8-1993

Incorporation of Traffic Collision Alert System (TCAS) Advisories on Heads-up Displays: Enhanced Pilot Response

Hamadeh A. Nureddine
Embry-Riddle Aeronautical University - Daytona Beach

INCORPORATION OF TRAFFIC COLLISION ALERT SYSTEM (TCAS) ADVISORIES ON HEADS-UP DISPLAYS: ENHANCED PILOT RESPONSE

by
Hamadeh A. Nureddine

A Thesis Submitted to the

Office of Graduate Programs

in Partial Fulfillment of the Requirements for the Degree of

Master of Aeronautical Science

Embry-Riddle Aeronautical University

Daytona Beach, Florida

August, 1993

INCORPORATION OF TRAFFIC COLLISION ALERT SYSTEM (TCAS) ADVISORIES ON HEADS-UP DISPLAYS: ENHANCED PILOT RESPONSE

by

Hamadeh A. Nureddine

This thesis was prepared under the direction of the candidate's thesis committee chairman, Dr. John T. McGrath, Department of Aeronautical Science, and has been approved by the members of the thesis committee. It was submitted to the Office of Graduate Programs and was accepted in partial fulfillment of the requirements for the degree of Master of Aeronautical Science.

THESIS COMMITTEE:

Dr. John T. McGrath
Chairman

Mr. Leslie Kumpula
Member

Mr. James Blanchard
Member

Department Chair, Aeronautical Science

Dean of Faculty, Daytona Beach Campus

6-22-93
Date

To the memory of my father,

he taught me to work hard,
and have a kind heart

ACKNOWLEDGMENTS

In preparing this thesis, I worked with, and came to appreciate a group of people to whom I am greatly indebted. I wish to thank Dr. John T. McGrath, chairman of my thesis committee, for his invaluable inputs and, above all, his friendship. My special thanks go to Mr. Leslie Kumpula, and to Mr. James Blanchard for their inputs, critique, and support throughout the process of preparing this work.

I would also like to express my gratitude for the creative work of Mr. Charles Frederick and Mr. Christopher Scarborough. Their software development skills were indispensable, and their perseverance in the face of numerous technical difficulties was a source of much needed comfort and support.

Finally, my deepest gratitude goes to my wife and children. Without their backing and encouragement all of this would not have been possible.

ABSTRACT

Author: Hamadeh A. Nureddine

Title: Incorporation of Traffic Collision Alert System (TCAS) Advisories on Heads-up Displays: Enhanced Pilot Response

Institute: Embry-Riddle Aeronautical University

Degree: Master of Aeronautical Science

Year: 1993

This study evaluated the effects that heads-up mounted TCAS displays had on pilot response and workload. Pilot response was evaluated by: (a) response time to a traffic advisory, and (b) number of missed traffic alerts. Workload assessment was accomplished in accordance with NASA's Task Load Index (TLX). Subjects were all licensed pilots with a minimum of a private pilot license and an instrument rating. A total of 32 subjects were randomly assigned to experimental and control groups utilizing HUD-mounted, and conventional, TCAS displays respectively. Performance data was collected during computer-simulated flights, while subjective workload levels were reported at the end. It was found that HUD-mounted TCAS displays yielded better performance results ($p=0.05$), while resulting in significantly less workload.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADI	Attitude Director Indicator
AGL	Above Ground Level
CDTI	Cockpit Display of Traffic Information
EF	Effort Measure in the TLX Scale
EFIS	Electronic Flight Instrument System
FAA	Federal Aviation Administration
FPM	Feet per Minute
FR	Frustration Measure in the TLX Scale
HUD	Heads-up Display
ID	Identification
IVSI	Instantaneous Vertical Speed Indicator
MD	Mental Demand Measure in the TLX Scale
NASA	National Aeronautics and Space Administration
ND	Navigation Display
OP	Own Performance Measure in the TLX Scale
OTA	Office of Technology Assessment
PD	Physical Demand Measure in the TLX Scale
PFD	Primary Flight Display
RA	Resolution Advisory
TA	Traffic Alert
TD	Temporal Demand Measure in the TLX Scale
TCAS	Traffic Collision Alert Sysytem
TLX	Task Loading Index

V1	Takeoff Decision Speeed
VA	Visual Angle
Vr	Takeoff Rotation Speed

Statistical Abbreviations:

M	Mean
N	Number of Subjects in a Group
P	Probability
Q	Semi-interquartile Range
Q_1	First Quartile
Q_3	Third Quartile
s^2	Estimate of the Pooled Variance
SD	Standard Deviation (N - 1)
$S_{(M1 - M2)}$	Estimate of the Standard Error of the Difference between the Means

Introduction

With the increased incorporation of TCAS equipment on board U.S. civil transport aircraft, a growing volume of feedback is now available on its relative effectiveness and the needed modifications. Because of the system's novelty, the design of TCAS displays was not bound by traditional display formats. Unlike other advances in aviation which were the results of mature technologies, TCAS was mandated by federal ruling. Therefore, it had to contend with a rapid introduction of systems and displays that were largely unproven and still under development. This meant that TCAS displays had to be adapted to existing cockpit layouts. The air carriers, for whom the system was mainly earmarked, resisted integrating TCAS displays into existing primary flight displays (PFD) largely due to cost. The result was a trend to locate TCAS displays autonomously, but outside the field of central vision.

Because of the high closure rates of jet aircraft, traffic advisories have to be acted upon almost instantly. By using the current system, the pilot makes a cognitive effort to consult the separate TCAS display once a warning is sounded. This process obviously leads to protracted pilot response and increased workload.

Generally, natural pilot reaction involves visual attempts to locate traffic. This is a learned response that is ingrained in pilots from their earliest training days. Exploiting this tendency by displaying TCAS advisories on HUD reduces pilot response time, workload, and instrumentation clutter.

Statement of the Problem

The purpose of this study was to investigate the effects of HUD displayed TCAS advisories on pilot response and workload levels. Pilot response was evaluated on two dependent measures: pilot response time to a traffic advisory, and the number of traffic advisories missed. For the purposes of this study, pilot response time was the time from the triggering of a resolution advisory (RA) until the pilot responded to that advisory by initiating a pitch change in the direction of that advisory. The number of missed advisories denotes the number of non-threat traffic that the subject failed to call out. Workload assessment was measured subjectively in accordance with the National Aeronautics and Space Administration (NASA) Task Load Index (TLX).

It must be noted here that, while a subject's **response** is detectable by external cues, such as actions, movements, etc. . ., subject **reactions** are more cognitive in nature, and would need complex physiological measures to detect them. Therefore, the researcher opted to measure response times, since ultimately, they are the critical criterion in determining the efficacy of any aviation display.

Review of Related Literature

The current TCAS installation in transport aircraft provides two kinds of data: (a) traffic alert (TA) data in the form of a plan view of own-ship with conflicting traffic displayed in relative position, and (b) resolution advisories (RAs) indicated both aurally and visually on the instantaneous vertical speed indicator (IVSI). There are minor differences in symbology and color between the three major types currently in the market.

Many studies were carried out to determine the frequency and effectiveness of alerts provided by the system. Delta airlines was particularly concerned with the distracting effects of the warnings, especially below 2500 feet above ground level (AGL). It has been reported that over 50% of all RAs and TAs were experienced within a terminal area and below this altitude (Klass, 1991). Understandably, Delta pilots feel that these warnings come at an awkward time in the approach phase.

On the other hand, Fokker has taken a different approach to the problem of displaying TCAS information (Mecham, 1991). Rather than displaying RAs on the IVSI, they have opted to provide the pilot with a pitch cue on the PFD. Some of the reasons cited are that pilots normally fly pitch angles rather than the vertical speed, and that the PFD is in the central vision field, thus providing the simplest and most instinctive instructions possible. Prior studies have explored problems with the display coloration and pilot preferences as to the location of the RAs (Tuttell, McNally, & Chappell, 1989;

Chappell, 1989). Some pilots showed preference for receiving their TCAS information from the Attitude Director Indicator (ADI) rather than the IVSI. Other pilots found difficulty in responding to the lighted segments on the IVSI, especially when an RA to climb or descend was reversed as a result of maneuvering by the intruding traffic.

In response to pleas by several professional organizations and aircraft manufacturers, the Office of Technology Assessment (OTA) conducted a study to evaluate the implications of using the system (Office of Technology Assessment [OTA], 1989). It concluded that the human factors aspects of the system needed further attention, and that the full effect of TCAS on other pilot duties and workload was still unclear. Prior to the release of the OTA findings, a NASA sponsored workshop identified major issues that needed to be addressed in the process of implementing the TCAS system on board (Chappell, 1988):

1. The optimum format for TCAS advisories.
2. Where and how to present advisories to the crew.
3. Whether the PFD is an appropriate location for TCAS advisories.
4. The effects of displaying traffic information on the behavior and performance of the crew.
5. Where and how to present traffic information to the crew.

6. The overall effects that such a system would have on pilot workload.

Some research in these areas indicates that a threat activated display produced the lowest workload effects on pilots, as opposed to full time displays (Battiste & Bertolussi, 1989). In contrast, Tillotson (1988) reported no significant increase in workload due to the use of TCAS, or any problems with the prioritization of tasks with RAs during the approach maneuvers. Concurrent research by Chappell, Scott, and Billings (1987) concluded that, for differing levels of traffic information, no significant change in performance took place. The only exception was that the greatest overshoot in vertical velocity took place in cockpits where traffic location was not displayed. The least amount of vertical overshoot occurred where a threat-activated display of traffic was used.

A proposed form of traffic display that is very similar to the TA portion of the TCAS is referred to as the Cockpit Display of Traffic Information (CDTI). This type of display has been in consideration since the early sixties, and has undergone many conceptual changes (Pryor, 1991; Stokes, Wickens, & Kite, 1990). Stokes et al. (1990) report that this advisory display can increase situational awareness in pilots by providing them with predictor information about their environment. This allows for more optimal corrections in conflict situations. The researchers are, however, concerned about several possible shortcomings: (a) misuse by pilots contradicting or overriding air traffic control (ATC) commands, (b) increased workload by the addition of yet another monitoring task, and (c) undue fascination on the part of the crew, which would result in detracting from their out-the-window scanning patterns. They continue to report that CDTI would not be located in the central field of vision, and therefore may be overlooked completely by pilots as tunneling of vision occurs during stressful, or high workload situations. This view is reinforced by Battiste and Bertolussi (1989), who showed that cluttered displays can cause higher workloads.

Further research into CDTI by NASA (Burgess, Davis, Hollister, & Sorensen, 1991) proposes a combined CDTI-TCAS display that could be shown on the PFD, Navigation Display (ND), or HUD. Hawkins (1987) reports that, although HUD was first intended for use during low visibility approaches in civil aircraft, it could have other safety applications as well. The FAA has, as of March 1992, formally accepted the use of HUD-mounted landing guidance systems, down to a visibility minimum of a quarter-mile. Air Alaska, and Northwest Airlines are the leading air carriers in this application, and have been working with different manufacturers to further enhance the system (Daly, 1992). The main benefit in current civil applications of HUD is the proximity of the needed information sources to the outside view. This is especially true where monitoring of the outside view is essential, and minimum transition time from one source of information to the othe is an advantage. Hawkins further adds that the need for scanning is reduced by the use of HUD, as most information is concentrated in the central field of vision. Since capturing and tracking traffic is mostly a look up activity, Hawkins' opinions about central vision are very useful. Other reports seem to reinforce Hawkins' opinions about the possible uses of HUD, especially in reducing pilot workload and time required to locate traffic (Edelman, 1990; Long, 1990). It must be pointed out, however, that none of these opinions were based on empirical or field data relating TCAS displays on the HUD.

More recently, there have been calls for experimenting with the HUD as a primary flight instrument. Oliver (1990) believes that PFDs and NDs have done little to improve the crew's ability to analyze and stay aware of the vertical situation and vertical flight path. By any measure, these are important parameters of situational awareness. Taylor (1990) is of the opinion that some technical and format problems, as well as some human aspects, are the main obstacles to using the HUD as a primary flight instrument. Oliver (1990) further points out that the HUD offers an improvement over conventional instruments in two ways:

1. Like the Electronic Flight Instrument System (EFIS), a pictorial display can be integrated from different sources, making the HUD an intuitive display.

2. Unlike the EFIS, the information is presented **in conjunction with** the real world scene, allowing for simultaneous assessment of both frames of reference.

Some researchers do not agree with these views on the possible effectiveness of HUD. Stokes and Wickens (1988) feel that HUD tends to compete with real world images for attention and cognitive resources. They feel that sampling of the outside world may actually break down as HUD captures the central vision, and HUD clutter intervenes with the visual process. Stokes and Wickens (1988) continue to cite some problems with the current civilian HUDs:

1. Standardization of symbology by the different manufacturers. The three major civil HUD manufacturers in the western world, Honeywell, Bendix, and King, do not have a standardized set of symbols. Instead, critical information like heading, attitude, and vertical speed representation differ from one supplier to the other.

2. Display clutter. Since HUD data is collated from many sources and projected mostly in alpha-numeric codes and monochrome tones, the information tends to be abstract and difficult to interpret. However, research has shown that pilots show stronger preference for symbols over the alpha-numeric method of display (Fischer, 1979). Often, pilots reported that they had to turn off the HUD display at critical times in the flight, because they felt that it interfered with their performance. In contrast, other researchers found that subjective pilot opinions on the use of HUD during approach were favorable (Fischer, Haines, & Price, 1980). Whenever it was available, pilots reported that they preferred to use the HUD for primary control of flight path, as it afforded them more clues. Meanwhile, they elected to use the outside-world information for monitoring purposes only. Although a higher workload level was generally reported, pilots surveyed

in the above-mentioned study felt that they controlled the aircraft better when utilizing HUD. The higher workload was attributed to the larger amount of data that was available for processing.

3. HUD formats tend to differ from those used on the back-up panels (NDs and PFDs). Therefore, the pilot may encounter some difficulty in transitioning between traditional instruments, and the HUD (cross-consultation).

Another problem that is very applicable to this study involves the effect of HUD on pilot attention to the out-the-window scene. In essence, the superimposed symbology allows consultation of **both** sources of information: the real world and aircraft instruments, without the need for scanning, as both sources are in the same field of vision. Ideally, the HUD symbology is projected at the same apparent depth, or distance, as the real world, which is infinity. This is achieved by the use of a Fresnel lens which collimates the light rays from the HUD and projects them as parallel, or infinity rays. As long as this process is not disturbed, the pilot will not need to re-accommodate his focus as he gleans the real world or the HUD symbology, something he has to do whenever he consults the traditional instrument panel. Unfortunately, some factors do interfere with this ideal process and tend to "pull" the pilot's visual accommodation from infinity to the HUD surface. Research has shown that such factors as screen surface dirt, dust, or scratches, and such perceptual cues as binocular convergence and relative motion do interfere with the infinity accommodation by signaling the closeness of the HUD (Stoke & Wickens, 1988). Other investigations have reported that HUD symbology was found to be more compelling than the outside-world scene, because there simply is more perceivable change taking place on the HUD screen (Fischer et al., 1980).

If the advantage of the HUD is that it allows viewing of two information sources concurrently, then, to what extent is attention affected by such a format? This basic question was the focus of a research program conducted by NASA-Ames Research Center (Fischer, 1979; Fischer, Haines, & Price, 1980). The program focused on three important issues that dealt with head-up/head-down position, and the accommodative distance of the HUD screen domain relative to the far domain outside the aircraft:

1. The ability to focus on one source, or domain, without interference or distraction from the other.
2. The ability to process information **simultaneously** from both sources.
3. The ability to switch attention between sources, or domains.

Some key findings of this extensive research program are summarized below:

1. The presence of HUD symbology did not harm the pilot's ability to extract required information from the external scene (Fischer, 1979), although a slight decrease in pilot monitoring of the outside view may have been evident (Fischer et al., 1980).

2. The presence of the external scene reduced the pilot's ability to extract information from the HUD, but only by a small degree.

3. Paying attention to both fields simultaneously did not appear to change performance significantly on either the HUD or external-scene extraction of information. This finding contrasts sharply with other research involving selective looking and switching between two visually superimposed fields. Neisser and Becklan (1975) found that , while it is not difficult to follow a specific scene when another one is superimposed on it, simultaneous monitoring of both visual sources was a difficult task. Performance in this context tended to be severely degraded. This view is reinforced by later research by Broadbent (1982). Furthermore, these latter findings are in line with generally accepted theories of information processing. The human ability to attend to several sources of information simultaneously is believed to be very restricted (Kantowitz & Casper, 1988). The final word in this area seems to be that there are no known ways to increase the processing capacity of human beings, which stands at approximately 10 bits per second. It is argued that the best that human-factors scientists can do is to arrange the format and content of the tasks to be performed, so that they are most compatible with the processing capabilities of human beings.

It is clear from the above discussion that the merits of HUD need more investigation. It is equally clear that heads-down displays do place serious limitations on pilot performance. This is especially true where visual locating of a target and concurrent consultation of instruments is required, which is how TCAS presently operates. The instrument panel (or the TCAS display), and the external visual field are in different spatial locations. Vertically, the two sources may be separated by approximately 45 degrees of arc. Unfortunately, the human eye is limited to one or two degrees of arc for central vision (Fischer, 1979). Furthermore, the two sources of information are at different focal distances: approximately 60 centimeters for the instrument panel, and optical infinity for the outside scene. Since the human eye is not capable of accommodating both scenes simultaneously, the pilot must continuously shift his gaze, or scan, both sources while changing his focal distance. The transition time can take as long as 2-5 seconds (Naish, 1964), and may also reflect the time needed by the human brain to perceive and react, as well as the physical act of seeing (Weintraub, Haines, & Randle, 1984). In a time -critical situation, such as that of two jets closing in for a possible collision, the heads-up and heads-down transition time can be critical time that is lost. This realization often leads to increased temporal and cognitive demands on the pilot, thus increasing overall workload.

As currently arranged, TCAS displays (RAs & TAs) are not spatially close to each other. One, the RA, is most often located on the IVSI, while the TA is mostly superimposed on the weather radar screen. This separation is a further cause for us to question the adequacy of the current displays. Wickens and Flach (1988) argue that the sampling of displays depends, among other things, on the correlation of displays both cognitively and spatially. They continue to say that spatial proximity of a stimulus is a critical dimension that determines whether it can be processed simultaneously or in series with other stimuli. In other words, their argument is that the closer any two sources of information are located, the greater is the likelihood of them being simultaneously, or nearly simultaneously, acted upon. Although the above argument seems to conform with most information processing theories, the fact remains that all displays cannot be simultaneously co-located.

Alternative theories such as that proposed by Weinberg (1975) argue that people tend to process meaningfully related material together, forming a "set" of information sources that are sampled whenever a certain task is carried out. For example, a pilot climbing out after takeoff uses a "climb set" of instruments and visual cues, like the altimeter, VSI, and airspeed indicator. This set of instruments is the one most frequently sampled during that particular phase, while a different set is established for other phases of the flight.

Given this idea of information sets, and the belief by many researchers that closer proximity of related information sources does enhance processing speed, it appears logical to assume that an integrated TCAS indicator would be beneficial. Indeed some later versions of TCAS instruments have superimposed TA & RA on the IVSI itself, thus doing away with the duality of sources. By including this integrated display into the central vision field, greater advantages could accrue by reducing the transition time and enhancing the rate of information processing. This would translate into reduced pilot reaction times and workload levels, both of which are important parameters to consider when a traffic conflict is imminent.

Statement of the Hypothesis

Research into the current display technology of TCAS suggests a strong need to develop better displays that take into consideration both cognitive and workload issues. The major concerns can be adequately addressed by exploiting the advantages offered by HUD technology. Based on the natural tendency of pilots to attempt to locate traffic visually, and the fact that integrated information sources that are located in the central visual field place lower cognitive demands on the pilot, it was hypothesized that incorporating TCAS advisories on HUDs will augment pilot response, while concurrently reducing perceived pilot workload.

Method

Subjects

The target population for this study will be all the licensed pilots of transport and commuter class aircraft, and any pilot population likely to utilize sophisticated aircraft that are equipped with TCAS. The available population for selection were all the licensed pilots who were enrolled at Embry-Riddle Aeronautical University (E-RAU) Air Science department between January and May of 1993. To ensure a minimum level of piloting skills, subjects were required to hold, as a minimum, a single-engine, private pilot license with instrument rating.

As a result of recruitment efforts, a total of 45 subjects volunteered for the experiment. Of those who did volunteer, the researcher could contact and schedule only 32 subjects. To ensure the randomness of subject assignment, the researcher used a random-number table in determining whether a subject was part of the experimental or control group. Initially, each volunteering subject was issued a serial identification number (ID), then, using a random number table as mentioned above, a random number was assigned to this subject. It was arbitrarily decided that, if the random number was even, then the subject was assigned to the control group. If however the assigned random number was odd, then the subject was relegated to the experimental group. The main concern in this procedure was to give the subjects an equal chance of being randomly assigned to either group.

As a result of this random assignment, the control group totaled 17 subjects, while the experimental group totaled 15. Prior to recruitment and data gathering, the researcher sought and obtained permission to use the subjects in this experiment from the proper university authorities via the committee chairman. A copy of this permission is provided in Appendix A.

Since the accessible population for this study was based on volunteer students, there were no direct means of controlling for variations in skill levels. Therefore, a pre-treatment questionnaire was applied to screen the subjects on the following personal data:

1. Age and gender.
2. License type and ratings held.
3. Previous experience, if any, with TCAS.
4. Total flight time experience.

Although not an integral part of this study, these personal variables were later correlated with the collected data on the dependent variables. Thus, most sources of sample bias are expected to be controlled for.

As Table 1 shows, all subjects were male pilots with a minimum of a single-engine pilot's license and instrument rating. On average, their age was 22.9 years, ranging from 20 to 29. Average total flight experience was 333.8 hours, ranging from a minimum of 160 hours to a maximum of 1020 hours. Although all subjects indicated that they had some knowledge of the TCAS system, none had operated or seen a TCAS display before.

Table 1

Subject Data Summary

ID #	Age	Total Hours	Licenses and Ratings[a]	Assignment[b]
1	28	550	CASMEL-I, CFII	E
2	29	300	CASMEL-I, CFII	C
3	21	244	CASMEL-I	C
6	22	176	PASEL-I	C
7	20	160	PASEL-I	E
8	22	290	CASMEL-I	C
9	24	1020	CASMEL-I, CFII	C
10	21	225	PASEL-I	C
11	20	180	CASEL-I	C
12	20	200	CASEL-I	C
13	23	184	PASEL-I	E
14	22	237	PASEL-I	C
15	22	193	PASEL-I	C
19	22	400	CASMEL-I	C
20	28	320	CASMEL-I	E
21	23	350	PASMEL-I	C
23	20	200	CASMEL-I	C
24	21	300	CASMEL-I, CFII	E

(table continues)

ID #	Age	Total Hours	Licenses and Ratings[a]	Assignment[b]
25	21	200	CASMEL-I	C
28	21	200	CASMEL-I	E
29	22	230	CASMEL-I	E
30	28	650	CASMEL-I, G	E
31	22	360	CASMEL-I, G	E
32	23	250	CASMEL-I, CFI	E
33	21	175	CASEL-I	C
35	27	1009	CASMEL-I, G	E
36	21	170	PASEL-I	E
37	25	1000	CASMEL-I, CFI	E
40	28	300	CASEL-I	C
42	22	201	CASEL-I	C
43	20	176	CASEL-I	E
44	24	230	PASEL-I	E
Averages:	22.9	333.8		
SD:	2.77	245.819		

Note. Subject identification numbers are listed in ascending order of assignment, but do not reflect the order in which subjects conducted the experiment. Missing identification numbers denote that subjects who did volunteer did not show up for the experiment. The large SD value for the total hours column was a result of a few extreme values. If those extreme values were to be ignored, the SD value would be 112.650.

[a]Licenses and Ratings: C- Commercial L- Landplane
P- Private M- Multi-engine
A- Airplane G- Glider
S- Single Engine I- Instrument Rating
CFI- Certified Flight Instructor
CFII- Certified Flight Instructor- Instrument

[b]Assignment: E- Experimental Group
C- Control Group

Instruments

The two major areas that were evaluated in this study were pilot response and workload assessment. Pilot response was observer-evaluated on two related measures: (a) response time to a traffic advisory in seconds, and (b) number of non-threat traffic advisories that were missed by the subject. Since the experiment was carried out in a computer-driven simulator, the observer was able to directly access pilot response time for each warning event from a special subroutine in the simulation program. A more complete description of the simulator is given in the procedures section of this paper. Subjects were also asked to call out the relative bearing (in clock coordinates) and approximate distances of all non-threat traffic that appeared on the TA display. If the traffic was not called out by the subject within three seconds of its appearance, it was counted as a missed traffic. The number of missed non-threat traffic advisories was collected by the researcher during the simulated flight session. The significance of this latter measure is that it indicates the extent of workload due to the task of flying the aircraft and scanning the different displays. Research into this area of "divided attention" and cognitive resource allocation has shown that this secondary-task loading is rather accurate in depicting increased mental workload in the primary task (Kantowitz & Casper, 1988).

The second area of evaluation was the individual assessment of workload by the subjects themselves. This was achieved immediately following the simulated flight. Since **perceived** workload levels are found to have a direct effect on performance vis a vis **real** workload levels, it was decided to use a self-reported evaluation by the subjects themselves (Hart & Staveland, 1988). For the purpose of this research, workload was defined as a hypothetical construct that is intended to measure, in some way, the cost to a human operator of performing a task at a particular level or standard. Further still, subjective workload is taken to mean the operator's impression of the requirements of a task in a given situation, and at a certain skill level. Therefore, subjective workload is taken to summarize the influence of many factors in addition to such objective measures as cognitive and physical demands imposed by the task.

The significance of subjective workload is its close correlation with operator behavior. If a human operator **considers** the workload generated by a given task to be excessive, he may behave as if he really is overloaded, even if the task demands are objectively measured and found to be at a lower level (Hart & Staveland, 1988). Even though subjective workload measures have been criticized for their relative lack of sensitivity to certain variables in any given task, their greatest redeeming values are their non-intrusive nature and ease of implementation. Over the years, researchers have developed several subjective measures such as the SWAT and Cooper-Harper scales (Kantowitz & Casper, 1988), however, some problems persisted. Most important was the difficulty in comparing results between experiments using different rating scales.

In an effort to overcome the more salient deficiencies of subjective workload ratings, NASA undertook a long-term research program to identify and isolate the factors that caused variations within and between these ratings. As a result, 10 different workload related factors were isolated, and these in turn were condensed into six variables. The researchers then developed an easy-to-administer scale that took into consideration both the weight and the magnitude of a given variable in determining overall subjective workload. This made the Task Loading Index, or TLX as it came to be known, a very practical tool for application in operational environments. As a result, several important workload studies have utilized it in one form or the other since its development (Battiste & Bortolussi, 1989; Hart & Hauser, 1987).

The TLX index which was used in this study measures six dependent workload factors:

1. **Mental Demand**. How much mental and perceptual activity was required to perform the task.

2. **Physical Demand**. How much physical activity was required.

3. **Temporal Demand**. How much time pressure was felt due to the rate or pace at which the tasks occurred.

4. **Own Performance**. How successful the subject thought he was in accomplishing the tasks set by the experimenter.

5. **Effort**. A measure of how hard a subject had to work to accomplish his own level of performance.

6. **Frustration Level**. A measure of how insecure, discouraged, irritated, or annoyed versus secure, gratified, content, or relaxed the subject felt during the experiment.

To reduce bias between subjects, a simple weighting system especially developed for the TLX was utilized. The subjects were asked to rate the contribution of each factor to the overall task workload on a bipolar continuum. The philosophy behind this is to provide the subjects with anchor, or end points that have natural psychological meaning rather than some arbitrary values. Some researchers have found that this graphic, as opposed to numerical, representation is superior in that it avoids non-linearity and bias for extreme values. The responses were quantified during the data analysis phase, and were assigned a value from zero (low) to 100 (high). After that, the weighted scores for each factor were summed, and simple averaging yielded a weighted subjective workload assessment. Given the intricate definitions of each factor, the subjects were briefed about the intended definitions prior to initiating the assessments. No special skills were required in scoring the data other than average mathematical knowledge. As a reference, a complete description of the TLX application procedure is included in Appendix B.

Although the NASA TLX index was released rather recently, it was the result of extensive research that spanned three years. Its major appeal is its simplicity and relative brevity. Reported validity values, when compared with equivalent tests, are high (r-squared values range from 0.78 to 0.90). Test/retest reliability coefficients averaged 0.83. The index and scoring directions have been published (Hart & Staveland, 1988), and are accessible to the public at no cost.

Experimental Design

For this study, the post-test only group experimental design will be used (Figure 1). This design was selected because it controls for most sources of invalidity, and is relatively easy to administer.

As figure 2 shows, the experimental variable was the method of displaying TCAS information, i.e. in a heads-up or a heads-down position. The dependent variables were: (a) the timeliness of response to the resolution advisories, (b) the number of non-intrusive traffic advisories that were missed, and (c) an assessment of workload reported by the subjects themselves in accordance with the NASA TLX. Both the control and experimental groups were assigned by random sampling.

	Subjects	Treatment	Post-test
Design:	R	X_1	O
	R	X_2	O

Symbols: R = Random Assignment of Subjects.
X_1= Experimental Treatment.
X_2= Control Treatment.
O = Post-test.

Figure 1: Post-test Only Control Group Experimental Design (Gay, 1987).

The question of eye accommodation from the heads down to the heads-up view was considered in the design of this experiment. However, the fact that the heads down display was located approximately the same distance away from the eye position as the HUD, minimizes any effects that may arise in this experiment. Further, since the experiment was conducted on a display screen where the HUD symbology and the outside view were co-located at the same focal distance (30 inches), no accommodation effects were anticipated on that display. In this respect, the latter condition is similar to that found in a real-world HUD. Since both the symbology and the outside view are focused on infinity, there too, no accommodation effects are expected to exist.

The only viable threat to validity was mortality (see Appendix C). However, all factors affecting internal and external validity were considered. To control for learning and adaptation effects, subjects were allowed some learning time on the simulator. The time allotted for adaptation was determined by means of a four-subject pilot study as described in the procedures section of this report. Personal data was collected prior to testing in order to guard against sampling bias.

Group	Assignment	N	Treatment	Post-test
Experimental	Random	15	HUD-mounted TCAS Display	(a) Response Time (b) Missed Warnings (c) Self-reported Workload Levels
Control	Random	17	Conventionally Mounted TCAS Display	Same Post-tests as for the Experimental Group

Figure 2: Experimental Design, Treatment, and Post-test.

Procedure

In designing the apparatus that was used in the experiment, the researcher ensured that the following principles were adhered to:

1. The flight simulation model had to reflect, as accurately as possible, the flight scenery and dynamics that would be available in an advanced graphic simulation.

2. The relative eye distances of the various displays, as well as their viewing angles, had to reflect those distances and angles normally found in a modern jet transport aircraft.

3. The actual display symbology must be synonymous with current aircraft symbology, and must remain as uncluttered and intuitive as possible.

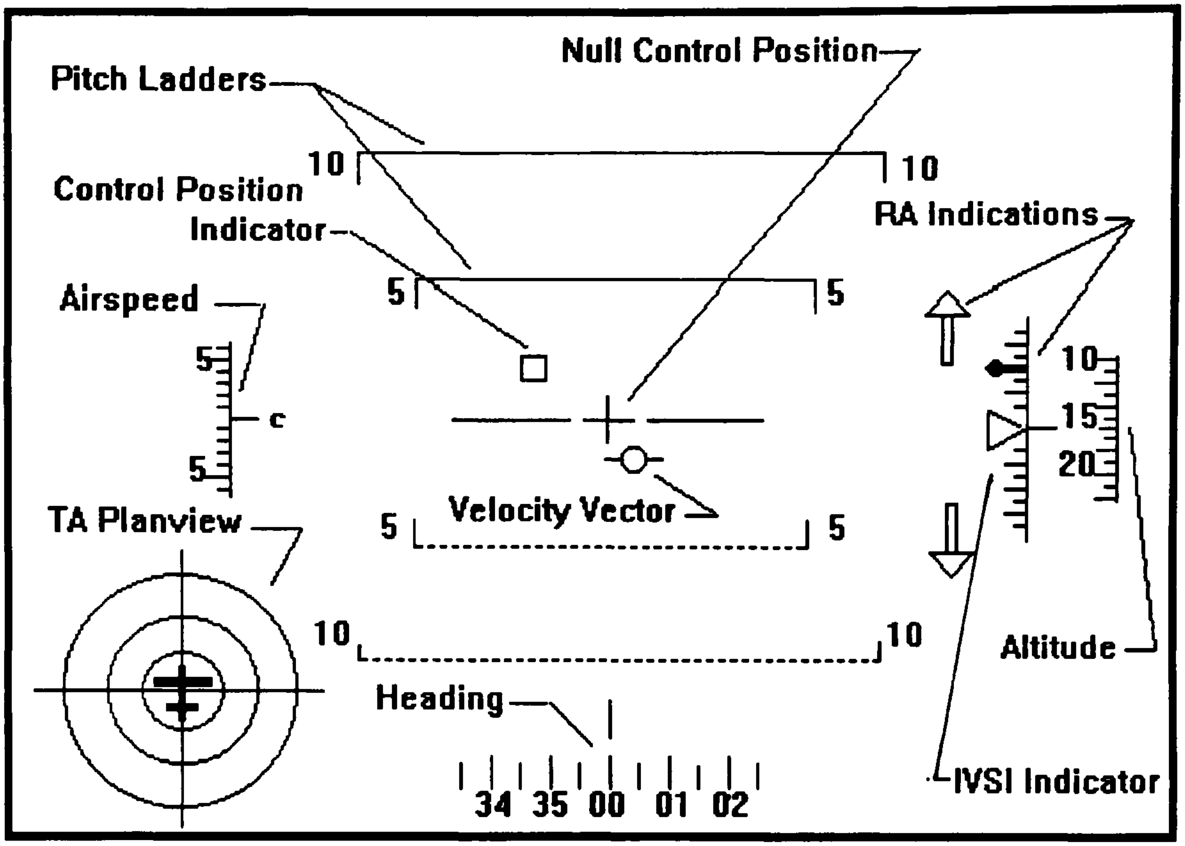
Given these operational constraints, the researcher elected to utilize a flight simulation program that was authored by Silicon Graphics Incorporated as the base computer model. With the aid of specialist software designers, this model was then tailored to include the TCAS scenario, the requisite warnings, and the response-time measurement subroutine.

This development phase resulted in a TCAS scenario that is driven in conjunction with either one of two display setups:

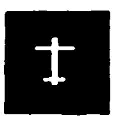
1. A single screen display with TCAS information shown directly on a HUD that, in turn, is superimposed on the outside view. This display was utilized for the experimental group. As Figure 3 shows, the HUD symbology was maintained at a minimum, showing

only airspeed, altitude, heading, pitch, and vertical speed information. In the left-hand corner of the display, a plan-view traffic alert (TA) display is also shown. This TA is where traffic bearing and distance from own-ship is displayed using standard TCAS symbols. RAs, on the other hand, are displayed in conjunction with the IVSI tape on the right-hand side of the screen. the RA is made up of three parts: (a) an intuitive **climb** or **descend** arrow that flashes on and off for the duration of an RA, (b) a red dot and line displayed on the IVSI tape itself to direct the pilot to the required rate of climb or descent in feet per minute, and (c) a continuous aural beep that cannot be silenced, except at the end of an RA.

2. A dual-screen setup that is utilized for the control group, and designed to replicate the heads-up / heads-down of the outside view, and cockpit instruments respectively. Figure 4 shows the heads-up screen, which is identical to the one used by the experimental group (Figure 3), except for the for the absence of any TCAS information. Instead, TCAS is displayed on a separate 9-inch monitor (Figure 5) that shows both the TA, and RA/IVSI co-located together. This dedicated TCAS monitor is positioned so that it reflects both the eye distances and visual angles of a comparable TCAS display on a jet transport aircraft.



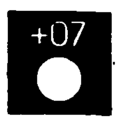
Note: The Following TCAS Symbols are Displayed on the TA Plan View:



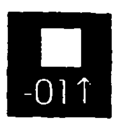
OWN AIRCRAFT: AIR PLANE SYMBOL, WHITE OR CYAN

NON-INTRUDING TRAFFIC ALTITUDE UNKNOWN OPEN DIAMOND, WHITE OR CYAN

PROXIMITY TRAFFIC 200 FEET BELOW, DESCENDING SOLID DIAMOND, WHITE OR CYAN



TRAFFIC ADVISORY (INTRUDER) 700 FEET ABOVE, LEVEL SOLID AMBER CIRCLE



RESOLUTION ADVISORY (THREAT) 100 FEET BELOW, CLIMBING SOLID RED SQUARE

Figure 3: Experimental Group Heads-up Display with TCAS Information Included.

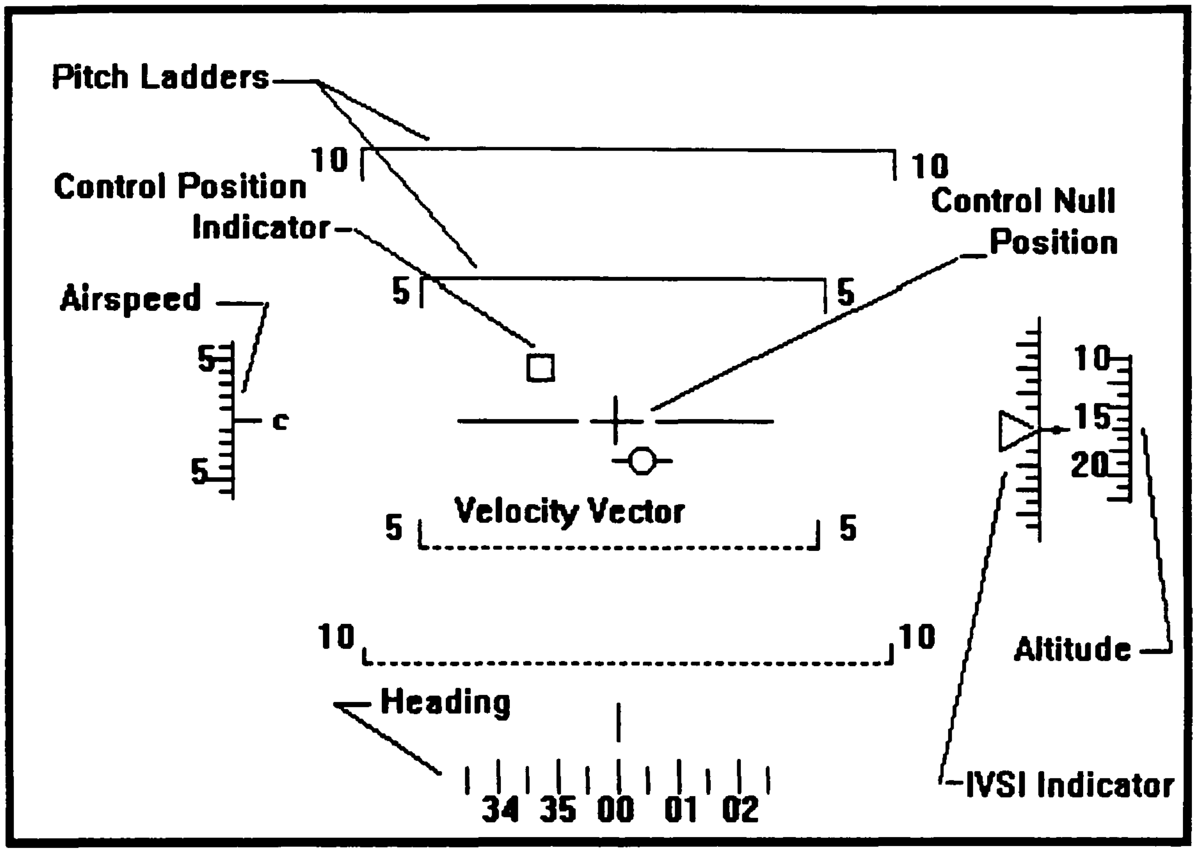
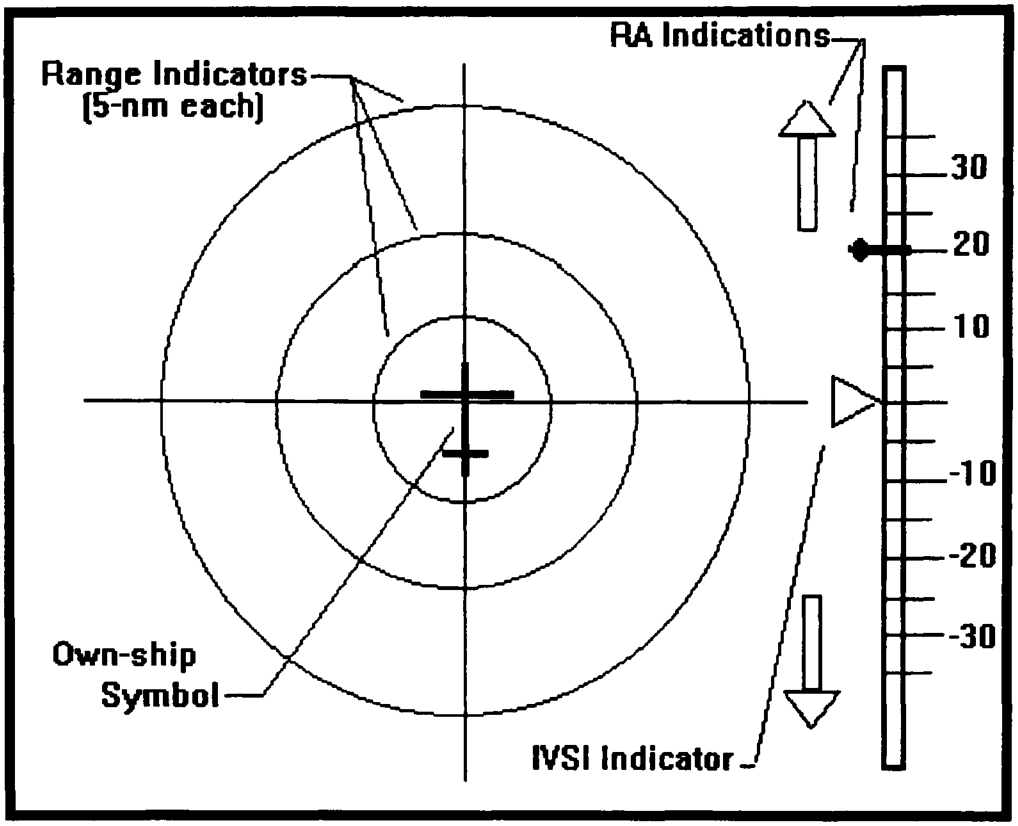


Figure 4: Control Group Heads-up Display.



Note: The Following TCAS Symbols are Displayed on the TA Plan-view:

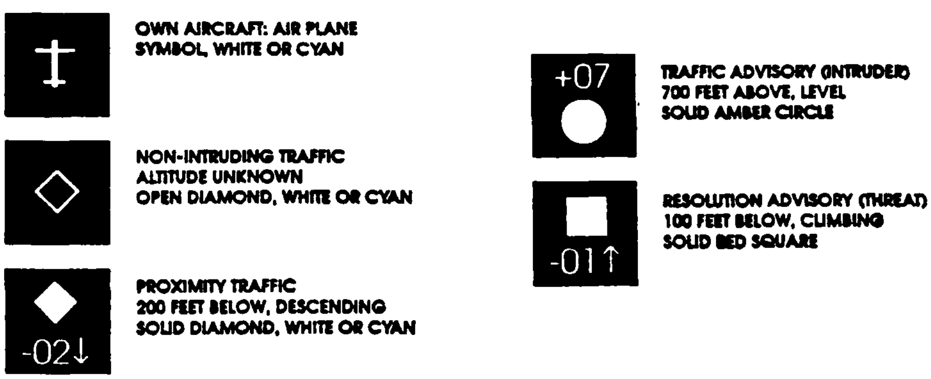


Figure 5: Control Group TA Display.

In both setups, the subject controls the flight path of the aircraft by manipulating an optically driven mouse device. This manipulation would result in movement of a red box that denotes control stick position in relation to the null control position, indicated by a cross in the center of the display. Those two symbols were located on the heads-up display for both groups. In order to minimize the learning time required to "fly" this simulation, the software was manipulated to include automatic turn coordination. The end result was a simulation model that was relatively easy to fly, with no keyboard inputs required by the subject, and only an X-Y movement of the mouse needed to obtain the desired flight-path.

In order to accurately reflect the eye distance and visual angles that would be representative of a typical airliner cockpit, the researcher obtained front wind screen and central pedestal panel distances for a Boeing 707 cockpit by actual measurement. The distances and angles that were obtained were based on anthropometric measures for a 50th percentile male in the relaxed sitting position (Saunders & McCormick, 1987). Figure 6 shows a sitting height of 34.1 inches, and an eye distance to the primary display area (out-the-window view) of 28-30 inches. The secondary display area, representing the central pedestal panel, was found to be at a 50-degree angle downward, and an eye distance of 45 inches. The primary display area was 46 inches above the cockpit floor, while the center of the secondary display area was measured at 20 inches of height, and

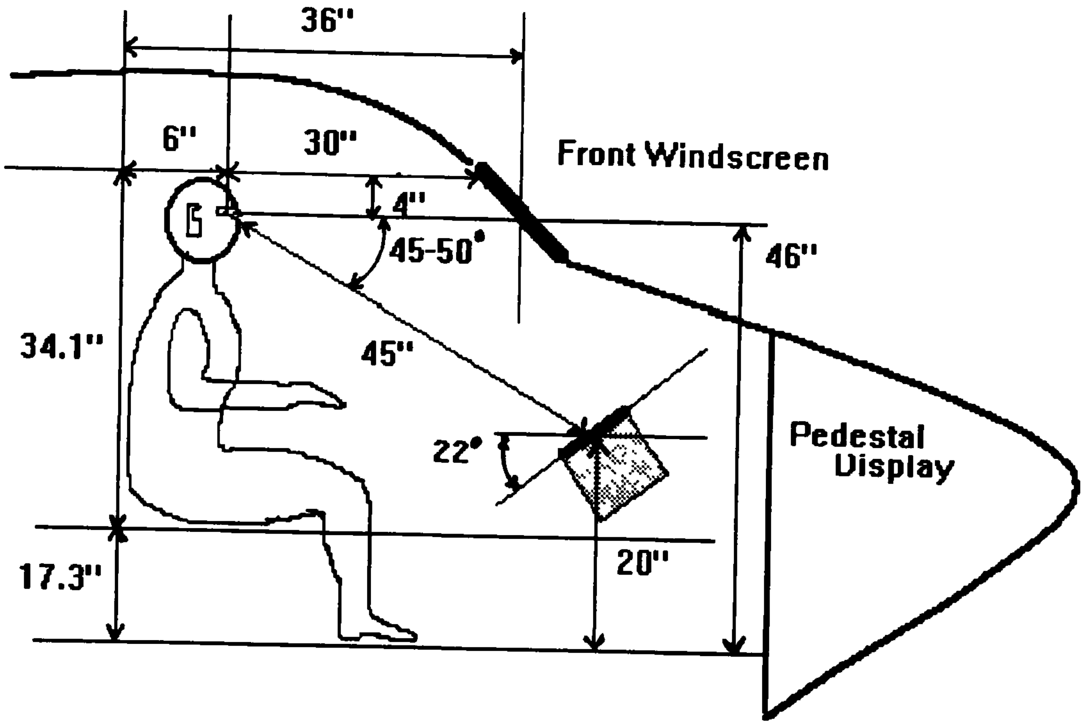


Figure 6: Anthropometric Measures of a 50th Percentile Male in a Typical Airliner Cockpit.

approximately 22-degrees of inclination. In Figure 7, a top view of the setup shows the secondary display area to be at 36.5 degrees to the right of the viewer. this is representative of the angle from the left pilot's seat.

An important factor in deciding on the distances and sizes of objects on both displays, was the total visual angle (VA) that is normally available to the pilot. By direct measurement, the researcher determined that the visual angle through the front

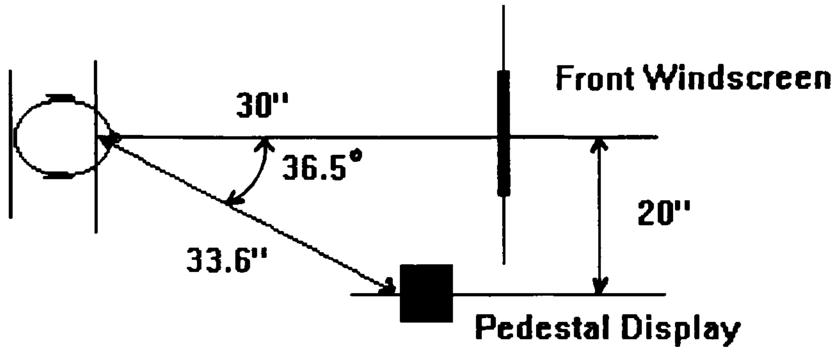


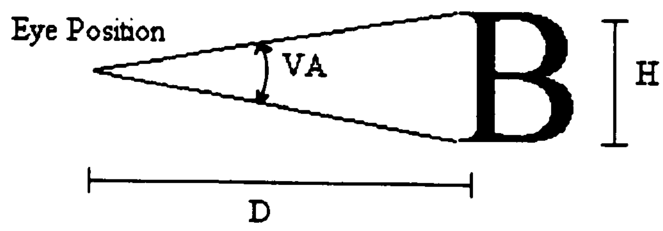
Figure 7: A Top View of Anthropometric Measures of a 50th Percentile Male in a Typical Airliner Cockpit.

wind screen was 18.7 degrees of arc. Since the main display screen represents the out-the-window view in this experiment, the actual size of a TA display that would "fit" into the central vision field could thus be determined. In Figure 8 below, the concept of visual angle (VA) is shown. Given a VA of 18.7 degrees of arc, and a screen height (H) of 10 inches, the distance D from the eye position to the screen could be determined by the following formula (Saunders & McCormick, 1987):

$$D = \frac{3438\ H}{VA}$$

By substitution, we get:

$$D = \frac{3438\ (10)}{18.7 \times 60} = 30 \text{ inches}$$



H= Height of Visual Stimulus
D= Distance from the Eye

Figure 8: Illustration of the Concept of Visual Angle.

As a result of these measurements, the main display screen, which measures 13 x 10 inches was placed 30 inches away from the relaxed eye position of the representative pilot, at a height of 46 inches. Also, the secondary display screen, which measured 9 x 9 inches, was placed at an inclination of 22 degrees, and a distance of 45 inches from the eye position (Figure 9).

Since the main purpose of this study was to investigate the effects of displaying TCAS information in the central field of vision, it was necessary to estimate the size of any TCAS display accordingly. Simply, for any TCAS display to be effective, its size must not exceed that which would be included in the central vision field , which would preclude the need for scanning. Human Factors scientists have approximated this central vision field at 2-4 degrees of arc. Therefore, the TCAS display, or more specifically, the TA must not exceed 2-4 degrees of arc when displayed in the heads-up position. Given a VA of 18.7 degrees of arc on a display that is placed 30 inches away from the eye position, the size of the central vision field can easily be calculated to be between 1-2 inches in size As a result, the actual size of the TA display for the experimental group was designed to be 2 inches in diameter

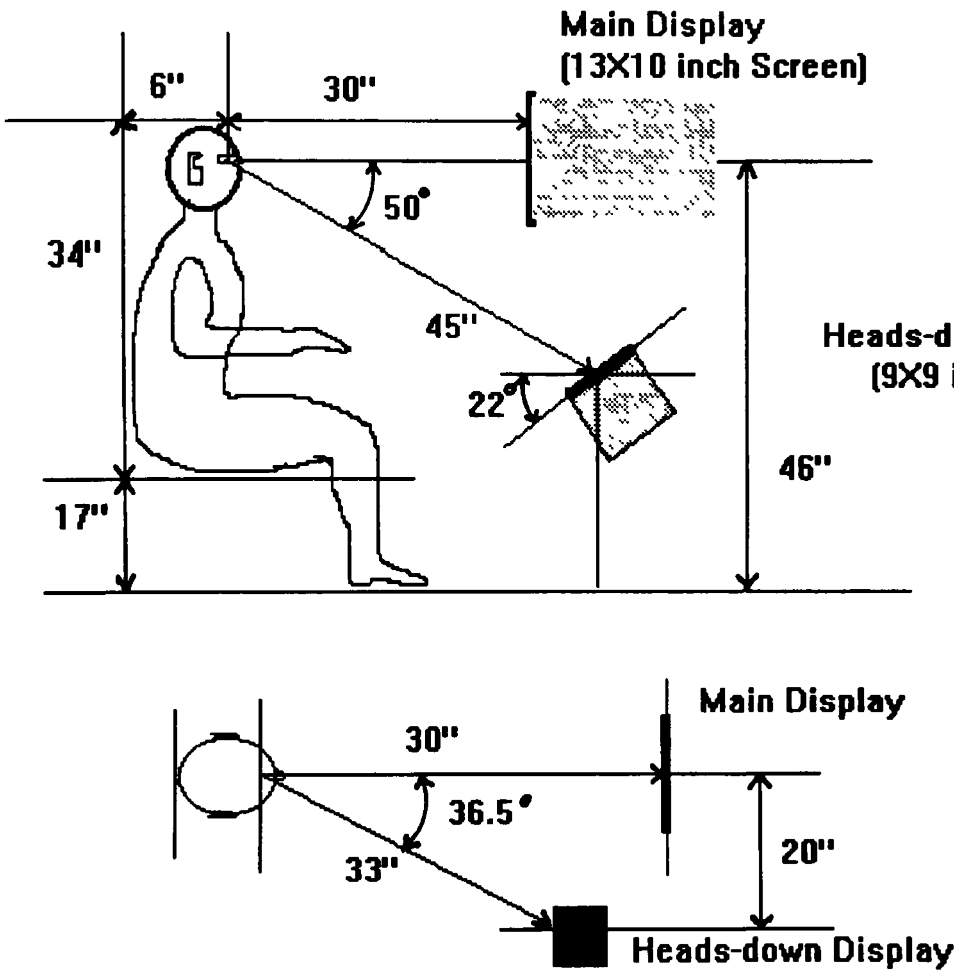


Figure 9: Display Measurements for the Simulation Study.

Since the subjects were to fly a computer based simulation, some learning, or adaptation time was naturally expected. In order to control for this learning, the researcher conducted a pilot study on four subjects who did not participate in the main study. In this pilot study, the researcher asked the subjects to fly a fixed scenario that included several attitude and heading changes. Throughout the flight sessions, the researcher observed the error values in both altitude and heading, and compared them to the acceptable standards set for private pilot practical tests set by the FAA (FAA, 1985). This exercise was carried out as a function of time, and it was naturally expected that the deviation values would decrease with time. The researcher noticed early on in the study that subjects required a short period of time (Average = 2.2 seconds) to accurately control heading. However, altitude control appeared to be more demanding, requiring longer learning times. The results of this pilot study showed that, in order for subjects to be able to "fly" this simulator within the prescribed standards, it was necessary to allow an adaptation time of between 6-8 minutes (see Appendix D). In order to be conservative in estimating adaptation time, the researcher decided to allow a learning time of 10 minutes for each subject in the main study.

Following their random assignment to either the experimental or control group, subjects were contacted by the researcher to set a convenient time for the experiment, which, on average, lasted 40-45 minutes. Each data-gathering experiment consisted of five steps listed in order of occurrence below:

1. An initial briefing step where the general purpose of the experiment is explained. The subjects were briefed on the HUD symbology and the controls for altitude and heading. Subjects were then briefed on the training session, which was intended to allow them the needed adaptation time.

2. The training flight consisted of a 10-minute flight in which the subject took off from an 11,000-foot runway in a B-747 aircraft. As per the briefing, the researcher acted as co-pilot in that he set takeoff power and flaps, called out the speeds (V1 and Vr), and managed the flaps, gear, and power subsequently. The researcher also requested that the subject fly given altitudes and headings, providing guiding remarks as necessary. A list of the standard maneuvers required during the training session is included in appendix E. During this session, the TCAS system was not activated, nor were any traffic advisories activated on the TA display.

3. Following the allotted training flight, the researcher halted the simulator and provided the subject with a prepared briefing on the TCAS system. This briefing included a general system overview, including operation, symbols, warnings, and the location of the TA and RA displays that apply to the subject's group. The subject was then asked to perform the following tasks during the data-gathering flight: (a) fly the aircraft at the

altitudes and headings requested by the researcher, (b) give the relative bearing (in clock coordinates) and distance of each non-intrusive traffic that appears on the TA display, and (c) respond as quickly, yet as accurately as possible to any RAs that might occur. The subjects were not told at this stage that the TCAS warnings were actually scripted, but were rather left under the impression that these warnings were randomly generated by the computer.

4. The data gathering flight was similar to the training flight in both procedure and required maneuvers (see Appendix E). However, in the data-gathering simulation, the TCAS system was activated, resulting in 15 proximity traffic indications (on the TA), and four resolution advisories. The TCAS scenario was identical for both experimental and control groups; the only difference being in the location of the TCAS information display. The four RAs involved a gradual approach of of a non-intrusive target from the left or right front quadrants. The traffic would, with time, change symbols as it became a more viable threat, resulting in an RA advisory that needed to be acted upon. The four RAs were spaced out throughout the flight scenario to ensure that the subjects were not overly tasked. The following is a summary of the TCAS RAs that were activated during the data-gathering flight. They are shown as a function of scenario time.

T+ 0'00" Simulation starts with aircraft at the departure end of the runway. The subjects takeoff, aided by the researcher. Subjects then follow the headings and altitudes provided by the researcher (appendix E).

T+ 4'00" Proximity traffic appears on the TA, approaching from the left quadrant. Traffic is 1500 feet below own aircraft and climbing, 10 miles away.

T+ 4'05" Traffic symbol changes to Threat, 2 miles away, 700 feet below and climbing. Aural Warning + RA to climb at 1500 feet per minute (fpm).

T+ 4'10" TCAS alert disappears.

T+ 7'00" Proximity traffic appears on TA, approaching from the right front quadrant. Traffic is 1000 feet above, and descending, 5 miles away.

T+ 7'05" Traffic symbol changes to a Threat, 1 mile away, 300 feet above and descending. Aural Warning + RA to climb at 2000 fpm.

T+ 7'10" TCAS alert disappears.

T+11'30" Proximity traffic appears on TA, approaching from the right front quadrant. Traffic is 1600 feet below, and climbing, 6 miles away.

T+11'35" Traffic symbol changes to Threat, 1 mile away, 700 feet below, and climbing. Aural Warning + RA to climb at 1500 fpm.

T+11'40" TCAS alert disappears.

T+13'30" Proximity traffic appears on TA, approaching from the right front quadrant. Traffic is 1500 feet below, and climbing, 5 miles away.

T+13'35" Traffic symbol changes to Threat, 1 mile away, 1000 feet below, and climbing. Aural Warning + RA to climb at 500 fpm.

T+13'40" TCAS alert disappears.

T+14'00" Simulation is stopped by the researcher.

During the flight, the computer that drives the simulation recorded the reaction time interval for each RA. For the purposes of this study, this interval is defined as as the time from the issuance of an RA warning until the pilot establishes a pitch setting that provides the needed vertical rate that is directed on the IVSI. At the end of the flight, the computer would provide a printout of the reaction times for each of the four RAs issued.

Another piece of data that was gathered during the flight was the number of proximity traffic alerts that were missed by the pilot. Here, the researcher observed and recorded the number of traffic targets that were not called out by the subject within the first three seconds of their appearance on the TA display. At the end of the flight, the tally was recorded by the researcher on the subject's data sheet (see Appendix F).

5. The final step in the data gathering process involved the application of the TLX subjective workload index (see Appendix B). After this procedure was completed, the subject was debriefed on the specific purpose of the experiment. The reaction times were obtained, rounded off to three decimal places, and recorded.

Analysis

The results of this study were found to support the hypothesis that incorporation of TCAS advisories on the HUD would augment pilot response, while concurrently reducing perceived pilot workload. In analyzing the data, three null hypotheses (H_0) and three related alternate (research) hypotheses (H_a) were tested using the applicable statistical tools. The first pair of hypotheses was concerned with response times to Resolution Advisories by the two groups:

H_{01} : There is no significant difference in the respective response times between the experimental and control groups.

H_{a1} : The response times of the experimental group are significantly lower than those of the control group.

The second pair of hypotheses addressed the number of non-threat traffic alerts that were missed by the subjects of the two groups:

H_{02}: There is no significant difference in the number of missed non-threat alerts between the two groups.

H_{a2}: The number of non-threat alerts missed by the experimental group is significantly lower than the number missed by the control group.

Finally, in addressing the subjective workload ratings, the following pair of hypotheses was evaluated:

H_{03}: There is no significant difference between the subjective workload levels reported by the two groups.

H_{a3}: The subjective workload levels of the experimental subjects are significantly lower than those of the control subjects.

Furthermore, correlation studies of workload levels, missed alerts, and reaction times were analyzed for any significant findings. The three measures were also correlated with pilot total time to test for any sampling biases. No significant correlation was evident, suggesting that sampling bias was not a significant threat to the validity of this study.

1. Pilot response times to an advisory. There were four RA warnings that were triggered for each subject at the prescribed intervals mentioned in the procedures section. Although the computer provided response times to five decimal places, the researcher rounded them off to only two. Table 2 below shows each group's average response times for each warning. A complete listing of the raw response times for both groups is listed in Appendix G. Since the sample sizes were relatively small and only two groups were used, the researcher chose to use the t-test method for independent means to investigate the hypotheses to a confidence interval of P= 0.05. The data was evaluated on the basis of

each successive warning. The response times for warning number 1 by the experimental group were tested against the response times for that same warning by the control group, etc. . .

Table 2

Average Response Times to Resolution Advisories (in seconds)

	Experimental Group		Control Group	
Advisory #	M	SD	M	SD
1	1.88	0.571	2.71	0.960
2	1.80	0.499	2.91	0.633
3	1.71	0.598	2.61	0.744
4	1.63	0.407	2.78	1.075

As a first step, the sum of squares for each group was obtained by the raw score method (Elzey, 1971). This was used to obtain an estimate of the pooled variance (s^2). Then, an Estimate of the Standard Error of the difference between means (S_{M1-M2}) was derived by the pooled variance method:

$$S_{M1-M2} = \sqrt{\frac{s^2}{N_1} + \frac{s^2}{N_2}}$$

Where s^2 = Pooled Variance
N_1 = Size of Sample 1
N_2 = Size of Sample 2

The next step was to determine the probability of obtaining a difference between the means ($M_1 - M_2$) that is equal to the difference at hand. This probability is expressed as:

$$t = \frac{M_1 - M_2}{S_{M1-M2}}$$

Based on the calculated t-ratio for each response to an RA, it may be stated with a 95% confidence level that the difference between the means for the two groups was statistically significant. It may further be stated, with equal confidence, that this difference was **not** due to a sampling error between the two groups. Thus, the probability of committing a Type-I error is equal to 0.05. The above analysis allows the researcher to therefore reject the null hypothesis H_{01} . As a result of conducting a one-tailed t-test, the H_{a1} was accepted by a comfortable margin (Table 3). Thus, the data was supportive of the research hypothesis that displaying TCAS advisories in a heads-up position would lead to reduced response times vis-a-vis displaying that information in a heads-down position.

Table 3

One-tailed t-ratio Tests of Response Times to Resolution Advisories ($P = 0.05$, $d_f = 30$)

Advisory#	S_{M1-M2}	Obtained t-ratio	Statistical t-ratio
1	0.284	2.919	1.697
2	0.204	5.415	1.697
3	0.241	3.730	1.697
4	0.295	3.878	1.697

2. Number of missed Traffic Alerts. By the same reasoning used above, the researcher hypothesized that displaying non-threat traffic alerts in the foveal field of vision by means of a HUD would cause a lower number of missed warnings. This can be attributed to the fact that, in a heads-up position, the TA information is being displayed in the same visual field as the outside view, thus requiring less divided attention, and less scanning than if it were displayed otherwise.

In collecting the data for this measure, the researcher tallied the number of TA alerts that were not called out by the subject within three seconds of their appearance on the TA plan-view. This 3-second interval, though arbitrarily chosen, was deemed to be a conservative approximation of the detection times that are representative of similar alerts. Table 4 summarizes the mean number of missed alerts for both groups. A complete listing of the missed-alert scores for both groups is supplied in Appendix H.

Table 4

Summary of Mean Values of Missed Traffic Alerts for both Groups

Group	M	SD
Experimental	1.13	1.187
Control	2.59	1.50

The same statistical analysis was utilized to test both the H_{02} and H_{a2} : namely, the t-ratio test. Table 5 shows the calculated data for the one-tailed test of the alternate hypothesis.

Table 5

Calculated Data for One-tailed t-ratio Test of Missed Traffic Alerts ($P = 0.05$, $d_f = 30$)

$S_{M1 - M2}$	Obtained t-ratio	One-tailed Ratio
0.483	3.012	1.697

Based on the above test, it may be stated with 95% confidence that the difference between the means of the two groups was statistically significant, and was not due to sampling error. This allows the researcher to support the research hypothesis by stating that subjects flying a HUD-mounted TCAS had a statistically significant lower rate of missing non-intrusive traffic alerts than did subjects who flew a conventionally-mounted TCAS display.

3. Subjective workload reporting. In analyzing the TLX data, the researcher divided the overall workload scale into 10 equal intervals. There were two types of analysis that were conducted on the data: (a) measures of central tendencies and spread of overall workload scores, and (b) an analysis of the specific workload measures that, when integrated together, make up the overall TLX score. The significance of this latter step was that it shed some light on the **specific** sources of workload for each group, and the differences, if any, between the two groups. A complete listing of the subjective workload scores for both groups is listed in Appendix I.

As a first step, a frequency distribution histogram was constructed for both groups (Figure 10). It clearly shows that the control group workload values tend to occur more frequently at the higher-value intervals than do the experimental group values. The data in Table 6 further illustrates this tendency.

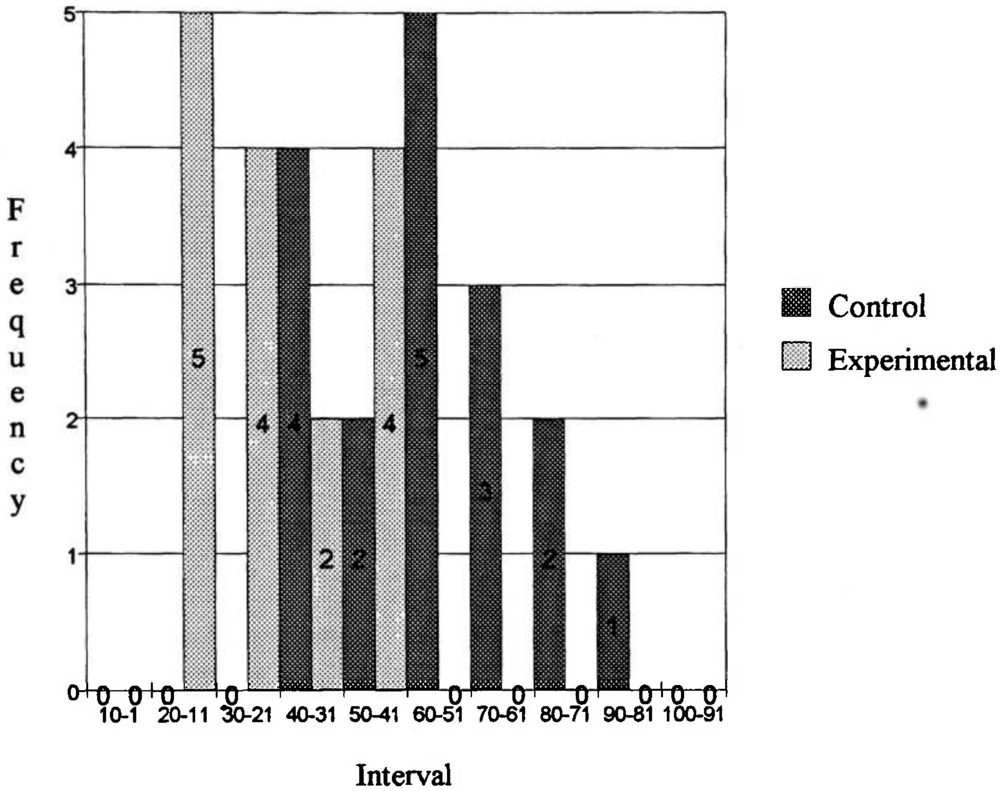


Figure 10: Subjective Workload Frequencies for the Two Groups.

Table 6

Summary of TLX Results for the Experimental and Control Groups

Parameter	Experimental	Control
N	15	17
Modal Interval	20-11	60-51
Median	26.75	55.50
Q1	18.00	41.75
Q3	41.13	66.33
Q	11.56	12.29

Note. Q1 denotes the First Quartile
Q3 denotes the Third Quartile
Q denotes the Semi-interquartile Range

Figure 11 graphically depicts the medians for the two groups, while showing the semi-interquartile ranges. It can be clearly seen that the central tendencies for both groups are significantly different, with the control group having a marked tendency to report workloads of higher values than the experimental group. The above data does support the research hypothesis that subjective workload values reported by pilots utilizing HUD-mounted TCAS would be lower than those reported by pilots who utilized conventionally-mounted TCAS. Since subjective workload assessment is closely associated with **perceived** workload levels, the researcher is of the opinion that the higher workload levels reported by the heads-down group were mainly due to the increased mental and cognitive activities that were needed to process information from the two sources concomitantly.

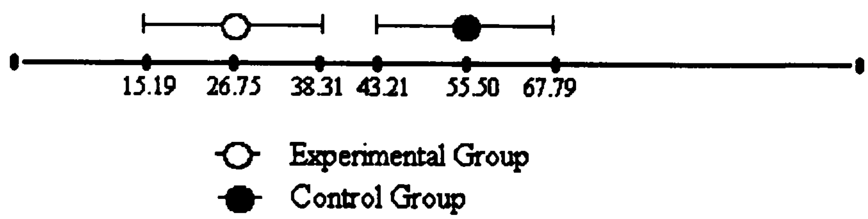


Figure 11: Depiction of Semi-interquartile Range and Median Values for the Reported TLX Workload Measure.

A discussion of the TLX measure in the Instruments section of this report showed that the overall TLX score was actually a weighted average of six different sources, or components of workload:

1. **Mental Demand** (MD). How much mental and perceptual activity was required to perform the task.

2. **Physical Demand** (PD). How much physical activity was required.

3. **Temporal Demand** (TD). How much time pressure was felt due to the rate or pace at which the tasks occurred.

4. **Own Performance** (OP). How successful the subject thought he was in accomplishing the tasks set by the experimenter.

5. **Effort** (EF). A measure of how hard a subject had to work to accomplish his own level of performance.

6. **Frustration Level** (FR). A measure of how insecure, discouraged, irritated, or annoyed versus secure, gratified, content, or relaxed the subject felt during the experiment.

One unique attribute of the TLX index is its ability to differentiate, upon further analysis, between the the sources of workload from one task to the other. Table 7 summarizes the central tendencies of reported magnitudes for each individual source of workload. A more detailed listing of these measures is listed in Appendix I.

Table 7

Calculated Median Values for the Specific Workload Measures

Measure	Experimental Group	Control Group
MD	26.33	66.75
PD	8.84	28.36
TD	28.84	68.84
OP	25.50	38.00
FR	21.50	65.50
EF	8.84	45.50

Upon analyzing the median values for each measure in table 7, several trends are noticed:

1. Even though the median values differ between the two groups, measures of the TD for both groups tend to have the highest reported values. This indicates that there is, understandably, a perceived urgency to the tasks at hand.

2. The median values for MD show this measure as the second greatest source of workload for both groups, although the median values differ greatly. Having to divide one's attention to obtain TCAS information from a heads-down display may have further increased the magnitude of this measure for the control group.

3. Although the experimental group shows that concern for performance (OP) is the third greatest source of workload, it is interesting to note that frustration and effort measures (FR & EF) were the more salient sources of workload for the heads-down control group.

4. Since the simulation took place on a computer-based setup, it is understandable that the physical demands (PD) imposed operating the pointing device would not be perceived as excessive. As a result, both groups ranked this measure last. However, the greater physical act of scanning (eyeball and head movement) in the control group may have partially contributed to the higher PD median in that group.

In summary, the preceding analysis of the self-reported workload index supports the research hypothesis that the display of TCAS information in a heads-up position does create a lower workload than the heads-down position. An in-depth analysis of the individual measures of workload indicates that, although TD and MD are the prominent sources of workload in both groups, frustration (FR) and effort (EF) tend to have more bearing on the overall workload level in the control group.

4. Correlation studies. Two type of correlation studies were carried out in this study. They each served a distinct purpose:

1. A correlation coefficient was obtained between the total flight time experience (TT), and response times to RAs, overall workload, and the number of TA alerts missed. The purpose of this correlation was to test for any sampling bias that may have resulted from flight experience. Table 8 summarizes the correlation coefficients for these measures. It can be seen that total flight time did not correlate significantly with any of the measures mentioned above. This means that flight experience had little or no bearing on the subject performance during the experiment.

2. A correlation coefficient was obtained between the subject-reported workload and: (a) the number of missed alerts, and (b) average subject response time to RAs. As can be seen in Table 9, the correlation between workload and missed alerts for the control group was moderate, while it was insignificant (at -0.310) for the experimental

Table 8

Correlation between Total Flight Time, the Number of Missed Alerts, Workload, and Average Response Times to RAs.

	Correlation Coefficient	
Measure	Experimental	Control
Missed Alerts	-0.429	0.118
Workload	0.035	-0.091
Average Response Time	0.156	0.056

group. This lack of any positive correlation in the latter group can be attributed to the small sample size in that group, and to the fact that the majority of subjects missed one or no traffic alerts during the exercise. Average response time showed a neutral correlation with workload. Again, the small sample sizes may have partially contributed to this lack of correlation between the two measures.

Table 9

Correlation between Reported Workload and Number of Missed Traffic Alerts, and Average Subject Response Time to RAs.

	Correlation Coefficient	
Measure	Experimental	Control
Missed Alerts	-0.310	0.487
Average response Time	0.189	-0.006

Conclusions

A general review of pilot appraisals of the current TCAS system would show that many refinements are suggested. Specifically, airline pilots are concerned about the high frequency of warnings in low-altitude, congested airspace (Klass, 1991). Understandably, the last few thousand feet of approach are busy times for any crew. Operation in dense traffic areas requires that pilots dedicate more time scanning the outside view than they would, say, during cruise. Given this requirement, such a high frequency, high priority warning as a TCAS RA would certainly be obtrusive when the pilot has to consult its associated heads-down display. The re-accommodation and transfer times needed contribute more to an already task-loaded flight deck.

With the introduction of HUD displays into the civil airliner cockpit, the possibility arises that, whenever the pilot's visual attention is mainly required outside the cockpit, TCAS advisories can be superimposed there. Initially, HUD was introduced on a limited scale to provide guidance for ILS approaches. As more and more research is done on the possible uses of civilian HUD, the list of its possible applications grows. With this in mind, the researcher had hypothesized that superimposing TCAS advisories on the HUD would reduce pilot workload, while enhancing overall performance.

The data collected in this simulation study supports the research hypothesis on all three measures investigated. First, it was shown that a considerable difference existed in the response times to RAs. This difference was in favor of the HUD-mounted TCAS. Also, the average number of missed traffic alerts was considerably lower for the subjects flying HUD-mounted TCAS. Finally, subject-reported workload levels showed a marked tendency for conventionally mounted TCAS to cause a higher higher workload overall. Also, high frustration levels and a greater perceived effort were required to respond to the conventionally-mounted system.

In designing this experiment, the researcher sought to replicate, as realistically as possible, the conditions and levels of task loading that exist in an aircraft cockpit. However, given the fact that this simulation was conducted on a computer terminal, further research is warranted. Specifically, this study should be replicated in a simulator that provides a large degree of realism both in the visual and motion cues, and the procedural complexities encountered in a typical airline operation. Given the wide differences that were encountered between the control and experimental groups, further investigation of the questions that were explored here takes on added importance.

Currently, the concepts and philosophies governing cockpit controls and displays are in a state of flux. Multifunction displays , data uplinks, and real-time computers are finding their way into the cockpit. Researchers are thus freed from traditional constraints that have hampered the development of more versatile cockpits. As the industry is poised to design and manufacture the next generation of transport aircraft, the technology associated with TCAS has to be accurately defined in order to be integrated into their flight decks. Hopefully, the results of this study are a step in that direction.

References

Battiste, V., & Bortolussi, M. (1989). Assessment of pilot workload with the introduction of an airborne threat alert system. Proceedings of the Seventh Aerospace Behavioral Technology Conference and Exposition (pp. 55-61). Warrendale, PA: Society of Automotive Engineers.

Broadbent, D. E. (1982). Task combination and selective intake of information. Acta Psychologica, 50, 253-290.

Burgess, M., Davis, D., Hollister, W., & Sorensen, J. (1991). Traffic alert and collision avoidance system (TCAS): Cockpit display of traffic information (CDTI) investigation Phase 1: Feasibility study (Report No. NASA-TM-104956). Hampton, VA: National Aeronautics and Space Administration, Langley Reseasrch Center.

Chappell, S. L. (Ed.). (1988). The traffic alert and collision avoidance system (TCAS) in the glass cockpit (NASA Technical Memorandum 101036). Proceedings of a workshop sponsored by the National Aeronautics and Space Administration (NASA) Ames Research Center, Moffet Field, CA, January, 1988. Washington, DC: NASA.

Chapell, S. L. (1988). Avoiding a maneuvering aircraft with TCAS. In R. S. Jensen (Ed.), Proceedings of the Fifth International Symposium on Aviation Psychology (pp. 338-343). Columbus, OH: Ohio State University Press.

Chapell, S. L., Scott, B. C., & Billings, C. E. (1987). Information transfer in pilots' use of a collision avoidance system. Proceedings of the Human Factors Society 31st Annual Meeting (pp. 428-431). Santa Monica, CA: Human Factors Society.

Daly, K. (1992, July 22-28). Looking ahead. Flight International, pp. 22-23.

Edelman, D. (1990). Heads up display applications for wide body transport aircraft: An operational point of view (SAE Report No. 901832). Warrendale, PA: Society of Automotive Engineers.

Elzey, F. F. (1971). A programmed introduction to statistics (2nd ed.). Belmont, CA: Brooks/Cole.

Fischer, E. (1979). The role of cognitive switching in head-up displays (Report No. NASA-CR-3137). Washington, DC: National Aeronautics and Space Administration.

Fischer, E., Haines, R., & Price, T. (1980). Cognitive issues in head-up displays (Report No. NASA-TP-1711). Washington, DC: National Aeronautics and Space Administration.

Federal Aviation Administration (1985). Private pilot practical test standards (FAA-S-8081-1AS). Washington, DC: U.S. Government Printing Office.

Gay, L. R. (1987). Educational research (3rd ed.). Columbus, OH: Merrill Publishing.

Hart, S. G., & Hauser, J. R. (1987). Inflight application of three pilot workload measurement techniques. Aviation, Space, and Environmental Medicine, 58, 402-410.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), Human Mental Workload (pp. 139-183). Amsterdam, The Netherlands: Elsevier Science Publishers B. V.

Hawkins, F., (1987). Human factors in flight. Brookfield, VT: Gower.

Kantowitz, B., & Casper, P. (1988). Human workload in aviation. In E. L. Wiener & D. C. Nagel (Eds.), Human Factors in Aviation (pp. 157-187). San Diego, CA: Harcourt Brace Jovanovich.

Klass, P. (1991, April 8). Airlines initial use of TCAS suggests need for minor changes. Aviation Week & Space Technology, pp. 36-37.

Long, H. A. (1990). HUD potential for narrow bodied air carrier aircraft (SAE Report No. 901831). Warrendale, PA: Society of Automotive Engineers.

Mecham, M. (1991, June 17). Human factors play large role in Fokker TCAS cockpit adaptation. Aviation Week & Space Technology, pp. 165-166.

Naish, J. M. (1964). Combination of information in superimposed visual fields. Nature, 202, 641-646.

Neisser, U., & Becklan, R. (1975). Selective looking: Attending to visually specified events. Cognitive Psychology, 7, 480-494.

Office of Technology Assessment. (1989). Safer skies with TCAS: Traffic alert and collision avoidance system special report (Report No. PB89-169221). Washington, DC: Author.

Oliver, J. G. (1990). Improving situational awareness through the use of intuitive pictorial displays (SAE Report No. 901829). Warrendale, PA: Society of Automotive Engineers.

Pryor, D. (1991, February). TCAS finally moves into the cockpit. Aerospace America, pp. 36-38.

Sanders, M. S., & McCormick, E. J. (1987). Human factors in engineering and design. New York: McGraw Hill.

Stokes, A., & Wickens, C. (1988). Aviation displays. In E. L. Wiener & D. C. Nagel (Eds.), Human Factors in Aviation (pp. 387-422). San Diego, CA: Harcourt Brace Jovanovich.

Stokes, A., Wickens, C., & Kite, K. (1990). Display technology: Human factors concepts. Warrendale, PA: Society of Automotive Engineers.

Taylor, C. N. (1990). The HUD as a primary flight instrument (SAE Report No. 901833). Warrendale,PA: Society of Automotive Engineers.

Tillotson, D. (1988). Results of the in-service evaluation of the traffic alert and collision avoidance system industry prototype. American Institute of Aeronautics and Astronautics/Institute of Electrical and Electronic Engineers. Eighth Digital Avionics Systems Conference (part 1, pp. 277-282). Washington, DC: AIAA.

Tuttell, R. J., Mcnally, K. L., & Chappell, S. L. (1989). Pilot response to avoidance regions depicted on alternate TCAS II resolution advisory displays. In R. S. Jensen (Ed.), Proceedings of the Fifth International Symposium on Aviation Psychology (pp. 344-348). Columbus, OH: Ohio State University Press.

Weinberg, G. (1975). An introduction to general systems thinking. New York: John Wiley and Sons.

Weintraub, D., Haines, R., & Randle, R. (1984). The utility of head-up displays: Eye focus vs. decision times. Proceedings of the Human Factors Society 28th Annual Meeting (pp. 529-533). Santa Monica, CA: Human Factors Society.

Wickens, D., & Flach, J. (1988). Information processing. In E. L. Wiener & D. C. Nagel (Eds.). Human Factors in Aviation (pp. 111-155). San Diego, CA: Harcourt Brace Juvanovich.

APPENDIX A

PERMISSION TO USE STUDENTS

MEMORANDUM

To: Dr. Tom Connolly, **Date:** 15 Jan 1993
Chairman, Aeronautical Science Dept.

From: Dr. J.T. McGrath
Chairman, Thesis Committee for H. Nureddine

Subject: Permission to use Air Science students in thesis research.

Mr. Nureddine is conducting a research study of the effects of alternative TCAS display formats on pilot workload and performance. As the attached abstract shows, the project calls for 40 pilot-subjects flying one-hour sessions each. The subjects will be randomly selected from the Air Science student body on a volunteer basis.

This is to request your approval to conduct the above mentioned research on E-RAU's Air Science students.

Thank You.

I concur to the utilization of Air Science students as volunteers in the above research.

X ______________________________

Dr. Tom Connolly
Chairman, Aeronautical Science

APPENDIX B

APPLICATION PROCEDURE FOR THE TLX SCALE

Application Procedure for the TLX Scale

Step 1. Ensure that the following materials are at hand:

1. Blank subject workload report sheet.
2. TLX rating scales.

Step 2. Pair-wise Comparison.

Read: You will be presented with several pairs of demands that were placed on you during the task that you have just completed. You will also be given easy to understand descriptions of what each demand signifies. For each pair presented to you, select that demand which, in your opinion, contributed the most to the overall workload you felt during the experiment. Remember, if you do not understand the description of any demand, please ask me for additional explanation.

Read: **Mental Demand:** Refers to how much mental activity was required (for example, thinking, calculating, remembering, looking, deciding). This refers to how easy or demanding the task was, how simple or complex, how exacting or forgiving it was.

Physical Demand: Refers to how much physical activity was required of you during the task. For example, pushing or pulling the controls, or activating switches. This measure also refers to how demanding the task was from a physical point of view; Was it a slow task or a fast one, was it restful or very active?

Time Demand: Refers to how much time pressure you did feel due to the pace at which the task occurred. Was the pace slow and leisurely, or rapid and frantic?

Performance Level: Refers to how successful you think you were in accomplishing the goals set by the experiment or by yourself. It also refers to how satisfied you were with your performance in accomplishing these goals.

Effort Measure: Refers to how hard you had to work both mentally and physically to accomplish the level of performance acceptable to **YOU**.

Frustartion Measure: Refers to how secure or insecure, discouraged or gratified, irritated or content you felt during the task.

Read: For each pair of measures that will be presented to you now, select the one that you feel contributed more to the overall workload that you felt during this experiment.

Read: Pair-wise comparison, section 1.

Step 3. Rating Scales.

Read: For each of the six measures that follow, indicate on the scale shown to you the magnitude of each measure in the task just completed.

Show Scales. Note: Use the applicable scale for each measure.

Read: Section 2, Rating Scales.

Step 4. Scoring.

1. Tally importance measure on weight scale.
2. Complete Section 2 by multiplying weight by rating and adding the total. Use overlay to determine the rating value.
3. Divide total by 15 to obtain the average workload rating.
4. Record rating in the box at the bottom of the score sheet.

Subjective Workload Report

Date Reported:______________________ ID#:________________

Age:__________________ Gender: M / F

Ratings Held:____

T/T:______________

Section 1: Pair-wise Comparisons.

PD / MD	TD / PD	TD / FR
TD / MD	OP / PD	TD / EF
OP / MD	FR / PD	OP / FR
FR / MD	EF / PD	OP / EF
EF / MD	TD / OP	EF / FR

Tally of Importance:

MD =	**PD =**	**TD =**
OP =	**FR =**	**EF =**

Section 2: Rating Scales.

Demand	Task Rating	Rating X Weight =
MD	________	_____ X _____ =
PD	________	_____ X _____ =
TD	________	_____ X _____ =
OP	________	_____ X _____ =
FR	________	_____ X _____ =
EF	________	_____ X _____ =
		SUM = ______

Sum/ 15= ________________ (Mean Workload Score).

TLX Rating Scale

Mental Demand:

X____X

Mentally not demanding	**Extremely demanding**

Physical Demand:

X____X

Physically not demanding at all	**Extremely demanding**

Time Demand:

X____X

Low or nonexistent time pressure	**Extremely high time pressure**

Performance Level:

X____X

It was an excellent performance	**It was a very poor performance**

Effort:

X____X

Very low effort was needed	**Extremely high effort was needed**

Frustration:

X____X

Not at all frustrated	**Extremely frustrated**

APPENDIX C

EXPERIMENTAL DESIGN AND SOURCES OF INVALIDITY

Sources of Invalidity for the Experimental Design

Internal Sources

1. History	+
2. Maturation	+
3. Testing	(+)
4. Instrumentation	(+)
5. Regression	(+)
6. Selection	+
7. Mortality	-
8. Selection Interaction	+

External Sources

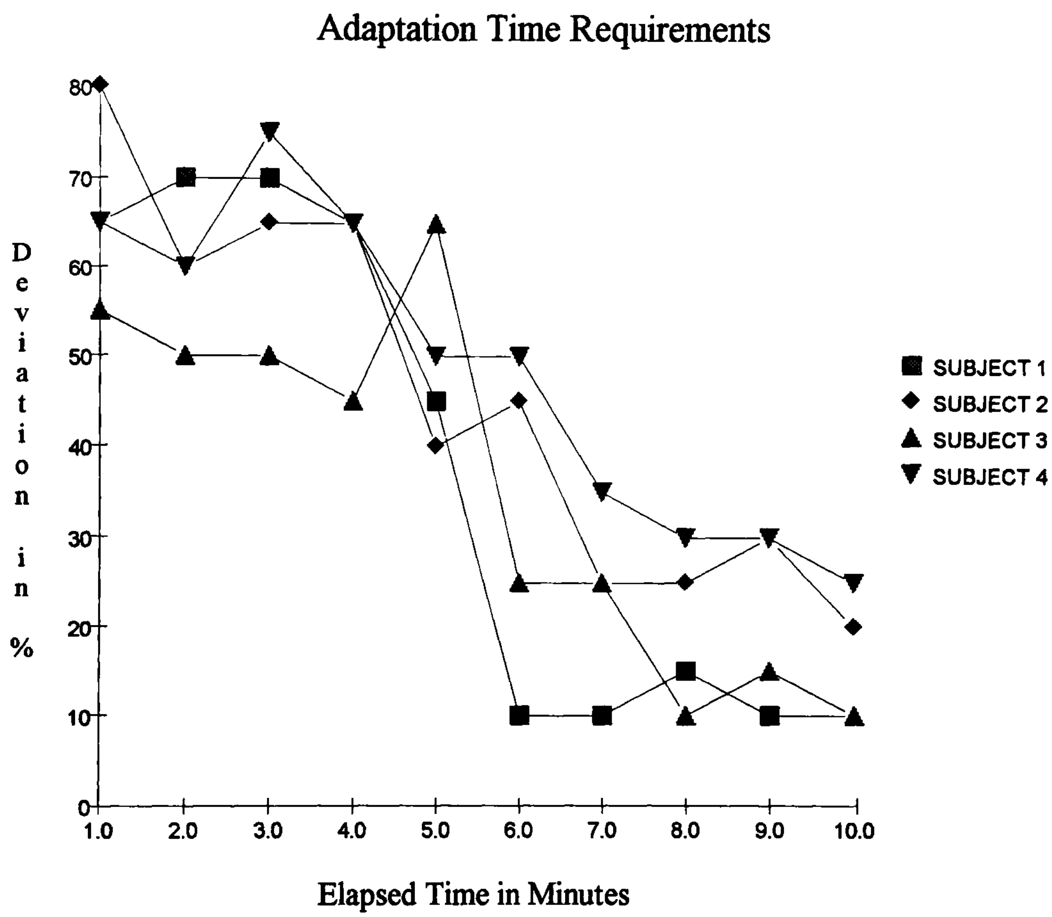
1. Pretest-X Interaction	(+)
2. Multiple-X Interaction	(+)

Symbols: + = Factor controlled for
(+)= Factor controlled for because not relevant
- = Factor not controlled for

Note: Adapted from Educational research (p. 285) by L.R. Gay, 1987, Columbus, OH: Merrill Publishing.

APPENDIX D

FLIGHT SIMULATION ADAPTATION TIMES


Adaptation Time Requirements
Deviation in %
80
70
60
50
40
30
20
10
0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
Elapsed Time in Minutes
SUBJECT 1
SUBJECT 2
SUBJECT 3
SUBJECT 4

APPENDIX E

SUMMARY OF TRAINING SESSION MANEUVERS

1. Bring up the HUD on runway.

2. Explain controls and indications

3. Explain the Takeoff maneuvers.

4. Takeoff, retract gear and flaps on schedule. Adjust throttles on schedule.

5. Maintain runway heading, climb and maintain 1500 feet.

6. Turn left, heading 270, climb and maintain 6000 feet.

7. Turn left, heading 200.

8. Turn left, heading 180, climb and maintain 8000 feet.

9. Turn left, heading 150, descend to 7000 feet.

10. Turn left, heading 130.

11. Turn left, heading 090, descend to 6000 feet.

12. Turn left, heading 010, decsend to 3000 feet

13. Turn right, heading 045, maintain altitude.

14. Turn left, heading 340, descend to 2500 feet.

APPENDIX F

SAMPLE DATA CARD

Subject Data Card

Date Reported:______________________ ID#:________________

Age:__________________ Gender: M / F

Ratings Held:____

T/T:______________

Section 1: Pair-wise Comparisons.

PD / MD	TD / PD	TD / FR
TD / MD	OP / PD	TD / EF
OP / MD	FR / PD	OP / FR
FR / MD	EF / PD	OP / EF
EF / MD	TD / OP	EF / FR

Tally of Importance:

MD =	**PD =**	**TD =**
OP =	**FR =**	**EF =**

Section 2: Tally of Missed Alerts =

Section 3: Reaction 1 = **Reaction 2 =**

Reaction 3 = **Reaction 4 =**

APPENDIX G

RESPONSE TIMES TO TRAFFIC ADVISORIES

Heads-down Control Group Results Summary

ID No.	React #1	React #2	React #3	React #4	Subject Average
3	4.05	2.38	2.01	2.08	2.63
9	2.15	3.24	2.31	4.50	3.05
8	2.54	3.01	1.97	3.17	2.67
2	2.78	2.75	3.38	2.98	2.97
11	1.84	2.68	2.68	1.02	2.06
10	2.38	4.27	2.79	2.45	2.97
15	2.90	2.31	2.15	4.02	2.85
19	1.25	3.10	3.18	1.61	2.28
6	2.12	3.01	2.31	3.61	2.76
21	2.28	1.98	2.04	1.48	1.94
12	2.98	3.53	2.44	1.98	2.73
25	4.87	3.16	2.92	4.89	3.96
14	4.61	3.54	4.98	3.32	4.11
23	2.03	2.31	2.65	2.95	2.49
42	2.36	1.71	1.88	1.84	1.95
33	2.51	3.21	2.45	2.55	2.68
40	2.46	3.20	2.28	2.75	2.67
Averages	2.71	2.91	2.61	2.78	2.75
SD	0.96	0.63	0.74	1.08	

Heads-Up Experimental Group Results Summary

ID No.	React #1	React #2	React #3	React #4	Subject Average
24	1.90	1.72	1.44	1.25	1.58
30	1.46	1.41	1.33	1.56	1.44
28	1.70	1.74	1.19	1.38	1.50
20	1.44	1.36	1.25	1.49	1.39
31	2.08	1.54	2.61	1.54	1.94
7	1.39	1.24	1.05	1.30	1.24
1	1.91	2.50	1.63	2.74	2.20
13	1.32	1.52	1.43	1.51	1.44
29	2.42	2.51	1.63	1.92	2.12
32	2.28	1.90	1.46	1.48	1.78
35	3.02	1.87	3.38	2.26	2.63
44	2.77	2.63	1.89	1.61	2.23
37	0.89	0.92	1.67	1.18	1.17
36	1.68	1.99	1.74	1.55	1.74
43	1.96	2.20	2.01	1.71	1.97
Averages	1.88	1.80	1.71	1.63	1.76
SD	0.57	0.50	0.60	0.41	

APPENDIX H

SUMMARY OF MISSED NON-THREAT ALERTS

Heads-down Control Group Results Summary

ID No.	Warnings Missed
3	4
9	4
8	2
2	4
11	3
10	4
15	1
19	1
6	6
21	1
12	3
25	2
14	3
23	1
42	1
33	3
40	1
Averages	2.59
SD	1.50

Heads-Up Experimental Group Results Summary

ID No.	Warnings Missed
24	0
30	1
28	0
20	2
31	1
7	0
1	0
13	1
29	3
32	3
35	0
44	3
37	0
36	1
43	2
Averages	1.13
SD	1.19

APPENDIX I

SPECIFIC AND OVERALL SUBJECTIVE WORKLOAD SCORES

Heads-down Control Group Results Summary

ID No.	Workload	MD	PD	TD	OP	FR	EF
3	49	70	20	60	50	65	25
9	52	70	25	70	35	30	15
8	58	55	10	35	75	55	50
2	82	90	80	90	80	80	80
11	69	85	55	15	50	80	65
10	58	70	55	65	40	65	75
15	44	80	20	90	25	90	10
19	37	75	75	90	20	75	10
6	71	35	10	90	35	95	85
21	66	85	50	70	30	60	90
12	58	70	30	75	40	70	70
25	36	35	40	10	35	40	25
14	65	60	85	85	60	50	60
23	72	80	10	75	60	80	70
42	35	55	15	55	35	70	15
33	53	50	20	80	50	65	5
40	33.6	40	30	60	20	55	5

Heads-Up Experimental Group Results Summary

ID No.	Workload	MD	PD	TD	OP	FR	EF
24	25	25	15	10	30	20	25
30	18	20	20	15	30	20	10
28	17	30	15	10	5	25	10
20	32	45	5	30	30	40	20
31	43	55	5	55	50	70	30
7	45.6	25	0	90	25	25	20
1	26	30	5	25	35	30	5
13	32	20	25	45	45	15	5
29	18	20	10	5	25	10	20
32	22	20	10	40	10	30	5
35	47.6	55	30	70	40	35	10
44	23.6	30	5	20	45	15	10
37	15	15	20	50	5	10	5
36	38	75	10	50	20	25	5
43	14	10	5	30	10	5	15