanced display designs and an analysis of its growth capabilities* (Master's thesis, Naval Postgraduate School).

Tsang, P. S., & Vidulich, M. A. (2006). Mental workload and situation awareness. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics*. Hoboken, NJ: John Wiley & Sons, Inc.

Uhlarik, J., & Comerford, D. A. (2002). *A review of situation awareness literature relevant to pilot surveillance functions* (No. DOT/FAA/AM-02/3). Washington, DC: Federal Aviation Administration.

Weinstein, L. F., & Wickens, C. D. (1992). Use of nontraditional flight displays for the reduction of central visual overload in the cockpit. *The International Journal of Aviation Psychology*, *2*(2), 121–142.

Wickens, C. D. (2002a). Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science*, *3*(2), 159–177.

Wickens, C. D. (2002b). Situation awareness and workload in aviation. *Current Directions in Psychological Science*, *11*(4), 128–133.

Wickens, C. D., & Carswell, C. M. (2006). Information processing. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics*. Hoboken, NJ: John Wiley & Sons, Inc.

Wickens, C. D., Lee, J. D., Liu, Y., & Gordon-Becker, S. E. (2004). *An introduction to human factors engineering* (2nd ed.). Upper Saddle River, NJ: Pearson Prentice Hall.

Wickens, C. D., Vidulich, M. A., & Sandry-Garza, D. (1984). Principles of SCR compatibility with spatial and verbal tasks: The role of display-control location and voice-interactive display-control interfacing. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *26*(5), 533–543.

Wilson, J. R., & Rutherford, A. (1989). Mental models: Theory and application in human factors. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *31*(6), 617–634.

Woodson, W. E. (1981). *Human factors design handbook*. New York: McGraw Hill.

Yantis, S. (1993). Stimulus-driven attentional capture. *Current Directions in Psychological Science*, *2*(5), 156–161.

THIS PAGE INTENTIONALLY LEFT BLANK

INITIAL DISTRIBUTION LIST

1.   Defense Technical Information Center
      Ft. Belvoir, Virginia

2.   Dudley Knox Library
      Naval Postgraduate School
      Monterey, California

3.   John K. Schmidt, PhD
      Naval Postgraduate School
      Monterey, California

4.   Robert S. McCann, PhD
      NASA Ames Research Center
      Moffett Field, California

5.   John O'Connor
      US Naval Test Pilot School
      Patuxent River, Maryland