Advanced Integrated General Aviation Primary Flight Display User Interface Design, Development and Assessment

By

Charles Brent Campbell

B.S. Mechanical Engineering
The United States Military Academy, 1991

SUBMITTED TO THE DEPARTMENT OF AERONAUTICS AND ASTRONAUTICS IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF ENGINEERING IN AERONAUTICS AND ASTRONAUTICS

AT THE

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2001

Signature of Author: ______________________
Charles Brent Campbell
Department of Aeronautics and Astronautics
May 11, 2001

Certified by: ______________________
Charles W. Boppe
Senior Lecturer, Department of Aeronautics and Astronautics
Thesis Supervisor

Certified by: ______________________
James K. Kuchar
Associate Professor, Department of Aeronautics and Astronautics
Thesis Supervisor

Accepted by: ______________________
Wallace E. Vander Velde
Professor, Department of Aeronautics and Astronautics
Chair, Committee on Graduate Students

Advanced Integrated General Aviation Primary Flight Display User Interface Design, Development and Assessment

By

Charles Brent Campbell

Submitted to the Department of Aeronautics and Astronautics
on May 11, 2001, in partial fulfillment of the
requirements for the degree of Master of Engineering

ABSTRACT

This thesis describes work performed during a project in the Master of Engineering degree program in the Department of Aeronautics and Astronautics of the Massachusetts Institute of Technology. It was performed in close coordination with the Avidyne Corporation of Bedford, Massachusetts and involved design, development and assessment of the user interface for a primary flight display/horizontal situation indicator. The effort began with a Quality Function Deployment analysis of needs and requirements. Next the hardware interface was developed through two trade study iterations. Software interfaces were developed using various techniques including the Goals, Operators, Methods, Selection Rules (GOMS) Keystroke-Level Model. Two iterations of software interface development were conducted to accommodate evolving corporate business strategy. A human subject evaluation using a personal computer based simulation resulted in quantitative and qualitative results that indicate significant gains over a recent prototype. Improvements to the user interface were made in several areas including task execution time, accuracy and a subjective comparison of ease of use. Over the six common tasks, the mean task execution time for the baseline display was 37.6 seconds compared with 23.6 seconds and 22.2 seconds for two alternative user interfaces. In addition the accuracy of setting the standby NAV format task was significantly better in the new user interfaces. In a redundant paired comparison of the three interfaces based upon ease of use, the new interfaces were significantly better than the baseline. The application of the GOMS Keystroke-Level Model to primary flight display user interface design was validated through the human subject evaluation. Project outcomes support the Avidyne product development goal of fielding the first 'Highway-in-the-Sky' (HITS) flight display for general aviation applications.

Thesis Supervisor: Charles W. Boppe
Title: Senior Lecturer, Department of Aeronautics and Astronautics

Thesis Supervisor: James K. Kuchar
Title: Associate Professor, Department of Aeronautics and Astronautics

ACKNOWLEDGEMENTS

I would like to thank the Avidyne Corporation for providing such outstanding support and making me feel like part of the team. It was a great and enjoyable opportunity to work with such a talented group of people. Everyone was extremely nice, taught me a lot and made me feel right at home. A special thanks to Rick Willard, Steve Jacobson and Simon Mathews who put up with me on a daily basis.

I would also like to thank my thesis supervisors, Charlie Boppe and Jim Kuchar. They provided continuous support and were always available. I really appreciate all of their time, effort and positive feedback.

Thanks to the United States Army for allowing me the opportunity to attend graduate school and thanks also to MIT Lincoln Laboratory for providing support and allowing me the flexibility to pursue my thesis with Avidyne.

Finally I would like to thank my family and especially Ashley for believing in me.

Table of Contents

List of Figures and Tables

List of Figures and Tables (Continued)

FIGURES **PAGE**

TABLES

ACRONYMS

AGATE	Advanced General Aviation Transport Experiments
AHP	Analytic Hierarchy Process
ALT	Altitude
ANOVA	Analysis of Variance
A/S	Airspeed
AUX	Auxiliary
Baro	Barometer setting
BRT	Brightness
CDI	Course Deviation Indicator
CRS	Course
D-Bar	Deviation Bar
ESC	Escape
FAA	Federal Aviation Administration
FMS	Flight Management System
FMT	Format
GA	General Aviation
GAMA	General Aviation Manufacturing Association
GID	Graphical Input Device
GOMS	Goals, Operators, Methods, Selection rules
GPS	Global Positioning System
H	Homing
HDG	Heading
HSI	Horizontal Situation Indicator
HITS	Highway In The Sky
ILS	Instrument Landing System
IRDA	Infra-Red Data Adapter
K	Keying
KLM	Keystroke-Level Model
M	Mentally Preparing

MFD	Multifunction Display
NASA	National Aeronautics and Space Administration
NAV	Navigation
N/A	Not Applicable
NRST	Nearest
OBS	Omni-Bearing Sensor
PFD	Primary Flight Display
PRI	Primary
QFD	Quality Function Deployment
R	Responding
RMI	Radio Magnetic Indicator
RMS	Root Mean Square
RNG	Range
SAE	Society of Automotive Engineering
SRC	Source
Stby	Standby
T	Turning
T_{DC}	Turning of a dual concentric knob
TLX	Task Load Index
TNG	Training
T_S	Turning of a single knob
UI	User Interface
VLOC	VOR/Localizer
VSI	Vertical Speed Indicator
VSPD	Vertical Speed
Wx	Weather

1 INTRODUCTION

The Advanced General Aviation Transport Experiments (AGATE) Alliance is a unique partnership of public and private interests committed to improving safety in general aviation and to revitalizing the U. S. small aircraft industry. Under the leadership of the National Aeronautics and Space Administration (NASA) and the Federal Aviation Administration (FAA), more than 70 companies have joined forces and shared resources to establish new standards and to validate emerging technologies for single-engine/single-pilot airplanes.

As part of this effort, the AGATE alliance has awarded the 'Highway in the Sky' (HITS) contract to Avidyne, Avrotec and their partners to design the general aviation (GA) aircraft cockpit of the future. The goal of HITS technology is to significantly increase utility, safety and ease-of-flying. Affordable glass cockpit technology will provide pilots with direct access to all information needed to safely determine their routes, speeds, proximity to adverse weather conditions, terrain and other aircraft.

There are two stages to the AGATE Highway in the Sky program. Phase I will focus on the design of the primary flight display/horizontal situation indicator (PFD/HSI) that is planned for certification by September 2001. This phase of the program will establish the certification basis for the underlying architecture. Phase II will build on the Phase I infrastructure by implementing an easy-to-understand and simple-to-fly HITS display.

The focus of this effort was the design, development and assessment of the user interface for the Phase I primary flight display/horizontal situation indicator. This included both hardware and software aspects of the user interface.

The Avidyne Corporation sponsored this effort and all work was done in close coordination with Avidyne engineers. Avidyne is an avionics industry leader with innovative products that greatly enhance pilots' situational awareness and safety during every phase of flight. Avidyne's goal is to provide an affordable glass cockpit to the general aviation market so that pilots may benefit from the increased safety and ease-of-use brought about by Avidyne's technology and effective design.

In addition Avidyne contracted with IDEO, an industrial design firm, to conduct the industrial design of the PFD/HSI bezel once the user interface was determined. Close coordination with the efforts of IDEO as well as the various groups within Avidyne was critical to the success of the design effort.

2 BACKGROUND

The targeted segment for this product is the part of the aviation industry known as General Aviation or GA. The airline industry in the United States has been classified into three classes: the military, the airlines and everybody else [14]. "Everybody else" comprises general aviation. It includes everyone from the newest licensed pilots to professional corporate pilots who fly complicated jet aircraft. It also includes a wide variety of aircraft. Many more expensive corporate aircraft have advanced and expensive avionics, however the low end of the GA spectrum cannot typically afford advanced avionics such as "glass cockpits" or computer screen displays. By creating a less expensive "glass cockpit" display, Avidyne wants to provide all GA users the opportunity to have the latest technology as well as the functionality that comes with it. Typically pilots at the low end of the GA spectrum are less experienced and fly less often than the corporate pilots at the high end. Flying skills tend to degrade quickly when not in frequent use, especially for new pilots. Advanced displays must be easy to use to accommodate this lack of experience and perhaps infrequent use.

Two types of electronic flight displays were considered during this effort, the primary flight display (PFD) and the multi-function display (MFD). The primary flight display includes the information most critical for safe flight. It includes the flight instruments that are located on the screen in the same traditional "T" configuration of conventional analog flight instruments. The PFD is normally centered directly in front of the pilot since this is the most critical and often referenced display. Typical user interface tasks performed on the PFD include setting various reference markers or bugs for desired parameters such as heading or altitude, selecting navigation sources and formats and adjusting the map display clutter and range scale. Figure 2.1 depicts the primary flight display that was demonstrated to NASA and the AGATE alliance by Avidyne.

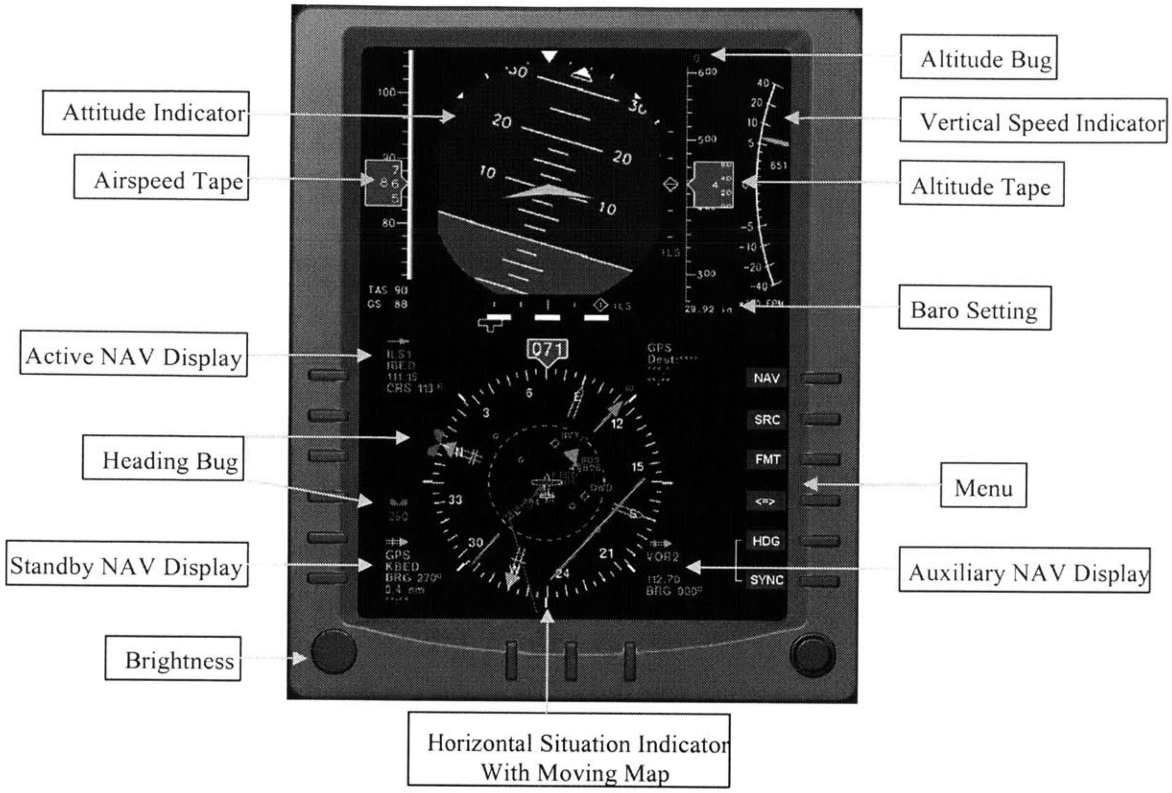


Figure 2.1: Demonstration Primary Flight Display

Buttons along the frame (or bezel) of the PFD are used to select functions displayed on the screen next to each button. In some cases, a rotary knob (lower right) is used to set parameter values (e.g. barometer setting). It is also possible to navigate through a menu of functions using the buttons – in this case, pressing a button may change the labels on the screen. Thus, a given button need not be restricted to a single function. The tradeoff between the number of buttons and the number of functions (menu depth vs. breadth) is an important human factors issue.

The multi-function display integrates a wide variety of sensor data onto a single computer screen for improved situational awareness. Types of data include navigation information (moving map), weather radar information, traffic advisory information, lightning strikes and ground proximity information. In some cases the MFD might provide a backup display to provide redundancy for the primary flight display in case of a failure. Another use of a multi-function display is the integration of flight management functions such as programming routes, changing radio frequencies or calculating fuel consumption. A typical MFD, the Avidyne

Flightmax 750, is shown in Figure 2.2. Future Avidyne PFD and MFD products will share a common hardware platform.

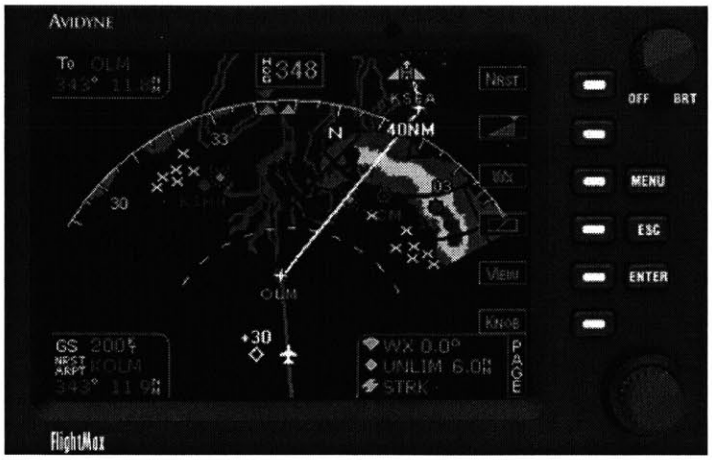


Figure 2.2: Avidyne FlightMax 750 Multi-Function Display (MFD)

Prior to this effort the prototype primary flight display depicted in Figure 2.1 was developed and demonstrated to NASA. This demonstration took place in February 1999. Since that time the demonstration unit was flown and evaluated, however no changes were made to the original configuration user interface. This effort began in January 2001 with the goal of having an improved software user interface on a newly designed hardware platform ready for demonstration at the Oshkosh Air Show in July 2001. The development schedule to meet this goal is shown in Figure 2.3.

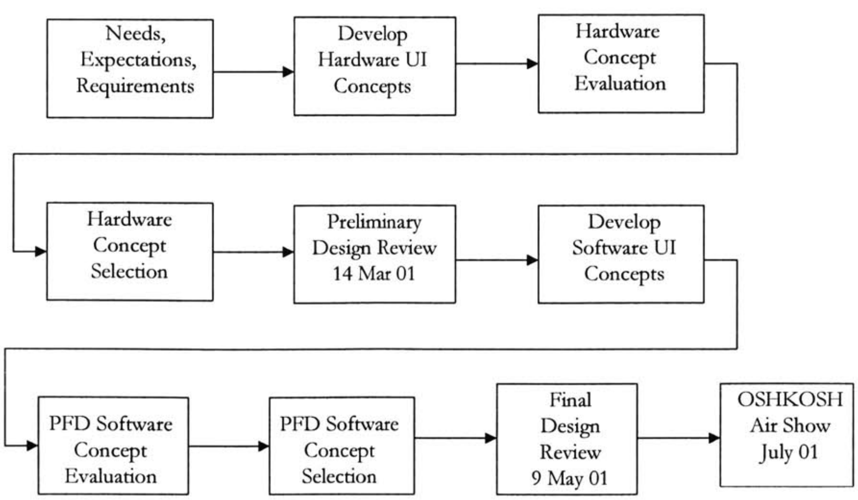


Figure 2.3: User Interface Development Schedule

3 METHODOLOGY

During the course of the design effort, a process was used to structure the effort and provide a logical framework. The formal methodology used during the user interface development is described in the following sections.

3.1 DEFINE PROJECT SCOPE

The first step was to define the scope of the project. Initially the effort of the project was to focus on the design, development and assessment of a remote control panel that could be used to control both a primary flight display and multi-function display. As this area was investigated it became apparent that the primary user interface with the PFD must be defined before serious consideration could be given to the design and utility of an additional remote control panel. For this reason the scope of the project was redefined to the design, development and assessment of the user interface for the Phase I primary flight display/horizontal situation indicator. This included both hardware as well as software aspects of the user interface. The hardware platform is the basis for many of Avidyne's applications to include the primary flight display and multifunction display. It must also incorporate future capability including flight management and highway-in-the-sky functions. For these reasons, much more than the primary flight display requirements were considered in the development of the hardware interface. Assessment was limited to that of the Phase I PFD/HSI functions.

3.2 LITERATURE/TECHNOLOGY SEARCH

Literature and technology research provided the basic, necessary information for understanding the problem. Also an attempt was made to take full advantage of previous work in this field in order to improve upon current designs. The introduction of "glass cockpits" is not new to commercial aviation, however they are relatively new to general aviation aircraft. In addition, advances in technology and computing power continually offer the potential for improvements. Background information and research results were collected via library and journal publications. The "state of the industry" was assessed through a review of competitor products on corporate websites. Also current industry magazines were a source of the latest competitor products and advances. Knowledge of industry standards such as those of the General Aviation Manufacturing Association (GAMA) and the Society of Automotive Engineering (SAE) were critical to the design process. This literature and technology information was essential during the Quality Function Deployment process and for generating new ideas and alternative concepts.

3.3 Needs and Requirements Analysis

To be successful, a product should meet the most important needs of the customer. The design process begins with understanding those needs. From these needs flow the technical requirements and design implementations that make up the design. Quality Function Deployment (QFD) was the method used to translate these needs into appropriate requirements and implementations [3]. QFD provides a formal means of deriving the most important technical requirements while minimizing individual bias. It also provides traceability for the derived implementations through the use of a Requirements Matrix. In addition QFD provides a means of identifying requirement conflicts.

3.4 Product Matrix

The next step in the QFD process was to translate the technical requirements identified in the Requirements Matrix into design implementations in a product matrix. Prioritized design implementations are derived from the product matrix and were used to determine alternative concepts.

3.5 Concept Development, Evaluation and Selection

Ideally, the user interface hardware and software should be developed concurrently. This would allow total integration of both the hardware and software in order to optimize the interface. The hardware platform in this case is common to both the primary flight display and the multifunction display. In addition, significant software upgrades are planned over the next several years that will add tremendous functionality but are not yet fully defined. The above constraints and long hardware lead times dictated that the hardware interface be developed first.

The design process was then repeated for the development of the software interface given the common hardware platform.

3.6 Hardware Interface Development

Evaluation criteria were developed to evaluate alternative concepts. These alternative concepts were developed based upon the most significant design implementations from the QFD process. The alternatives were evaluated using a Pugh matrix to identify the preferred concept [16]. A Pugh matrix was used since meaningful, numerical weightings could not be determined at this early conceptual design phase. Relative assessments in the Pugh matrix were made using engineering judgment.

Due to the variety of possible implementations, an additional iteration of concept development, evaluation and selection was performed to further refine the desired hardware. A final preferred hardware interface concept was selected based on the generated alternatives and Avidyne's schedule and customer goal constraints. This concept was then passed to an industrial design firm. The industrial design firm had several tasks. These tasks included developing an Avidyne "design language" or distinctive look, meeting certain mechanical attachment requirements, and preparing the hardware platform for manufacturing.

3.7 Software Interface Development

The software interface was developed by first determining the required tasks that the interface must support. These tasks were identified from requirement documents. Next implementations of these tasks were developed taking into consideration feedback from flight tests using an early prototype. Analyses of these implementations were made using a Goals, Operators, Methods, Selection Rules (GOMS) Keystroke Level Model [2]. Quantitative results from the GOMS analyses along with other human factors considerations were used to select a proposed configuration. After a review by the Avidyne president, an additional iteration of development was conducted taking into consideration corporate-level goals. This resulted in an additional alternative that was included in computer-based simulations for pilot assessments.

4 Needs, Expectations and Requirements

A good first step in product design would be to identify the needs of the customer, but in this case the first step was identifying the customer. For this product there were actually several customers whose needs had to be addressed. This complicated the process. An additional complicating factor is that Avidyne did not wish to convey to specific aircraft manufacturers the impression that they were tailoring their product for them. In other words, Avidyne wanted to address the aircraft manufacturers' needs but not allow any single aircraft manufacturer to dominate the design.

4.1 Customer Needs

Initially three categories of customers, or stakeholders, were identified. The first category included end users such as pilots who would actually fly using the Avidyne products and maintenance personnel responsible for maintaining the equipment. Aircraft manufacturers and owners who would install the Avidyne products either as original equipment or as upgrades comprised the second category. The final category included sub-groups within Avidyne itself. It was important that the hardware interface meet the needs of the various groups developing software applications including the Primary Flight Display Group, Multifunction Display Group and Flight Management System Group as well as the Hardware Group itself.

Determining needs and weightings for this disparate collection of customers that all could agree on was a difficult task. The strategy employed was to ask the various customers about their needs and then identify recurring needs as the most important. Survey data collected by the AGATE alliance reflected the views of pilots, aircraft owners and aircraft manufacturers. Avidyne marketing personnel were also interviewed since they had close working relationships with potential customers. Also site visits were conducted to three prospective aircraft manufacturers. Finally a group made up of select members from each of the various Avidyne sub-groups was assembled. This group synthesized the various needs into a single list of the thirteen most important needs and weighted them accordingly. Table 4.1 lists these needs and their weightings on a scale from 1 (not important) to 10 (very important).

Table 4.1: Customer Needs and Weightings

Full Functionality	10
Reduce Pilot Workload	10
Ease of Certification	10
Low Cost of Ownership	10
Ease of Use	9
High Adaptability	8
Aesthetic Appeal	8
High Reliability	8
Ease of Manufacture	6
High Maintainability	6
Low Weight	6
Error Tolerant	5
High Upgradeability	3

Ultimately the needs were also divided into three categories. They were classified as user needs, industry needs and user/industry needs. It was determined that the needs of the various aircraft manufacturers and Avidyne were similar enough that they were consolidated as industry needs. The identified needs are as follows:

USER NEEDS

- Full Functionality (10) – Having the full range of features customers expect in a state-of-the-art "glass cockpit" configuration. As a minimum the glass cockpit must perform the functions of a traditional cockpit and meet FAA certification requirements. Ideally the combined configuration will enable higher-level integrated functions not available in current non-integrated configurations.

- Ease of Use (9) – Must be intuitive to the user and require little initial and refresher training. Must be useable without extensive reference to supporting documentation.

- Reduce Pilot Workload (10) – Makes pilots' tasks easier and avoids high-demand tasks during critical phases of flight.

- Error Tolerant (5) – Enables user to quickly discover and recover from mistakes.

INDUSTRY NEEDS

- Ease of Certification (10) – Must meet all FAA certification requirements in a timely manner.

- High Adaptability (8) – Must be useable in various cockpits and configurations. Able to be sold and installed in various configurations to meet customer needs.

- Ease of Manufacture (6) – Ability to be manufactured quickly and without significant rework of current manufacturing capabilities.

USER / INDUSTRY NEEDS

- Low Cost of Ownership (10) – Initial and operational costs of products must be lower than that of comparable products currently in use by commercial airlines and high-end business aircraft. This is central to Avidyne's philosophy of providing high performance at a price affordable to General Aviation users.

- Aesthetic Appeal (8) – Appeals to both users and industry customers alike. Reflects Avidyne's "design language."

- High Upgradeability (3) – Ability to accommodate future software changes and incorporate new functions and features including Highway-in-the-Sky (HITS) displays and operation with a Flight Management System (FMS).

- High Reliability (8) – High probability of long duration of failure-free performance under normal conditions as stated in Requirements and Technical Concepts for Aviation (RTCA) DO-160D: Environmental and Test Conditions for Airborne Equipment.

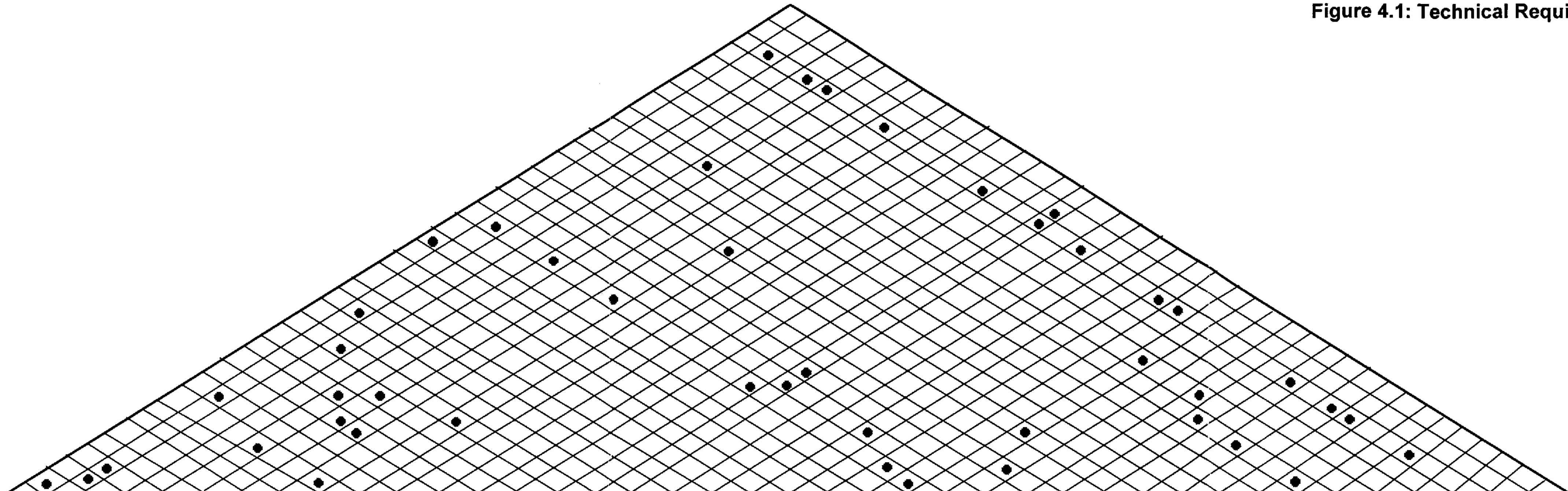
- High Maintainability (6) – High probability of long duration of maintenance-free performance. Must be easy to install/replace and repair.

- Low Weight (6) – Minimize weight consistent with program target goal of 3.5 lbs.

4.2 QUALITY FUNCTION DEPLOYMENT REQUIREMENTS MATRIX

Once the customer needs were identified and weighted, the QFD process was used to identify technical requirements for the cockpit display interface. In order to minimize biases of individuals and the various groups at Avidyne, a cross-disciplinary group was assembled to apply the QFD technique. These individuals represented each of the engineering sub-groups at Avidyne. This group started with the customer needs or "whats" and translated them into technical requirements or "hows." The complete matrix is presented in Figure 4.1. Rows represent customer needs and columns represent technical requirements that fulfill those needs. Numbers at the intersection of the needs and technical requirements represent how well the corresponding requirement helps satisfy the corresponding need. A nine represents a strong positive correlation, a three is moderate and a one is weak. How well a requirement satisfies a need was based upon the engineering judgment of the group, research conducted and statements of work from various aircraft manufacturers. In addition to technical requirements, various constraints were also identified and listed in the columns. These constraints were not assessed relative to the customers' needs but rather were listed in order to identify potential conflicts with the technical requirements. The most significant of these constraints is size as defined by the various aircraft manufacturers. In order to be marketable to all aircraft manufacturers, the interface must conform to the smallest size constraint. Also industry engineering standards as well as FAA regulations must be adhered to. Conflicts are identified in the upper portion of the Requirements Matrix.

Figure 4.1: Technical Requirements Matrix



	Weighting	Functional Integration	Dedicated Controls	Multi-function Controls	Pointing Devices	Large # of Controls	Auto Context-Dependent Functions	Voice Activated Functions	Distributed Interface	Single Interface	Software Upgradeable	Reconfigureable Hardware	Reduce Functions	Configure for Convenient Location	Intuitive Controls	Low # of inputs / action	Ergonomic Design	GUI Best Practices	Visual Feedback	Audible Feedback	Tactile Feedback	"Undo" Capability	Modular Design	Adjustable Geometry	Minimize Number of Parts	Use COTS Components	Use Components Common to all Interfaces	Avidyne Design Language	High MTBF	Redundancy	High Quality Parts / Components	Easy to Install / Replace	Easy to Repair	Durable Materials	Light Materials	Include "Growth" Controls		CONSTRAINTS	Size Constraints	FAA Regulations	Engineering Standards
User Needs																																									
Full Functionality	10	9	9	9	9	9	9	9	9	9	9	9			1									3																	
Reduce Pilot Workload	10	9	9		3	3	9	9			3	1	9	9	9	9	9	9	3	3	3	3		3																	
Ease of Use	9	9	9		3	3	9	9			3	1	9	9	9	9	9	9	9	9	9	3		3						3											
Error Tolerant	5	3					3				3		3	3	3	3	9	3	9	9		9																			
Industry Needs																																									
Ease of Certification	10										9	1	3						9	9	9								3		1										
High Adaptability	8			9					3														9	3						3											
Ease of Manufacture	6	3		9						3	3		1					1							9		3				3										
User / Industry Needs																																									
Low Cost of Ownership	10	1		3						9	3		1					1						3	9	9	3		3		3	9	9			3					
Aesthetic Appeal	8				1		1	1			1			3			9	1	9	9	9						3	9													
High Reliability	8		1						3		1		3			1		1							3				9	9	9			3							
High Maintainability	6										1		3			3		1							9	3					9	9	9	3							
Low Weight	6	1		1						3	1		1					1							9										9						
High Upgradeability	3			9	3	1	3	3	3		9	3						3					3	3						3						9					
Total		310	269	279	164	150	293	278	147	216	355	128	280	210	196	212	288	239	318	318	273	102	81	150	276	108	72	72	132	132	184	144	144	42	54	57					
Ranking		17	15	15	8	7	16	15	7	11	20	5	15	11	10	11	16	13	18	18	15	4	2	7	15	4	2	2	6	6	9	7	7	0	1	1					
																							$350 for hardware				20,000 Hours						3.5 lbs						

Based upon the correlation between the various requirements and needs, a score is determined for each technical requirement. A relative importance score is calculated using the raw scores and the technical requirements can be prioritized with this score. Relative scoring of the technical requirements on the Requirements Matrix will be used as the weighting in the Product Matrix. The top ten of the thirty-three technical requirements can be seen in Table 4.2.

Table 4.2: Top Ten Technical Requirements

Rank	Technical Requirements	Score
1	Software Upgradeable	355
2	Visual Feedback	318
3	Audible Feedback	318
4	Functional Integration	310
5	Auto Context-Dependent Functions	293
6	Ergonomic Design	288
7	Reduce Functions	280
8	Multi-Function Controls	279
9	Voice Activated Functions	278
10	Minimize # of Parts	276

Although all of the technical requirements will be carried over into the Product Design Matrix, it is important to note the highest priority technical requirements. These requirements have the highest impact in meeting the various customer needs. In this case the highest priority requirement, that the design be software upgradeable, is significantly higher than the second highest (355 vs. 318). From the second requirement on down, the requirements are grouped rather closely. If possible, conflicts between the top priorities should be resolved. Fortunately the top four priorities do not conflict with each other. The most significant conflict is that of reducing the number of functions with using automatic context-dependent functions and voice-activated functions. It turns out that reducing the number of functions is an unacceptable approach from a marketing and sales standpoint. Also in this particular case the addition of automatic context-dependent functions and voice activated functions is not practical due to schedule, cost and complexity issues; however these requirements should certainly be considered in future designs.

4.3 Product Design Matrix

After several iterations of the Technical Requirements Matrix, the group placed the technical requirements from the Requirements Matrix along the rows of the Product Design Matrix and again translated the "whats" which are now requirements into the "hows" or design implementations. Although some of Avidyne's earlier decisions and schedule constraints limited certain design implementations, a deliberate attempt was made to ignore these limitations in order to discover the highest priority implementations. This makes the analysis more valuable for future design work that may not be constrained. Also this allows for the consideration of innovative ways to resolve conflicts and improve designs. The Product Design Matrix is shown in Figure 4.2.

Figure 4.2: Product Design Matrix

Technical Requirements	Weighting	Common Remote Control Panel for PFD & MFD	Dedicated Knobs	Dedicated Buttons	Dedicated Switches	Multi-Function Knobs	"Soft" Keys	Touchscreen	Touchpad	Mouse	Joystick	Trackball	Bezel Functions and Remote Control Panel	Bezel Only Interface	State Indicator Buttons	External Control Display Screen	Label Change to Acknowledge	Color Changes	Descriptive Labels	Sounds to Acknowledge	Detented Controls	Escape Function	Group Like Functions	Pictorial Representations (icons)	Shape Controls Like Functions	Distinctive Control Shapes	Large Controls	Anchor Point for Hand	Open GL Architecture	Size for Instrument Panel	Size for Radio Rack	High Tolerance Reliable / Durable Controls	Mount with Dzus Fasteners	Quick Disconnect Cable	Easily Visible Controls	Divide PFD & MFD Control Panels	Incorporate Redundant Controls on Bezel	Common Material on all Interfaces	Space-age colors	Cutting-edge shaped controls	Contoured Panel	Incorporate Extra Controls	Modular Circuit Board	Durable, Plastic Material	Durable, Composite Material	Durable, Metal Material	Menus	Multi-Windows
Functional Integration	17	9											9																9																			
Dedicated Controls	15		9	9	9								9	1	9				9																	9												
Multi-Function Controls	15	9				9	9	9	9	9	9	9	1	9															9																		9	9
Large # of Controls	7		9	9	9								3																																			
Pointing Devices	8							9	9	9	9	9																	9																			
Auto Context-Dependent Functions	16																												3																			
Voice Activated Functions	15																																															
Distributed Interface	7	9						9	9	9	9	9	9																							9												
Single Interface	15													9																																		
Reduce Functions	15																																															
Visual Feedback	18						3	9	9	9	9	9			9	9	9	9	9					9	9	9			3						3												9	9
Audible Feedback	18																			9									3																			
Tactile Feedback	15		9	3	9	1									3						9				9	9						1																
"Undo" Capability	4																					9																										
GUI Best Practices	13							3	3	3	3	3					9	9		9				9					9																		9	9
Intuitive Controls	10		9	9	9	1	3	9	9	9	9	9			9		3	3	9		9		9	9	9																						9	9
Ergonomic Design	16		9	3	3	9							9		3						9		9			9	9	9		3	1				9													
Minimize # of Parts	15	9				9	9	3	3	3	3	3		9																	3																	
Software Upgradeable	20		3	3	3	9	9	9	9	9	9	9					9	9	9	9				9					9																		9	9
Configure for Convenient Location	11					3	3						3	3																																		
Low # of Inputs / Action	11		9	9	9			1	1	1	1	1	9		9																																	
Redundancy	6												9																							9	9											
High Quality Parts / Components	9																				9											9												9	9	9		
High MTBF	6		9	9	9																											9																
Easy to Install / Replace	7		9		9																									3	9		9	9														
Use COTS Components	4					9		1	1	1	1	1				1											1				3	3	3															
Reconfigurable Hardware	5		9	9	9																																											
Easy to Repair	7					9																										9	9	9														
Modular Design	2		9	1	9	9	9	3					3		1																					9	9											
Use Components Common to all Interfaces	2		9	1	9	9	3	3					9	9																		3				9												
Avidyne Design Language	2		1	1	1			1	1	1	1	1			3		1	1							1	3				3								3	9				3	9	3			
Durable Materials	1																															9								9			9	9	9			
Include "Growth" Controls	1												3																9												9	9						
Adjustable Geometry	7																																										9					
Light Materials	1		1	9	1	1	1																																				9	9				
Total		486	864	591	768	788	592	815	803	803	803	803	723	471	587	166	491	491	567	459	450	36	234	549	389	447	148	144	822	75	136	240	138	126	198	288	72	6	18	9	9	9	87	117	96	81	684	684
Ranking		6	10	7	9	9	7	9	9	9	9	9	8	5	7	2	6	6	7	5	5	0	3	6	4	5	2	2	10	1	2	3	2	1	2	3	1	0	0	0	0	0	1	1	1	1	8	8

Design implementations that addressed specific technical requirements were developed. These implementations were based on the engineering experience of the group as well as information derived from the literature and technology search. Again correlation scores were assigned as to what extent the design implementations satisfied the technical requirements. The top fourteen design implementations are listed in Table 4.3. The top fourteen were selected because implementations four through eight all scored the same and can be conveniently grouped together as pointing devices.

Table 4.3: Priority Design Implementations

Rank	**Design Implementations**	Score
1	Dedicated Knobs	927
2	Dedicated Switches	831
3	Open GL Architecture	822
4-8	Pointing Devices	803
9	Multi-Function Knobs	788
10	Bezel Functions and Remote Panel	744
11	Menus	684
11	Multi-Windows	684
13	Dedicated Buttons	654
14	Soft Keys	592

Just as in the Requirements Matrix, all of the implementations are important, however the highest priority implementations represent those features that best satisfy the most highly weighted technical requirements. These implementations will be key in determining the various alternative concepts. It is also important to note the conflicts inherent in these implementations, primarily dedicated controls (knobs, buttons or switches that always perform the same function) versus multi-function controls (software programmable "soft keys" and multiplex knobs).

5 Hardware Interface Development

The common hardware platform for the PFD and MFD required longer lead times for manufacturing than did the software. In order to meet the production timeline, the hardware configuration was developed first and then passed to an industrial design firm for further development.

5.1 Trade Studies

During the QFD process, all ideas were encouraged in order to gain a broad understanding of the problem and to encourage innovative design implementations. The task of selecting among the various implementations was a very difficult one due to the almost infinite number of variations. A trade study was used during this conceptual design phase to select the best direction for the design.

5.2 Selection Criteria

Prior to developing the alternative concepts, selection criteria were chosen with which to evaluate the various concepts. These criteria closely mirrored the original customer needs although in some cases the needs were broken down into more measurable qualities. Significant constraints were also included. The selection criteria used are listed below:

- Number of Inputs per Action
- Potential for Mode Confusion
- Certification
- Size
- Manufacturing Cost
- Engineering and Development Cost
- Complements Future Upgrades, HITS, FMS
- Reliability
- Maintainability
- Weight
- Schedule

5.3 Assumptions

Prior to developing the alternative concepts, some simplifying assumptions had to be made in order to narrow the scope of the design trade space. Throughout this period of the design process, the possibility of including a remote control panel in the design was very much in question. The concept behind the remote control panel was that it would be an optional piece of equipment that would therefore provide redundant functions but would be designed and located to permit easier access to PFD and/or MFD functions. There were numerous unknowns regarding this panel including whether customers wanted it, if cockpit space was available, and if it would be worth the significant development cost. Due to these significant unknowns, this trade study deferred the decision on the remote control panel and focused on the primary interface attached to the chassis in the instrument panel. In addition, this interface took the form of bezel controls in keeping with Avidyne's design concept for the chassis with quick removable cards located behind the screen with bezel. Also due to schedule constraints and design decisions made prior to this effort, certain high priority technical requirements were not considered such as automatic context-dependent functions and voice-activated functions.

Since the final list of user interface tasks for the primary flight display was still undetermined during the hardware interface development, the tasks from the demonstration configuration were used. These tasks were as follows:

- T-1. Set the baro altimeter setting
- T-2. Sync the baro setting to 29.92
- T-3. Set the altitude (ALT) bug
- T-4. Sync the altitude bug to the current altitude
- T-5. Set the heading (HDG) bug
- T-6. Sync the heading bug to the current heading
- T-7. Set the active NAV display course (CRS)
- T-8. Sync the active NAV display course to the current course
- T-9. Set the standby (Stby) NAV display course
- T-10. Sync the standby NAV display course to the current course
- T-11. Select the source for the active NAV display (4 settings)
- T-12. Select the format for the active NAV display (1 settings)

T-13. Select the source for the standby NAV display (4 settings)

T-14. Select the format for the standby NAV display (4 settings)

T-15. Select the source for the auxiliary NAV display (4 settings)

T-16. Select the format for the auxiliary NAV display (3 Settings)

T-17. Swap the active and standby NAV displays

T-18. Select the declutter setting (4 Settings)

T-19. Select the view setting (2 settings)

T-20. Adjust the map range (9 settings)

T-21. Select GPS hold (prevents the current GPS waypoint from automatically sequencing)

T-22. Select the mode (Normal, Backup or Composite)

T-23. Turn the unit on or off

T-24. Adjust the brightness

Four tasks were not yet implemented in the demonstration configuration but are to be incorporated into the Phase I PFD design. These include:

T-25. Set the vertical speed (VSPD) bug

T-26. Sync the vertical speed bug to the current vertical speed

T-27. Set the airspeed (A/S) bug

T-28. Sync the airspeed bug to the current airspeed

5.4 Alternative Concepts

Based upon the above assumptions, three alternative concepts were developed in addition to the demonstration configuration. These concepts were developed to evaluate the trades between dedicated controls and multi-use controls as highlighted in the Product Matrix. In the near term this hardware platform must support the primary flight display (without HITS) and the multifunction display (without a FMS). During this initial investigation, the concepts were developed primarily with the capabilities of demonstration PFD software configuration in mind. The highly successful Avidyne multifunction display currently has six soft keys (whose functions can be specified in software), three fixed keys (whose functions cannot be modified) and one dual concentric knob (for coarse/fine selection of heading, altitude, etc.). Based on feedback from the MFD group, any configuration that had at least this number of controls would be satisfactory. During development of the concepts, the sketches were kept abstract to represent only the number, operation and the general location of the controls as opposed to the size, shape or exact location. These details would be determined through the industrial design firm's efforts. In addition ideas were not discounted due to perceived problems that could possibly be overcome later, i.e. the perception that dual concentric knobs would not mechanically fit on the bezel.

5.4.1 Demonstration Unit Configuration

The AGATE / NASA demonstration unit utilized hardware designed by Avrotec that consisted of a portrait-oriented bezel with six keys on the left and right sides and three keys centered on the bottom. A dual concentric knob was located in the bottom right corner and a simple rotary knob was located in the bottom left corner. Although the MFD functions have never been coupled with this hardware, a likely mapping of the MFD features to this configuration was determined. A schematic of this configuration is shown in Figure 5.1 although it is drawn in a landscape orientation for consistency with the other concepts. Both sketches are displayed to give a better indication of how these units will be oriented in the cockpit.

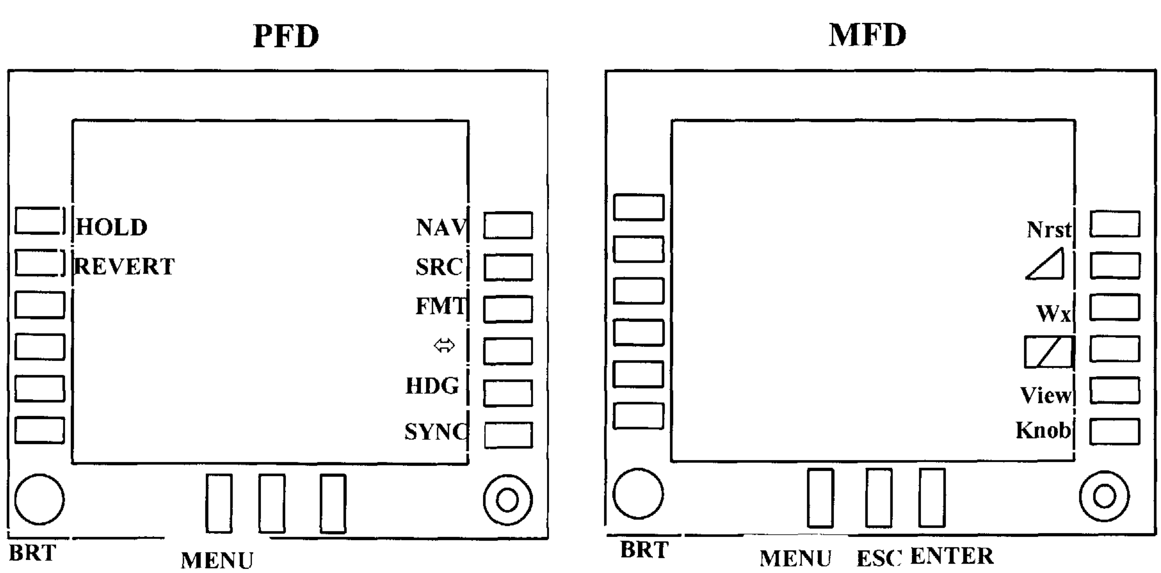


Figure 5.1: Demonstration Unit Hardware Configuration

The demonstration unit configuration was included to provide a known baseline for comparison of the other concepts. Each of the other concepts was generated with the intent of improving upon the demonstration unit. A new configuration was desirable not only to improve the usability of the products but also to create an Avidyne hardware platform that reflected a distinctive Avidyne "design language."

5.4.2 Single Dual Concentric Knob Configuration

The single knob configuration was very similar to that of the AGATE / NASA demonstration unit. Controls were laid out similarly except that six instead of three buttons were located across the bottom. The concept behind this configuration is that six parameters would best be adjusted using the single dual concentric knob. These six parameters include the heading, airspeed, vertical speed and altitude bugs as well as the barometer setting and course. Due to the expected frequency of their use, they are located next to the most easily accessible right hand buttons. (Many of Avidyne's customers' aircraft have left side stick controllers in the pilot-in-command's seat making the right side buttons more easily accessible.) One of the three map adjustment features, range, also benefits from the use of a dual concentric knob due to the large number of settings (nine) available. For this reason the three map features are located on the bottom right buttons with the range closest to the knob. The seven buttons mentioned corresponding to the knob are used to set the mode for the use of the knob. Six buttons on the left side are used for the remaining six navigation display parameters. Three of the buttons along the bottom are available for growth or additional flexibility. This configuration is shown in Figure 5.2.

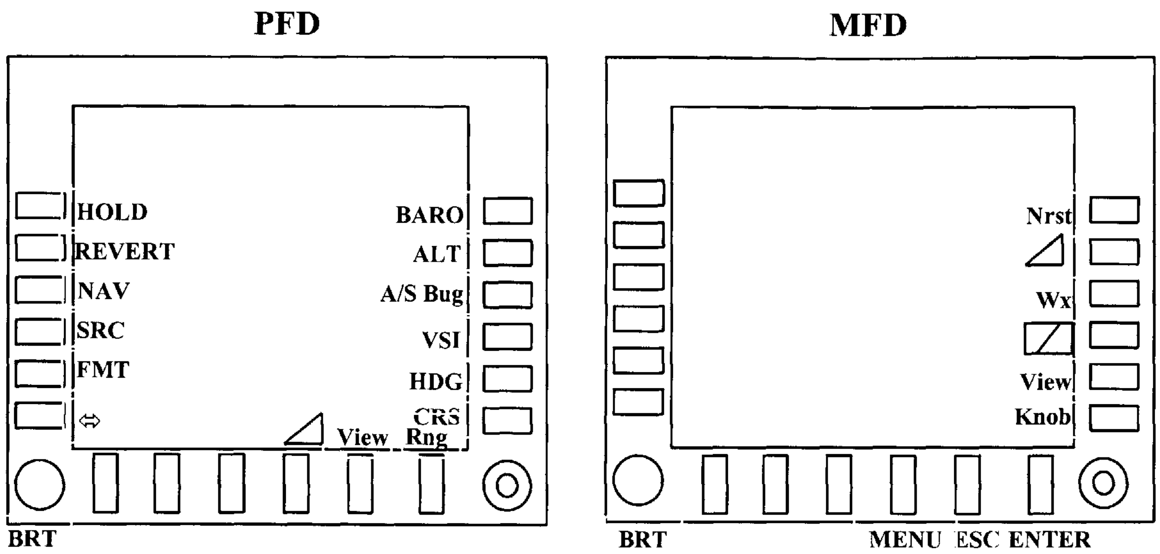


Figure 5.2: Single Dual Concentric Knob Configuration

There is more visibility of the available functions and this configuration removes the need to toggle through multiple options. Navigation display interaction remains the same as the demonstration unit with only the location changing. Map adjustment features have been brought to the top level to provide continuous visibility and to reduce the number of inputs per action. A comparison of the number of inputs per action for each of the new configurations as well as the demonstration unit can be found in Appendix A.

5.4.3 Double Dual Concentric Knobs Configuration

This configuration adds a second dual concentric knob on the bottom left corner of the bezel. With the addition of the second knob, functions that were previously assigned to one knob can be split between the two. This allows functions to be more spatially associated and allows two different knob functions that are frequently used to be continuously available without constantly changing the mode of a single knob. For example, if one were receiving frequent heading and altitude changes from air traffic control, the left knob could be set to heading and the right knob could be set to altitude. The six modes of the knob are displayed across the bottom with the three on the left associated with the left knob and the three on the right corresponding to the right knob. Map functions are located on the left side where the range function can be spatially associated with the left knob. Navigation display functions are located on the top right side buttons as in the original configuration as shown in Figure 5.3.

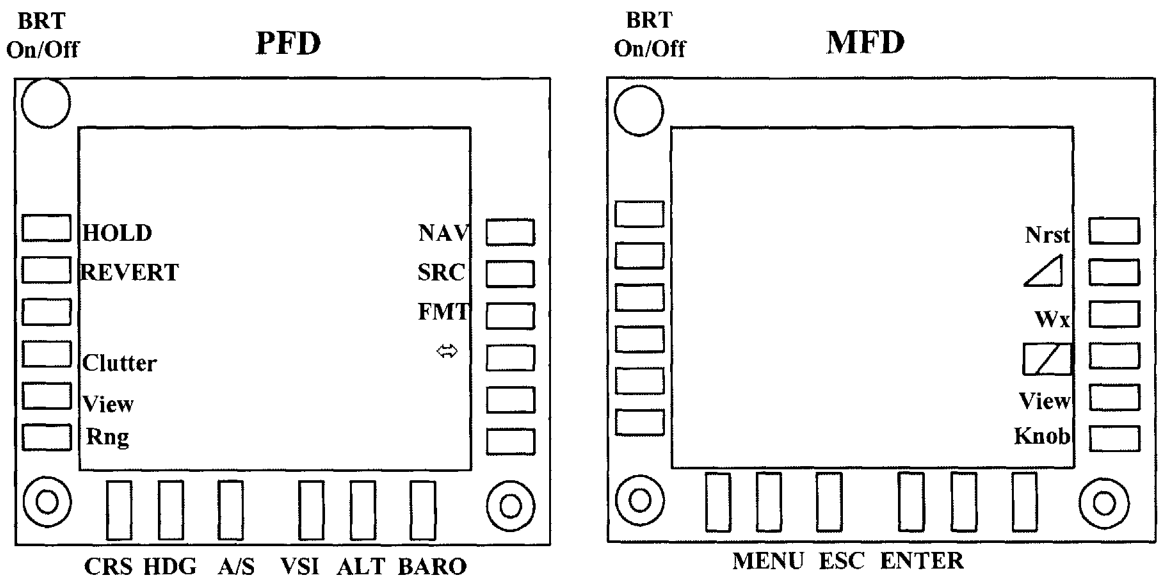


Figure 5.3: Double Dual Concentric Knobs Configuration

Like the single knob configuration, there is more visibility of the available functions and there is less toggling through multiple options. Having a second knob halves the number of modes per knob, however it causes the users to not only have to decide on the mode of the knob but also which knob is appropriate. The number of inputs per action is identical to the one knob configuration.

5.4.4 QUAD DUAL CONCENTRIC KNOBS CONFIGURATION

In the trade between dedicated and multi-function controls, one extreme is to have dedicated controls for each function. This is the idea behind quad dual concentric knobs. The original configuration has four knob modes. Although attaching four knobs is thought to be difficult mechanically, this idea was considered so as not to unnecessarily constrain the solution set based on perceived difficulty. Although the original configuration has four knob modes, an analysis of requirements resulted in seven functions that could best be accomplished through the use of a knob. In order to accommodate these seven functions, modes had to be established for three of the four knobs. Each knob was associated with two buttons that selected the knob mode. The range function was left as a dedicated knob and the other map functions of view and map (or declutter) were grouped near the range knob so that all of the map features were grouped together. The arrangement is illustrated in Figure 5.4.

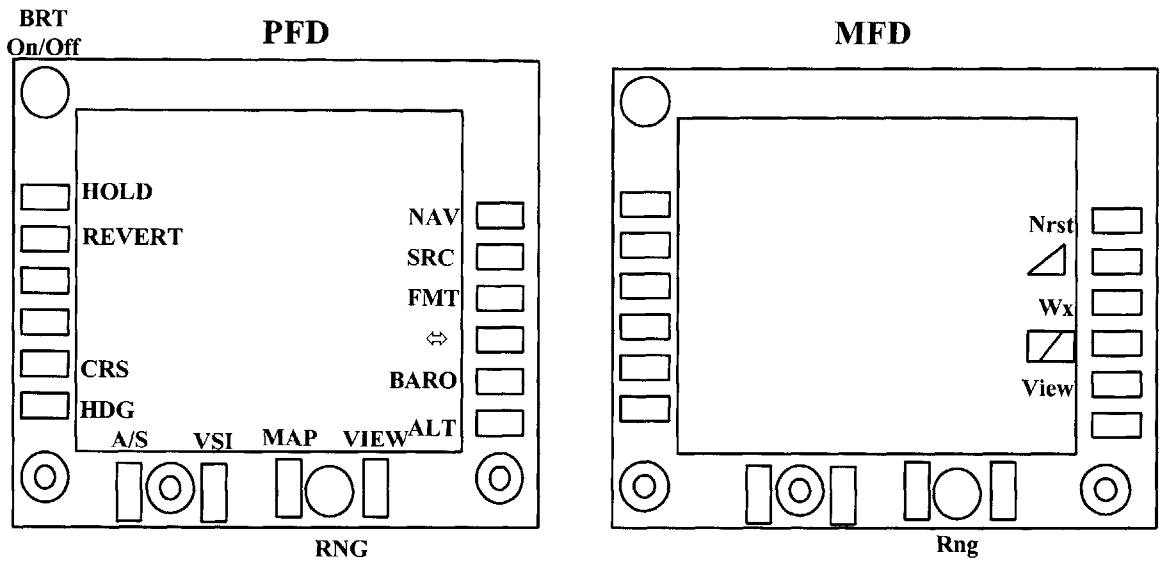


Figure 5.4: Quad Dual Concentric Knobs Configuration

Four dual concentric knobs allow for more visibility of available functions and the need for less toggling. The need for added functionality requires the continued modal use of the knobs. With the introduction of future software upgrades, HITS and a FMS, this trend is likely to continue necessitating the modal use of knobs regardless of the number of knobs used. There is only a very slight improvement in the number of inputs per action over the single and dual knob configurations.

5.5 Pugh Matrix/Selection

Once the four configurations were defined, they were compared based on the identified selection criteria. Due to the preliminary nature of the concepts and the inability to accurately quantify specific attributes, a Pugh Matrix was used to select the preferred concept. The demonstration unit was used as a baseline for comparison since it was in prototype form and a known quantity while the other configurations were simply paper concepts. Each of the three new concepts was judged against the baseline as being the same (S), better (+) or worse (-) as shown in Figure 5.5.

Configuration / Criteria	Demonstration Unit	Single Knob	Double Knobs	Quad Knobs
Fewer Inputs/Action		+	+	+
Less Potential for Mode Confusion		+	+	+
Certification	Baseline	S	S	S
Manufacturing Cost		S	-	-
Develop Cost		-	-	-
Weight		S	-	-
Size		S	S	-
Schedule		S	S	-
Complements Future Upgrades		+	+	-
Reliability		S	-	-
Maintainability		S	-	-

S ~ Same
+ ~ Better
- ~ Worse

Figure 5.5: Pugh Matrix Comparison

Based on this analysis, the single dual concentric knob configuration was judged to be the preferred configuration. Significant differences are discussed below. The impact of adding more knobs to the hardware display was discussed with the mechanical engineers in the hardware group. In general the more knobs that are added to the display the more the cost of manufacturing and the higher the weight. Also knobs were judged to be less reliable than buttons and to increase maintenance requirements. Incorporation of more than two knobs on the bezel would require an increase in the bezel size and would likely lead to a slip in the current schedule due to mechanical engineering challenges. The quad knob configuration was also judged to be less complementary to upgrades since there is a decrease in the total number of buttons and the buttons are necessary for delineating the modes of the knobs.

5.6 CONCEPT REFINEMENT

After selection of the preferred concept, the single dual concentric knob configuration was presented to the various Avidyne groups for feedback. There was concern that this concept would not have the necessary growth potential to accommodate future upgrades, particularly the addition of the flight management system, and that it was optimized around the needs of the near term primary flight display. A second design iteration was conducted to address these concerns and ensure that the design would meet future growth requirements. For this design iteration three concepts were compared. Three dedicated functions were added to each concept. These included an on/off button, a mode switch for selecting between normal and back up mode and a dedicated brightness control.

5.6.1 BASELINE CONFIGURATION

The first concept was the preferred configuration from the first design iteration. For this design iteration it was established as the baseline. This concept is shown in a portrait orientation in Figure 5.6.

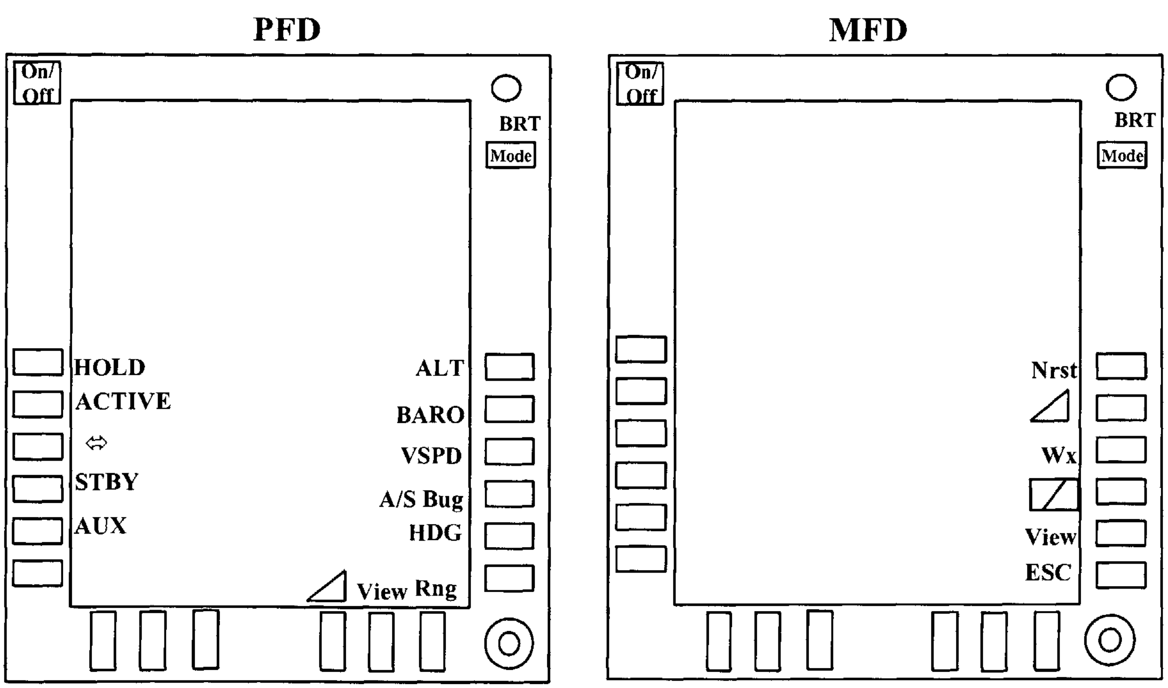


Figure 5.6: Baseline Configuration

An advantage of the baseline configuration is that it minimizes the incorporation of "extra" unused buttons, minimizing cost and a cluttered appearance. A disadvantage is that all controls are on the lower half of the display. This could possibly limit the capability of future upgrades by not having any buttons on the top half of the display that can be spatially associated with functions on the top half of the screen. Also the labeling of functions is limited to the space adjacent to the corresponding buttons.

5.6.2 Growth Configuration

The growth configuration attempted to overcome some of the shortcomings of the baseline configuration by adding four buttons to the upper half of the bezel. Also a second dual concentric knob was added in the lower left corner since it offered more flexibility for a slight penalty in cost, weight, reliability and maintainability. This configuration is shown in Figure 5.7.

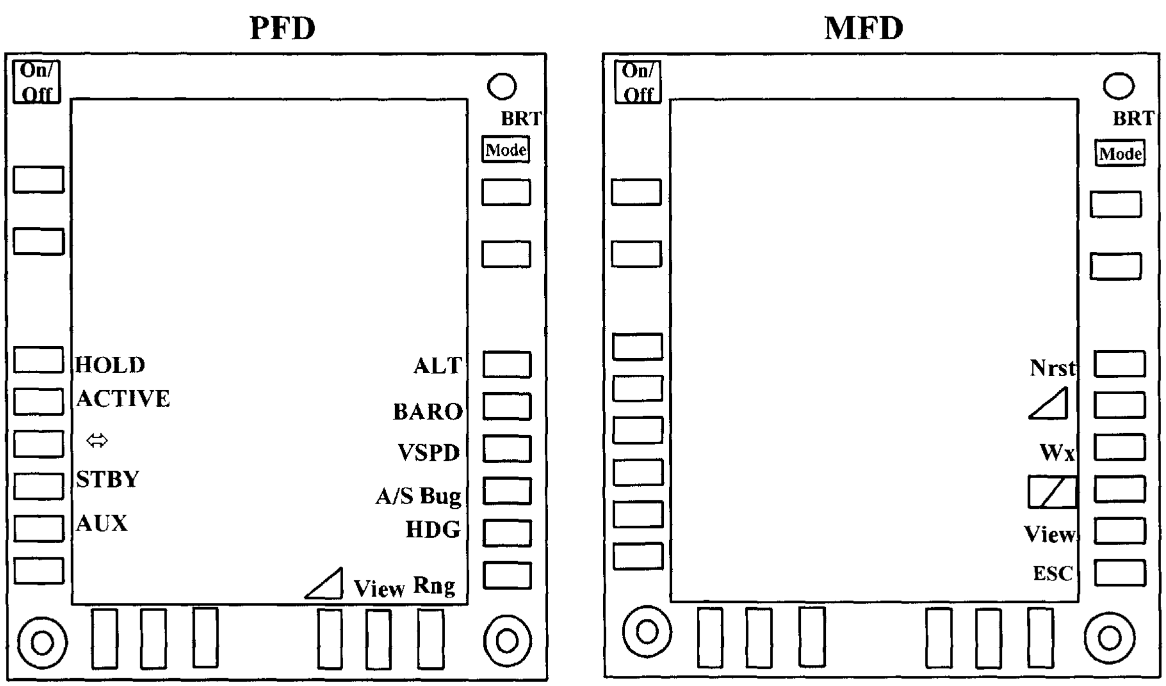


Figure 5.7: Growth Configuration

The obvious advantage of this configuration is that it allows for spatially associated buttons on the top half of the display for the MFD and future upgrade. Also the additional knob could be very useful for as yet undetermined functions. Disadvantages include the cost of having more controls, the inclusion of many unused buttons in some applications such as the Phase I PFD. This display could be potentially confusing due to the sheer number of buttons and its cluttered appearance. Labels are also limited in that they must correspond to button locations although this factor is not as limiting in this configuration as it is in the baseline.

5.6.3 Clean Configuration

A somewhat radical, new concept was considered as the third alternative. This is known as the clean configuration. This concept minimizes the number of controls while maximizing the flexibility of software. The idea is that all of the functions can be labeled on the screen using software labels and a simple rotary knob is used to highlight the desired function. A dual concentric knob is then used to adjust the desired function. An escape button is included to aid in navigation of any sub menus or to correct errors. This concept is illustrated in Figure 5.8.

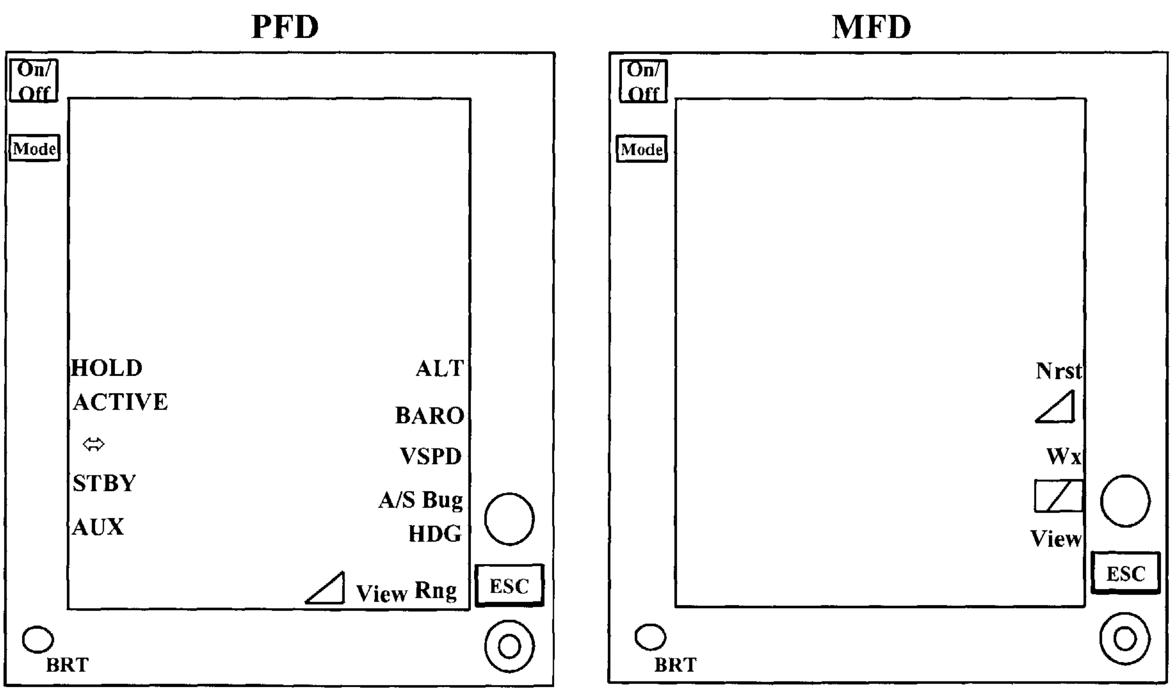


Figure 5.8: Clean Configuration

This configuration presents an extremely simple, clean appearance. There is maximum flexibility for incorporating future functionality and spatially associating labels with functions. This concept has a low cost compared to the other concepts and maps well to a possible simple remote control panel that could have the same rotary knob, escape button and dual concentric knob located for more convenient use. This concept will likely not be as easy to use as conventional controls. There are risks associated with this concept in that it is much different than current competitor products and it is not what consumers are familiar with. Also there could be certification issues as well as issues with adapting the software from its current configuration to this different concept.

5.7 PUGH MATRIX/SELECTION

The three new concepts were evaluated based upon the original selection criteria. One new selection criterion of "Proximity of Controls" was added to account for the desire to have controls that can be spatially associated with all areas of the display. Again a Pugh Matrix was used and can be seen here as Figure 5.9.

Criteria / Configuration	Baseline	Growth	Clean
Proximity of Controls		+	+
# Inputs / Action		+	-
Potential for Mode Confusion		S	-
Certification		S	-
Manufacturing Cost	Baseline	-	+
Engr & Develop Cost		-	+
Weight		S	+
Size		S	+
Schedule		S	-
Complements Future Upgrades, HITS, FMS		+	+
Reliability		-	+
Maintainability		-	+

S ~ Same
+ ~ Better
- ~ Worse

Figure 5.9: Second Iteration Pugh Matrix Comparison

Based upon the Pugh Matrix, the clean configuration is the preferred concept. The potential flexibility, simplicity and cost savings of the clean configuration are substantial, however the negatives of the clean configuration were unacceptable to Avidyne. In particular, certification of the clean configuration could potentially be very difficult and jeopardize the schedule and the program. To reduce the risks associated with potential certification and thus schedule problems, a compromise between the baseline and growth configurations was developed as a final concept.

5.8 Final Concept

The final selected concept was an adaptation of the growth configuration. Due to space constraints in the landscape configuration, only seven buttons were placed on each side in addition to the on/off, mode and brightness controls. Two dual concentric knobs were included, however the bottom row of buttons was reduced from six to five. Although groupings of the buttons and association with the various knobs were considered, the buttons were evenly spaced and not physically associated with the knobs since the future groupings and associations were unknown. A schematic of the final concept is shown in Figure 5.10.

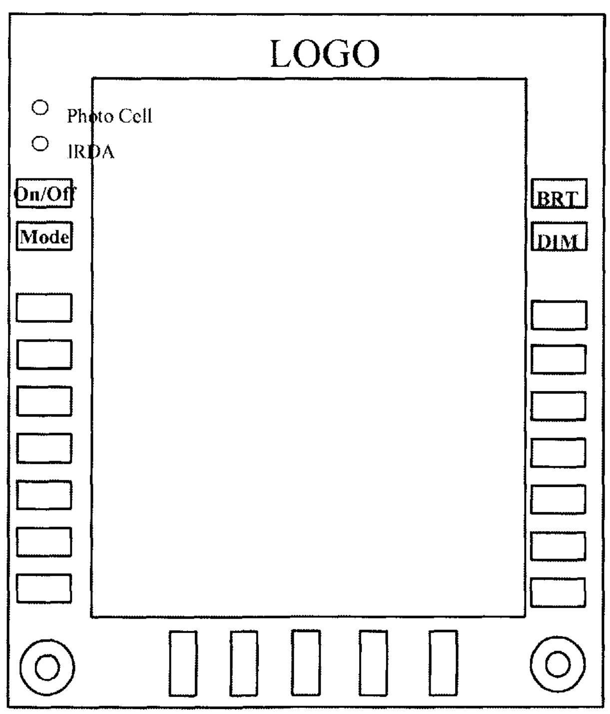


Figure 5.10: Schematic of Final Hardware Concept

This schematic was passed to an industrial design firm to further develop the desired Avidyne "design language" and determine the exact location, size, shape, color, tactile feel, etc of the bezel and controls. A rendering of the results of this effort is shown in Figure 5.11.

Figure 5.11: Rendering of Final Hardware Concept

6 SOFTWARE INTERFACE DEVELOPMENT

A challenge of the software interface development is to incorporate the best possible implementations in the interface within short schedule constraints. The difficulty is in determining the best features without multiple, time-consuming human evaluations. One tool that is being adapted to do this is a variation of the Goals, Operators, Methods and Selection Rules (GOMS) Keystroke-Level Model (KLM) [2]. This model can be used to predict the relative execution times of various implementations with first order accuracy. These execution times, with a careful consideration of potential errors, can provide useful information that leads to the best implementations. In addition, a measure of control input efficiency was developed to provide a further method of analysis. Once a preferred concept was determined, a human subject experiment was conducted to validate the GOMS results and provide feedback on the success of the effort.

6.1 TASK IDENTIFICATION

Software interface development of the Phase I Primary Flight Display began with identification of the tasks to be performed by the PFD. Based upon Avidyne system requirements documents for the PFD, a list of tasks were identified. The tasks are as follows:

- T-1. Set the baro altimeter setting
- T-2. Sync the baro setting to 29.92
- T-3. Set the altitude (ALT) bug
- T-4. Sync the altitude bug to the current altitude
- T-5. Set the heading (HDG) bug
- T-6. Sync the heading bug to the current heading
- T-7. Set the active NAV display course (CRS)
- T-8. Sync the active NAV display course to the current course
- T-9. Set the standby (Stby) NAV display course
- T-10. Sync the standby NAV display course to the current course
- T-11. Select the source for the active NAV display (4 settings)
- T-12. Select the format for the active NAV display (1 settings)
- T-13. Select the source for the standby NAV display (4 settings)

T-14. Select the format for the standby NAV display (4 settings)

T-15. Select the source for the auxiliary NAV display (4 settings)

T-16. Select the format for the auxiliary NAV display (3 Settings)

T-17. Swap the active and standby NAV displays

T-18. Select the declutter setting (4 Settings)

T-19. Select the view setting (2 settings)

T-20. Adjust the map range (9 settings)

T-21. Select GPS hold setting

T-22. Select the mode (Normal, Backup or Composite)

T-23. Turn the unit on or off

T-24. Adjust the brightness

Four tasks are not currently implemented in the demonstration configuration but will be incorporated into the Phase I PFD design. They include:

T-25. Set the vertical speed (VSPD) bug

T-26. Sync the vertical speed bug to the current vertical speed

T-27. Set the airspeed (A/S) bug

T-28. Sync the airspeed bug to the current airspeed

6.2 Demonstration Unit User Interface

In order to provide a baseline for analysis and explain the various tasks performed by the user interface of the primary flight display, the demonstration unit portrait configuration is described. This demonstration configuration has been flight evaluated and the feedback from the flight evaluation will be incorporated into the new design. The top menu level of the demonstration unit configuration is shown in Figure 6.1.

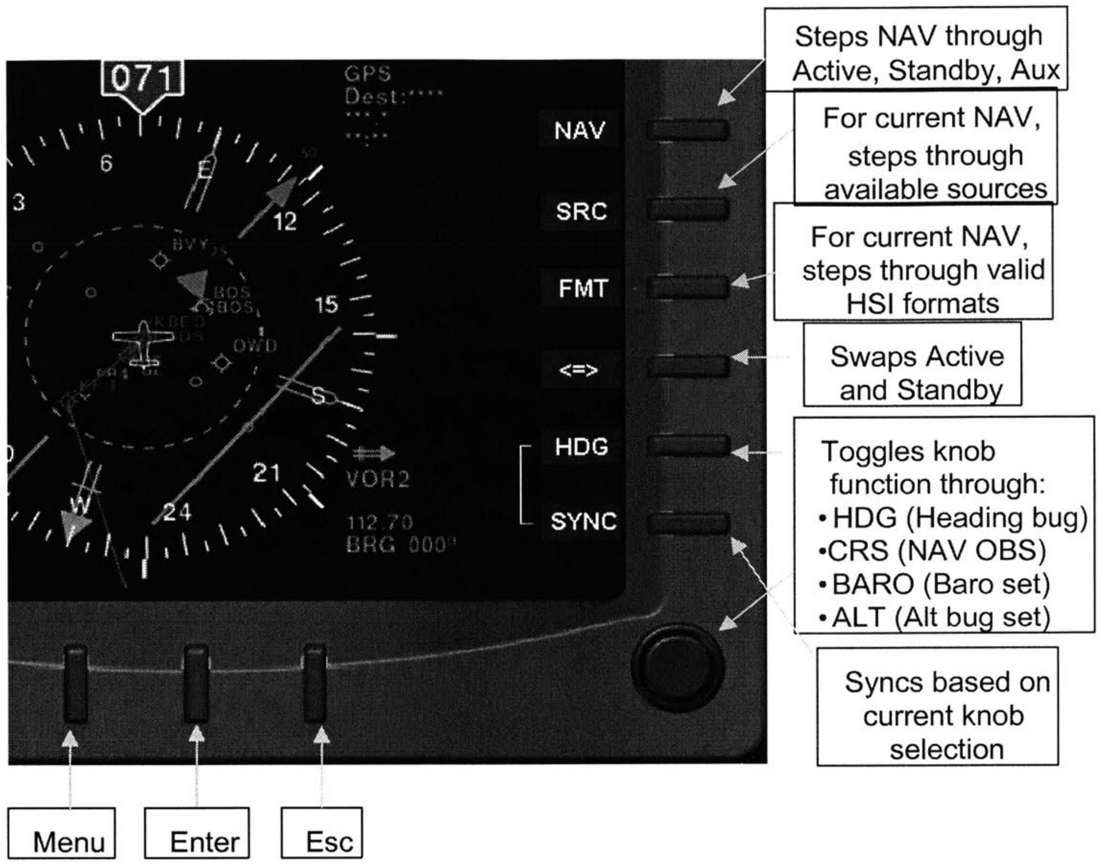


Figure 6.1: Top-Level Menu of the Demonstration Configuration

Several of the salient features of this menu level are discussed below.

- Navigation display adjustment – In order to select the desired navigation display for inputs, the NAV button is used to toggle through the four choices of active, standby, auxiliary and none. Once the desired NAV display is selected, the source and format buttons may be used to modify the selected NAV display. The source and format buttons each toggle through from one to four choices as shown in the hierarchical task analysis in Figure 6.2. If no adjustments are made within five seconds, the display automatically defaults to no NAV display selected. If the Source, Format or Course features are used with no NAV display selected, the interface will default automatically to adjusting the active NAV display.
- Heading, Course, Baro, and Altitude settings – These parameters may be set by first selecting the proper mode using the button that corresponds to the HDG label in Figure 6.1. This button toggles through these four choices and the dual concentric knob is used

to set the desired value. In order to adjust the course of the standby or AUX NAV display, the standby or AUX NAV as well as course must be selected.

- Sync feature – Sync is a dedicated button that affects the setting of the parameter displayed immediately above the sync label on the display. For example, when HDG is displayed, the sync button will cause the heading bug to be set to the current aircraft heading. Synchronizing while baro is displayed will cause the baro setting to change to 29.92.

A hierarchical task analysis of the demonstration configuration is shown in Figure 6.2.

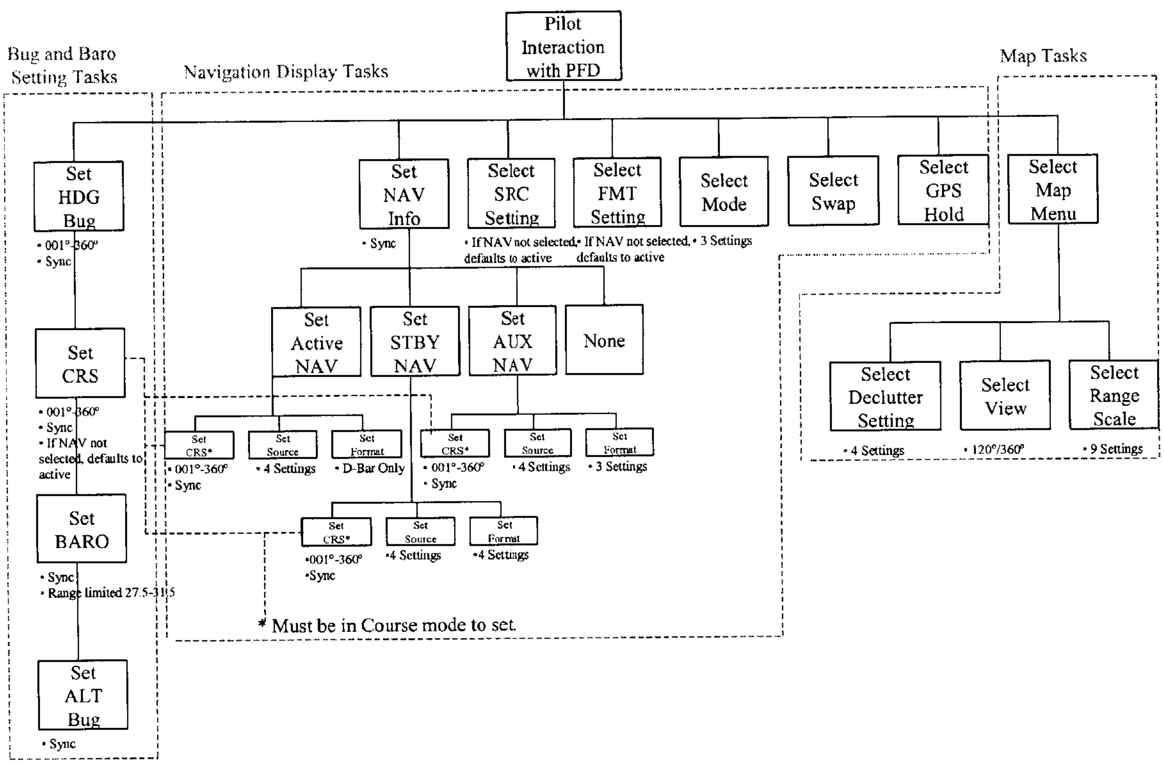


Figure 6.2: Demonstration Configuration Hierarchical Task Analysis

The map sub-menu is reached by pressing the “menu” button at the bottom of the bezel. The map sub-menu is shown in Figure 6.3.

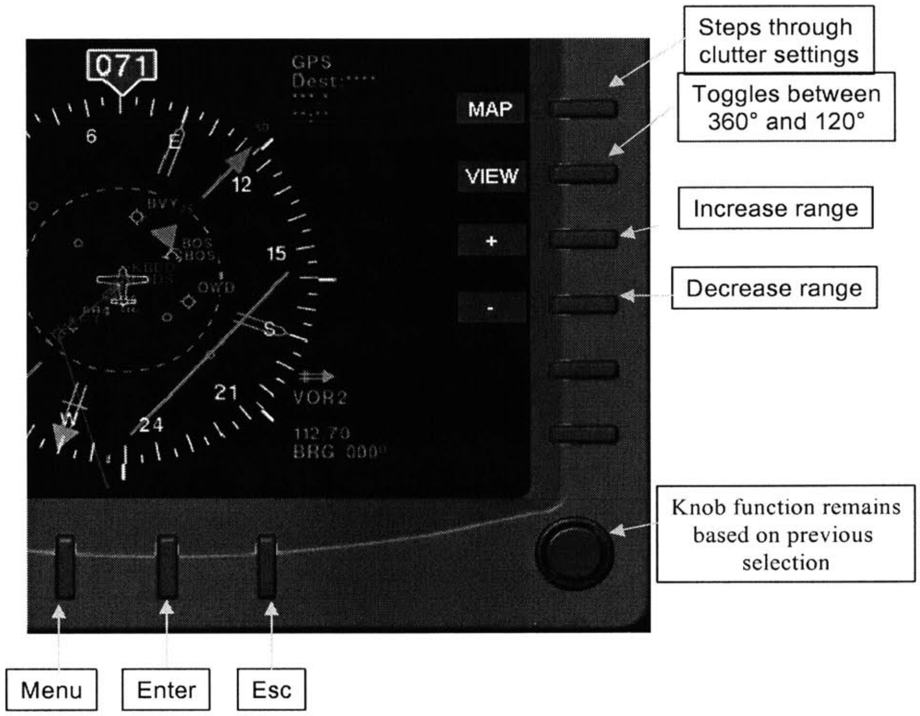


Figure 6.3: Map Sub-Menu of the Demonstration Configuration

The map sub-menu allows the user to select between four declutter settings, change the HSI from a 360° compass rose to a 120° arc and increase or decrease the range. Selecting the menu button again while on the map sub-menu returns the user to the top-level menu.

6.3 Flight Test Feedback of the Demonstration Unit

Avidyne conducted flight tests of the demonstration configuration primary flight display using several pilots of varying experience. Below is a list of some of the comments pertaining to difficulties in using the demonstration unit. Solutions to these difficulties will be implemented in the proposed user interface.

- Avoid toggling between numerous choices with a single button, i.e. HDG, CRS, Baro and ALT features on a single button leads to mode confusion and high workloads. Also overstepping of settings when toggling through choices is common.

- The PFD adds to the typical setup time for ILS approaches. In some cases excessive times setting up for ILS approaches were noted.
- Use of the buttons on right side of the display is adequate during turbulence, however the bottom buttons are more difficult to use.
- Current NAV display setup is workable with sufficient training. Increased training requirements should not be required to overcome a poor interface.
- It is more difficult to go to a separate map menu.
- Dedicated map controls are desirable.
- The +/- buttons associated with map range are confusing.
- The requirement for course to be selected when adjusting the standby and AUX displays is not obvious and leads to mode confusion.

6.4 Identifying Functional Groups of Tasks

Rather than analyzing each task individually, it is useful to identify functional groups of tasks that can be logically grouped together. Functional groups of tasks make it easier for the user to find a particular task to be performed. It is also helpful if like tasks within a functional group can be accomplished in a similar manner in order to have user interface consistency. Based on this desire, the following functional groups were identified:

<u>Map Adjustment Tasks</u>

- Adjust the range (RNG)
- Select the declutter setting
- Select the view setting

<u>Navigation Display Tasks</u>

- Set the active NAV display course (CRS)
- Sync the active CRS to the current heading
- Set the standby (Stby) NAV display course
- Sync the standby CRS to the current heading
- Select the source for the active NAV display

- Select the format for the active NAV display
- Select the source for the standby NAV display
- Select the format for the standby NAV display
- Select the source for the auxiliary NAV display
- Select the format for the auxiliary NAV display
- Swap the active and standby NAV displays
- Select GPS hold

Bug and Baro Setting Tasks

- Set the baro altimeter setting
- Sync the baro setting to 29.92
- Set the altitude (ALT) bug
- Sync the ALT bug to the current ALT
- Set the heading (HDG) bug
- Sync the HDG bug to the current HDG
- Set the vertical speed (VSPD) bug
- Sync the VSPD bug to the current VSPD
- Set the airspeed (A/S) bug
- Sync the A/S bug to the current A/S

In accordance with the system requirements, the following tasks will have dedicated controls and so will not be placed into functional groups:

- Select the mode (Normal, Backup or Composite)
- Turn the unit on or off
- Adjust the brightness

6.5 Goals, Operators, Methods and Selection Rules (GOMS) Keystroke-Level Model (KLM)

The GOMS Keystroke-Level Model provides a means of analyzing the relative time it takes to perform a task [2]. This quick and simple analysis tool does not purport to determine the absolute task performance time, however applied uniformly to various interfaces it can determine the relative task performance time. These relative times can also be compared to the ideal case that is defined as the time the task would take if there were a control dedicated to that particular function only. It is unrealistic to dedicate a unique control to every task due to space constraints, however this technique provides a goal by which to measure the optimization of the implementation. Also in this ideal case the number of control inputs required can be determined. Again this ideal case can be compared to the number of control inputs required for a particular task implementation. From this information a measure of control input efficiency can be determined. Control input efficiency is defined as the minimum number of control inputs required for a task given that there is a dedicated control for that task divided by the number of control inputs required for a particular interface implementation. This provides a quantitative measure by which various interface implementations can be compared for speed and efficiency. Unfortunately this analysis will not provide a means to estimate the error rate associated with a given interface nor the degree of visibility or feedback of a particular implementation. However, it will provide a rank order of implementations that can be evaluated more thoroughly [17].

When the GOMS model was developed, it was observed that the time it takes the user-computer system to perform a task is the sum of the times it takes to perform the serial elementary operations that make up the task. Although there is wide variability in these times for each user, typical times can be used to make a *comparative* analysis of tasks involved in using a keyboard and graphical input device (GID). Typical times for different operations are given in Table 6.1 based on the research of Card, Moran and Newell 1983 [2]. The original nomenclature is used where each of the times is designated by a one-letter mnemonic.

Table 6.1 Typical Times for Different Operations

K = 0.2 sec	Keying: The time it takes to tap a key on the keyboard
P = 1.1 sec	Pointing: The time it takes a user to point to a position on a display
H = 0.4 sec	Homing: The time it takes a user's hand to move from the keyboard to the GID or from the GID to the keyboard
M = 1.35 sec	Mentally Preparing: The time it takes a user to prepare mentally for the next step
R	Responding: The time a user must wait for a computer to respond to input

These numbers vary widely and this simple model does not reflect the absolute times required for performing a task, however the correct ranking of the performance times of two interfaces is usually obtained. The operations described above are not exactly the same as that of using the bezel controls of the primary flight display, but they are analogous. For this analysis, keying will be used to represent the pressing of buttons on the bezel and homing will represent the time it takes a user's hand to move to a button or knob. One operation for which there is no time data is the turning of a knob to select the desired value. Since this value is likely to vary widely for each application, the variable "T" will represent the time it takes to select the desired value. The dual concentric knobs on the bezel could conceivably be used in two different ways. One is when both knobs are used to adjust the same value with the outer knob being used for coarse adjustments and the inner knob for fine adjustments. The other is for a single knob (outer or inner) to set a value. Since the adjustment times for these two cases could be different, they will be represented by T_{DC} and T_S respectively. The adaptation of the times for different operations is shown in Table 6.2.

Table 6.2 Adapted Typical Times for Different Operations

K = 0.5 sec	Keying: The time it takes to press a button on the bezel. (This time was selected based upon the time necessary to type random letters as determined by Card and Moran [2].)
H = 0.4 sec	Homing: The time it takes a user's hand to move to the bezel.
M = 1.35 sec	Mentally Preparing: The time it takes a user to prepare mentally for the next step
$T = T_{DC}$ or T_S	Turning: The time it takes to adjust a dual concentric knob or single knob (outer or inner) to the desired value
R	Responding: The time a user must wait for a computer to respond to input

To calculate the relative time to perform a task, a list of operations necessary to perform the task must be made from the GOMS list of operations (K, H and T). This analysis assumes that the user's hand does not start on the bezel so each task will begin with an H operation. Once the user has homed to the bezel, any further movement between controls is accounted for in the keying, turning or mental preparation time. Next it must be determined at what points the operator will stop and perform mental operations (M). The basic rules developed by Card, Moran and Newell 1983, p. 265 [2] for deciding where mental operations will occur are shown in Table 6.3. In these rules, a string is a sequence of characters. A delimiter is a character that marks the beginning or the end of a meaningful string of text. An argument is the information that must be supplied to a command. In this particular application, strings of text are not entered and most keystrokes execute pre-defined commands so that most of the time, many of these rules will not be necessary.

Table 6.3 Heuristics for Placing Mental Operators

Rule 0	**Initial insertion of candidate Ms**
	Insert Ms in front of all Ks (keystrokes). A modification to this rule for this analysis is to place Ms in front of all Ts (turning of knobs).

Rule 1	**Deletion of anticipated Ms**
	If an operator following an M is fully anticipated in an operator just previous to that M, then delete that M. For example, if you move the GID with the intent of tapping the GID button when you reach the target of your GID move, then you delete, by this rule, the M you inserted as a consequence of rule 0. In this case, P M K becomes P K.

Rule 2	**Deletion of Ms within cognitive units**
	If a string of M Ks belongs to a cognitive unit, then delete all the Ms but the first. A cognitive unit is a contiguous sequence of typed characters that form a command name or that is required as an argument to a command. In this application, multiple presses of the same key will be treated as a cognitive unit.

Rule 3	**Deletion of Ms before consecutive terminators**
	If a K is redundant delimiter at the end of a cognitive unit, such as the delimiter of a command immediately following the delimiter of its argument, then delete the M in front of it.

Rule 4	**Deletion of Ms that are terminators of commands**
	If a K is a delimiter that follows a constant string - for example, a command name or any typed entity that is the same every time that you use it – then delete the M in front of it. (Adding the delimiter will have become habitual, thus the delimiter will have become part of the string and not require a separate M.) But if the K is a delimiter for an argument string or any string that can vary, then keep the M in front of it.

Rule 5	**Deletion of overlapped Ms**
	Do not count any portion of an M that overlaps an R – a delay, with the user waiting for a response from the computer.

6.6 MAP ADJUSTMENT TASKS

6.6.1 IMPLEMENTATIONS

Three alternative methods (or implementations) of performing the map adjustment tasks were conceived taking into consideration the feedback from the flight evaluations of the demonstration configuration. Since having a map sub-menu was more difficult to use and sufficient controls are available, the requirement to first select the map menu was eliminated. This also eliminated the need for a dedicated menu button. The three implementations of the map adjustment tasks were as follows:

Map Adjustment Tasks First Implementation (RNG Defaults to "Nothing Selected")

The map adjustment tasks first and second implementations are illustrated in Figure 6.4. Tasks T-18 (declutter) and T-19 (view) were assigned to line select keys. These keys toggle through four and two choices respectively, i.e. 120° and 360° for the view task. Note four options were already identified as difficult to manage when toggling in this manner.

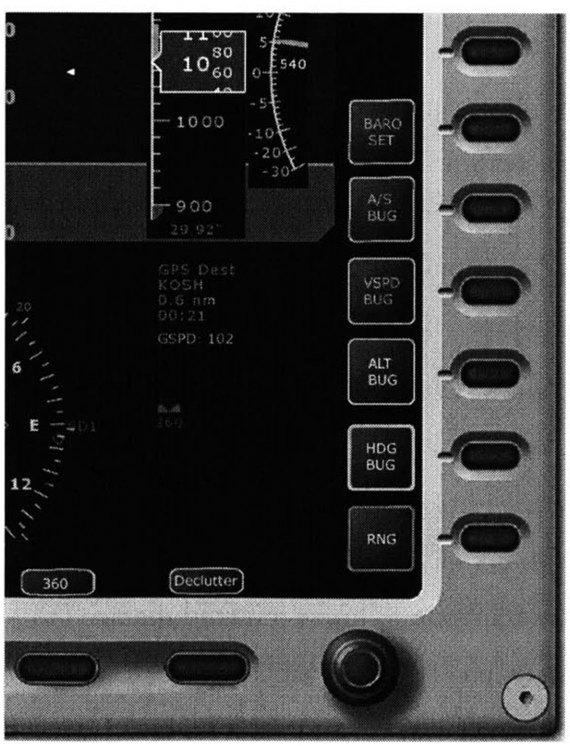


Figure 6.4: Map Adjustment Tasks First and Second Implementations

Task T-20 (range) had nine possible settings (1, 2, 5, 10, 25, 50, 100, 250, 500 nm). Due to the large number of choices it was preferable to adjust this using a knob. RNG was associated with a line select key on the right side that sets the function of the right knob to adjust the HSI range. There was no need for coarse and fine adjustment for the RNG function so both the inner and outer knob adjusted the RNG setting. RNG was one of six functions available that sets the mode of the right knob. (The other functions were BARO, A/S, VSPD, ALT and HDG). When RNG was selected by pressing the associated line select key, it was highlighted or otherwise indicated to show the current function of the knob. This selection of RNG "timed out" after five seconds without control inputs. The interface defaulted to no functions being selected and the knob performed no actions until one of the six available functions was again selected. Reasons this defaults to "nothing selected" were as follows:

- Use of the right knob was subject to mode errors in that the knob executed one of several different possible responses depending upon the state.
- Requiring the user to always select the desired mode immediately prior to the use of the knob placed the mode in the user's short-term memory reducing the probability of mode errors [17].
- This configuration takes advantage of the human trait of habit development allowing the user to always perform the same action (select a mode prior to turning the knob) [17].
- This allows the user to develop habits that smooth the flow of work. The user always performs the same action rather than having to sometimes press a key before turning the knob and sometimes not pressing a key before turning the knob to accomplish the same task [14].

<u>Map Adjustment Tasks Second Implementation (RNG Remains Selected)</u>

Tasks T-18 (declutter) and T-19 (view) were accomplished using line select keys the same as the first implementation.

Task T-20 (range) was one of six menu choices associated with line select keys on the right side of the display similar to the first implementation. The difference is that once a function was selected it remained active until a different function was selected.

<u>Map Adjustment Tasks Third Implementation (Dedicated to Left Knob)</u>

The map adjustment tasks third implementation is shown in Figure 6.5. Task T-18 (declutter) was dedicated to the left, outer, dual concentric knob. This eliminated the need to toggle through four options using a button (a problem noted earlier). Task T-19 (view) was dedicated to the push button of the left dual concentric knob. When pressed it toggled between 120° and 360°.

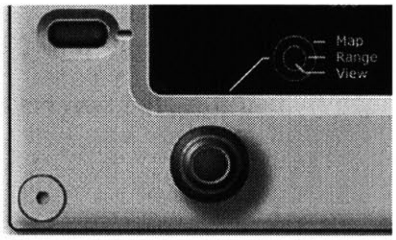


Figure 6.5: Map Adjustment Tasks Third Implementation

Task T-20 (range) was dedicated to the left, inner, dual concentric knob. This is similar to the current Avidyne FlightMax products that all have the map range function dedicated to the inner dual concentric knob.

This implementation had the advantage of all of the available map features being functionally grouped together on the left, dual concentric knob. Additionally these same functions were similar on the multi-function display (MFD) and could be assigned to the MFD left, dual concentric knob creating consistency between the displays and unloading the burden on the current MFD menu structure.

6.6.2 Example GOMS Analysis of Task T-20 (Range)

This analysis is described below in detail in order to illustrate the application of the GOMS Keystroke-Level Model.

<u>Analysis of Task T-20 (RNG) First Implementation (RNG Defaults to "Nothing Selected")</u>

In order to adjust the range the user must first move their hand to the RNG line select key, press the key, move their hand to the right dual concentric knob and turn to the desired setting. This series of actions appears as follows using the GOMS model:

$$H \; K \; T_S$$

Next the heuristics in Table 6.3 are applied. In this case only Rule 0 is applicable. Using Rule 0 the actions become: $H \; M \; K \; M \; T_S$

Substituting the times and variables from Table 6.2 yields

$$0.4 + 1.35 + 0.5 + 1.35 + T_S = 3.6 + T_S \text{ sec}$$

$3.6 + T_S$ seconds is the relative time it will take to accomplish task T-20, Adjust the Horizontal Situation Indicator Range.

The control input efficiency was determined by dividing the minimum number of control inputs (i.e. a dedicated knob or one) by the number of control inputs in this implementation. The number of control inputs in a particular implementation is determined by counting the number of Ks and Ts. For this example control input efficiency = ½ or 0.5 .

Analysis of Task T-20 (RNG) Second Implementation (RNG Remains Selected)

There are two scenarios in this implementation. If some other function is selected (i.e. HDG), the user performed the same actions as above. Sometimes the RNG function will already be selected. In this scenario the relative task time was determined as follows:

$$H \ M \ T_S = 0.4 + 1.35 + T_S = 1.75 + T_S \text{ sec}$$

The assumption was made that at any given time each of the six alternatives are equally likely. Taking into consideration the two scenarios the overall relative task time was calculated as follows:

$$(1/6 * (1.75 + T_S)) + (5/6 * (3.6 + T_S)) = 3.29 + T_S$$

This relative time is less than that of the first implementation, however, because there are two scenarios, this implementation was difficult to operate automatically. Unfortunately the error rate that this difficulty may cause could not be quantified using the GOMS KLM.

Control input efficiency for the second implementation was (1/6 * 1) + (5/6 * 0.5) = 0.6 .

Analysis of Task T-20 (RNG) Third Implementation (Dedicated to Left Knob)

The third implementation was the ideal case from a task time and input efficiency point of view. The analysis was the same as for the first scenario of the second implementation. This resulted in a relative task time of $1.75 + T_S$ seconds and a control input efficiency of one, the highest possible. A difficulty inherent in this implementation was the labeling of the dedicated knob.

Results of Task T-20 (RNG) Analysis

The third implementation with the left dual concentric knob dedicated to the RNG function was the preferred implementation. It was accomplished more quickly then the other two implementations and eliminated the possibility of mode errors. It also avoided the potentially frustrating requirement to always press a key before adjusting the knob.

6.6.3 Analysis of Remaining Map Adjustment Tasks

Task T-18 (declutter) was accomplished using a line select key to toggle through four options in both the first and second implementations. Assuming each option was equally likely and mental preparation was only required before the first key press, the relative task time was 2.75 sec with an efficiency of one. The third implementation had a relative task time of 1.75 + T_S sec also with an efficiency of one. It was difficult to discriminate between the two based on a GOMS analysis alone since the time T_S is unknown. In this case the third implementation was selected in order to avoid toggling through four options which was previously identified as undesirable in flight tests and was likely more error prone.

The relative task time and control input efficiency for Task T-19 (view) was the same for all proposed implementations. Again the third implementation was the preferred choice as it enabled all of the HSI adjustment features to be functionally grouped together. It also allowed for the possibility of consistency between the PFD and the MFD if the map adjustment features are located similarly on the MFD.

6.7 Setting Bug and Baro Tasks

6.7.1 Implementations

Two alternative implementations of the bug and baro setting tasks were conceived taking into consideration the feedback from the flight evaluations of the demonstration configuration. It was already difficult to toggle through four choices and this difficulty would increase with the addition of the VSPD and A/S bugs. In order to avoid this problem, the various tasks were displayed on the right side corresponding to the line select keys. This provided simultaneous visibility of all of the tasks rather than just the selected task. The two implementations of the bug and baro setting tasks were similar to the first two implementations of the RNG task above.

<u>Bug and Baro Setting Tasks First Implementation (Mode Defaults to "Nothing Selected")</u>

The following five tasks were included in this implementation: T-1 (BARO), T-3 (ALT), T-5 (HDG), T-25 (VSPD) and T-27 (A/S). The first implementation is illustrated in Figure 6.6.

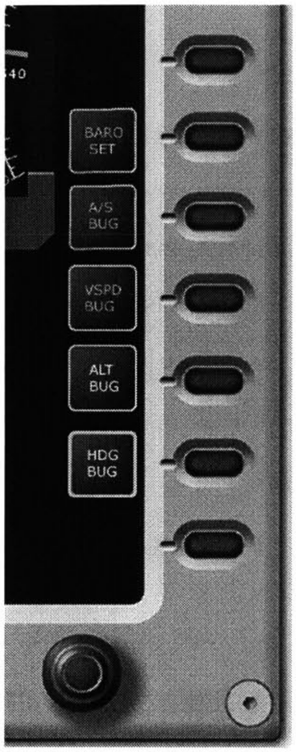


Figure 6.6: Bug and Baro Setting Tasks First Implementation

All five choices were displayed on the right side with each corresponding to line select keys. Corresponding keys set the function of the right knob. The outer knob was used for coarse adjustment and the inner knob was used for fine adjustment. When a function was selected by pressing the associated line select key, it was highlighted or otherwise indicated to show the current function of the knob. Selections "timed-out" after five seconds without control inputs. The interface defaulted to no functions being selected and the knob performed no actions until one of the five available functions was again selected. Reasons the default was to "no selection" are discussed in section 6.6.1.

Bug and Baro Setting Tasks Second Implementation (Mode Remains Selected)

This implementation was similar to the first except that once a function was selected it remained active until a different function was selected.

6.7.2 Analysis of Bug and Baro Setting Tasks

Assuming that each of the five tasks was equally likely, the result of a GOMS analysis is shown in Table 6.4.

Table 6.4: Bug and Baro Setting Tasks GOMS Analysis Results

Tasks	Ideal Case			Implementation One				Implementation Two			
	Sequence	Inputs	Time	Sequence	Inputs	Time	Efficiency	Sequence	Inputs	Time	Efficiency
Bug and Baro Settings	H M T	1	$1.75 + T_{DC}$	H M K M T	2	$3.6 + T_{DC}$	0.5	H M T (1/5) H M K M T (4/5)	1.8	$3.23 + T_{DC}$	0.6

Based on the GOMS analysis, the second implementation was preferred. The advantage in time and efficiency was slight, however a closer inspection reveals greater potential time and efficiency savings. If the number of available modes was reduced, there was a corresponding decrease in the time required to perform the tasks. The number of modes could potentially be reduced in future upgrades through the use of greater automation. Also this analysis assumed the vertical speed and airspeed bugs would be implemented. Simplifying the user interface may be a consideration in whether or not to implement these additional features.

This analysis also assumed that all of the modes were equally likely. An informal survey of pilots suggests that some modes are used more often than others. This suggests that further time would be saved especially when a mode was repeatedly used. This seems likely for example when using the heading bug while being vectored by air traffic control.

Although the first implementation takes advantage of the human trait of habit development and appears likely to reduce mode errors, an informal survey of pilots suggests that pilots may be irritated by always having to press a key before setting a bug. In many current cockpits, pilots are accustomed to simply turning a dedicated knob.

6.7.3 Analysis of Bug and Baro Synchronization Tasks

Synchronization tasks associated with the various bugs and baro settings were always accomplished through the use of a single button, therefore a GOMS analysis was unnecessary. Location of the button was the only consideration. The feature being synchronized corresponded to the selected mode of the knob, therefore it was logical to place the synchronization feature in the push button on the lower right dual concentric knob. A concern with this placement was that if, in some other application, that push button was to be used as the enter feature for a desired action to take effect. It was undesirable to have that push button action differ significantly between the PFD and MFD. If this were the case, it was preferable to associate the synchronization function with a dedicated key along the bottom or upper right.

6.8 Navigation Display Tasks

6.8.1 Implementations

It was identified in previous flight evaluations that it was not obvious to the user that CRS must be selected when adjusting the standby and AUX NAV displays. To correct this problem, the CRS mode was only displayed after a NAV display was selected (active, standby or auxiliary). This way the user had to consciously decide which NAV display to adjust prior to using the CRS feature thereby avoiding mode confusion. Once a NAV display was selected, a CRS label corresponding to the bottom right line select key appeared. The CRS label was highlighted indicating that setting the NAV course was the current function of the right dual concentric knob. In other words, when a NAV display was chosen, the knob function defaulted to setting the course of the selected NAV display. This feature was common to all implementations and created the tasks of selecting the active, standby and auxiliary navigation displays.

<u>Selecting Active, Standby and Auxiliary NAV Displays First Implementation (Direct Selection)</u>

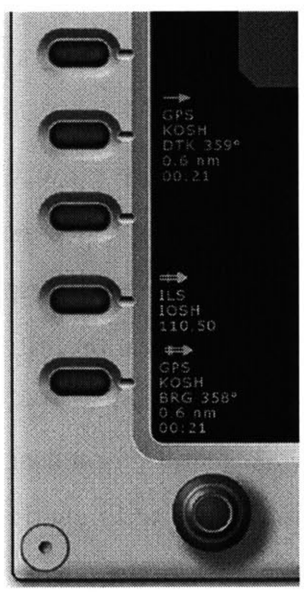


Figure 6.7: Selecting the NAV Displays First Implementation

It was intuitive to the user to directly select the desired information that was to be changed. For this reason the information for the three NAV displays was arranged on the left corresponding to three line select keys. This is illustrated in Figure 6.7.

Selecting Active, Standby and Auxiliary NAV Displays Second Implementation (Toggling)

The second implementation was the same as the demonstration configuration where a NAV label corresponding to a right line select key toggled through the available options. This is shown in Figure 6.8.

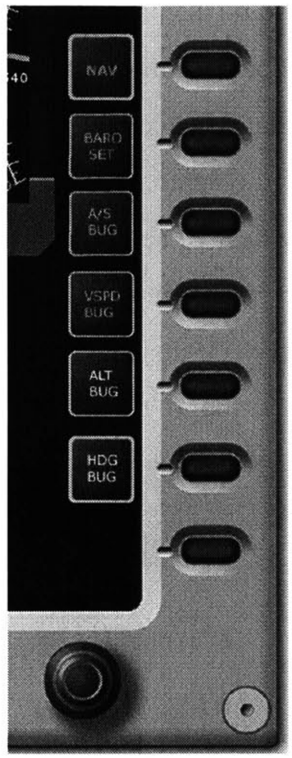


Figure 6.8: Selecting the Navigation Displays Second Implementation

Results of Selecting the Active, Standby and Auxiliary Navigation Displays Analysis

Implementation one was preferred since it resulted in dedicated keys for each task and therefore a minimum relative time and maximum efficiency. Disadvantages of this implementation were that the keys are located on the left side where they were more difficult to use and the possible need for a small font size to ensure that all of the information would fit.

It can easily be seen that the second implementation takes relatively longer and was less efficient than the first implementation. It required toggling through four choices, which was identified as something to be avoided. In addition it resulted in a key on the right side that was functionally different than the other keys on the right that all set the mode of the knob.

Tasks T-11 thru T-16 (Source and Format) First Implementation (SRC and FMT Toggling)

Once one of the three NAV displays was selected, a SRC and FMT label appeared corresponding to the right two line select keys on the bottom row. These keys were chosen so that all of the other labels can continue to be displayed. This allowed the user to select any other option at any time during this process if performing a higher priority task was necessary. Pressing the SRC or FMT keys caused the selected NAV information to be superseded by the available source or format settings as in the demonstration configuration. The user had to toggle through the available options. Once the desired selection was made, all other options were still available for immediate selection. If no control inputs were made for five seconds, the interface defaulted to the top most level where no NAV displays were selected. The first implementation of the source and format tasks is shown in Figure 6.9.

Figure 6.9: Tasks T-11 thru T-16 (Source and Format) First Implementation

Tasks T-11 thru T-16 (Source and Format) Second Implementation (SRC and FMT Sub-Menu)

The SRC and FMT labels again appeared corresponding to the right two line select keys on the bottom row when a NAV display was selected. When the SRC or FMT keys were pressed, the available choices were displayed on the right corresponding to the right line select keys. The CRS feature remained selected on the bottom key on the right side dictating the action of the knob, but all other labels were superseded by the SRC or FMT choices. An ESC (escape) label corresponded to the middle line select key on the bottom row. This key enabled the user to return to the original choices of knob function in case the user needed to execute a different task. Once all NAV adjustments were complete and no control inputs were made for five seconds, the interface defaulted to the top most level where no NAV displays were selected. This implementation is depicted in Figure 6.10.

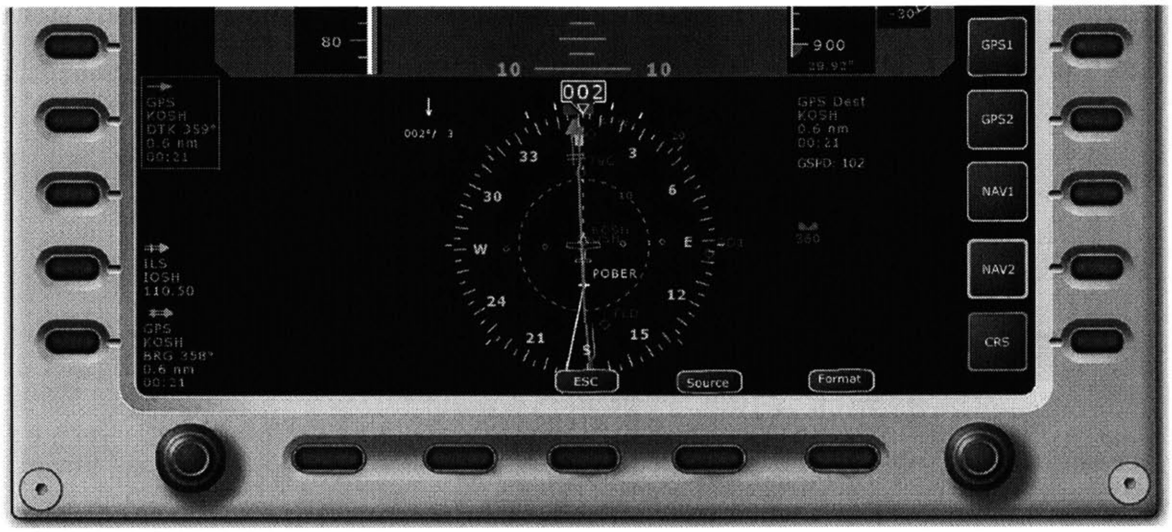


Figure 6.10: Tasks T-11 thru T-16 (Source and Format) Second Implementation

6.8.2 Analysis of Navigation Display Tasks

Tasks T-11 thru T-16 (Source and Format) Analysis

The second implementation resulted in the interface with the smallest relative time and higher efficiency. Completing the SRC or FMT tasks always took three key presses whereas the first implementation took from three to five. Also selecting the desired SRC or FMT choice was less error prone than toggling through multiple choices. A disadvantage of the second implementation was the hiding of the knob function menu, however the addition of the escape feature allowed this menu to be quickly restored.

Tasks T-17 (Swap) and T-21 (GPS Hold) Analysis

Both tasks T-17 (Swap) and T-21 (GPS Hold) were implemented with dedicated buttons as in the demonstration configuration. Since the NAV displays were arranged vertically on the left of the screen, the swap function was logically placed between the active and standby NAV displays. This function corresponded to the third key from the bottom on the left side. GPS

Hold enabled the user to prevent the automatic sequencing of GPS waypoints. It was located directly above the active NAV information on the left side corresponding to the fifth button from the bottom. There was a time delay between selecting GPS Hold and the response of the GPS. During this time delay the green outline of the GPS Hold label flashed until confirmation was received from the GPS. Once positive confirmation was received, the outline was steady green until GPS Hold was turned off. The implementation of these two tasks is shown in Figure 6.11.

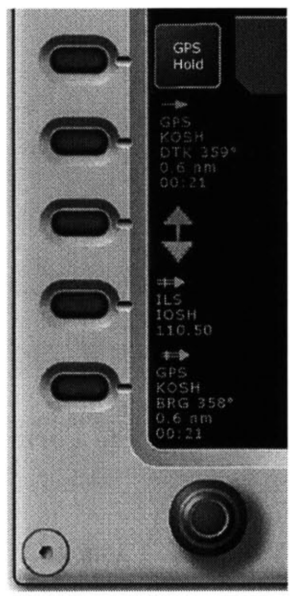


Figure 6.11: Swap and GPS Hold Implementation

7 PROPOSED INTEGRATED HARDWARE AND SOFTWARE INTERFACE

7.1 CONFIGURATION

The results of the hardware and software interface development are combined into the proposed interface configuration shown below. Figure 7.1 shows the top level of the interface.

Figure 7.1 Primary Flight Display Proposed User Interface (Top Level)

An example is shown of how the pilot would select the source of the active NAV display. First the pilot would press the button corresponding to the active NAV display (fourth from the bottom on the left). Figure 7.2 shows the proposed user interface after this button is pressed. The SRC, FMT and CRS labels now appear and the CRS label is highlighted indicating the function of the knob. All of the other functions are still visible however and can be selected if needed.

Figure 7.2 Primary Flight Display Proposed User Interface (Active NAV Selected)

Next the button labeled "source" is pressed. This causes the menu on the left to change to the choice of four sources. A label called "ESC" also appears enabling the pilot to return to the previous menu. The resulting display is shown in Figure 7.3.

Figure 7.3 Primary Flight Display Proposed User Interface (SRC Selected)

A task analysis of the proposed user interface reveals a flatter hierarchy with more tasks visible on the top level. A hierarchical task analysis of the proposed configuration is shown in Figure 7.4.

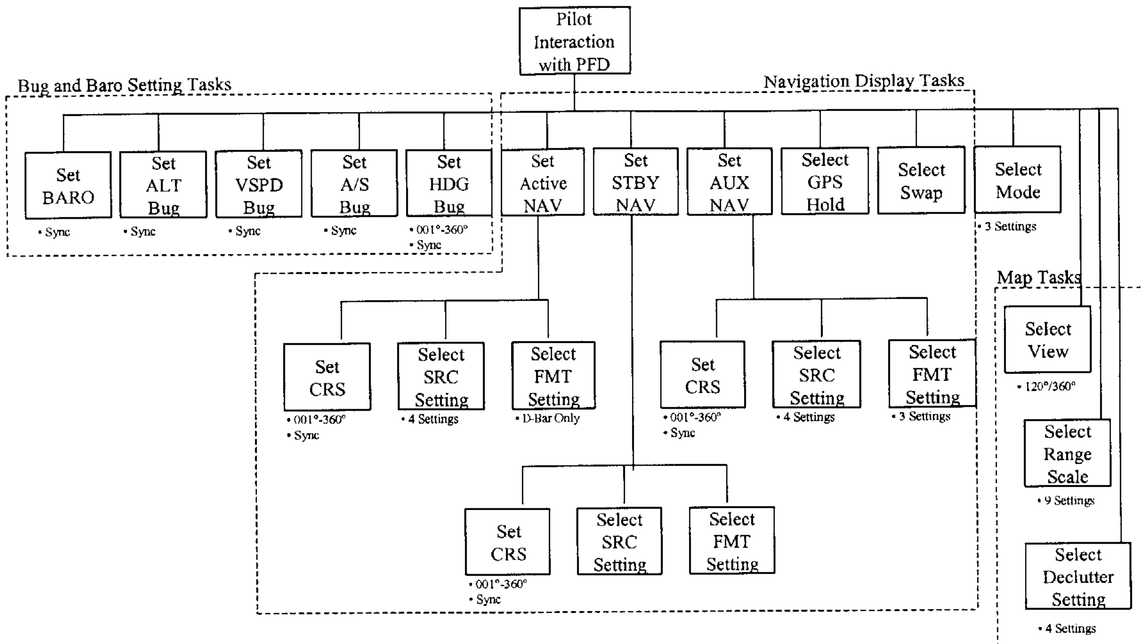


Figure 7.4: Proposed User Interface Hierarchical Task Analysis

7.2 Proposed Interface Review

The proposed interface configuration and its underlying rationale were presented to the president of the Avidyne Corporation for approval. This presentation resulted in a corporate-level decision to simplify the product as much as possible given the context of a customer's aircraft. Installation specific simplification goals are summarized below:

- Design for installation in a specific aircraft taking into consideration the other installed equipment in that aircraft.
- Perform all tasks on other installed equipment whenever possible in order to simplify the PFD operation and reduce the number of bezel controls needed.
- Functionality can be reduced if necessary to simplify operation and interaction.
- Consider landscape format only.
- Reduce the number of controls as much as possible. The basic hardware design shall essentially remain the same, however controls can be depopulated or removed from that design.
- Automate features whenever feasible in order to simplify operation.

These simplification goals are consistent with the top ten technical requirements derived earlier during the Quality Function Deployment process. In particular taking into consideration the other equipment installed in the aircraft enhances opportunities for functional integration. Also reducing functions was previously considered untenable from a marketing standpoint, but is now open for consideration in the name of reducing complexity and simplifying use.

8 PHASE I PRIMARY FLIGHT DISPLAY AND USER INTERFACE REDESIGN

8.1 TASK ANALYSIS

A redesign of the Phase I primary flight display was conducted utilizing the new simplification goals. A task analysis was performed taking into consideration the larger context of the associated equipment installed in a customer's aircraft. The functions and tasks related to the user interface were analyzed and opportunities for simplification through the reduction of functionality and closer integration with other installed equipment was explored. Also redundant tasks performed by other installed equipment were eliminated. The relevant equipment included dual Garmin Global Navigation System (GNS) 430s and a STEC S55 autopilot. A diagram of inputs/outputs for this equipment is shown in Figure 8.1.

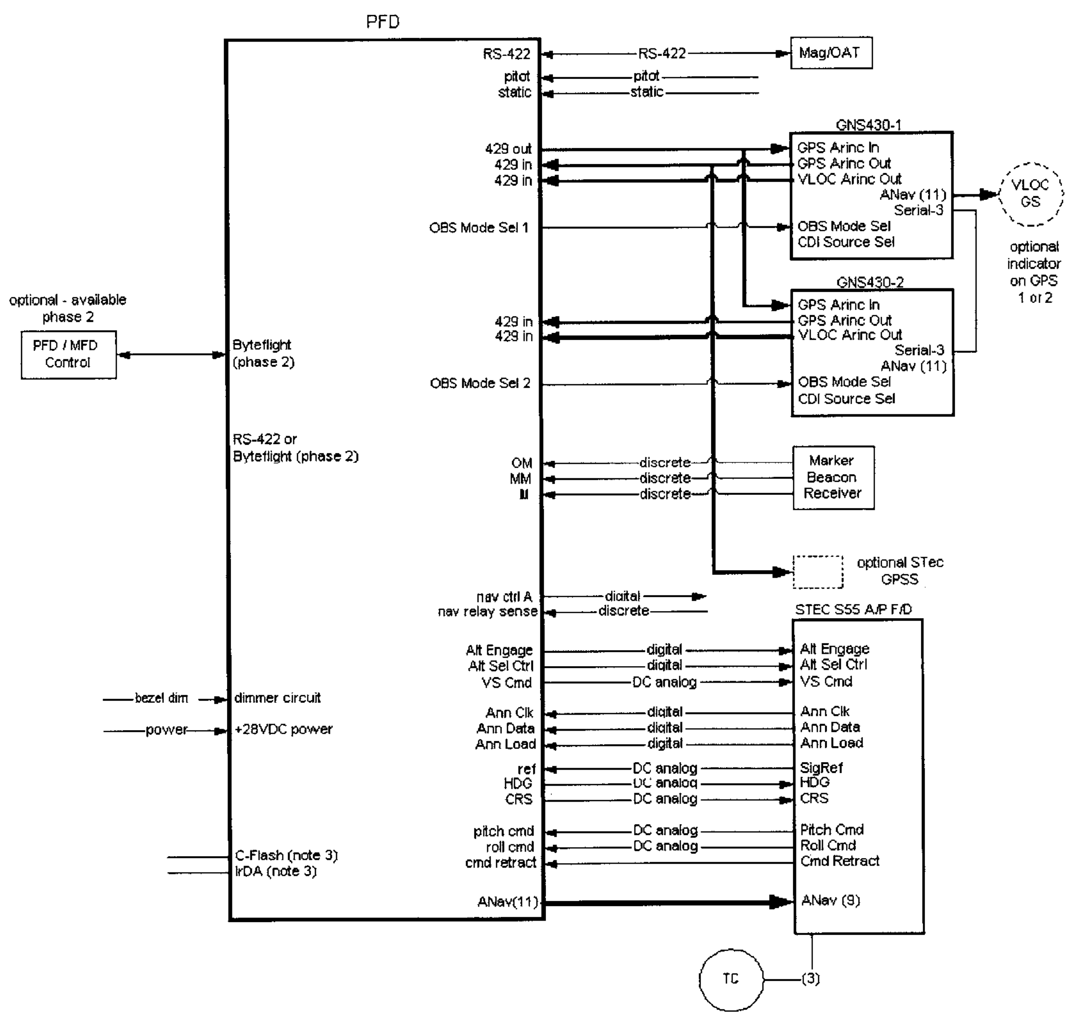


Figure 8.1: Input/Output Diagram for the PFD with Dual GNS 430s and STEC S55

8.2 Eliminated Tasks

Based on analysis, nine of twenty-eight tasks were eliminated. A brief description of the justification for eliminating the tasks follows.

T-12. Select the format for the active NAV display – This function was incorporated with task T-19 (Select the View Setting). A third setting was created that displays the HSI map without the navigation needles being displayed. This provides additional functionality in that it was previously not an option to turn off the deviation bar (D-Bar) of the active NAV source. The deviation bar normally appears on the HSI as a directional needle whose center indicates the relative deviation from the desired course. The ability to turn the D-Bar off accommodates those users who wish to have a clear view of the HSI map for GPS navigation.

T-13. Select the source for the standby NAV display – The standby NAV source will automatically default to the standby source of the selected GNS 430. For example, if GNS 430-1 is selected as the active NAV source and it is currently in GPS mode, the standby NAV source will default to the active VLOC on the GNS 430-1. The VLOC, or VOR/Localizer, can receive information from a VOR, Localizer or receive glideslope information.

T-16. Select the format for the auxiliary NAV display – The auxiliary NAV display is always a bearing pointer.

T-21. Select GPS hold setting – This is accomplished on the GNS 430 using the OBS button.

T-23. Turn the unit on or off – There will be no on/off button on the hardware. The user can only turn the unit on or off by removing power.

The following tasks identified as future capabilities will not be implemented:

T-25. Set the vertical speed (VSPD) bug

T-26. Sync the vertical speed bug to the current vertical speed

T-27. Set the airspeed (A/S) bug

T-28. Sync the airspeed bug to the current airspeed

8.3 Modified Tasks

Several of the original tasks were modified in order to reduce complexity as follows:

T-11. Select the source for the active NAV display – Selects between NAV1 or GNS 430-1 (normally on top) and NAV2 or GNS 430-2 (normally on the bottom). For example, if the GNS 430-1 is selected for primary navigation, the state selected by the CDI button on the GNS 430-1 or Select control on the PFD will set the active NAV source (either GPS1 or VLOC1). The standby NAV source will default to the unselected state (GPS1 or VLOC1). See Figure 8.2.

Figure 8.2: Dual Garmin GNS 430s

T-15. Select the source for the auxiliary NAV display – Five settings are available including GPS1, GPS2, VLOC1, VLOC2 and off. The aux NAV display always appears as a bearing pointer.

T-17. Swap the active and standby NAV displays - This task selects between the GPS and VLOC of the primary GNS 430 and acts the same as the CDI select button on the GNS 430. Although this is a redundant feature it was deemed highly desirable to retain this feature on the bezel. The bezel implementation should also slave to the state of the corresponding GNS 430 CDI select button.

T-19. Select the view setting – An additional setting called "map only" has been added which turns off all of the navigation needles for better viewing of the HSI map. The 360° and arc views are still available.

T-22. Select the mode (Normal, Backup or Composite) – Only Normal and Backup mode will be available for the Phase I PFD.

8.4 Task Functional Groups

The user interface tasks can still be logically grouped by function. The new groups after the above changes are applied are as follows:

Map Adjustment Tasks

- Adjust the range (RNG) (9 settings)
- Select the declutter setting (5 settings)
- Select the view setting (3 settings)

Navigation Display Tasks

- Set the active NAV display course (CRS)
- Sync the active CRS to the current heading
- Set the standby (Stby) NAV display course
- Sync the standby CRS to the current heading
- Select the source for the active NAV display (4 settings)
- Select the format for the standby NAV display (4 settings)
- Select the source for the auxiliary NAV display (5 settings)
- Swap the active and stby NAV displays

Bug and Baro Setting Tasks

- Set the baro altimeter setting
- Sync the baro setting to 29.92
- Set the altitude (ALT) bug
- Sync the ALT bug to the current ALT
- Set the heading (HDG) bug
- Sync the HDG bug to the current HDG

In accordance with the system requirements, the following tasks will have dedicated controls and so will not be placed into functional groups:

- Select the mode (Normal and Backup)
- Adjust the brightness

8.5 REDESIGNED INTERFACE

Implementations of the bug and baro setting tasks and HSI map adjustment tasks remain virtually unchanged from the original proposed configuration. Differences are simply the elimination of the airspeed and vertical speed bugs and the addition of another setting (for a total of three) to the view task. Navigational display tasks have been modified in order to simplify their operation and eliminate the need for submenus. On the left side, the bottom button selects the source for the auxiliary NAV display or turns it off. The button above it selects one of the available four settings for the format of the standby NAV display. Standby NAV format settings include D-Bar, RMI, Text only and none. The third button from the bottom on the left side swaps the active and standby NAV displays. This button corresponds to the CDI button on the Garmin GNS 430. The NAV source button corresponds to the fourth button from the bottom on the left side and selects between NAV1 and NAV2. The redesigned interface is shown in Figure 8.3.

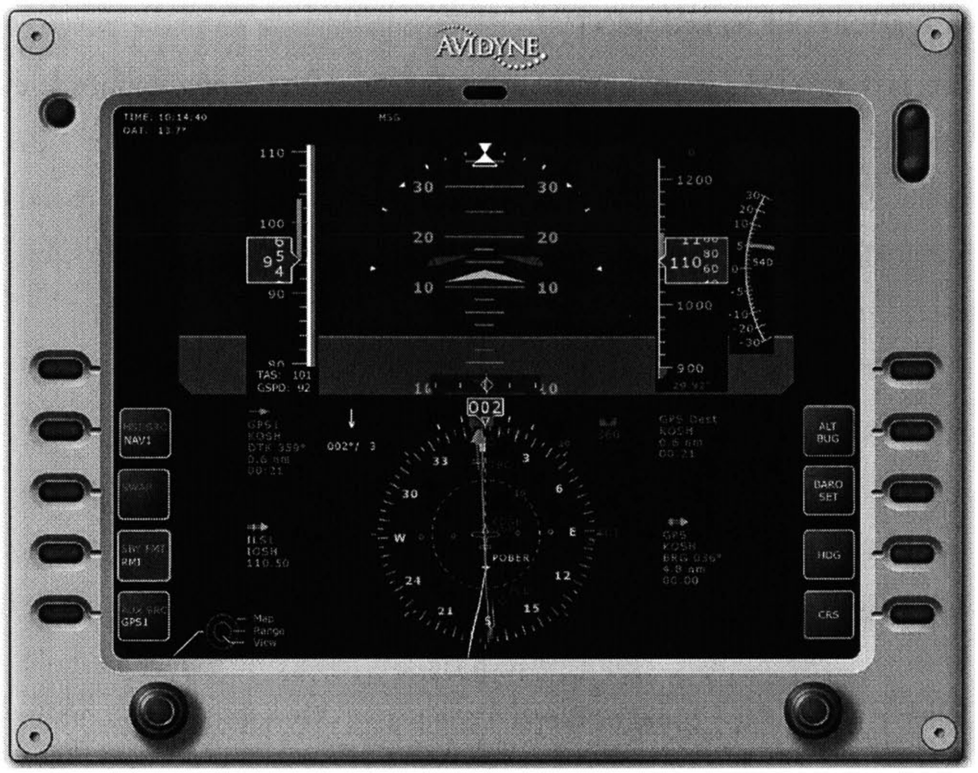


Figure 8.3: Redesigned Primary Flight Display User Interface

Since a goal of the redesign effort was to eliminate any unnecessary controls, the unused bottom bezel buttons were removed. The minimum number of controls necessary for MFD operation on the sides was determined to be five; therefore two buttons were eliminated from each side. This allows a limited growth capability on the PFD and cleans up the appearance of the bezel considerably.

9 INTERFACE DESIGN EXPERIMENT

Test subjects completed a series of tasks using three interface configurations to evaluate the configurations and validate the GOMS Keystroke-Level Model time predictions. Both the prototype configuration (Version A) and two alternative configurations (Versions B and C) were evaluated in order to determine if there were any performance gains and to capture subjective feedback for design improvements.

9.1 EXPERIMENTAL SETUP

The landscape-oriented bezel discussed earlier was not yet operational so the new configurations were adapted to the original portrait-oriented bezel. Unfortunately the portrait-oriented bezel has only one dual concentric knob without a push button capability. The use of this bezel prevented all tasks from being fully evaluated. Six tasks common to all three configurations were evaluated and could be directly compared. Six additional tasks were evaluated three of which were directly comparable between A and B and three were directly comparable between A and C. The synchronization tasks could not be accomplished in versions B or C and so were not evaluated. The same electronic flight instrument displays were used in all three versions to eliminate any bias from the displays.

Version A: The original NASA/AGATE demonstration user interface is shown in Figure 9.1. This interface was used as a baseline for comparison to determine if any significant improvements were made.

Figure 9.1: User Interface Version A (Demonstration Configuration)

Version B: This version was the redesigned user interface optimized for operation with dual Garmin GNS 430s and a landscape orientation. Minor changes were made to adapt the newly designed interface to the portrait-oriented Avrotec bezel as shown in Figure 9.2. In general the simplified navigation display tasks are on the left and the bug and baro setting tasks are on the right. Map adjustment tasks are located on the left dual concentric knob (not evaluated due to a hardware limitation).

Figure 9.2: User Interface Version B (Map Controls on Left Knob)

Version C: This configuration was similar to Version B except that the map adjustment tasks are on the right side buttons and the bug and baro setting tasks are split between the two dual concentric knobs. Heading and course tasks are accomplished with the left knob while the altitude bug and baro setting tasks are accomplished with the right knob. The left and right knobs default to heading and baro respectively. Course or altitude bug mode can be selected by pressing the button immediately above the knob. The knobs will remain in this mode as long as there are inputs to the knob, however they will default to heading and baro after five seconds without inputs. The concept is that the heading and baro tasks will be used more frequently and it will enable the user to have dedicated knobs for these tasks similar to a conventional cockpit. Version C is pictured in Figure 9.3.

Figure 9.3: User Interface Version C (Map Controls on the Right Side Buttons)

Protocol: To mitigate the effects of learning, the interface configurations were presented in a counterbalanced order. Before the experiment, the pilots were trained on the use of the interface and the flight displays. They were given an opportunity to fly the displays and conduct practice tasks until they felt comfortable with the displays and what was expected of them. The Avrotec hardware was networked with a computer running Microsoft Flight Simulator 2000 to provide flight inputs. Subjects flew the simulator with their left hand using a standard joystick. Subjects were required to use their left hand since the design aircraft has a left side-stick controller and to induce additional workload. Subjects were required to maintain a constant heading, airspeed and altitude while completing the user interface tasks to induce workload and distract them from the primary task. Subjects completed 13 tasks in Version A and 9 tasks each in Versions B and C in random order. Each task was evaluated twice for a total of 26 in Version A and 18 in Versions B and C. The simulator setup is shown below. (The Microsoft Flight Simulator screen was not visible to the subjects while conducting the experiment.)

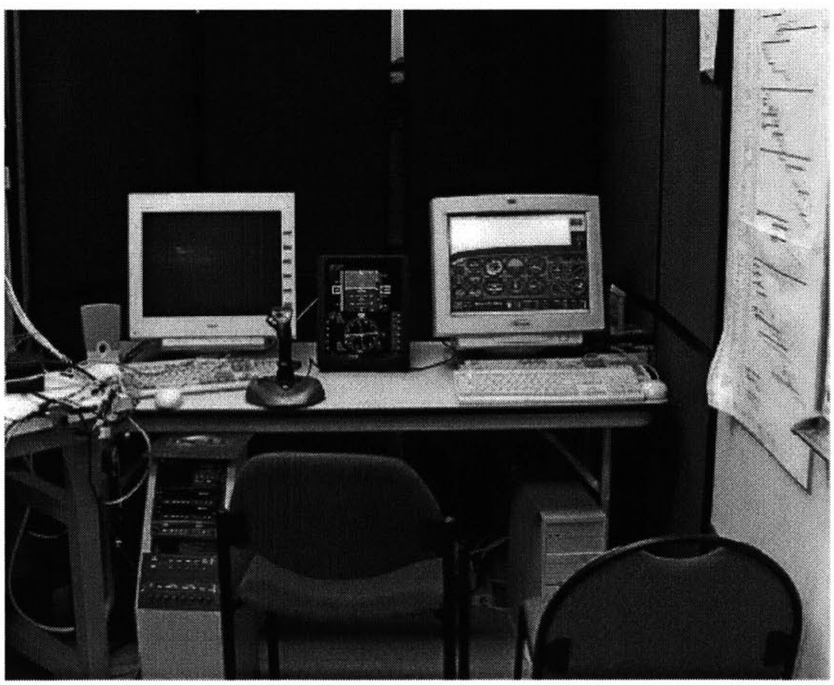


Figure 9.4: Simulator Setup

The complete protocol including a list of the evaluated tasks is contained in Appendix B.

10 Results of the Interface Design Experiment

The experiment was conducted from the end of April to early May 2001. The subject profile is discussed first, followed by the results of the experiment.

10.1 Subject Profile

A total of six instrument rated pilots from the General Aviation community took part in the study. The age of the group ranged between 26 and 53, with an average of 37 and a standard deviation of 9.6 years. Two were female, representing 33% of the subject pool.

Three of the six pilots were Certified Flight Instructors qualified to teach instruments. The six pilots had an average of 1686 flight hours. Only one pilot had rated military experience (two had non-rated military flight engineer experience). None of the subjects had any significant prior glass cockpit display experience.

10.2 Performance Metrics

Performance metrics include the task execution time, accuracy, corrected error rate, uncorrected error rate, control input efficiency and subjective workload. In addition, the secondary flying tasks of maintaining assigned heading, airspeed and altitude were also analyzed for statistical significance. Finally the predicted mean execution time based on the GOMS Keystroke-Level Model was compared to the actual mean execution time.

10.3 Task Execution Time

Execution times were recorded for each task by the computer. Time began when the subject moved their hand from the start position on the table (based on a key input from the evaluator) and ended with the last control input. Actual values are contained in Appendix C. Mean execution times for each task shown in Figure 10.1 include any errors that occurred while performing the task.

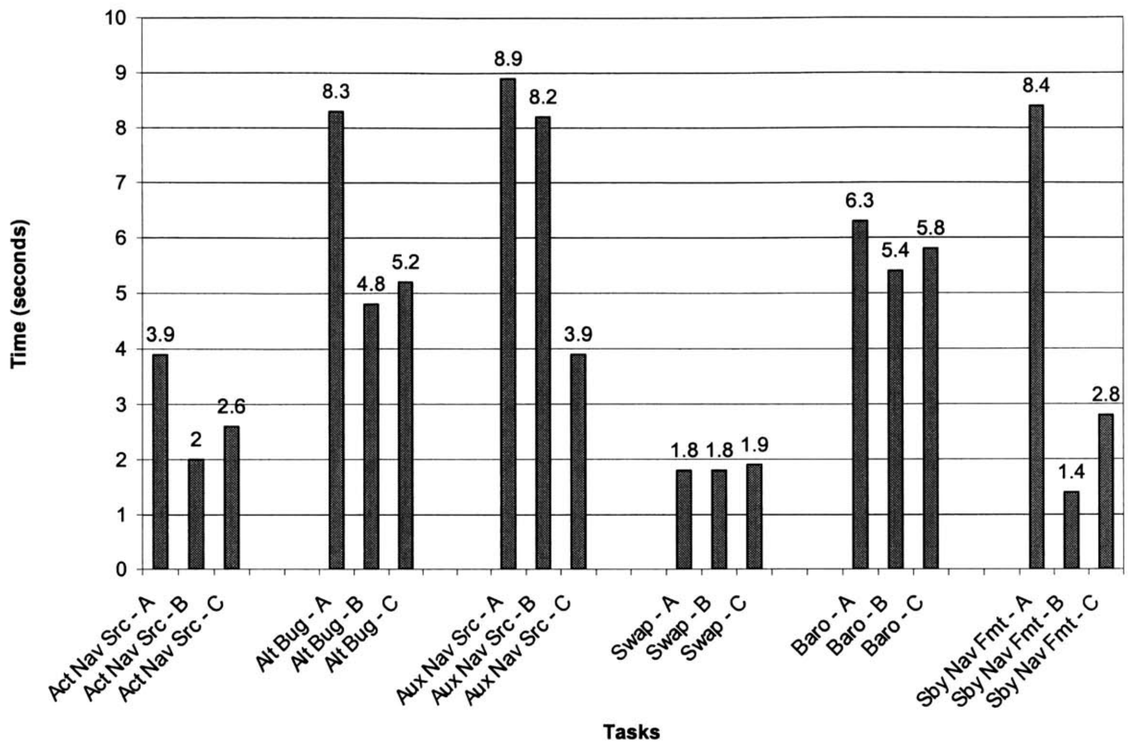


Figure 10.1: Mean Execution Time Including Errors for Versions A, B and C

An ANOVA of the mean execution times was conducted to determine if the mean execution times were significantly different. The mean execution time of the altitude bug setting task in Versions B and C was significantly shorter ($p < 0.05$) than in Version A. There was no significant difference in the altitude bug setting task between Versions B and C. Similarly the standby NAV format task was significantly shorter ($p < 0.05$) in Versions B and C than in Version A. Again there was no significant difference for this task between Versions B and C.

The tasks that can be directly compared between Versions A and B and A and C are shown in Figures 10.2 and 10.3 respectively.

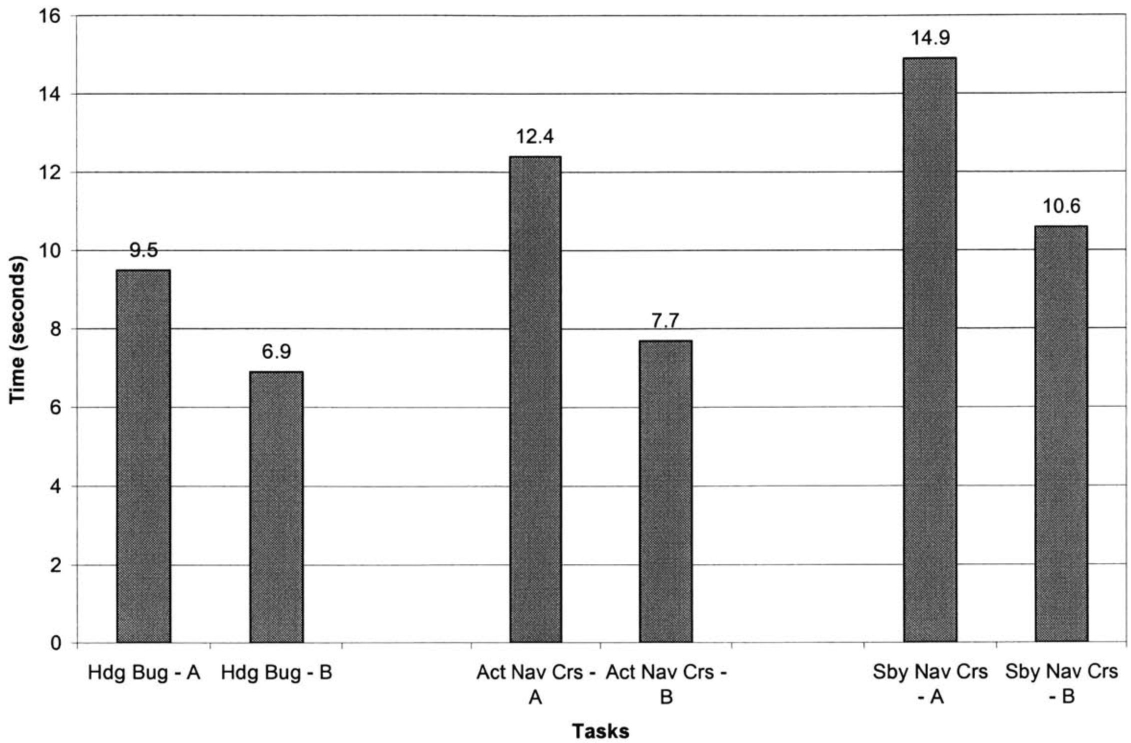


Figure 10.2: Mean Execution Time Including Errors for Versions A and B

An ANOVA of the tasks common to Versions A and B revealed a significantly shorter mean execution time in version B of the heading bug task ($p < 0.05$) and the active NAV course task ($p < 0.05$).

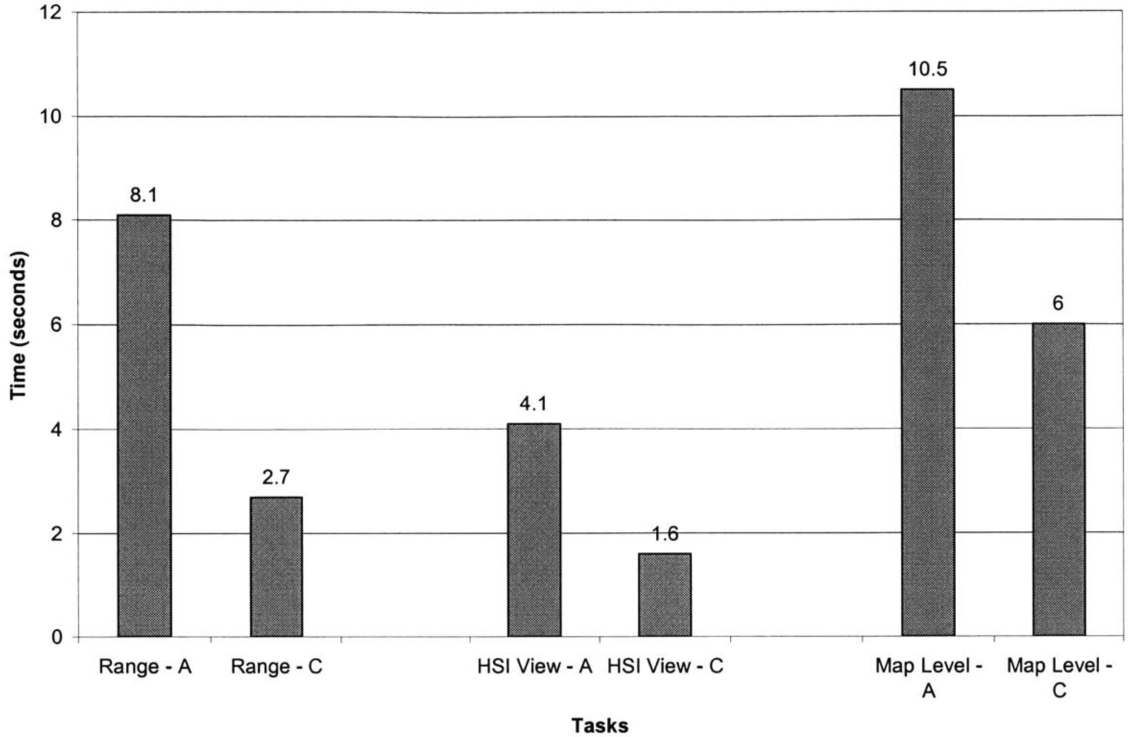


Figure 10.3: Mean Execution Time Including Errors for Versions A and C

The differences in the mean execution times between Versions A and C were not statistically significant at the 5% level ($p > 0.05$).

10.4 TASK ACCURACY

Accuracy of the various tasks was determined by dividing the total number of times a task was correctly performed by the total number of times the task was performed. A correctly performed task was considered one where no extra control inputs were made. Tasks are organized into three groups: the tasks performed in Versions A, B and C (Figure 10.4); the tasks performed in Versions A and B (Figure 10.5); and the tasks performed in Versions A and C (Figure 10.6).

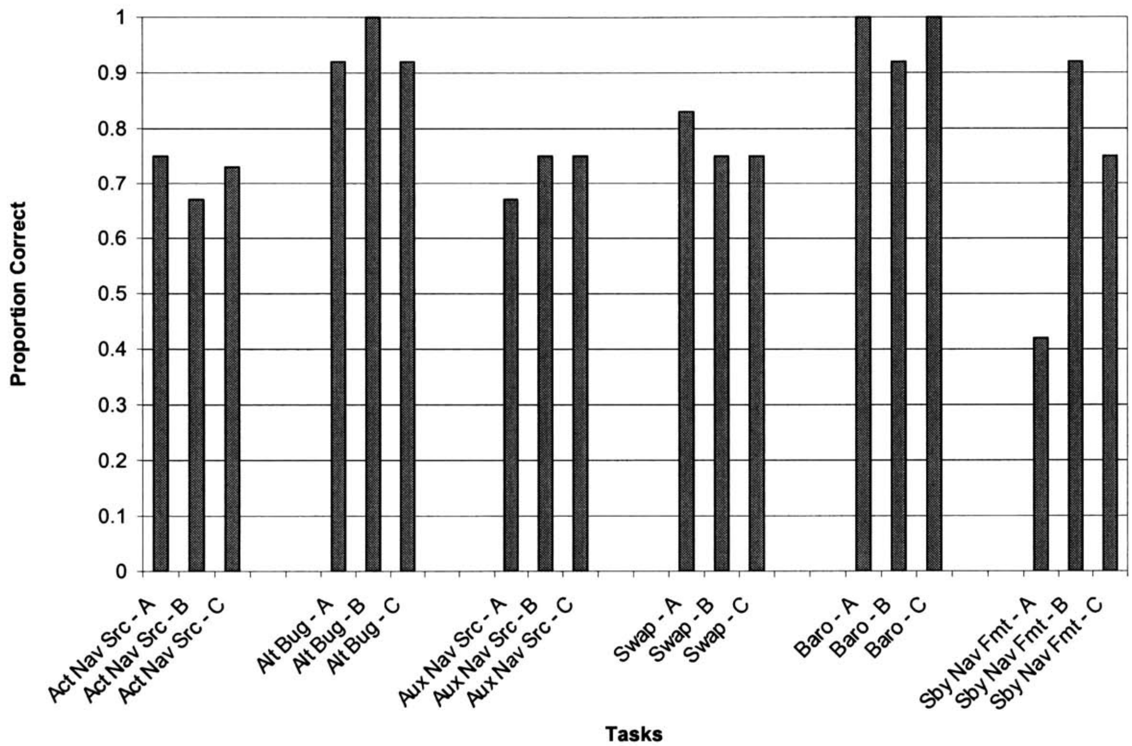


Figure 10.4: Accuracy Comparison Between Versions A, B and C

Accuracy differences between the versions were tested for significance by determining the standard deviation using the following formula: $p\ (1-p)\ /\ n^{1/2}$ where p is the proportion correct and "n" is the sample size of each task. The largest standard deviation among the three common tasks was then used. Differences were considered significant if the difference between versions was greater than three times the largest standard deviation ($p < 0.01$). Based on this criterion, the standby NAV format task was significantly more accurate in Versions B and C than in Version A.

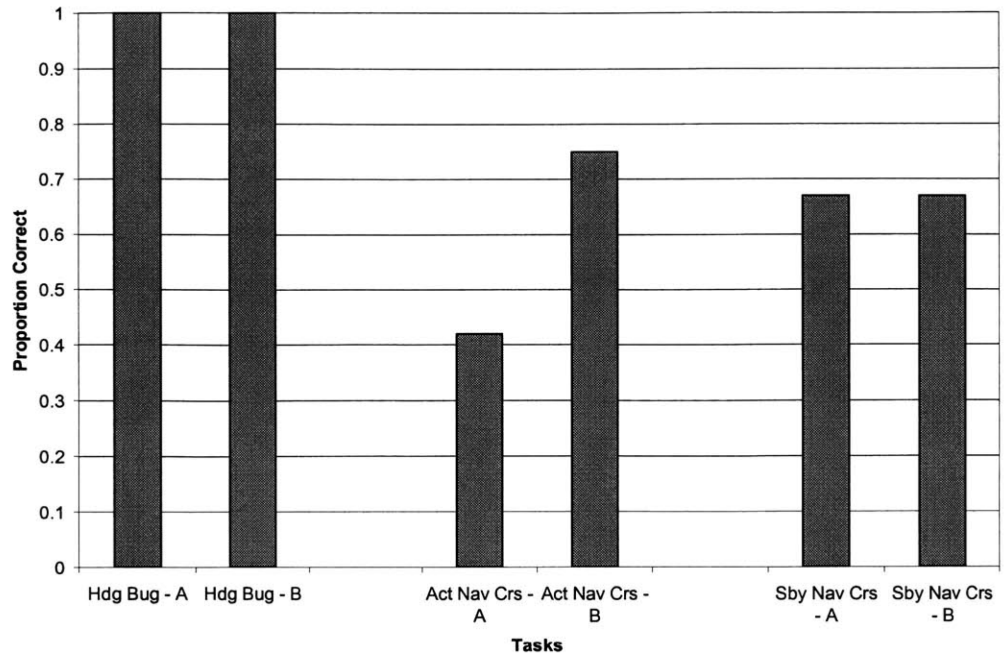


Figure 10.5: Accuracy Comparison Between Versions A and B

The implementation of the active NAV course task was significantly more accurate in Version B than in Version A as shown in Figure 10.5.

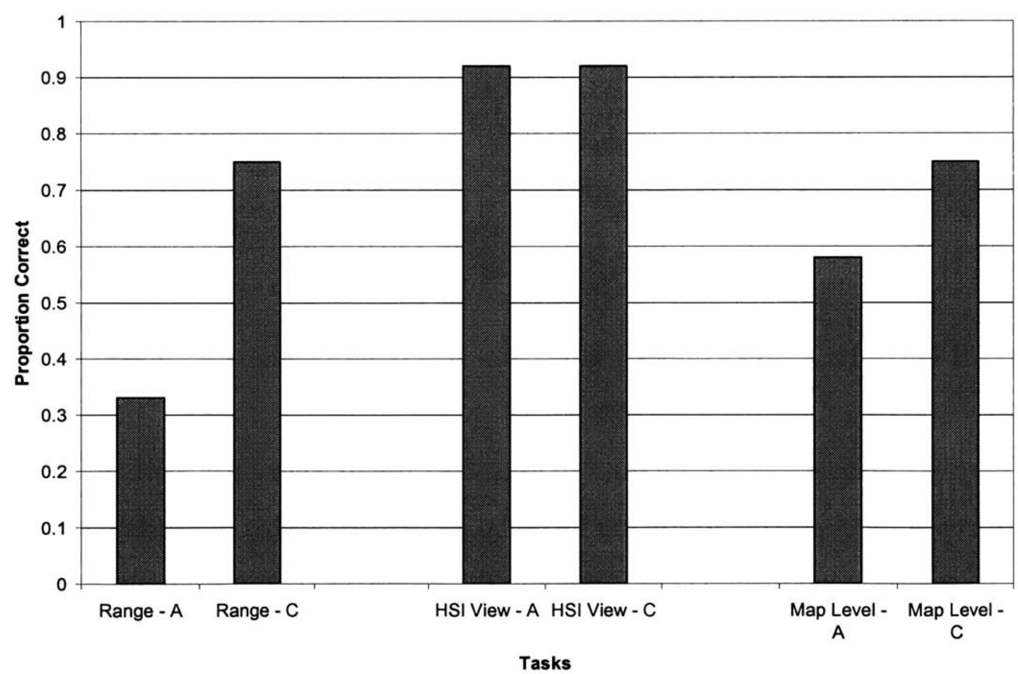


Figure 10.6: Accuracy Comparison Between Versions A and C

Execution of the range task was significantly more accurate in Version C than in Version A as shown in Figure 10.6. The only difference between these two tasks was that in Version A the user must first move to a submenu.

10.5 ERROR RATE

There were two types of errors recorded during the experiment, corrected and uncorrected errors. A corrected error is one in which the user recognized the error and eventually correctly performed the task. The most common example of this was toggling past the desired setting. Error rates are broken out by tasks. Task error rates that are statistically different are the same ones determined in the accuracy analysis. Error rates for tasks performed in Versions A, B and C are shown in Figure 10.7.

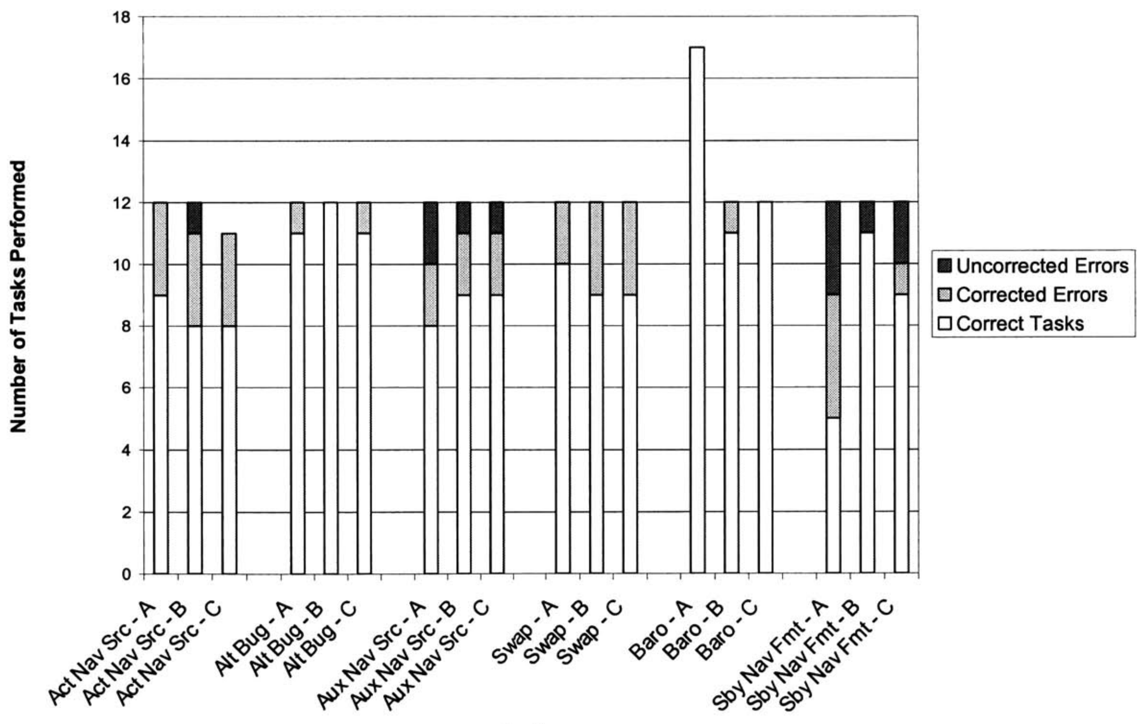


Figure 10.7: Error Rates for Versions A, B and C

The error rates for the tasks in Versions A and B are shown in Figure 10.8.

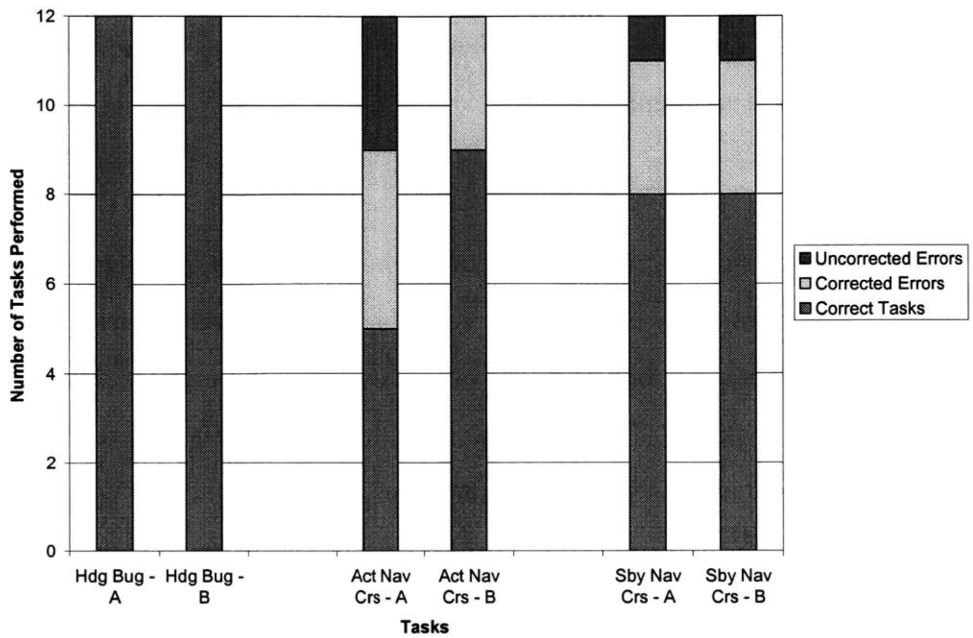


Figure 10.8: Error Rates for Versions A and B

The error rates for the tasks in Versions A and C are shown in Figure 10.9.

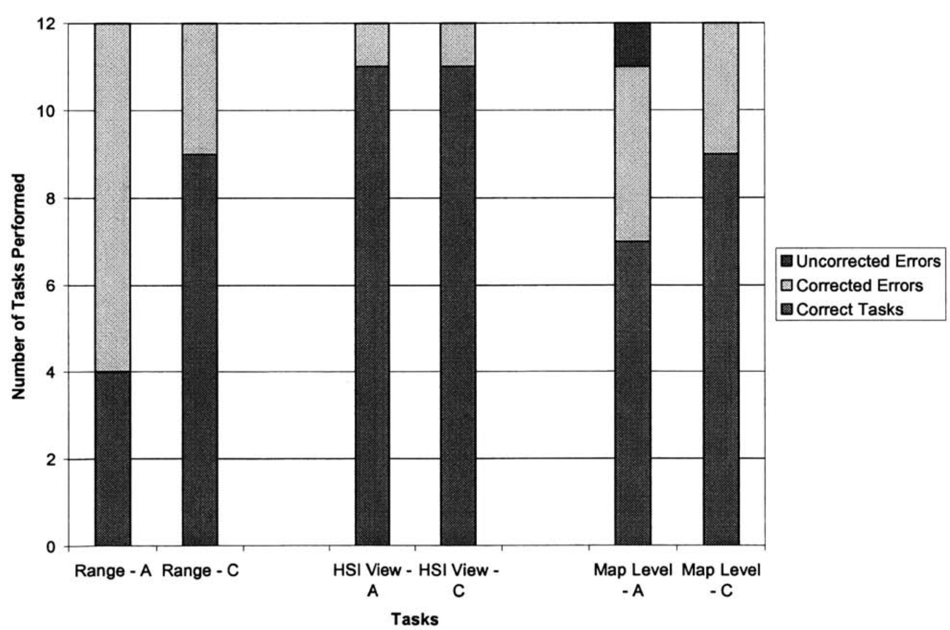


Figure 10.9: Error Rates for Versions A and C

10.6 CONTROL INPUT EFFICIENCY

Control input efficiency was determined by dividing the minimum number of control inputs required to perform a task by the actual number of control inputs used. Setting a dual-concentric knob to a desired value was treated as one control input for this calculation. In order to make direct comparisons, only the six tasks evaluated in all three versions were considered. The results are shown in Figure 10.10.

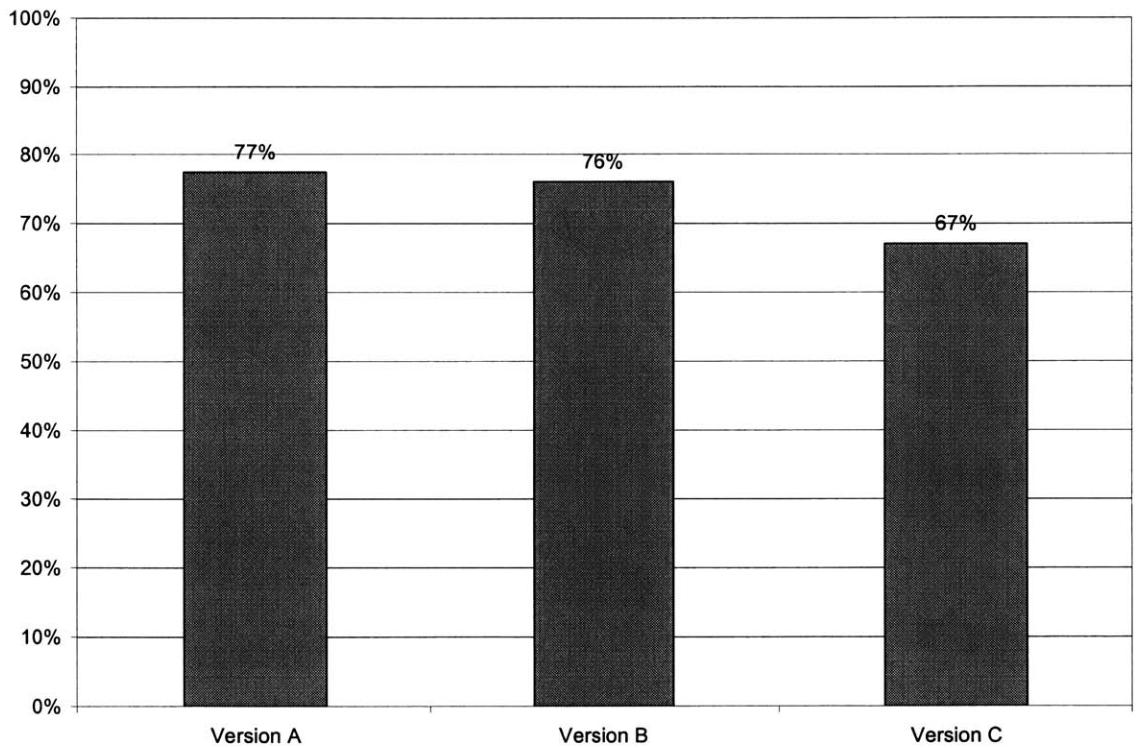


Figure 10.10: Control Input Efficiency

Versions A and B had an efficiency nearly 10% higher than that of Version C although this difference was not statistically significant. Efficiencies of Versions A and B were nearly the same, although it took over twice the number of control inputs in version A (323) to accomplish these same tasks as in Version B (150).

10.7 RESULTS OF PAIRED COMPARISON

The Analytic Hierarchy Process (AHP) utilizing redundant paired comparisons yielded a relative ranking between the alternatives as shown in Figure 10.11 [8].

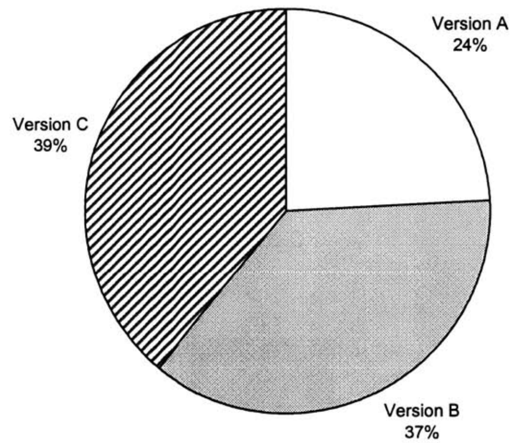


Figure 10.11: Results of Paired Comparison

Versions B and C each dominated Version A however there was no statistically significant difference between Versions B and C.

10.8 SUBJECTIVE FEEDBACK

An analysis of the subjective feedback for common comments revealed the following trends:

- All six subjects reported confusion between the "HSI SRC" and "SRC SEL" labels of Versions B and C. The labels did not convey any meaning to the subjects and they felt they had to memorize the function of the buttons rather than rely on the labels to remind them of their functions.
- Four subjects preferred having all of the knob mode options displayed as in Version B. All of the knob functions were easily visible at all times. In Version A the pilots

had to toggle through options and were "surprised" as each option appeared and often would toggle pass the desired option.

- Four subjects expressed a preference for having the map controls on the top level. Again having map controls on the top level provided continuous visibility of these tasks.
- Half of the subjects thought the "Map" label was confusing and recommended "Declutter." "Map" did not convey any useful meaning to the subjects. They had to memorize the function of the button rather than rely on the label as an indicator.

A complete transcript of the subjective comments is contained in Appendix D.

10.9 SUBJECTIVE WORKLOAD

The subjects were required to fly straight and level at a constant speed while performing the user interface tasks to better simulate the conditions under which these displays will be used. The NASA Task Load Index (TLX) was used to determine the subjects' workload that included both the flying and user interface tasks. Since the flying tasks and flight displays were identical with the exception of the user interface, the difference in workload can be attributed to the user interface. Additional information regarding the NASA Task Load Index can be found in Appendix E. Version C was rated as having the lowest workload although all three displays were within two standard deviations indicating no statistically significant difference among the displays as shown in Figure 10.12.

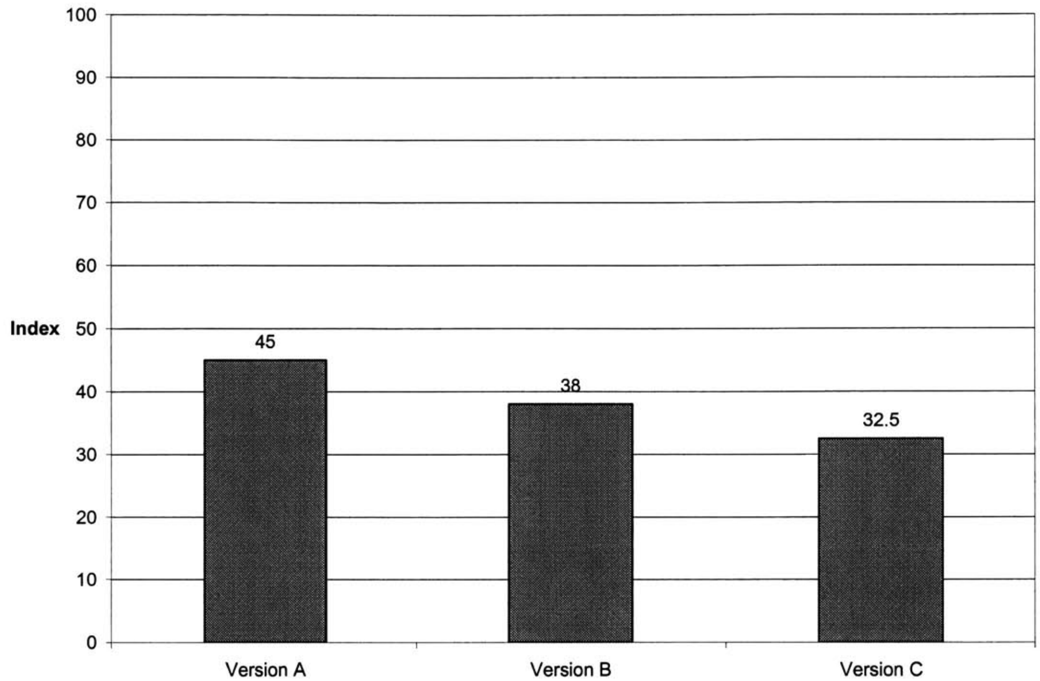


Figure 10.12: Mean Composite NASA-TLX Rating

10.10 Flying Tasks

Each subject was assigned to fly 295 degrees, 120 knots and 3000 feet while simultaneously performing the required user interface tasks. The subjects' heading, airspeed and altitude were recorded throughout the evaluation and a statistical analysis was performed to determine if there was any statistical difference between the displays. No correlation was found ($p > 0.05$), indicating that although performance in completing tasks on the interface varied between formats, this did not translate into an effect on flying performance.

10.11 GOMS Key Level Model Predictions

To validate the use of the GOMS key level model in this application, the predicted execution times were compared to the observed execution times. Only error free tasks were compared since errors require more control inputs preventing direct comparison. Predicted execution time was calculated using the following values from Chapter 5:

- Keying (K) = 0.5 sec

- Homing (H) = 0.4 sec
- Mentally Preparing (M) = 1.35 sec

The time required to select a value using a dual-concentric knob was determined from repeated trials by different individuals to determine an average value.

- Turning of Dual Concentric Knob (T) = 3.8 sec

The accuracy of the predictions can be seen graphically in Figure 10.13, which plots the predicted vs observed data from Table 10.1.

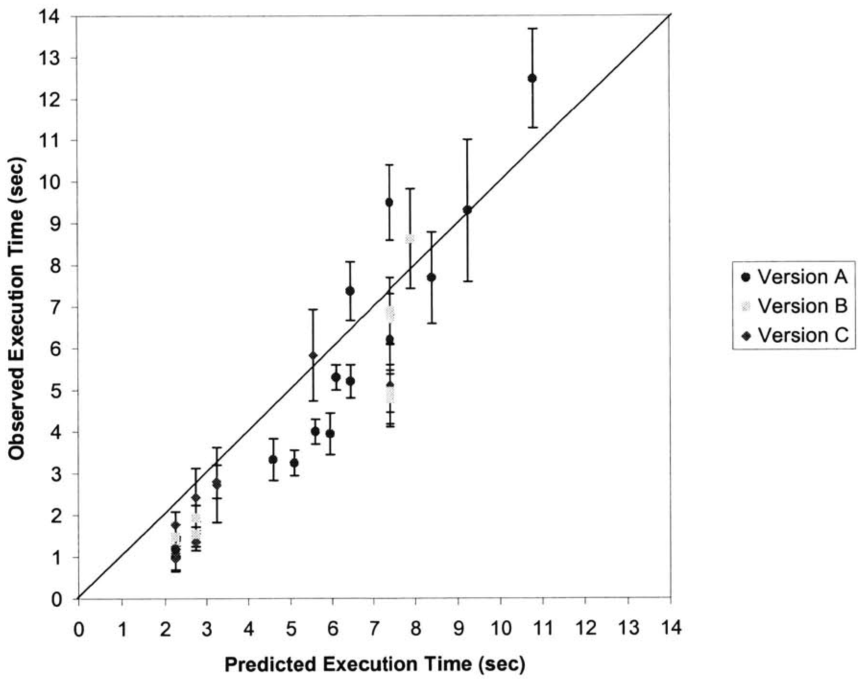


Figure 10.13: Predicted vs. Observed Execution Times

The error bars indicate one standard error from the mean observed execution time. For each task the standard error of estimation of the population mean for samples of size "n" was determined as SE = SD / $(n)^{1/2}$. Standard error values are listed in Table 10.1.

Table 10.1: Calculated and Observed Execution Times

Task - Version	GOMS Analysis	Calculated (sec)	Observed (sec) M	+/- SE	(n)	Pred. Error
RNG - A	H M K M K K M K	6.5	5.2	0.4	4	19%
RNG - C	H M K K K	3.3	2.8	0.4	9	14%
Act NAV Src - A	H M K M K K	4.6	3.3	0.5	9	28%
Act NAV Src - B	H M K	2.3	1.2	0.2	8	46%
Act NAV Src - C	H M K	2.3	1.8	0.3	8	21%
Alt Bug - A	H M K K K M T	8.4	7.7	1.1	11	8%
Alt Bug - B	H M K M T	7.4	4.8	0.6	12	35%
Alt Bug - C	H M K M T	7.4	5.1	1.0	11	31%
HSI View - A	H M K M K M K	6.0	3.9	0.5	11	34%
HSI View - C	H M K	2.3	1.1	0.4	11	52%
Aux NAV Src - A	H M K K K M K K K	6.1	5.3	0.3	8	13%
Aux NAV Src - B	H M K K	2.8	1.9	0.3	9	30%
Aux NAV Src - C	H M K K	2.8	1.4	0.2	9	51%
Hdg Bug - A	H M K M T	7.4	9.5	0.9	12	-28%
Hdg Bug - B	H M K M T	7.4	6.9	0.8	12	7%
Swap - A	H M K	2.3	1.2	0.2	10	47%
Swap - B	H M K	2.3	1.5	0.3	9	34%
Swap - C	H M K	2.3	1.0	0.3	9	58%
Act NAV Crs - A	H M K M K M T	9.3	9.3	1.7	5	-1%
Act NAV Crs - B	H M K M T	7.4	6.7	0.6	9	9%
Baro - A	H M K M T	7.4	6.2	0.6	17	16%
Baro - B	H M K M T	7.4	5.0	0.5	11	33%
Baro - C	H M T	5.6	5.8	1.1	12	-5%
Sby NAV Src - A	H M K K M K K	5.1	3.3	0.3	6	36%
Sby NAV Crs - A	H M K K K M K K M T	10.8	12.5	1.2	8	-16%
Sby NAV Crs - B	H M K K M T	7.9	8.6	1.2	8	-9%
Map Level - A	H M K M K K M K	6.5	7.4	0.7	7	-14%
Map Level - C	H M K K K	3.3	2.7	0.9	9	16%
Sby NAV Fmt - A	H M K K M K K K	5.6	4.0	0.3	5	29%
Sby NAV Fmt - B	H M K K	2.8	1.5	0.3	11	44%
Sby NAV Fmt - C	H M K K	2.8	2.4	0.7	9	12%

In addition, the percentage of prediction error for each task was calculated as follows:

$$\text{Pred. Error} = (T_{calc} - T_{obs}) / T_{calc}$$

where T_{calc} is the calculated or predicted time and T_{obs} is the mean observed time for the task. These values are contained in the right most column in Table 10.1. Comparing the time per task calculated from the model with the observed times gives an RMS (root mean square) error of 30% of the average predicted execution time. RMS error is determined as follows:

$$RMS\ (e) = (\Sigma e_i^2/N)^{1/2}$$

where e_i is the prediction error on the *i*th unit task and "N" is the number of unit tasks [2]. A value of 30% is comparable to the RMS error measurements achieved by Card, Moran and Newell [2] during their initial validation of the keystroke-level model as applied to text editors. This error can be interpreted as the average model error. 30% is high; however predicting execution times for individual tasks is a stringent test. If the unit of prediction were all of the tasks rather than the unit tasks the overall error would be less since the high and low predictions of the unit tasks would tend to cancel each other [2]. The predicted execution time is generally slightly larger than the observed time.

With one exception, the relative execution times of the tasks remained the same for the observed execution times. In other words, if the implementation of a task in Version A was predicted to take longer than the implementation of the same task in Versions B and C, it usually did when actually used. The exception was the implementation of the Baro setting task in Version C. It is possible that the added complexity of the automatic mode change contributed to this increase in execution time.

The GOMS KLM proved useful in determining the relative execution time for various implementations of a task. It was also fairly accurate at determining the actual task execution time. These can be important discriminators when choosing between implementations, however the limitation of the model is its inability to predict error rates. Often error rates are more critical than execution times. The KLM must be applied in conjunction with judgment concerning potential sources of error for various implementations. The advantage of the KLM is that it can be used quantitatively to evaluate design ideas early in the design process without the need for a running system. Also it can be easily applied in conjunction with other methods to fully evaluate a design. The Keystroke-Level Model proved to be an effective, complementary design tool.

11 RECOMMENDATIONS

In all cases the statistically significant results indicate that the implementations of Versions B and C were improvements over Version A. Unfortunately, discrimination between Versions B and C could not be made during this evaluation due to the implemented hardware constraints that prevented full implementation of the configurations.

An analysis of the errors and subjective comments from the test participants reveals potential improvements to Versions B and C. These recommended improvements are as follows:

- Highlight the standby and auxiliary navigation sources with a box when changing the format or source (as in Version A)
- Provide the options when changing the standby format and auxiliary sources (as in Version A)
- Highlight and enlarge the parameter being changed by the knob i.e. the heading bug digital readout when in heading mode
- Increase the font size of the range scale for improved visibility and readability

In addition, several label changes are recommended. These label changes implemented in Version B are illustrated in Figure 10.

- Change the "HSI SRC" label to "NAV SRC" and annunciate the selected source i.e. "NAV1"
- Change the "SRC SEL" label to "SWAP"
- Annunciate the current format of the "SBY FMT" i.e. "RMI"
- Change the "AUX BRG" label to "AUX SRC" and annunciate the current source i.e. "GPS1"
- Change the "MAP LVL" label to "Declutter" and add a state level indicator such as 1 through 5

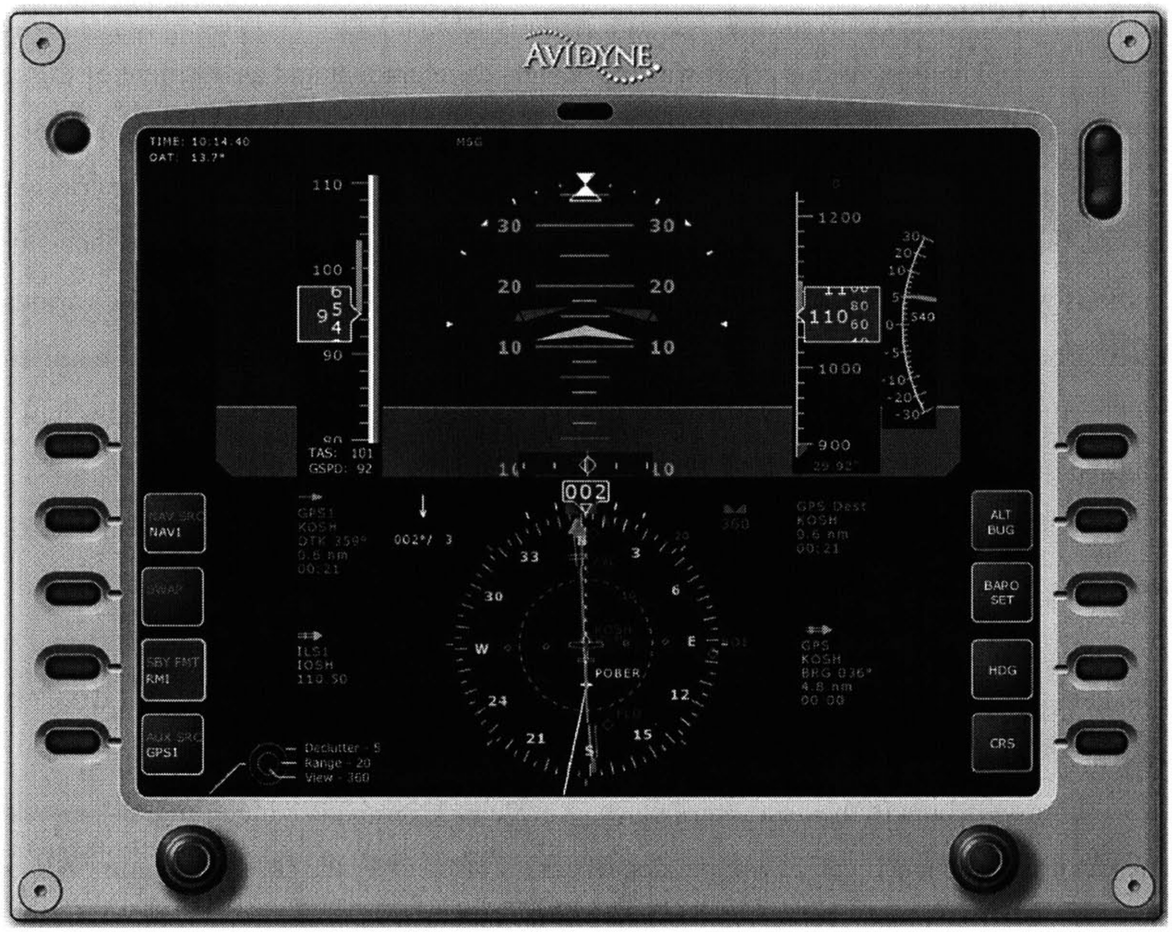


Figure 11.1: Recommended Label Changes Implemented in Version B

The above recommendations apply to both Versions B and C. In order to discriminate between these similar versions and explore more thoroughly the best use of a dual-concentric left knob, this experiment should be repeated once a fully functional bezel is available. In any case both Versions B and C are demonstrably better than the original NASA-AGATE demonstration user interface.

12 CONCLUSIONS

The goal of this effort was the design, development and assessment of the user interface for the Phase I primary flight display/horizontal situation indicator. This included both hardware and software aspects of the user interface. Through the use of a structured design process, this goal was achieved with a new integrated hardware-software user interface. Multiple iterations were key to satisfying the requirements and gaining concurrence from the various design groups. Flexibility was also needed to achieve a balance between technical requirements and the corporate top-level business strategy for a competitive product. The generation of unbiased quantitative and qualitative data through human subject evaluation was an indispensable part of the design process. This can prevent costly oversights and modifications later in the process as well as aid in certification. Ultimately it is hoped that this effort will positively affect the first practical general aviation Highway-in-the-Sky flight display and serve as an example of multiple best practices that can be implemented in future product development efforts.

Improvements to the user interface were made in several areas including task execution time, accuracy and a subjective comparison of ease of use. Over the six tasks common to all three versions, the mean task execution time for the demonstration configuration (Version A) was 37.6 seconds compared with 23.6 seconds and 22.2 seconds for Versions B and C respectively. In addition the accuracy of setting the standby NAV format task was significantly better in Versions B and C than in Version A. In a redundant paired comparison of the three versions based upon ease of use, Versions B and C were significantly better than Version A.

Several general principles can be drawn from the superior performance of Versions B and C over Version A that may be useful in future design efforts.

- Submenus should be avoided whenever possible. They can increase complexity and execution time. Also submenus can hide functions from the user.
- Hidden or non-intuitive interactions should be avoided such as the interaction between the knob mode selection and NAV display selection when setting a course in Version A.
- Confine tasks to as few different controls as possible. In Versions B and C tasks were executed with a maximum of two different controls. Versions A often required three different controls to be manipulated.

- Group like functions together whenever possible.
- Less options and/or functionality can reduce complexity.

The application of the GOMS Keystroke-Level Model to primary flight display user interface design was validated through a human subject evaluation. The GOMS KLM proved useful in determining the relative execution time for various implementations of a task. It was also fairly accurate at determining the actual task execution time. These can be important discriminators when choosing between implementations, however the limitation of the model is its inability to predict error rates. Often error rates are more critical than execution times. The KLM must be applied in conjunction with judgment concerning potential sources of error for various implementations. The advantage of the KLM is that it can be used quantitatively to evaluate design ideas early in the design process without the need for a running system. Also it can be easily applied in conjunction with other methods to fully evaluate a design. The Keystroke-Level Model proved to be an effective, complementary design tool.

13 REFERENCES

[1] "Advanced Technology Aircraft Safety Survey Report," Department of Transport and Regional Development, Bureau of Air Safety Investigation, June 1998.

[2] Card, S., Moran, T., Newell, A., The Psychology of Human-Computer Interaction, Lawrence Erlbaum Associates, Inc., 1983.

[3] Clausing, D., Hauser, J., "The House of Quality," *Harvard Business Review*, Vol 66, No. 3, pp. 63-73, 1988.

[4] Crane, J., Bang, E., Hartel, M., "Standardizing Interactive Display Functions on the 777 Flight Deck," SAE Technical Paper 942093, 1994.

[5] Fan T.P., Kuchar J.K., "Evaluation of Interfaces for Pilot-Air Traffic Control Data Link Communications," 18th Digital Avionics Systems Conference, St Louis, MO, October 27-28, 1999.

[6] General Aviation Manufacturers Association, Recommended Practices and Guidelines for Part 23 Cockpit/Flight Deck Design, GAMA Publication No. 10, September 2000.

[7] Gray, W., John, B., Atwood, M., "Project Ernestine: Validating a GOMS Analysis for Predicting and Explaining Real-World Task Performance," Human-Computer Interaction, 8(3), (pp. 237-309), 1993.

[8] Hansman, R., Yang, L., Application of the Analytic Hierarchy Process for Making Subjective Comparisons Between Multiple Automation/Display Options," Preprint, 6th IFAC/IFIP/IFORS/IEA Symposium, June 27-29, 1995, Cambridge, MA.

[9] Hutchins, Edwin, "How a Cockpit Remembers its Speeds," Distributed Cognition Laboratory, Department of Cognitive Science, University of California, San Diego, May 31, 1995.

[10] John, B., "Why GOMS?" Interactions, (pp. 80-89), October 1995.

[11] Lee, E., MacGregor, J., "Minimizing User Search Time in Menu Retrieval Systems," Human Factors, 27(2), (pp. 157-162), 1985.

[12] Lintern G., Talleur D.A., Waite T., "Functional Interface Design for the Modern Aircraft Cockpit," The International Journal of Aviation Psychology, 9(3), (pp. 225-240). Lawrence Erlbaum Associates, Inc., 1999.

[13] Miller, D., "The Depth/Breadth Tradeoff in Hierarchical Computer Menus," Proceedings of the Human Factors Society, 25th Annual Meeting, (pp. 296-300), 1981.

[14] Nagel, D., Wiener, E., Human Factors in Aviation, Academic Press, San Diego, CA, 1988.

[15] Prabhu, P., Prabhu, G., "Human Error and User-Interface Design," Handbook of Human Computer Interaction, Rochester, NY, 1997.

[16] Pugh, S., Creating Innovative Products Using Total Design: The Living Legacy of Stuart Pugh, Addison Wesley Longman, Inc., 1996

[17] Raskin, J., The Humane: Interface New Directions for Designing Interactive Systems, Addison-Wesley, 2000.

[18] Rasmussen J., "Ecological Interface Design for Reliable Human-Machine Systems," The International Journal of Aviation Psychology, 9(3), (pp. 203-223). Lawrence Erlbaum Associates, Inc., 1999.

[19] Riley V., "Developing a Pilot-Centered Autoflight Interface," American Institute of Aeronautics and Astronautics, 2000-01-5598, 2000.

[20] Sanders, M., McCormick, E., Human Factors in Engineering and Design, McGraw-Hill, Inc., 7th Ed, 1993.

[21] Wickens, C., Engineering Psychology and Human Performance, Harper Collins Publishers, 1999.

[22] Williams K. W., "GPS User Interface Design Problems: II," in R.S. Jensen & B. Cox (Eds.), Proceedings on the Tenth International Symposium on Aviation Psychology (pp. 152-157). Ohio State University, Columbus, OH, 1999.

Appendix A Knob Configuration Inputs / Action Comparison

User Interface Tasks	Demonstration Configuration	One Knob Configuration	Two Knobs Configuration	Four Knobs Configuration
Set Baro Setting	4	3	3	3
Set Baro 29.92	3	2	2	2
Set Altitude Bug	5	3	3	3
Sync Alt Bug to Current Alt	4	2	2	2
Set Airspeed Bug*	6	3	3	3
Sync Airspeed Bug	5	2	2	2
Set Vertical Speed Bug*	7	3	3	3
Sync VSI Bug	6	2	2	2
Set Heading Bug	2	3	3	3
Sync Heading Bug	1	2	2	2
Set Course (Active NAV OBS)	4	4	4	4
Sync Course (Active)	3	3	3	3
Set Course (Stby NAV OBS)	5	5	5	5
Sync Course (Stby)	4	4	4	4
Active Source Select GPS1	2	2	2	2
Active Source Select GPS2	3	3	3	3
Active Source Select NAV1	4	4	4	4
Active Source Select NAV2	5	5	5	5
Active Format Select D-Bar	2	2	2	2
Standby Source Select GPS1	3	3	3	3
Standby Source Select GPS2	4	4	4	4
Standby Source Select NAV1	5	5	5	5
Standby Source Select NAV2	6	6	6	6
Standby Format Select D-Bar	3	3	3	3
Standby Format Select RMI	4	4	4	4
Standby Format Select Text	5	5	5	5
Standby Format Select None	6	6	6	6
Auxiliary Source Select GPS1	4	4	4	4
Auxiliary Source Select GPS2	5	5	5	5
Auxiliary Source Select NAV1	6	6	6	6
Auxiliary Source Select NAV2	7	7	7	7
Auxiliary Format Select RMI	5	5	5	5
Auxiliary Format Select Text	6	6	6	6
Auxiliary Format Select None	7	7	7	7
Swap Active and Standby	1	1	1	1
Select Map Menu	1	0	0	0
Set Clutter Setting 1	1	0	0	0
Set Clutter Setting 2	2	1	1	1
Set Clutter Setting 3	3	2	2	2
Set Clutter Setting 4	4	3	3	3
View 360	1	0	0	0
View 120	2	1	1	1
Increase Range	2	2	2	1
Decrease Range	2	2	2	1
GPS Hold	1	1	1	1
Revert	1	1	1	1
On/Off	1	1	1	1
Brightness Increase	1	1	1	1
Brightness Decrease	1	1	1	1
Avg # inputs / action	3.4	2.9	2.9	2.8

Notes:
*Denotes tasks to be implemented in future configurations.
For the Demonstration Configuration, assumes the knob function always starts in HDG mode and the top menu is displayed.
Adjusting both the outer knob and the inner knob counts as two inputs.

Appendix B User Interface Experiment

Experimental Design

1.1 User Interface Versions

User Interface Version A – NASA/AGATE Demonstration Unit
User Interface Version B – Proposal B
User Interface Version C – Proposal C

1.2 Counterbalancing

Counterbalancing of the experiment to account for learning, fatigue and other order effects.

	Subjects					
Display	**1**	**2**	**3**	**4**	**5**	**6**
1	A	A	B	B	C	C
2	B	C	A	C	A	B
3	C	B	C	A	B	A

1.3 Subject Pool

Six rated pilots with instrument experience.

2. Subject Instructions

2.1 The purpose of this experiment is to evaluate three different user interfaces for a Primary Flight Display. You will perform a series of user interface tasks while performing straight and level flight at an assigned heading, altitude and airspeed. The entire experiment lasts approximately 2 hours and 30 minutes.

2.2 Read the consent form and sign upon agreement.

2.3 Subject completes pilot background and experience questionnaire.

2.4 User Interface 1 training. The subject is familiarized with all of the available user interface options, the PFD flight display and the joystick. The subject may ask any questions and has the opportunity to practice interacting with the display to include flying until the subject feels comfortable. Time to complete training ____________.

2.5 Subject performs flight profile (approx 20 minutes) using UI 1.

2.6 Subject completes subjective questionnaire part I (questions relating specifically to User Interface 1).

2.7 Subject completes NASA Task Load Index (TLX) Sources of Workload Evaluation.

2.8 User Interface 2 training. The subject is familiarized with all of the available user interface options, the PFD flight display and the joystick. The subject may ask any questions and has the opportunity to practice interacting with the display to include flying until the subject feels comfortable. Time to complete training ____________.

2.9 Subject performs flight profile using UI 2.

2.10 Subject completes subjective questionnaire part I (questions relating specifically to UI 2).

2.11 Subject completes NASA Task Load Index (TLX) Sources of Workload Evaluation.

2.12 User Interface 3 training. The subject is familiarized with all of the available user interface options, the PFD flight display and the joystick. The subject may ask any questions and has the opportunity to practice interacting with the display to include flying until the subject feels comfortable. Time to complete training ____________.

2.13 Subject performs flight profile using UI 3.

2.14 Subject completes subjective questionnaire part I (questions relating specifically to UI 3).

2.15 Subject completes NASA Task Load Index (TLX) Sources of Workload Evaluation.

2.16 Subject completes subjective questionnaire part II (overall).

Informed Consent Statement

User Interface Experiment

Student Researcher:
Brent Campbell
781-274-8432

Principle Investigator:
Prof. J. Kuchar
MIT Rm. 33-305
77 Massachusetts Ave.
Cambridge, MA 02139

Your participation in this experiment is voluntary. You may halt the experiment at any time and withdraw from the study for any reason without prejudice. You will remain anonymous in any report, which describes this work. If you have any questions concerning the purpose, procedures, or risks associated with this experiment, please ask them. The computer will record the flight data. These recordings will be used only for the purposes of this study and you will not be identified in the analysis or presentation of results.

The purpose of this experiment is to perform human factors evaluations of three Primary Flight Display user interfaces. Each pilot will be given the opportunity to practice using each of the displays directly before performing the experiment. The pilot will be required to fly a specified altitude, heading and direction while performing specified UI tasks. The performance of the flight will be gauged by a subjective questionnaire, a pilot determined NASA task load index for each user interface and a variety of parameters recorded by Microsoft Flight Simulator.

CONSENT

In the unlikely event of physical injury resulting from participation in this research, I understand that medical treatment will be available from the MIT Medical Department, including first aid, emergency treatment and follow-up care as needed, and that my insurance carrier may be billed for the cost of such treatment. However, no compensation can be provided for medical care apart from the forgoing. I further understand that making such medical treatment available, or providing it, does not imply that such injury is the Investigator's fault. I also understand that by my participation in this study, I am not waiving any of my legal rights.*

I understand that I may also contact the Chairman of the Committee on the Use of Humans as Experimental Subjects, MIT 253-6787, if I feel I have been treated unfairly as a subject.

I volunteer to participate in this experiment that involves flying a PC based flight simulator with several user interface configurations, for a total of approximately 1.5 hours. I understand that I may discontinue my participation at any time. I have been informed as to the nature of this experiment, and agree to participate in the experiment.

____________________ ____________________

Date Signature

* Further information may be obtained by calling the Institute's Insurance and Legal Affairs Office at 253-2822.

Pilot Background and Experience Questionnaire

Personal Data

Name:

Age:

Gender M/F:

PILOT EXPERIENCE

Total Hours:

Ratings:

Primary A/C:

Jet Experience:

Military Flight Experience:

Experience with glass cockpit displays:

Subjective Questionnaire Part I (display specific) Part II (overall)

Part I

1. What are the best features of the User Interface?

User Interface A

User Interface B

User Interface C

2. What are the worst features of the User Interface?

User Interface A

User Interface B

User Interface C

3. What specific things would you like to see different?

User Interface A

User Interface B

User Interface C

Part II

1. Using the scale below rate the three interfaces based on their ease of use?

User Interface A | User Interface C

Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

User Interface B | User Interface C

Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

User Interface A | User Interface B

Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

2. Do you have any general comments?

User Interface Tasks Worksheet – Version A

Tasks	Control Inputs
1. Set the HSI map range to 50 nm.	Min: 4, Actual_____
2. Set the active NAV display source to GPS2.	Min: 3, Actual_____
3. Set the altitude bug to 3000'.	Min: 4, Actual_____
4. Set the HSI view to ARC.	Min: 3, Actual_____
5. Set the Aux NAV display source to GPS2.	Min: 6, Actual_____
6. Set the Heading bug to 120°.	Min: 2, Actual_____
7. Swap the active and standby NAV sources.	Min: 1, Actual_____
8. Set the active course to 175°.	Min: 3, Actual_____
9. Set the active NAV source to GPS1.	Min: 5, Actual_____
10. Set the HSI map range to 10 nm.	Min: 4, Actual_____
11. Set the Baro to 30.16.	Min: 2, Actual_____
12. Set the standby NAV source to ILS1.	Min: 4, Actual_____
13. Set the standby NAV course to 311°.	Min: 6, Actual_____
14. Set the HSI map to display no icons.	Min: 6, Actual_____
15. Set the Baro to 29.92.	Min: 2, Actual_____
16. Set the HSI view to 360°.	Min: 4, Actual_____
17. Set the standby NAV format to text only.	Min: 5, Actual_____
18. Set the heading bug to 317°.	Min: 3, Actual_____
19. Set the HSI map to display all icons.	Min: 3, Actual_____
20. Set the standby NAV format to D-bar.	Min: 5, Actual_____
21. Set the standby NAV course to 060°.	Min: 4, Actual_____
22. Swap the active and standby NAV sources.	Min: 1, Actual_____
23. Set the active NAV course to 210°.	Min: 2, Actual_____
24. Set the Aux NAV display source to ILS1.	Min: 5, Actual_____
25. Set the altitude bug to 2000'.	Min: 3, Actual_____
26. Set the Baro to 30.01.	Min: 4, Actual_____

User Interface Tasks Worksheet– Version B

	Tasks	Control Inputs
1.	Set the active NAV display source to GPS2.	Min: 1, Actual_____
2.	Set the altitude bug to 3000'.	Min: 2, Actual_____
3.	Set the Heading bug to 060°.	Min: 2, Actual_____
4.	Set the Aux NAV display source to GPS2.	Min: 2, Actual_____
5.	Set the active NAV source to GPS1.	Min: 1, Actual_____
6.	Set the standby NAV course to 060°.	Min: 3, Actual_____
7.	Set the Baro to 30.16.	Min: 2, Actual_____
8.	Set the heading bug to 273°.	Min: 2, Actual_____
9.	Set the standby NAV format to off.	Min: 3, Actual_____
10.	Set the standby NAV format to D-Bar.	Min: 1, Actual_____
11.	Set the standby NAV course to 176°.	Min: 3, Actual_____
12.	Select ILS1 as the active NAV display source.	Min: 1, Actual_____
13.	Set the active course to 170°.	Min: 2, Actual_____
14.	Set the Aux NAV display source to ILS1.	Min: 1, Actual_____
15.	Set the active course to 235°.	Min: 1, Actual_____
16.	Select GPS1 as the active NAV display source.	Min: 1, Actual_____
17.	Set the altitude bug to 2000'.	Min: 2, Actual_____
18.	Set the Baro to 30.01.	Min: 2, Actual_____

User Interface Tasks Worksheet– Version C

	Tasks	Control Inputs
1.	Set the HSI map range to 100 nm.	Min: 3, Actual_____
2.	Set the active NAV display source to GPS2.	Min: 1, Actual_____
3.	Set the altitude bug to 3000'.	Min: 2, Actual_____
4.	Set the HSI view to ARC.	Min: 1, Actual_____
5.	Set the Aux NAV display source to GPS2.	Min: 2, Actual_____
6.	Set the active NAV source to GPS1.	Min: 1, Actual_____
7.	Set the HSI map range to 10 nm.	Min: 3, Actual_____
8.	Set the Baro to 30.16.	Min: 1, Actual_____
9.	Set the HSI map to display no icons.	Min: 4, Actual_____
10.	Set the HSI view to 360°.	Min: 2, Actual_____
11.	Set the standby NAV format to RMI.	Min: 1, Actual_____
12.	Set the HSI map to display all icons.	Min: 1, Actual_____
13.	Set the standby NAV format to D-Bar.	Min: 3, Actual_____
14.	Select ILS1 as the active NAV display source.	Min: 1, Actual_____
15.	Set the Aux NAV display source to ILS1.	Min: 1, Actual_____
16.	Select GPS1 as the active NAV display source.	Min: 1, Actual_____
17.	Set the altitude bug to 2000'.	Min: 2, Actual_____
18.	Set the Baro to 30.01.	Min: 1, Actual_____

Starting States of User Interfaces

Version A

Active NAV: GPS1
Standby NAV: ILS1
Standby NAV: Format: D-Bar
Aux NAV: Off
Range: 10 nm
View: 360°
Map Level: All icons displayed
Altitude Bug: 0
Baro Setting: 29.92
Knob Mode: Heading

Versions B and C

HSI Source: GPS1
Source Select: GPS1
Standby NAV Format: D-Bar
Aux NAV: Off
Range: 10 nm (N/A for Ver B)
View: 360° (N/A for Ver B)
Map Level: All icons displayed (N/A for Ver B)
Altitude Bug: 0
Baro Setting: 29.92
Knob Mode: Heading

Appendix C Observed Execution Times

Table C.1: Observed Execution Times Including Errors

Observed Execution Times Including Errors

Task - Version	Subject 1			Subject 2			Subject 3			Subject 4			Subject 5			Subject 6			Mean
RNG - A	7.2	6.5		8.3	10		5.9	5.2		11	4.4		6.3	4.9		16.8	11.2		8.1
RNG - C	2.1	2.5		3.4	2.5		1.3	3.5		3.6	2.5		1.9	5.2		2	1.9		2.7
Act NAV Src - A	3.1	4.5		7.8	3.8		1.5	2.3		2	4.2		2.3	4.7		5.4	4.7		3.9
Act NAV Src - B	2.1	1.8		0.8	0.9		7.3	1.8		0.7	0.5		2.8	1.4		1.8	1.5		2.0
Act NAV Src - C	0.4	4.1		2.1	1.3		3.2	2.1		2.1	0.7		8.2			2.3	2.1		2.6
Alt Bug - A	10.1	13		10.6	6.3		4.8	3.2		13.1	5.1		10.1	4		14.5	4.3		8.3
Alt Bug - B	5.6	2.6		5.7	1.7		4.8	2.3		7.5	2.9		6.6	6.6		6.9	4.2		4.8
Alt Bug - C	4.8	2.4		5.7	4.1		5.5	2.3		13.7	1.5		6.6	3.9		7.9	4.4		5.2
HSI View - A	2.1	7.2		1.8	3.8		4.8	4.1		3.9	5.7		2.9	3.2		3.9	5.7		4.1
HSI View - C	0.3	1		0.4	1.3		0.5	0.6		4.7	1.2		0.4	7.4		0.7	0.9		1.6
Aux NAV Src - A	5	5		39.4	4.8		3.9	5.5		5.9	6.8		10.3	8.8		5.7	5.3		8.9
Aux NAV Src - B	21	2.5		2.8	0.9		1.6	20.2		3	0.8		39.9	2.6		2	1.2		8.2
Aux NAV Src - C	1.8	0.7		1	6.4		3	1.2		1	1.7		23.3	1.3		5	0.5		3.9
Hdg Bug - A	15.3	12.7		10.6	7.3		5.1	6.2		11	13.4		7.5	7.3		8.7	8.8		9.5
Hdg Bug - B	5.7	11		3.5	6.6		5.7	4.7		8.9	11		3.1	7.9		8.6	6		6.9
Swap - A	1.1	1.7		1.3	0.8		0.8	0.3		1.7	0.4		5.8	2.2		3.7	1.6		1.8
Swap - B	0.9	1.4		3.4	0.9		3.7	2.7		1.6	1.4		1	2.1		1.1	1.6		1.8
Swap - C	0.9	0.7		0.3	2.8		0.2	0.3		0.6	1.1		2.7	1.8		8.5	2.7		1.9
Act NAV Crs - A	10.1	7.5		9.9	13.1		7.4	5.4		14.3	23.4		20.8	10.4		10.9	15.3		12.4
Act NAV Crs - B	7.6	9.1		5.8	4.8		5.2	4.2		13.4	6.7		7.4	7.2		9.6	10.9		7.7
Baro - A	7.7	11.5	3.4	4.1	4.4	3.2	4.3	5		10.2	5.6	6.9	5.6	9.8	3.1	9.2	6.7	4.8	6.3
Baro - B	4.4	3.2		9.9	3.7		3	4.8		8.9	5		3.6	5		6.1	6.9		5.4
Baro - C	3.5	6.8		4.1	8.7		3	4.4		16.7	4.1		3.8	4.2		5.1	5.6		5.8
Sby NAV Src - A	3.9			3.4			2.4			2.9			4			2.9			3.3
Sby NAV Crs - A	10.1	8.9		31.2	15.3		11.2	20.3		18.1	11.2		17	8.9		16.1	10.1		14.9
Sby NAV Crs - B	12.3	4.8		6.1	8.5		21	10.6		12.5	12.4		13.5	5.5		13.2	6.2		10.6
Map Level - A	17.8	8.6		6.8	4.1		11.2	5.6		8.4	8.7		26.8	9.8		8.6	9.4		10.5
Map Level - C	26.8	1.1		3.5	2.3		2.8	0.6		9.1	0.8		8.3	12.3		2.5	1.8		6.0
Sby NAV Fmt - A	4	5.3		20.3	3.7		18.4	15.3		4.4	12.7		3.4	4.5		3.6	5.4		8.4
Sby NAV Fmt - B	2.9	0.2		1.8	0.8		0.7	0.8		3.6	0.9		1.6	0.1		2.8	0.6		1.4
Sby NAV Fmt - C	1.1	1.7		2.1	3.2		0.3	1.3		7.9	2.8		3.5	1.4		4.9	3.6		2.8

Table C.2: Observed Execution Times Not Including Errors

Observed Execution Times Not Including Errors

Task - Version	Subject 1			Subject 2			Subject 3		Subject 4			Subject 5			Subject 6			Mean	Std Dev
RNG - A								5.2		4.4		6.3	4.9					5.2	**0.8**
RNG - C	2.1	2.5			2.5		1.3	3.5	3.6	2.5		1.9	5.2					2.8	**1.2**
Act NAV Src - A	3.1	4.5					1.5	2.3	2	4.2		2.3	4.7		5.4			3.3	**1.4**
Act NAV Src - B		1.8		0.8	0.9			1.8	0.7	0.5					1.8	1.5		1.2	**0.6**
Act NAV Src - C	0.4			2.1	1.3		3.2	2.1		0.7					2.3	2.1		1.8	**0.9**
Alt Bug - A	10.1	13		10.6	6.3		4.8	3.2	13.1	5.1		10.1	4			4.3		7.7	**3.7**
Alt Bug - B	5.6	2.6		5.7	1.7		4.8	2.3	7.5	2.9		6.6	6.6		6.9	4.2		4.8	**2.0**
Alt Bug - C	4.8	2.4		5.7	4.1		5.5	2.3	13.7	1.5			3.9		7.9	4.4		5.1	**3.4**
HSI View - A	2.1	7.2		1.8	3.8		4.8	4.1	3.9	5.7		2.9	3.2		3.9			3.9	**1.6**
HSI View - C	0.3	1		0.4	1.3		0.5	0.6	4.7	1.2		0.4			0.7	0.9		1.1	**1.2**
Aux NAV Src - A	5	5			4.8		3.9		5.9	6.8					5.7	5.3		5.3	**0.9**
Aux NAV Src - B		2.5		2.8	0.9		1.6		3	0.8			2.6		2	1.2		1.9	**0.8**
Aux NAV Src - C	1.8	0.7		1			3	1.2	1	1.7			1.3			0.5		1.4	**0.7**
Hdg Bug - A	15.3	12.7		10.6	7.3		5.1	6.2	11	13.4		7.5	7.3		8.7	8.8		9.5	**3.1**
Hdg Bug - B	5.7	11		3.5	6.6		5.7	4.7	8.9	11		3.1	7.9		8.6	6		6.9	**2.6**
Swap - A	1.1	1.7		1.3	0.8		0.8	0.3	1.7	0.4			2.2			1.6		1.2	**0.6**
Swap - B	0.9	1.4		3.4	0.9				1.6	1.4		1			1.1	1.6		1.5	**0.8**
Swap - C	0.9	0.7		0.3			0.2	0.3	0.6	1.1		2.7	1.8					1.0	**0.8**
Act NAV Crs - A		7.5					7.4	5.4							10.9	15.3		9.3	**3.9**
Act NAV Crs - B	7.6	9.1		5.8	4.8		5.2	4.2		6.7		7.4			9.6			6.7	**1.9**
Baro - A	7.7	11.5	3.4	4.1	4.4	3.2	4.3	5	10.2	5.6	6.9	5.6	9.8	3.1	9.2	6.7	4.8	6.2	**2.6**
Baro - B	4.4	3.2			3.7		3	4.8	8.9	5		3.6	5		6.1	6.9		5.0	**1.8**
Baro - C	3.5	6.8		4.1	8.7		3	4.4	16.7	4.1		3.8	4.2		5.1	5.6		5.8	**3.8**
Sby NAV Src - A	3.9			3.4			2.4		2.9			4			2.9			3.3	**0.6**
Sby NAV Crs - A	10.1	8.9			15.3		11.2		18.1	11.2			8.9		16.1			12.5	**3.5**
Sby NAV Crs - B	12.3	4.8		6.1	8.5					12.4			5.5		13.2	6.2		8.6	**3.5**
Map Level - A				6.8	4.1			5.6	8.4	8.7					8.6	9.4		7.4	**1.9**
Map Level - C		1.1		3.5	2.3		2.8	0.6	9.1	0.8					2.5	1.8		2.7	**2.6**
Sby NAV Fmt - A	4	5.3			3.7							3.4			3.6			4.0	**0.8**
Sby NAV Fmt - B	2.9	0.2		1.8	0.8			0.8	3.6	0.9		1.6	1		2.8	0.6		1.5	**1.1**
Sby NAV Fmt - C	1.1	1.7		2.1	3.2		0.3	1.3	7.9	2.8			1.4					2.4	**2.2**

Appendix D Subjective Data

Subjective Questionnaire Part 1 (display specific) Part II (overall)

Part 1

1. What are the best features of the User Interface?

User Interface A
Attitude indicator is life- like. (eg. Work's like mechanical indicator), coarse and fine knobs worked fine, multipress of NAV button, SRC & FMT button to select were intuitive. The HDG The box around the active, stand by, aux was a good interface. 5 second refresh was almost long enough.

User Interface B
Less confusing multiple use of selection buttons compared to A
Better distinction of function (e.g. less multiplicity of function) with one exception – e.g. much better alt bug, baro set, HDG & CRS
Lack of confusion

User Interface C
Having map functionality available full-time so as not having to use "menu" function of A interface same positives as user interface "B"

2. What are the worst features of the User Interface?

User Interface A
Having to go back to standard menu after selecting map for options was not as intuitive. The view and map repetitive hits on the button were not as intuitive as could be for declutter, upclutter and range.

User Interface B
"AUX BRG" was an inappropriate label for the function of that button. "HSI SRC" and "SRC SEL" require too much pilot compensation to figure out that "HSC SRC" was radio box 1 and "SRC SEL" was really for GPS vs ILS Source – overcome somewhat with TNG

User Interface C
Note: There is a learning curve in the order of displays such that annoying features of "B" are somewhat less annoying in "C" "Aux Brg" labeling still counter intuitive "Aux SRC" or, if you must "Aux BRG SRC" would be better, some confusions between "view" and "MAP LVL", maybe "declutter"

3. What specific things would you like to see different?

User Interface A
The map interface command selections were more difficult than the main menu selections. There should be more harmony in controls of selection processes between the two

User Interface B
As noted in comments for 1 & 2 ??
Labeling of "Aux Brg" should be more representative of what the button really does to a less strident extent, labeling of "HIS SRC" and "SRC SEL" could be somewhat more descriptive as well.

User Interface C
"Map LVL' not explicit enough "Declutter"
"Aux BRG" as already stated
Consideration of "HIS SRC" & "SRC SEL" for enhanced labeling clarity /accuracy

Part II

1. Using the scale below rate the three interfaces based on their ease of use?

User Interface A — User Interface C

				x
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

User Interface B — User Interface C

			x	
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

User Interface A — User Interface B

			x	
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

2. Do you have any general comments?

Learning curve has some influence, particularly B to C.
Labeling of buttons as stated could use some enhancements.

I think the general thrust of what is being implemented here for general aviation cockpits is excellent and you should be applauded for your efforts.

Subjective Questionnaire Part 1 (display specific) Part II (overall)

Part 1

1. What are the best features of the User Interface?

User Interface A
Only one "function" per button in NAV mode and only one "function" per button in map mode makes the memory exercise much easier and the less likely to make a mistake, although not having to switch modes would be better.

User Interface B
Easier because the heading, baro, alt functions are readily accessible, if the map selection functions are still available this is probably the best of the three

User Interface C
Easier than Interface A because there is no mode swapping, but having map level and view above the range buttons might have been easier for me.

2. What are the worst features of the User Interface?

User Interface A
The timeout feature for NAV characteristic selection is annoying. A nice "enter" key to signify the end of a sequence of keystrokes would be more natural for me, as well as an "abort" sequence key.

User Interface B
The position of the NAV CRS selection on the right side, opposite the other NAV functions might take some getting used to.

User Interface C
Notation for the Nav modes. Why make it more complex than it is. (see comments). As in UI A, having to select map features with several push buttons, that only select "in one direction" means that passing the desired setting accidentally requires a complete "circuit."

3. What specific things would you like to see different?

User Interface A
It would be easier for me to think in terms of primary, secondary, and auxiliary NAVs rather than your designations.

User Interface B
N/A

User Interface C

Instead of HSI SRC, thinking in terms of NAV1 and NAV2 would be more natural. Instead of SRC SEL, thinking in terms of "swap" would be more intuitive. Instead of "SBY FMT" using Primary and Secondary would be much more intuitive for me (see notes for UI A).

Part II

1. Using the scale below rate the three interfaces based on their ease of use?

User Interface A / User Interface C

Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use
			x	

User Interface B / User Interface C

Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use
	x			

User Interface A / User Interface B

Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use
				x

2. Do you have any general comments?

I'm surprised that I liked using these interfaces, I expected something much more obscure as in the KLN 89B multifunction buttons and knobs.

Subjective Questionnaire Part 1 (display specific) Part II (overall)

Part 1

1. What are the best features of the User Interface?

User Interface A
NAV source select very clear

User Interface B
Knob for Adjusting HDG bug, Altitude Bug, CRS, BARO

User Interface C
NAV adjustments at top level

2. What are the worst features of the User Interface?

User Interface A
All pretty nice. Confusion in the menu over two middle RHS buttons

User Interface B
Choosing NAV source is confusing without experience

User Interface C
NAV selection takes some learning, and have to look carefully to set ALT Bug or Barometer

3. What specific things would you like to see different?

User Interface A
Label "view" and "map" with more meaningful names

User Interface B
Make it clearer which NAV is displayed where. Maybe dedicated buttons so that whenever one presses button "A" you get consistent NAV display and SRC

User Interface C
Change all bug/barometer cueing to make it more obvious what is being set.

Part II

1. Using the scale below rate the three interfaces based on their ease of use?

User Interface A — User Interface C

Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use
			x	

User Interface B — User Interface C

Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use
	x			

User Interface A — User Interface B

Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use
				x

2. Do you have any general comments?
Nice Display!

Subjective Questionnaire Part 1 (display specific) Part II (overall)

Part 1

1. What are the best features of the User Interface?

User Interface A
Swap feature is nice

User Interface B
LCD is nice, tapes are nice (velocity and alt)...but need time learning

User Interface C
Same as B. Ring range is nice. Names of buttons on right –hand side are better than a left hand side.

2. What are the worst features of the User Interface?

User Interface A
Setting course for other than active NAV needs further learning. Really don't like this VSI.

User Interface B
Name for HSI-SRC & SRC SEL Buttons. VSI interface.

User Interface C
Declutter using view B nice, but needs learning.

3. What specific things would you like to see different?

User Interface A
Course set differently for other than active NAV

User Interface B
Different names for comments #1 in 2. Normal VSI.

User Interface C
Same as with B. Automatic decluttering with MAP LVL change (especially when zooming out.)

Part II

1. Using the scale below rate the three interfaces based on their ease of use?

User Interface A | User Interface C

		x		
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

User Interface B | User Interface C

		x		
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

User Interface A | User Interface B

		x		
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

2. Do you have any general comments?
Nice features of some were blended with less nice features.

Subjective Questionnaire Part 1 (display specific) Part II (overall)

Part 1

1. What are the best features of the User Interface?

User Interface A
Better designation of active/standby/aux during source/course changes; was usually apparent which thing I was changing.
Menu function for source data display is nice

User Interface B
Discrete buttons for HDG/CRS/baro/ALT are nice, but probably a waste

User Interface C
Visible feedback from user entries, menu selection buttons
Highlighting of selected input fields (ALT select, for example)
Display symbology is clear and easy to read
Controls for map display manipulation were clear; it was (usually) obvious what effect a button push might have

2. What are the worst features of the User Interface?

User Interface A
Still a little confusing remembering which NAV /source data block was Pri/ backup/aux
Selection of CRS/baro/ALT using concentric knobs is a little confusing
Too many better pushes to get to map data functions; should be on same page as other buttons

User Interface B
Selection "HSI SRC," "SRC SEL" and "SBY FMT" are still confusing: back to a multiple-level selection, took more time to think about what I was selecting

User Interface C
Correlation of "HSI SRC" button entries with NAV 1/2, source selection is less than clear; had to devote extra time to select right NAV unit, <u>then</u> select source from that unit; multiple-level task was a little harder.
Correlation of "SBY FMT" button pushes with changes in D-Bar/RMI/ text format was less than clear
Interaction of map display with selected NAV/SRC wasn't apparent

3. What specific things would you like to see different?

User Interface A
Pri/backup/AUX data blocks should be labeled as such

Change function of baro/HDG/ALT knobs so that inner knob doesn't increment other digits (i.e. only controls hundredths of inches in baro, and doesn't change tenths when small knobs is turned from "9" to "0"
Move map display controls back on primary page

User Interface B
Don't waste discrete button on ALT/baro etc – one button that toggles between selections (with visual feedback) was fine
Go back to Interface "A" selection at Pri/Backup/AUX data blocks, use labels

User Interface C
More direct labeling of NAV/source selections; i.e. some sort of clear text identification of primary NAV/source; color-coding isn't so apparent (to me, at least)
Make the "ALT Bug" function consistent – either baro or ALT select should be highlighted so that the selected field is apparent

Part II

1. Using the scale below rate the three interfaces based on their ease of use?

User Interface A				User Interface C
	x			
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

User Interface B				User Interface C
			x	
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

User Interface A				User Interface B
	x			
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

2. Do you have any general comments?
The amount of instability in the sim was good enough to make some selections a chore. Because I wasn't used to the display, making multi-level selections (NAV, then SRC) was more difficult due to the division of time between controlling the aircraft and figuring out which HSI select/source select button did what? Clear labeling (in addition to color-coding) of Pri/Backup/AUX data blocks would help.

Subjective Questionnaire Part 1 (display specific) Part II (overall)

Part 1

1. What are the best features of the User Interface?

User Interface A
Menu Pop up when different options were available
Swap function

User Interface B
I liked the new function options on this display were more useful
Map declutter and sale on knob ring were better placed

User Interface C
Easy to learn. Feel of buttons was excellent.

2. What are the worst features of the User Interface?

User Interface A
Two pages of menus – particularly for functions you may use quite regularly

User Interface B
Leaving default active after an input is made.
No fault tolerance if an inadvertent input is made.

User Interface C
Wording is unclear on button labels (HIS SRC vs, SRC SEL), how about "active" and "source"?
Range scale value is too small while adjusting; also, "SBY FMT" & "AUX BRG" = "STBY" & "AUX"
Altimeter Readout was unclear for current altitude. Range of display is distracting. White dashed bar blocks analog type heading info.
Digital Heading is too large and overpowering. Inconsistent w/normal flying tolerances.

2. What specific things would you like to see different?

User Interface A
"NAV" button has poor label – make "P/S/A" for primary/secondary/auxiliary
Single Button toggle between HDG/CRS/ALT
These do not intuitively go together by their function.
Grouping by interface B more intuitive when all else is grouped by function
Swap button in a poor location. It should be placed near the item it swaps on left side
Map display functions are poorly placed. These may be used quite frequently. Switching back and forth between menu pages is cumbersome.
Also, incorporate menu options windows info declutter function so it is clear what is happening

User Interface B
+ or - Hold for fast adjustment, single push for find adjustment
As a beginner, moving between function selection and range adjustment at the knob, took away from scan longer than necessary.
When making adjustments, enlarging in addition to highlighting.
Larger buttons, maybe more visible
I wouldn't leave a default adjustment active, inadvertently bumping a knob set to altimeter or course can have big impact. If not caught, immediately, push to accept if a change is inputed for Baro/ALT
Remove knob altogether as noted above, or increase tactile

User Interface C
Would like a bi-directional button vice a single loop.
Rename map level to "clutter", confused with VIEW button with current level.
RNG+
Consolidate to a two position toggle, for example.
RNG-
Attitude Indicator is imprecise. It covers 0-5 degrees in pitch. Where exactly is the aircraft? Where is straight and level? Make it a single ban speed indicator.
GPS ground track line should not run through RMI indication. It can block your heading or draw you off course.

Part II

1. Using the scale below rate the three interfaces based on their ease of use?

User Interface A | User Interface C

				x
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

User Interface B | User Interface C

	x			
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

User Interface A | User Interface B

				x
Much Easier to Use	Easier to Use	Same	Easier to Use	Much Easier to Use

2. Do you have any general comments?
HSI is generally too small for all the data

Buttons could be bigger and more visible (contrast), also backlit
Knob functions could be incorporated into push buttons
Menus should appear in all display options
Display is too small with all the available data. Very difficult to see any symbology with all of the overlays.
Adjusting course for secondary radio was not intuitive.

Appendix E NASA-TLX Information

SUBJECT INSTRUCTIONS: RATINGS (Mouse Version)

We are not only interested in assessing your performance but also the experiences you had during the different task conditions. Right now we are going to describe the technique that will be used to examine your experiences. In the most general sense we are examining the "Workload" you experienced. Workload is a difficult concept to define precisely, but a simple one to understand generally. The factors that influence your experience of workload may come from the task itself, your feelings about your own performance, how much effort you put in, or the stress and frustration you felt. The workload contributed by different task elements may change as you get more familiar with a task, perform easier or harder versions of it, or move from one task to another. Physical components of workload are relatively easy to conceptualize and evaluate. However, the mental components of workload may be more difficult to measure.

Since workload is something that is experienced individually by each person, there are no effective "rulers" that can be used to estimate the workload of different activities. One way to find out about workload is to ask people to describe the feelings they experienced. Because workload may be caused by many different factors, we would like you to evaluate several of them individually rather than lumping them into a single global evaluation of overall workload. This set of six rating scales was developed for you to use in evaluating your experiences during different tasks. Please read the descriptions of the scales carefully. If you have a question about any of the scales in the table, please ask me about it. It is extremely important that they be clear to you. You may keep the descriptions with you for reference during the experiment.

After performing the task, six rating scales will be displayed. You will evaluate the task by marking each scale at the point which matches your experience. Each line has two endpoint descriptors that describe the scale. Note that "own performance" goes from "good" on the left to "bad" on the right. This order has been confusing for some people. Move the arrow to the right or left with the mouse until it points at the desired location. When you are satisfied, press either button to enter your selection. Please consider your responses carefully in distinguishing among the task conditions. Consider each scale individually. Your ratings will play an important role in the evaluation being conducted, thus, your active participation is essential to the success of this experiment, and is greatly appreciated.

SUBJECT INSTRUCTIONS: SOURCES-OF-WORKLOAD EVALUATION (Mouse Version)

Throughout this experiment the rating scales are used to assess your experiences in the different task conditions. Scales of this sort are extremely useful, but their utility suffers from the tendency people have to interpret them in individual ways. For example, some people feel that mental or temporal demands are the essential aspects of workload regardless of the effort they expended or the performance they achieved. Others feel that if they performed well the workload must have been, low, and vice versa. Yet others feel that effort or feelings of frustration are the most important factors in workload; and so on. The results of previous studies have already found every conceivable pattern of values. In addition, the factors that create levels of workload differ depending on the task. For example, some tasks might be difficult because they must be completed very quickly. Others may seem easy or hard because of the intensity of mental or physical effort required. Yet others feel difficult because they cannot be performed well, no matter how much effort is expended.

The evaluation you are about to perform is a technique that has been developed by NASA to assess the relative importance of six factors in determining how much workload you experienced. The procedure is simple: You will be presented with a series of pairs of rating scale titles (for example, Effort vs. Mental Demands) and asked to choose which of the items was more important to <u>your</u> experience of workload in the task(s) that you just performed. Each pair of scale titles will appear separately on the screen<u>. Select- the Scale Title that represents the more important contributor to workload- for the Specific task(s) you performed in this experiment</u>.

Press the left button to select the top item in the pair, and the right button to select the bottom item. A pointer shows which title was selected. To enter that choice, press the button again, and a new pair of titles will appear. If you change your mind, press the other button to cancel your first choice, and then start over.

After you have finished the entire series we will be able to use the pattern of your choices to create a weighted combination of the ratings from that task into a summary workload score. Please consider your choices carefully and make them consistent with how you used the rating scales during the particular task you were asked to evaluate. Don't think that there is any *correct* pattern; we are only interested in your opinions. If you have any questions, please ask them now. Thank you for your participation.

RATING SCALE DEFINITIONS		
Title	Endpoints	Descriptions
MENTAL DEMAND	*Low/High*	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?
PHYSICAL DEMAND	*Low/High*	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?
TEMPORAL DEMAND	*Low/High*	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?
EFFORT	*Low/High*	How hard did you have to work (mentally and physically) to accomplish your level of performance?
PERFORMANCE	*Good/Poor*	How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?
FRUSTRATION LEVEL	*Low/High*	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?

MENTAL DEMAND

Low High

PHYSICAL DEMAND

Low High

TEMPORAL DEMAND

Low High

PERFORMANCE

Good Poor

EFFORT

Low High

FRUSTRATION

Low High