12-2005

Simulation for Evaluating the Usability of Integrated Flight Decks

Troy Ricardo Weeks

Simulation for Evaluating the Usability of Integrated Flight Decks

by

Troy Ricardo Weekes

Bachelor of Science
Aeronautical Science
Florida Institute of Technology
2003

A thesis
submitted to Florida Institute of Technology
in partial fulfillment of the requirements
for the degree of

Master of Science
Aviation Human Factors

Melbourne, Florida
December, 2005

We the undersigned committee hereby recommend
that the attached document be accepted as fulfilling in
part the requirements for the degree of
Master of Science of Aviation Human Factors.

"Simulation for Evaluating the Usability of Integrated Flight Decks,"
a thesis by Troy R. Weekes.

____

John E. Deaton, Ph.D.
Associate Professor and Program Chair, Aviation Human Factors
Committee Chair

____

Kenneth E. Crooks, J.D.
Assistant Professor, Aviation Law

____

Lars R. Jones, Ph.D.
Assistant Professor, Humanities and Communication

____

Michael K. Karim, Ph.D.
Associate Professor and Dean, College of Aeronautics

Abstract

Title: Simulation for Evaluating the Usability of Integrated Flight Decks

Author: Troy Ricardo Weekes, B.S.
Committee Chair: John E. Deaton, Ph.D.

There is a modern revolution of pilot interface designs as technological advancements force industrial progress in general aviation (GA). Recent accident statistics and safety reports indicate that computerized flight decks in Technically Advanced Aircraft (TAA) present novel safety hazards. Little empirical data have been used to measure the usability of integrated flight decks due to the complexity and time-consuming nature of flight data assessment. The objectives of this thesis are (1) to describe the human factors issues that are induced by the design of the integrated flight deck, (2) to demonstrate the evaluative capabilities of a TAA flight simulator, and (3) to validate the performance of the TAA flight simulator using an experimental approach. The theoretical and practical applications of a data-driven, evaluative-simulation technique, which measures pilot performance and the usability of integrated flight decks, are discussed. Two experiments are presented to validate the functionality of the TAA flight simulator. In the first experiment, 24 pilots were tested under conditions with and without the flight director while hand-flying two basic attitude instrument maneuvers. Each pilot flew the four possible trials in a different order. Pilot performance was quantified using five objective measures (standard deviation, root mean square error, number of deviations, time outside tolerance and mean time to exceed tolerance) for four flight parameters (altitude, heading, bank angle and vertical speed). Pilot dwell performance was analyzed using the total dwell time in four areas of interest in the flight deck (primary and secondary display, center console and outside view). In the second experiment, the same 24 pilots were tested to compare the usability of a Flight Management System (FMS) with and without the autopilot engaged. In one trial, the pilots flew a scenario with the autopilot while performing in-flight FMS tasks. Pilot IFD operation performance during in-flight FMS tasks was evaluated using the number of operations on the secondary display and keyboard, as well as the total time to complete the prescribed tasks. Pilot performance was measured using the five objective measures for two flight parameters (altitude and heading). In the other trial of the second experiment, the pilots flew the scenario without the autopilot and also performed in-flight FMS tasks. Results for Experiment 1 showed that altitude and heading contained less variability than vertical speed and bank angle. TD was the most sensitive derived performance measure; whereas, MTE and ND were the least sensitive of the derived performance measures. Within-subjects MANOVAs demonstrated significant differences for pilot performance on heading, vertical speed and bank angle across the flight director states. Generally, pilot performance is improved by the use of the flight director. The MANOVAs for dwell times did not reveal any significant changes across the flight director states. The results for Experiment 2 showed that altitude and heading were appropriate flight parameters for demonstrating the effect of autopilot

usage on pilot performance and IFD operation performance. All derived pilot performance measures yielded statistically significant results. Also, the number of operations on the secondary IFD and the task time yielded statistically significant results. However, the number of operations on the keyboard failed to reveal significant differences across the autopilot states. Overall, pilot performance was improved by the use of the autopilot. The evaluative-simulation technique has been used to determine whether or not the flight director, autopilot, and FMS achieve their purposes, to augment pilot performance. Evaluative simulations can be introduced during the system design-phase, then implemented during the system development-phase when prototype technology becomes available.

Acknowledgement

First and foremost, praises are given to the Most High Jah Rastafari for endowing me with the faith, knowledge and health to accomplish this thesis.

Special thanks are given to the thesis committee comprising Dr. John Deaton, Dr. Kenneth Crooks, Dr. Lars Jones, Mr. Mike Ingram and Mr. Simon Matthews. They all assumed responsibility for refining the fundamental ideas and facilitating the successful creation of this thesis.

This thesis could not have been produced without the commitment and excitement of the Avidyne Simulated Flight Environment (AviSAFE) development team members: Terry Reinert, Robert Johnston, Gaurav Saxena, Jarret Rock, Sebastein Giuliano, Patrick Lenihan and Michelle Ruder. One sterling commendation is due to Mr. Rick Willard, who is responsible for the functionality of AviSim. Much of my thinking has been influenced by work and conversations with colleagues at the Avidyne Corporation, namely, Gus Kyriakos, Steve Horning, Tom Landers, Stephen Lindsley, Linda Alflen, Miles Gazic, Steve Kroese, Geoffrey Shapiro, Ted Blumstein, Boris Goranovic, Mike Deroche and Richard Mankowski. I also deeply appreciate communicating with Dr. Esa Rantanen, Dr. David Wilder and Dr. Cecilia Knoll for consultation on algorithms, treatment integrity, experimental control and statistical analysis of the defined measures. I thank Mr. Bud Lowell, Dr. Donna Wilt, and the participants of the study for their involvement. I have great appreciation for their patience with the study's schedule.

Finally, my greatest thanks go to my caring mother, Harriet Davis. I give thankful love to my queen Sable Uzuri and my wonderful family that have inspired me all the way. I reserve a special recognition to my aunt, Ianthe Hill, for being a consistent exemplar of excellence in my life.

Dedication

To the memory of my father,

Captain Richard Davis.

Table of Contents

Abstract iii
Acknowledgement....vi
Dedication....vii
Table of Contents....viii
List of Figures....x
List of Tables....xii
List of Keywords....xiv
List of Abbreviations....xvii
1. Introduction 1
2. Literature Review 6
2.1 GA TAA Accident Analysis 6
2.2 TAA Accident Summaries 10
2.3 Addressing Human Factors Issues at the Pilot Interface 13
2.3.1 Use of Color 15
2.3.2 Symbology 15
2.3.3 Labels 16
2.3.4 System Status Indications, Modes, Annunciations and Messages 16
2.3.5 Controls 17
2.3.6 Display Placement and Readability 18
2.3.7 Warning, Caution and Advisory 19
2.3.8 Error Prevention, Detection and Recovery 19
2.3.9 Automation 19
2.4 Evaluating Usability of the Integrated Flight Deck 22
2.4.1 Objective Pilot Performance Measurement 26
2.5 AviSAFE – Real-time Evaluative-simulation 32
3. Experimental Methodology 36
3.1 Experiment 1 36
3.1.1 Participants 36

3.1.2 Procedure 40
3.1.3 Experimental Trials 42
3.1.4 Experimental Platform 44
3.1.5 Experimental Design 45
3.1.6 Results 48
3.1.7 Discussion 61
3.2 Experiment 2 63
3.2.1 Participants 64
3.2.2 Procedure 65
3.2.3 Experimental Platform 69
3.2.4 Experimental Design 70
3.2.5 Results 74
3.2.6 Discussion 93
4. Conclusion 96
5. List of References 99
6. Appendices 107
6.1 Appendix A – *AviSAFE Architecture* 108
6.2 Appendix B – *Electronic Survey Instrument* 120
6.3 Appendix C – *Experiment 1 Trial Sequence Randomization* 123
6.4 Appendix D – *Experiment 2 Trial Sequence Randomization* 124
6.5 Appendix E – *Flight Data Recorder Specifications* 125

List of Figures

Figure 1. Conventional steam gauge cockpit 1

Figure 2. Modern glass cockpit 2

Figure 3. Percentage of GA accidents by error category by year 7

Figure 4. Usability framework for product development 24

Figure 5. AviSAFE development configuration 34

Figure 6. E2L display with and without the flight director for straight and level........ 42

Figure 7. E2L display with and without the flight director for right climbing turn........ 43

Figure 8. AviSAFE simulation configuration 44

Figure 9. Graph of altitude performance during straight and level........ 50

Figure 10. Graph of heading performance during straight and level 52

Figure 11. Graph of dwell time performance during straight and level........ 54

Figure 12. Graph of vertical speed performance during the climbing turn........ 56

Figure 13. Graph of bank angle performance during the climbing turn 58

Figure 14. Graph of dwell time performance during the climbing turn........ 60

Figure 15. Flight plan and MAP for departure clearance........ 66

Figure 16. Flight plan and MAP showing amendments without the autopilot 68

Figure 17. Flight plan and MAP showing amendments with the autopilot 69

Figure 18. Graph of altitude SD performance under autopilot conditions........ 76

Figure 19. Graph of altitude RMSE performance under autopilot conditions........ 77

Figure 20. Graph of altitude ND performance under autopilot conditions........ 78

Figure 21. Graph of altitude TD performance under autopilot conditions 79

Figure 22. Graph of altitude MTE performance under autopilot conditions 80

Figure 23 Graph of heading SD performance under autopilot conditions 83

Figure 24. Graph of heading RMSE performance under autopilot conditions 84

Figure 25. Graph of heading ND performance under autopilot conditions 85

Figure 26. Graph of heading TD performance under autopilot conditions................. 86

Figure 27. Graph of heading MTE performance under autopilot conditions.............. 87

Figure 28. Graph of NOS performance under autopilot conditions............................ 90

Figure 29. Graph of NOK performance under different autopilot conditions 91

Figure 30. Graph of TT performance under different autopilot conditions 92

Figure 31. AviSim GUI with the E2L communications module expanded 110

Figure 32. Dual E2L instances running with PFD and moving MAP 111

Figure 33. FlightGear scenery displaying outside view of the cockpit..................... 113

Figure 34. The AviPlot GUI in the multiple plot dialog configuration 114

Figure 35. The AviTrace GUI in the tree configuration 116

Figure 36. The AviMedia GUI in the playback mode 118

List of Tables

Table 1. Five most frequent skill-based error categories 8

Table 2. Five most frequent decision error categories 9

Table 3. Top 5 automation issues ranked by overall supporting evidence 20

Table 4. Top 5 automation issues ranking highest in multiple criteria 20

Table 5. Directly measured flight parameters used in the literature 29

Table 6. Derivative measures used in the literature 31

Table 7. Participants of the research study by pilot certificate 38

Table 8. Frequency distribution of subjective rank on PC experience 39

Table 9. Frequency distribution of subjective rank on typing experience 39

Table 10. Within-subjects MANOVA design for altitude performance 45

Table 11. Within-subjects MANOVA design for heading performance 45

Table 12. Within-subjects MANOVA design for vertical speed performance 46

Table 13. Within-subjects MANOVA design for bank angle performance 46

Table 14. Within-subjects MANOVA design for dwell time performance 47

Table 15. Within-subjects MANOVA design for dwell time performance 47

Table 16. Means for altitude performance during straight and level 49

Table 17. Means for heading performance during straight and level 51

Table 18. Means for dwell time performance during straight and level 53

Table 19. Means for vertical speed performance during the climbing turn 55

Table 20. Means for bank angle performance during the climbing turn 57

Table 21. Means for dwell time performance during the climbing turn 59

Table 22. Two-factor mixed MANOVA design for altitude performance 71

Table 23. Two-factor mixed MANOVA design for heading performance................. 72

Table 24. Two-factor mixed MANOVA design for IFD operation performance....... 73

Table 25. Means for altitude performance under different autopilot states 75

Table 26. Means for heading performance under different autopilot states 82

Table 27. Means for IFD operation performance under different autopilot states...... 89

List of Keywords

Autopilot

An autopilot is the function in an automatic flight control system that computes and provides appropriate commands to servo mechanisms. These mechanisms manipulate the airplane control surfaces so that the airplane's attitude is automatically controlled within acceptable tolerances (GAMA, 2004).

Flight Director

A flight director is the software function in an automatic flight control system that computes and provides a visual guidance cue to the pilot. The flight director indicates a calculated trajectory to fly during autopilot and manual flight control. The flight director acts through an autopilot for automatic control of the airplane on a provided flight path (GAMA, 2004).

Flight Management System (FMS)

A Flight Management System (FMS) is a flight computer system that uses a large database that allows routes to be preprogrammed using a data loader (Jeppesen, 2004). The system uses conventional navigational aids, an inertial reference system or the satellite global positioning system to constantly update its position accurately. A typical FMS provides information for continuous automatic navigation, guidance, and aircraft performance management.

Integrated Flight Deck (IFD)

An Integrated Flight Deck combines flight guidance, airborne surveillance, airplane systems, engine systems, and situational awareness control and display functions into a number of interdependent electronic displays. At a minimum, an IFD must include electronic display and control of all primary airplane airspeed, altitude and attitude instruments, and all essential navigation and communication functions (GAMA, 2004).

Multi-Function Display (MFD)

A Multi-Function Display is any other electronic physical display unit, other than the PFD, which can be used to display various types of information, including a variety of required or supplemental information. The items and the location where they are displayed on an MFD may be selected by the pilot (GAMA, 2004).

Primary Flight Display (PFD)

A Primary Flight Display is a single physical unit that displays all of the following aircraft parameters: altitude, airspeed, attitude and direction and flight attitude (GAMA, 2004). Other information may be displayed on the PFD, but this additional information cannot obstruct any items required on the PFD. In accordance with 14 CFR Part 23.1321, the PFD must be located directly in front of the pilot in a fixed layout.

Technically Advanced Aircraft (TAA)

A Technically Advanced Aircraft is defined as an aircraft that has at a minimum, an IFR-certified GPS navigation equipment (navigator) with moving MAP or a multi-function display (MFD) with weather, traffic or terrain graphics, and an integrated autopilot. In general, TAA are aircraft with pilot interfaces provided by one or more computers in order to aviate, navigate, or communicate (TAA Safety Team, 2003).

Usability

Usability is the extent to which a product can be used by specific users to achieve particular goals with effectiveness, efficiency and satisfaction in a specified context of use (ISO, 1998). Effectiveness refers to the accuracy and completeness with which the user achieves intended goals. Efficiency refers to the relation of resources expended to the accuracy and completeness of the goal. Satisfaction refers to the freedom from discomfort and the positive attitudes towards the use of the product. Usability testing includes the methods of measuring user performance and satisfaction, and the study of the principles that may predict whether a product is found usable in practice.

List of Abbreviations

14 CFR: Title 14 Code of Federal Regulations
AC: Advisory Circular
AOI: Area of Interest
AOPA: Aircraft Owners and Pilots Association
ARINC: Aeronautical Radio, Incorporated
ASF: Air Safety Foundation
ATC: Air Traffic Control
AviSAFE: Avidyne Simulated Flight Environment
CDI: Course Deviation Indicator
CRT: Cathode Ray Tube
DTC: Dwell Time on the Center Console
DTP: Dwell Time on the Primary Display
DTO: Dwell Time on the Outside View
DTS: Dwell Time on the Secondary Display
DOT: Department of Transportation
EFIS: Electronic Flight Information System
FAR: Federal Aviation Regulations
FIT: Florida Institute of Technology (or Florida Tech)
FMS: Flight Management System
FOV: Field-of-View
fpm: Feet per minute
ft: Feet
FTE: Flight Technical Error
GA: General Aviation
GAMA: General Aviation Manufacturers Association
GDM: Gardner VORTAC
GPS: Global Positioning System
HCD: Human Centered Design
HF: Human Factors

HFACS: Human Factors Analysis and Classification System
HFDS: Human Factors Design Standards
HFT&E: Human Factors Test & Evaluation
HSI: Horizontal Situation Indicator
Hz: Hertz
ICAO: International Civil Aviation Organization
IFD: Integrated Flight Deck
IFR: Instrument Flight Rules
IMC: Instrument Meteorological Conditions
IP: Internet Protocol
ISO: International Organization for Standardization
KMHT: Manchester Airport
KORH: Worcester Regional Airport
KPWM: Portland International Jetport
LCD: Liquid Crystal Display
M: Mean
MANOVA: Multivariate Analysis of Variance
MAP: MAP
MFD: Multifunction Display
MSL: Mean Sea Level
MTE: Mean Time to Exceed Tolerance
MVFR: Marginal Visual Flight Rules
MVMC: Marginal Visual Meteorological Conditions
ND: Number of Deviations
NOS: Number of Operations on the Secondary IFD
NOK: Number of Operations on the Keyboard
NDB: Non-directional Beacon
NTSB: National Transportation and Safety Board
PFD: Primary Flight Display
PTS: Practical Test Standards
RMSE: Root Mean Square Error
SD: Standard Deviation (of a dependent variable)

SD: Standard Deviation (of a flight parameter)

SEM: Standard Error of the Mean

TAA: Technically Advanced Aircraft

TCP: Transmission Control Protocol

TD: Time Outside Tolerance

TSO: Technical Standard Order

TT: Task Time

UDP: Universal Datagram Protocol

USB: Universal Serial Bus

VFR: Visual Flight Rules

VHF: Very-high-frequency

VORTAC: VHF Omnidirectional Range and Tactical Air Navigation

VMC: Visual Meteorological Conditions

WGS: World Geodetic System

Λ: Wilks' Multivariate Criterion

1. Introduction

From the beginning of powered flight in 1903 through the 1980s, avionics systems comprised conventional “steam gauges” or “six-pack” instrument panels (see Figure 1). During the early 1970s, the military had already begun to integrate human-computer cockpits or “glass cockpits”. In the 1980s and early 1990s, simple glass cockpits were introduced to commercial jets. They comprised of cathode ray tubes (CRTs) that were capable of displaying electronic graphics of the conventional flight instruments. By the mid-1990s, CRT-displays were superseded by liquid crystal displays (LCDs) that presented larger graphics with considerable savings in weight and energy consumption.

Figure 1. Conventional steam gauge cockpit

Today, although the bulk of light general aviation (GA) aircraft still use steam gauges, virtually every newly designed GA aircraft is a technically advanced aircraft (TAA). According to AOPA (2005), many GA aircraft owners are retrofitting their classic aircraft to convert them to TAA with instrument flight rated (IFR) certified global positioning system (GPS) navigators and multifunction displays (MFDs). In the GA TAA Industry Safety Study (2003), the FAA asserted that TAA provide increased *available safety*, or the potential for increased safety. As more of these advanced electronics make their way into the GA cockpit, there is a greater need to study the effects that novel displays have on single-pilot operations (Williams & Ball, 2004). Figure 2 shows a modern glass cockpit of a TAA.

Figure 2. Modern glass cockpit

An important point for the GA community using TAA to notice is that a new mental model is required by pilots controlling TAA since manual flight control must now be synchronized with the management of automated systems. The Aircraft Owners and Pilots Association (2005) affirmed that there should be a modification of what constitutes GA flying to include airline-style operating procedures, regular use of the autopilot, and greater dependence on avionics for multiple tasks beyond sole navigation. With TAA, pilots must add systems management techniques to basic stick and rudder skills in order to process more information without compromising safety.

Automated aids and decision support tools are rapidly becoming indispensable tools in high-technology cockpits and are assuming increasing control of "cognitive" flight tasks, such as calculating fuel-efficient routes, navigating, or detecting and diagnosing system malfunctions and abnormalities (Mosier, Skitka, Heers, & Burdick, 1998). An inescapable facet of automation is that it feeds into the general human tendency to travel the road of least cognitive effort (Mosier et al, 1998).

This mental shift has proven to be challenging for some conventionally trained pilots. Pilots use a wide range of sometimes ineffective monitoring strategies, thus experience considerable problems with tracking the status and behavior of the automation on modern glass cockpits (Mumaw, Sarter & Wickens, 2001; Sarter, Wickens, Mumaw, Kimbal, Marsh, Nikolic & Xu, 2003). There is empirical evidence to suggest that pilots have insufficient knowledge of automation behavior to anticipate important automation state changes (Mumaw, Sarter & Wickens, 2001).

Automation in aircraft flight decks provide pilots with a new heuristic for decision-making and cognitive task performance through information synthesis, system monitoring, diagnosis, planning and prediction, in addition to controlling the physical placement of the aircraft. Even though these systems are designed specifically to decrease mental workload, the presence of automated cues diminishes the likelihood that pilots will either make the effort to seek other diagnostic information or process all available information in cognitively complex ways (Mosier et al, 1998). In addition, automated cues increase the probability that pilots will cut off situation assessment prematurely when prompted to take a course of action recommended by the automated aid (Mosier et al, 1998).

The future cockpits of airliners and business jets are expected to be more sophisticated with higher levels of automation and control input devices such as keyboards, joysticks and trackballs are expected to simplify data input (AOPA, 2005). GA TAA can expect Flight Management Systems (FMS) to become integrated into the user interfaces of the primary flight display (PFD) and multi-function display (MFD).

With this increasing level of automation and functionality, flight safety becomes a challenge to be considered and the application of human factors research in flight deck design becomes imperative. Therefore, in order to rationalize the need for addressing human factors issues in the design-phase of product development, this thesis focuses on the following three major objectives:

1. To describe potential human factors issues induced by the design of the Integrated Flight Deck (IFD), this thesis presents the historical trends in GA flight advancement and TAA accident analysis.
2. To demonstrate the evaluative capabilities of a TAA flight simulator, this thesis introduces an evaluative-simulation strategy as a solution for enhancing the IFD's human centered design (HCD).
3. To validate the performance of the TAA flight simulator using an experimental approach, the final portion of this thesis uses two formal experimental designs and statistical analysis techniques to test seemingly predictable hypotheses.

The hypotheses of the two experiments were constructed to demonstrate how evaluative-simulation can benefit flight deck designers who aim to design safe and usable products. The hypotheses incorporate commonly used features of TAA automation such as the flight director, autopilot and FMS. The hypotheses of the first experiment investigated the effect of the flight director on the pilot's hand-flying ability and visual dwell pattern during a straight and level maneuver, and a climbing turn maneuver. The hypotheses of the second experiment investigated the effect of the autopilot and pilot certification on the pilot's hand-flying ability and operation of the IFD during in-flight FMS amendment tasks. These hypotheses are relevant considering the recent trends that have been revealed in analysis of TAA accidents and incidents.

2. Literature Review

2.1 GA TAA Accident Analysis

Accident analyses have shown that pilots require adequate training and practice with specific TAA systems in order to derive the operational benefits of such systems. Some accident analysts have proposed that safety efforts should be redirected to address human factors issues in the system design process (Wiegmann & Shappell, 2001).

A review of the TAA accidents shows that the majority are caused by a chain of poor decisions and the pilot's lack of experience (TAA Safety Study Team, 2003). There are 14,436 GA accidents between 1990 and 2,000 that were recorded by the databases maintained by the National Transportation and Safety Board (NTSB) and FAA National Aviation Safety Data Analysis Center (Wiegmann, Shappell, Boquet, Detwiler, Holcomb & Faaborg, 2005). The Human Factors Analysis and Classification System (HFACS), which is used to identify the categorical causes of accidents, indicate that skill-based errors were associated with the largest portion (79.2 percent), followed by decision errors (29.7 percent), violations (13.7 percent), and perceptual errors (5.7 percent).

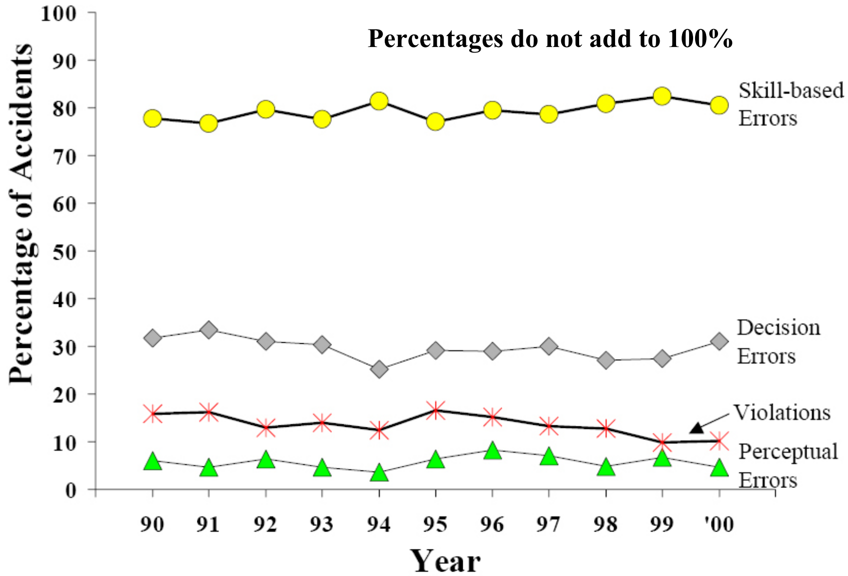


Figure 3. Percentage of GA accidents by error category by year (Wiegmann et al, 2005, p.7)

The proportion of accidents associated with at least one instance of each unsafe act category, as classified by HFACS, remained relatively unchanged over the 11-year period examined in this study (see Figure 3). This seems to suggest that safety efforts directed at GA over the last several years have had little effect on any specific category of human error (Wiegmann et al, 2005). The only exceptions seemed to be a small dip in the percentage of accidents associated with decision errors in 1994, and a gradual decline in violations observed from 1995 to 2,000. Further accident analysis and the experiments of this thesis will examine the contribution of pilot skill and decision-making to GA accidents. Also, the possibility of using evaluative-simulation during the system design process as an intervention technique will be explored.

Table 1. Five most frequent skill-based errors (Wiegmann et al, 2005, p.12)

Error Category	Frequency (percent)		
	FATAL	**NON-FATAL**	**TOTAL**
Directional Control	20 (0.50)	2018 (15.2)	2038 (11.8)
Airspeed	713 (17.9)	1127 (8.5)	1840 (10.6)
Stall/Spin	592 (14.9)	753 (5.7)	1345 (7.7)
Aircraft Control	654 (16.5)	665 (5.0)	1319 (7.6)
Compensation for winds	23 (0.6)	1046 (6.2)	1069 (6.2)

The most frequently occurring human error categories with skill-based errors are presented in Table 1. Approximately 12 percent of all skill-based errors involve errors in maintaining direction control, followed by airspeed (10.6 percent), stall/spin (7.7 percent), aircraft control (7.6 percent) and errors associated with compensating for wind conditions (6.2 percent). Together, these five cause factors accounted for 44 percent, nearly one half, of all the skill-based errors in the NTSB database. "Directional control" typically refers to control of the aircraft on the ground while "aircraft control" refers to control of the aircraft in-flight. The percentage of skill-based errors involving stall/spin, airspeed, and aircraft control were greater for fatal than non-fatal accidents.

Human error categories such as directional control and compensation for wind conditions were rarely associated with fatal accidents. GA TAA have design features such as the autopilot and flight director that can assist the pilot in maintaining "aircraft control". Experiment 1 attempted to show whether or not using the flight director has a significant impact on the pilot's performance maintaining aircraft control.

Experiment 2 attempted to show that the use of the autopilot can improve pilot control of the aircraft as well as general performance on secondary tasks.

Table 2. Five most frequent decision errors (Wiegmann et al, 2005, p.12)

Error Category	Frequency (percent)		
	FATAL	NON-FATAL	TOTAL
In-flight Planning	268 (22.9)	683 (17.0)	951 (18.3)
Ground Planning/Decision-making	115 (9.8)	349 (8.7)	464 (8.9)
Fuel Management	40 (3.4)	413 (10.3)	453 (8.7)
Unsuitable Terrain Selection	16 (1.4)	391 (9.8)	407 (7.8)
Go Around	22 (1.9)	291 (7.3)	313 (6.0)

The most frequently occurring human error categories with decision errors are presented in Table 2. Improper in-flight planning contributes for approximately 18 percent of all decision errors. Errors categorized as in-flight planning refer to planning or revision performed after the aircraft has taken off. The remaining decision errors, such as preflight planning/decision errors (8.9 percent), fuel management (8.7 percent), poor selection of terrain for takeoff/landing (7.8 percent), and go-around decision (6.0), all occurred at approximately the same frequencies. In total, these five causal categories accounted for roughly half (49.89 percent) of all decision errors in the NTSB database. The categories in-flight planning and planning/decision-making on the ground tended to be associated more often with fatal than non-fatal accidents. Whereas the categories unsuitable terrain, go around, and fuel management were associated more often with non-fatal accidents.

GA TAA have equipment such as GPS navigators or integrated FMS that can improve the pilot's ability to successful complete in-flight planning, and when coupled with the use of an autopilot, can make in-flight planning much safer. Experiment 2 incorporated in-flight planning as a secondary task and investigated the impact of using the autopilot while reprogramming the FMS.

To date, analysis of GA TAA accidents reported by the NTSB has failed to show a statistically valid link between distractions blamed on TAA and other distraction-caused accidents in the non-TAA fleet (AOPA, 2005). However, a review of TAA accidents shows that the majority of accidents are not caused by something directly related to the aircraft but by the pilot's lack of experience and a chain of poor decisions (AOPA, 2005).

If the FAA and the aviation industry are to achieve their goal of significantly reducing the aviation accident rate over the next ten years, human factors, which is the primary cause of aviation accidents, must be addressed (Wiegmann & Shappell, 2001).

2.2 TAA Accident Summaries

Some pilots develop a sense of over-reliance in their avionics and the aircraft, believing that the equipment will compensate for pilot deficiencies (AOPA, 2005). Wiegmann and Shappell (2001) emphasized the use of data-driven intervention strategies and stressed that a comprehensive analysis of existing databases needs to be conducted in order to determine the human factors responsible for aviation accidents and incidents.

The NTSB (2002b) database entry, #FTW03FA029, reports that in one specific cross-country accident a non-instrument-rated private pilot encountered heavy fog and poor visibility leading up to a fatal crash when the airplane was destroyed when it impacted the terrain. The private pilot had approximately 1884 hours total flight time, of which 82 hours were in the accident airplane. The airplane was a Cirrus SR20 equipped with a Garmin GNS 430 GPS. The NTSB (2002b) determined the probable cause of the accident as the pilot's inadvertent flight into instrument meteorological conditions (IMC) and failure to maintain clearance with the terrain. A contributing factor was the pilot's failure to obtain an updated preflight weather briefing. The AOPA Safety Foundation Report (2005) inferred that this pilot was tempted to continue flight after encountering IMC because he had TAA equipment on board.

In a similar cross-country accident, a non-instrument-rated pilot continued a visual flight rules (VFR) flight into IMC then encountered icing conditions that led to a fatal crash. The private pilot had approximately 837 hours of total flight experience. The NTSB (2002a) database entry, #NYC03FA015, reports the probable cause of the accident as the pilot's improper in-flight decision to continue flight into known adverse weather conditions. A factor related to the accident was the icing conditions. The AOPA Safety Foundation Report (2005) inferred that the pilot succumbed to the belief that the advanced avionics on board could compensate for the lack of qualifications to fly in IMC.

Other accidents prove that TAA pilots encounter significant difficulty not only with judgment and decision-making but also with the operation of TAA equipment. The NTSB (2000) database entry, #LAX01FA003, reports that on October 2004, an instrument-rated private pilot who took off from an airport with a 600-foot ceiling failed to maintain directional control and altitude. These conditions led to a right descending spiral then a fatal crash when the airplane was destroyed on impact with the terrain. The AOPA Safety Foundation Report (2005) inferred that the pilot became spatially disoriented while trying to use the avionics that he was unfamiliar with.

In another NTSB (2003) database entry, #LAX03FA072, reports that the probable cause of the accident was the pilot's failure to maintain the course for the published approach procedure due to his diverted attention. The distraction was an erroneous frequency assignment provided by ATC and the resultant task overload induced by this problem and the confusion surrounding the ATC clearances to get established on the final approach course, which likely involved repeated reprogramming of the navigation system (NTSB, 2003). The aircraft was equipped with an ARNAV ICDS 2,000 PFD electronic flight instrumentation system (EFIS). The pilot was an instrument-rated private pilot with 460.7 hours total flight time of which 334 hours were in the accident airplane. The AOPA Safety Foundation Report (2005) inferred that the pilot had lost situational awareness due to problems encountered while operating the avionics.

The probable causes of the accidents summarized here imply that there are certain mindsets that are adopted by the pilots of TAA as well as specific operational difficulties that cause or contribute to accidents and fatalities. The flaws seem to lie latent and dormant in the basic designs, operational procedures, training program effectiveness and risk acknowledgement of automated systems (Besco & Funk, 1999). The root cause of some of these accidents might be induced by the human centered design (HCD) of the TAA equipment. This thesis proceeds to show that evaluative-simulation can be used as a design-phase and development-phase technique that can reduce the number of TAA accidents by ensuring that human factors issues and usability criteria are adequately addressed.

2.3 Addressing Human Factors Issues at the Pilot Interface

> "The next step is obvious: we must include human factors requirements into the certification processes of people, procedures, and technology, so that human factors issues are considered at the time when we are defining the blueprint of our system, before it is operational and not after. This is, in my view, a cost-effective approach to anticipate human error rather than regretting its consequences," said Jack Howell, Director, Air Navigation Bureau, ICAO, while addressing the Opening Session of the Third Global Flight Safety and Human Factors Symposium (FAA, 1996).

The pilot interface receives input from the four following sources: pilot, automation, aircraft and environment. All sources of input to the interface require significant human factors consideration. To optimize usability of the interface, functional requirements of the automation must be derived from analysis of piloting tasks. The operations required by the pilot should be simple and not demand high workload; the automation's interaction should be visible to the pilot and not too complex for the pilot to infer the automation's future state.

Some research has claimed that the introduction of modern flight decks, which automate many piloting tasks, has reduced or eliminated some types of pilot errors, but has also introduced other types of errors (FAA, 1996; Sarter & Woods, 1995). Some causes of accidents, summarized previously, have emphasized difficulties of pilot interaction with flight deck automation. Other indicators of potential safety problems, such as pilot reports, training and operational difficulties, research studies, and surveys also point to interface vulnerabilities (Wiegmann et al, 2005).

FAA certification teams have developed a taxonomy of functional areas with reoccurring human factors issues found at the pilot interface during the review of new digital avionics (FAA, 2002a). The FAA taxonomy includes the following functional areas: use of color; symbology; labels; system status indications, modes, annunciations and messages; controls; display placement and readability; warning, caution and advisory; error prevention, detection and recovery; automation. The functional areas that are pertinent to the experiments are introduced briefly within the context of the experimental studies.

2.3.1 Use of Color

Color can greatly improve the usability and effectiveness of a visual display. During the display design phase, great care must be taken when choosing a color philosophy and applying it to a system (FAA, 2002a). The use of many different colors on a single display can decrease the effectiveness of the display. Industry has generally recommended that no more than six colors be used for a single application (FAA, 2002a). Several new Technical Standard Orders (TSOs) require information to be coded using a minimum of two techniques that may include color, shape, and location in order to accommodate the proportion of the pilot population that is colorblind or color deficient (GAMA, 2000). During all experimental trials, the use of color was kept constant by the experimenter.

2.3.2 Symbology

The number of symbols used in displays tends to increase with the need to display more information. In some cases, electronic symbols may be inconsistent with the symbols on paper charts and common flight deck symbology (FAA, 2002a). In other cases, the selected symbols may not be easily discernable, which can affect readability at long distances, off-center viewing angles and poor lighting conditions (FAA, 2002a). Issues involved with the consistency and legibility of symbology present a growing concern with the likelihood of confusion and interpretation errors, which could lead to inappropriate pilot action (FAA, 2002a). Symbology was kept constant by the experimenter within experiments. However, different functional pages containing different symbology were used between the experiments.

2.3.3 <u>Labels</u>

Title 14 CFR Part 23.1555 requires that each cockpit control, other than controls whose functions are obvious, must be plainly marked as to its function and method of operation. Electronic avionics that use multifunction controls are typically labeled with icons and abbreviated captions. It is important to note that the chance that a pilot may inadvertently activate the incorrect control is increased when the pilot does not know which system or function is being controlled (FAA, 2002a). Like symbology, labels were kept constant by the experimenter within experiments. However, different functional pages containing different labels were used between the experiments.

2.3.4 <u>System Status Indications, Modes, Annunciations and Messages</u>

When designers proliferate automation modes without supporting new cognitive demands, which are required by the operator, new mode-related errors form, and failure paths can result (Sarter & Woods, 1995). The effect of automated systems on situation awareness and the out-of-the-loop performance problem have been established as critical issues that can undermine the effectiveness of pilot-aircraft performance (Endsley, 1996).

Transitions between display states in the FMS can occur as an immediate consequence of pilot input, when a preprogrammed target (e.g. a target altitude) is reached or when the system changes its mode autonomously in order to prevent the pilot from putting the aircraft into an unsafe configuration (Sarter & Woods, 1995). The typical FMS exhibits functional flexibility when it provides multiple methods at different levels of automation for the pilot to change altitude. The pilot has to decide the most suitable method depending on the particular flight situation. Therefore, the pilot must know about the functions of the different modes, how to switch between modes, how each mode is set up to fly the aircraft, and how to keep track of the active mode. These are cognitive demands that can accumulate at high-tempo and high-criticality periods, thereby adding more mental workload when pilots are most in need of effective systems (Sarter & Woods, 1995).

Frequently, displays associated with complex automated systems involve electronic screens with information embedded in hierarchical displays associated with various system modes. There have been problems with the operator getting lost in menus, finding the desired display screen, and interpreting cluttered displays (Endsley, 1996). Both experiments required the pilot to interpret various indications, modes and annunciations in order to complete the trials.

2.3.5 Controls

Physical controls provide the primary pilot interface for making data input and selecting system options. Well-designed and placed controls are an absolute necessity for the safe operation of an aircraft. Conventional aircraft control devices include knobs, buttons, levers, switches, wheels, and keyboards (AOPA, 2005). Joysticks, touch pads and trackballs are being utilized in TAA (AOPA, 2005).

The limited instrument panel space forces designers and installers to make compromises when locating controls. The main human factors issue with controls becomes evident when controls are installed so close to each other that it is difficult to operate one control without inadvertently operating another (FAA, 2002a). Also, unrelated controls may be located in the same area of the instrument panel, leading to delays in response time finding the appropriate control, increasing workload and increasing potential pilot errors due to confusion (FAA, 2002a). The TAA flight simulator used in the experiments approximated the flight controls and IFD input devices with low-fidelity Universal Serial Bus (USB) controls and input devices. Despite the low fidelity of the flight controls, their orientation and placement were modeled after direct measurements of the Cirrus SR22 aircraft.

2.3.6 Display Placement and Readability

The introduction and use of digital electronic display systems provided opportunities to significantly change and improve display readability and usability (FAA 2002a). Such designs have implemented vertical tapes with predictive information and coded symbology. In some cases, displays had to be placed in locations outside of the pilot's normal viewing area, which significantly affects display effectiveness and readability (FAA, 2002a). Some displays cannot be read well when viewed from an angle (FAA, 2002a). Displayed symbols and colors viewed from angles may appear different (FAA, 2002a). Reflections caused by external light sources may also affect the readability of the displays (FAA, 2002a). Display placement and readability were kept constant by the experimenter for all experimental trials. Like the flight controls, the display's orientation and placement were modeled after direct measurements of the Cirrus SR22 aircraft.

2.3.7 Warning, Caution and Advisory

Warnings, cautions and advisories were not used during any of the experimental trials.

2.3.8 Error Prevention, Detection and Recovery

Error prevention, detection and recovery were not used during any of the experimental trials.

2.3.9 Automation

Despite the vast span of the eight areas aforementioned, there is a subtle area that contributes significantly at the pilot interface, namely the automation issue. Automation is the allocation of functions to machines that would be otherwise be allocated to humans (Wickens & Hollands, 2000). The modern IFD contains automated features that are physically altered by the pilot. Some of these automated features are flight directors, autopilots, autothrottles, flight management systems, and centralized warning and alerting systems (Funk et al, 1999). The introduction of automation is often intended to reduce workload and augment performance; however, this is not always the result as the arising human factors issues indicate.

Table 3. Top 5 automation issues ranked by overall supporting evidence (Funk et al, 1999, p. 120)

Rank	Issue ID	Abbreviated Issue Statement	Sum of Strengths
1	issue105	Understanding of automation may be inadequate	+63
2	issue083	Behavior of automation may not be apparent	+35
3	issue131	Pilots may be overconfident in automation	+33
4	issue092	Displays (visual and aural) may be poorly designed	+32
5	issue133	Training may be inadequate	+31

Table 4. Top 5 automation issues ranking highest in multiple criteria (Funk et al, 1999, p. 121)

Issue ID	Abbreviated Issue Statement	Cit	Agmt	Crit	Str	Rank Sum	Meta Rank
issue102	Automation may demand attention	1	2	10	18	31	1
issue108	Automation behavior may be unexpected and unexplained	3	23	18	8	52	2
issue131	Pilots may be overconfident in automation	2	32	23	5	62	3
issue025	Failure assessment may be difficult	16	6	17	26	65	4
issue083	Behavior of automation may not be apparent	7	20	34	6	67	5
Cit = rank number of (unsubstantiated) citations; Agmt = rank by mean expert agreement rating; Crit = rank by mean expert criticality rating; Str = rank by sum of evidence strengths							

The issues with the greatest overall supportive evidence (Table 3) and especially those issues ranking highest in multiple criteria (Table 4) are considered as problems that require solutions (Funk et al, 1999). These issues corroborate the human factors issues identified by other researchers. Pilots have reported significant difficulties in understanding what their automated flight management systems were doing and why (Funk et al, 1999; Sarter & Woods, 1992). Substantial empirical evidence from surveys and similar studies indicates that "glass cockpit" pilots sometimes lose track of the status and behavior of automated flight deck systems and, as a result, experience "automation surprises" (Sarter et al, 2003; Mumaw et al, 2001; Mumaw, Sarter, & Wickens, 2001; Funk et al, 1999).

Many of the factors that can lead to situation awareness problems–monitoring, passive decision-making, poor feedback, poor mental models–can be directly traced to the way that the automated systems were designed (Endsley, 1996). As such, it is essential that there be measures in place that minimize these problems during system design, thus allowing the potential benefits of automation to be realized without depriving the pilot of the situation awareness needed for good performance.

Automation bias, which refers to omission and commission errors resulting from the use of automated cues as a heuristic replacement for vigilant information seeking and processing, is a significant factor in pilot interaction with automation, and pilots are not utilizing all available information when performing tasks and making decisions in conjunction with automation (Mosier et al, 1998).

In 1996, the FAA Human Factors Team recommended that "the FAA should task an aviation industry working group to produce a set of guiding principles for designers to use as a recommended practice in designing and integrating human-centered flight deck automation." In 1999, the FAA issued revisions to AC 23.1309-1C, , and AC 23.1311-1A, and developed an Industry Guide to Product Certification that is to be used by aircraft and flight deck manufacturers to develop compliant products. In 2,000, the GAMA published its Recommended Practices and Guidelines for Part 23 Cockpit/Flight Deck Design to provide recommendations for the design of cockpits/flight decks to enhance overall aircraft safety.

Experiment 1 dealt with the impact of the flight director on pilot performance and dwell time performance. In the current implementation of the flight director in the Cirrus SR22, there is a two-position switch that allows the pilot to toggle the flight director on and off. During Experiment 1, the experimenter controlled the state of the flight director in order to investigate its effects. Experiment 2 dealt with the impact of the autopilot on pilot performance and IFD operation performance. In the current implementation of the autopilot in the Cirrus SR22, there is a two-position switch that allows the pilot to engage and disengage the autopilot. During Experiment 2, the experimenter controlled the state of the autopilot in order to investigate its effects.

2.4 Evaluating Usability of the Integrated Flight Deck

In system design, "human needs and abilities are the guiding forces" (Shneiderman & Plaisant, 2005, p.25). In a survey conducted by Tenny, Rogers, & Pew (1998), the average response of 132 pilots on 13 questions related to human-centeredness of flight deck automation was 3.53, in which 5 indicates a maximum human-centeredness response. This score showed that participants tended to endorse the human-centered view of automation development. The participants wanted flight deck automation to be simple, reliable and to produce predictable results (Tenny et al, 1998). In addition to clarifying pilot preferences for future design efforts, the survey results urged researchers to develop a scientific basis for a human-centered design (HCD) philosophy.

A vital foundation for the designers of interactive systems is an understanding of the cognitive and perceptual abilities of users (Hackos & Redish, 1998; Wickens & Hollands, 2000). In order to provide universal usability, novice-expert differences, age ranges and technological diversity guide design, and facilitate the plasticity of the interface by transforming its content (Shneiderman & Plaisant, 2005). In addition to accommodating different classes of users and skill levels, designers need to support a wide range of hardware and software platforms while ensuring backward compatibility through generations of software (Shneiderman & Plaisant, 2005).

Planning for usability as part of IFD system design and development involves the systematic identification of requirements for usability measures and verifiable descriptions of the context of use (ISO, 1998). In a summative usability evaluation, several measures are used for benchmarking the usability of a product (ANSI, 2001). Figure 4 illustrates a usability framework with measurable and verifiable components and relationships (ANSI 2001; ISO, 1998).

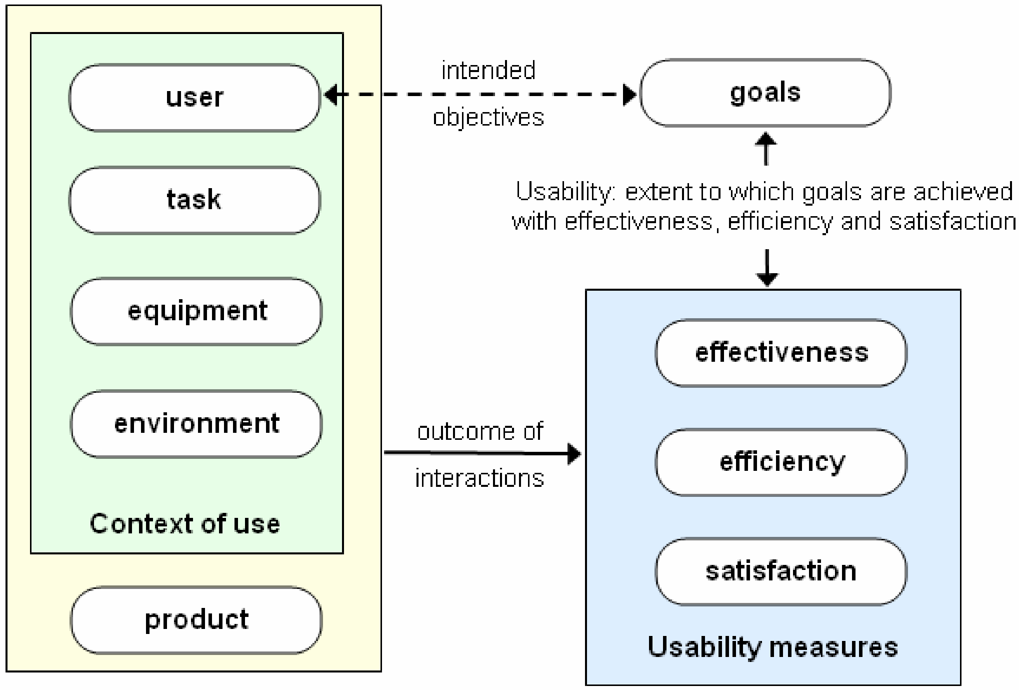


Figure 4. Usability framework for product development (ISO, 1998, p. 4)

There is general agreement from the standards boards ANSI (2001) and ISO (1998) as to what the dimensions of usability are (effectiveness, efficiency & satisfaction) and, to a lesser extent, which metrics are most commonly used to quantify those dimensions. Effectiveness includes measures for completion rates and errors, efficiency is measured from time on task, and satisfaction is summarized using any of a number of standardized satisfaction questionnaires (ANSI, 2001; ISO, 1998).

When implemented correctly, usability requirements provide measurable design targets which form the basis for verification of the resulting design (Hackos & Redish, 1998). Early involvement of users is crucial to the design process because the cost of making revisions increases dramatically as the development of the system proceeds (Clamann & Kaber, 2004). If usability problems with the system are detected by domain experts interacting with functional prototypes, problems can be corrected, hence reducing the possibility of producing a flawed product and placing the burden on users during the training process (Clamann & Kaber, 2004). The availability of powerful computer hardware and rapid prototyping tools has simplified the process of designing mockups.

The coding parameters of an IFD interface are used consistently to attract the pilot's attention to important information, changes in the state of a system, unusual situations, or potential problems that require user action. For example, in the Avidyne Entegra user interface, spatial coding is implemented in the use of information tiles that aggregate related parameters. Color coding is used for many purposes, for example, to determine the active flight leg in the flight plan on the FMS and moving map (MAP).

There are several interaction types used in the operation of the IFD depending on the task requirements and the preference of the pilot. The Avidyne Entegra user interface features function keys to enable IFD function selection. This type of interaction is appropriate since the tasks in a given functional mode involve choices within constrained sets of alternatives. When the pilot selects the MAP function, little training is required to predict the general outcome. When the pilot interacts with the FMS function, a direct manipulation of a flight leg could be visualized as the waypoint entry on the MAP function. Thus, the manipulations in the text-based FMS function are mimicked by the visual appearance of concrete waypoint objects and flight legs in the MAP function. Due to the direct nature of this type of interaction, users with little or no training may still be able to complete in-flight planning tasks.

Entering a waypoint identifier requires the use of a form completion. In this case data entry is flexible and input is possible through knob or keyboard controls. Knob controls for data entry require moderate practice even though computer algorithms may enhance form completion. Subjective assessments claim that, relative to the keyboard entry for form completion, knob control responses may be slow and require complex manipulation. This is a reasonable hypothesis but empirical validation is needed.

2.4.1 <u>Objective Pilot Performance Measurement</u>

Subjective evaluations of performance are advantageous due to the relative ease of implementation, high face validity, and simplicity in providing specific feedback, but are disadvantageous due to problems of inter and intra-rater reliability (Johnson & Rantanen, 2005). According to Johnson and Rantanen (2005), objective measures based on flight data recordings have the potential to alleviate these concerns.

The use of objective performance measures can provide an alternative and complimentary approach to subjective evaluation in both training and research environments (Johnson & Rantanen, 2005). Avidyne Simulated Flight Environment (AviSAFE) flight data recording and processing suite automatically collects the type of data that makes objective pilot performance measurement practical and inexpensive

(Appendix A). The objective performance measures, which are used in the experiments to measure pilot performance, are described below in detail.

Standard deviation (SD) describes the amount of variability around the mean of a given measure. For a flight maneuver, a small SD is indicative of good pilot performance. However, SD does not provide any information about the possible error relative to a given criteria (Rantanen & Talleur, 2001; Rantanen, Talleur, Taylor, Bradshaw, Emanuel, Lendrum & Hulin, 2001). One such criterion could be the tolerance of a practical test standard.

Root mean square error (RMSE) summarizes the overall error and describes tracking performance for a specified value of a flight parameter. For example, RMSE can describe how well the pilot maintained an altitude of 5,000 feet MSL over a five-minute period. For a flight maneuver, a small RMSE is indicative of good performance. However, RMSE does not contain information about the direction of deviations or the frequency of deviations from the criterion (Rantanen & Talleur, 2001; Rantanen et al, 2001).

The number of deviations (ND) is a count of the number of occurrences when the aircraft strays outside a predefined tolerance. ND compliments RMSE since it provides a measure of the frequency of deviations (Rantanen & Talleur, 2001; Rantanen et al, 2001). A small ND is indicative of good pilot performance. However, ND must be taken into consideration with the time that is spent outside the tolerance since a pilot can make a few deviations but spend a substantial amount of time outside the tolerance (Rantanen & Talleur, 2001; Rantanen et al, 2001). This might appear a good performance if the time spent outside the tolerance is not taken into account.

The time outside tolerance (TD) is the cumulative time that the aircraft spends outside a given tolerance. TD provides an indication of tracking performance beyond RMSE and ND (Rantanen & Talleur, 2001; Rantanen et al, 2001). A small TD indicates good performance.

The mean time to exceed tolerance (MTE) at any time is computed from the rate of change between successive data points and the aircraft's position relative to a given tolerance (Rantanen & Talleur, 2001; Rantanen et al, 2001). MTE provides additional information to the description of pilot performance since it provides an indication of tracking performance within the tolerance region as opposed to ND and TD, which provide tracking information outside the tolerance region.

Table 5. Directly measured flight parameters (Johnson & Rantanen, 2005, p.4)

Parameter	Frequency	percent	Cum. percent
Altitude	21	12.88	12.88
Airspeed	19	11.66	24.54
Roll	17	10.43	34.97
Control Inputs	17	10.43	45.40
Heading	16	9.82	55.21
Pitch	16	9.82	65.03
Vertical Speed	11	6.75	71.78
VOR Tracking	8	4.91	76.69
Yaw	5	3.07	79.75
Turn rate	5	3.07	82.82
Glide Slope Tracking	5	3.07	85.89
Flaps	4	2.45	88.34
Trim	4	2.45	90.80
Speed Brakes	3	1.84	92.64
Sideslip	3	1.84	94.48
Landing Gear	3	1.84	96.32
Acceleration	3	1.84	98.16
Position	2	1.23	99.39
NDB Tracking	1	0.61	100.00

Johnson and Rantanen (2005) described the frequency of use of several flight parameters and derivative measures such as SD, RMSE, ND, TD and MTE, and concluded that 55 percent of the studies in literature measure and record altitude, airspeed, roll, control inputs and heading as direct flight parameters. Table 5 shows the flight parameters that have been used to measure pilot performance ordered by their frequency of use. The experiments in this study utilized the following flight parameters with the respective ranks as depicted in Table 5: altitude (1), roll or bank angle (3), heading (5) vertical speed (7).

The raw data of the directly measured flight parameters need further processing in order to provide useful information about pilot performance. For example, the root mean square error (RMSE) for altitude with error bars ± 1 standard error of the mean (SEM) was a reliable measure for vertical flight technical error (FTE) (Oman, Kendra, Hayashi, Stearns, & Burki-Cohen, 2001; Levy, Som & Greenhaw, 2003; Prinzel, Comstock, Glaab, Kramer, & Arthur, 2004; Schnell, Kwon, & Merchant, 2004). Likewise, the RMSE for cross-track error with error bars ± 1 standard error of the mean (SEM) was a reliable measure for lateral FTE (Oman, Kendra, Hayashi, Stearns, & Burki-Cohen, 2001; Levy, Som & Greenhaw, 2003; Prinzel, Comstock, Glaab, Kramer, & Arthur, 2004; Schnell, Kwon, & Merchant, 2004). Table 6 shows the derivative measures that have been used throughout the literature to measure pilot performance.

Table 6. Derivative measures (Johnson & Rantanen, 2005, p.5)

Derivative Metric	Frequency	percent	Cum. percent
Root Mean Square Error (RMSE)	16	21.92	21.92
Standard Deviation (SD)	8	10.96	32.88
Max/Min	8	10.96	43.84
Mean	6	8.22	52.05
Frequency Analyses	5	6.85	58.90
Range	5	6.85	65.75
Deviation from Criterion	4	5.48	71.23
Time on Target	4	5.48	76.71
Mean Absolute Error	3	4.11	80.82
Autocorrelation	3	4.11	84.93
Time Outside Tolerance (TD)	3	4.11	89.04
Median	2	2.74	91.78
Number of Deviations (ND)	2	2.74	94.52
Boolean	1	1.37	95.89
Correlation	1	1.37	97.26
Moments	1	1.37	98.63
Mean Time Exceed Tolerance (MTE)	1	1.37	100.00

The experiments in this study utilize the following derivative measures with the respective ranks as depicted in Table 6: RMSE (1), SD (2), TD (11), ND (13) and MTE (14). The flight data processing suite incorporates flight parameter tolerances defined in the practical test standards for commercial pilots in order to measure pilot performance objectively. Tests for FAA pilot certificates and associated ratings are administered by FAA inspectors and designated pilot examiners in accordance with FAA-developed practical test standards (PTS). Title 14 CFR Part 61 specifies the areas of operation in which knowledge and skill must be demonstrated by the applicant and the PTS contains the standards to which maneuvers on FAA practical tests must be performed (FAA, 2002b). In training and testing, pilots are graded according to these practical test standards, with clear indications of pass or fail. During evaluation with the flight data processing capability of AviPlot, a pass or fail can be further analyzed by referring to the five pilot performance metrics (Johnson & Rantanen 2005; Rantanen & Talleur, 2001; Rantanen et al, 2001).

2.5 AviSAFE – Real-time Evaluative-simulation

This thesis asserts that human factors test and evaluation (HFT&E) of IFDs in TAA should rely greatly on various levels of simulation. Such simulation levels range from part-task evaluations for each major IFD function up to a comprehensive IFD scenario-based evaluation for the entire system. However, the results of a simulator can only be trusted if the simulator has been verified.

Content validity, or the degree to which the content of the simulation is representative of the simulated situation, is an important criterion for the overall validity of low-fidelity simulations. Whereas content validity is an important first step in establishing the validity of a PC-based simulation, the more important criterion in research is a simulation's construct validity (Jentsch & Bowers, 1998). In the study performed by Jentsch & Bowers (1998), researchers assessed the construct validity of flight simulations by demonstrating: (a) that one could reliably manipulate variables of interest to create a range of performance settings, (b) that performance in the range of settings was related to a range of behaviors, which was interpreted as evidence of convergence validity, and (c) that participant performance on specific behaviors correlated only with the underlying theoretical constructs but not with unrelated constructs, which was interpreted as evidence of discriminant validity.

The key to ensuring validity lies in the appropriate use of available hardware and software and in the design of simulator scenarios (Jentsch & Bowers, 1998). Well designed scenarios incorporate easily simulated details to increase realism, such as accurate radio communications and background radio chatter (Jentsch & Bowers, 1998). Also, avoiding attempts to simulate events that cannot be faithfully recreated in the simulator is important in ensuring the validity of simulations (Jentsch & Bowers, 1998). Realistic scenario-based simulation can comprise typical GA transportation operational scenarios, with an emphasis on situations that have traditionally caused fatal accidents, including abnormal operations. For example, a non-instrument rated pilot flying into marginal visual metrological conditions (MVMC) in mountainous areas that become instrument meteorological conditions (IMC), or a new instrument-rated pilot receiving last-minute approach clearance amendments coupled with an erroneous tower frequency. Realistic scenarios, such as those in Experiment 2, incorporate realistic radio communications which are typically described as a deficiency in flight simulation (Longridge, Burki-Cohen, Go & Kendra, 2001).

One of the issues identified by many studies is a lack of pilot understanding of the automation (FAA, 1996; Endsley, 1996; Sarter & Woods, 1994; Sarter & Woods, 1995). The most direct way to gain an understanding of the pilot skills required to use flight deck automation would be to observe pilots using the automation while they fly the aircraft. However, this may not be efficient because various strategies may be employed by different pilots using the automation as well as the inefficient use of valuable resources. Another method to identify and understand the skills necessary to fly automated aircraft is to observe and record pilot use of automation during structured scenarios flown in an aircraft simulator of reasonable fidelity.

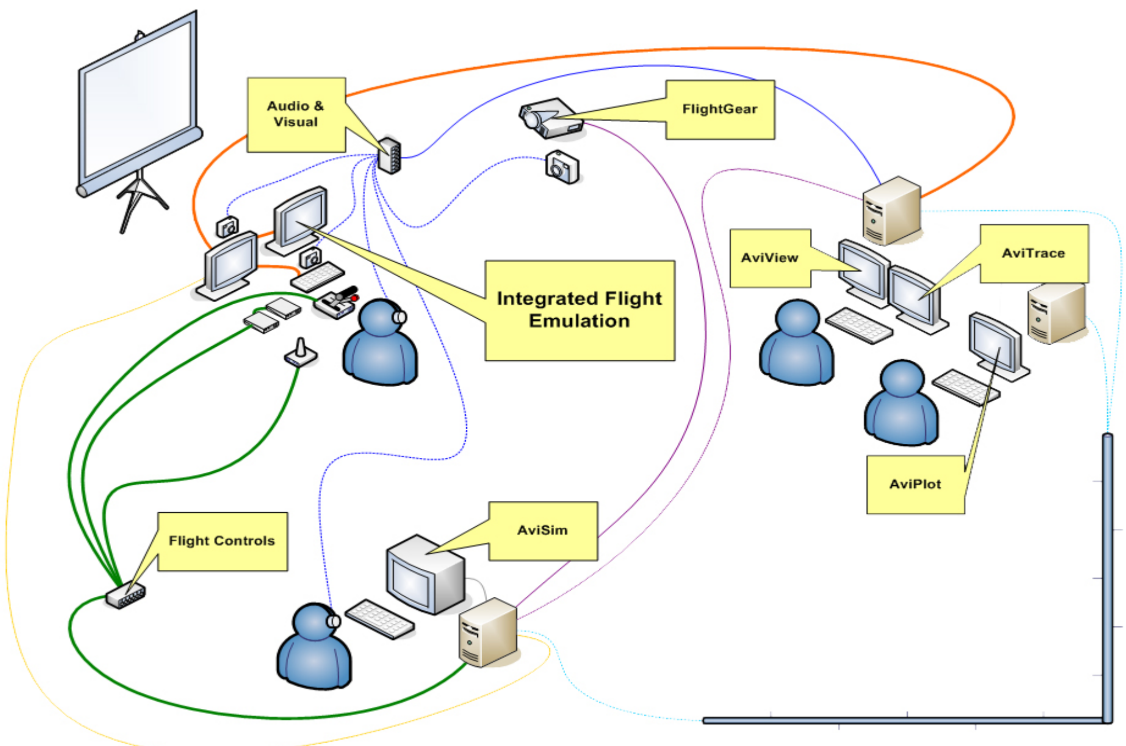


Figure 5. AviSAFE development configuration

Evaluative-simulation can be enhanced by the current technological capability of capturing pilot performance data in a digital form, ready for immediate processing. These data, once processed, can be analyzed in detail to fully understand the factors that influence the pilots' use of automation and the skills associated with their use of the automation. In general, performance measures provide the advantage of being objective and are usually non-intrusive (Endsley, 1996). AviSAFE comprises an integrated flight emulation and a flight data recording and processing suite that objectively quantifies pilot performance in real-time (see Figure 5).

Simulation data include all performance, stability, control and other necessary flight parameters digitally recorded to measure pilot performance and verify the efficacy of the simulator. Simulation data are validated by comparing them to a similar set of test parameters that are digitally recorded in an airplane using a calibrated data acquisition system of sufficient resolution. The third objective of this thesis, which concerns the experimental validation of AviSAFE's efficacy, precedes verification using planned comparisons with airplane flight data. The experiments showed the HFT&E of several IFD features in the TAA flight simulator.

3. Experimental Methodology

3.1 Experiment 1

In the first study, twenty-four pilots were trained on the purpose and use of the flight director then tested in randomly ordered trials to investigate the effect of using the flight director while hand-flying two scenario maneuvers. The following hypotheses were postulated for Experiment 1:

- While hand-flying, pilots experience better pilot performance on the straight and level maneuver when using the flight director compared to when not using the flight director
- While hand-flying, pilots experience better pilot performance on the climbing turn maneuver when using the flight director compared to when not using the flight director
- Flight director usage causes a different monitoring strategy as more time is spent focusing on the primary IFD when using the flight director

3.1.1 <u>Participants</u>

Twenty-four pilots from FIT Aviation flight school were the participants involved in both experimental studies (see Table 7). The pilots were randomly selected from a group of volunteers willing to participate. Volunteers registered for participation by using a survey instrument (Appendix B). The group of registered participants contained 5 women and 19 men of which 6 were flight instructors.

There were 8 student pilots with a range of 18 to 72 hours (M = 34.5 hours; SD = 20.036 hours). Only one of the student pilots had one hour of instrument time; the other student pilots had none (M = 0.125 hours; SD = 0.354 hours). There were 8 private pilots with a range of 60 to 182 hours (M = 125 hours; SD = 39.381 hours). The instrument time for the private pilots was within a range of 4 to 37 hours (M = 11.125 hours; SD = 10.842 hours). There were 8 commercial pilots with a range of 270 to 2500 hours (M = 617.25 hours; SD = 763.235 hours). The instrument time for the commercial pilots was within a range 43 to 350 (M = 109 hours; SD = 100.858 hours).

Table 7. Participants of the research study by pilot certification

SubjectID	Certificate	Instructor	Flight Hours	Instrument Hours	Sex
23	Student	No	72	0	Male
18	Student	No	60	0	Male
19	Student	No	31	0	Female
10	Student	No	26	0	Male
8	Student	No	23	1	Male
4	Student	No	23	0	Male
12	Student	No	23	0	Male
24	Student	No	18	0	Male
		M	34.5	0.125	
		SD	20.036	0.354	
1	Private	No	182	13	Female
13	Private	No	164	4	Male
21	Private	No	140	7	Male
11	Private	No	130	7	Male
20	Private	No	130	5	Male
22	Private	No	100	37	Male
17	Private	No	94	6	Female
16	Private	No	60	10	Male
		M	125	11.125	
		SD	39.381	10.842	
2	Commercial	Yes	2500	350	Female
15	Commercial	Yes	480	100	Male
6	Commercial	Yes	375	103	Male
5	Commercial	Yes	350	90	Male
7	Commercial	No	330	70	Male
3	Commercial	No	318	30	Female
14	Commercial	Yes	315	86	Male
9	Commercial	Yes	270	43	Male
		M	617.25	109	
		SD	763.235	100.858	

Other demographic information collected includes a subjective rank of personal computer experience and personal typing experience. Table 8 shows the frequency distribution for subjective rank of personal computer experience with "Average" as the modal category. The second highest category is "Above Average".

Table 8. Frequency distribution of subjective rank on PC experience

Rank of PC experience	Frequency
Basic	2
Below average	0
Average	12
Above average	9
Extensive	1

Table 9 shows the frequency distribution for subjective rank of typing experience with "Above Average" as the modal category. The second highest category is "Average". These subject variables, subjective ranks on PC experience and typing experience, are not analyzed using inferential statistics.

Table 9. Frequency distribution of subjective rank on typing experience

Rank of typing experience	Frequency
Basic	1
Below average	3
Average	7
Above average	10
Extensive	3

3.1.2 <u>Procedure</u>

The steps of this procedure were replicated for each participant in the study. After signing the consent form and providing demographic information, the participant was briefed on the arrangement of the study into four trials. A training presentation was conducted to demonstrate the purpose and use of the flight director to maintain aircraft control on a given flight path. During the training session other features and limitations of AviSAFE were discussed within the context of the study. During the briefing and training presentation, the participant was not told the flight parameters being measured or the order of the four trials.

The experimental conditions involved two flight maneuvers being performed with two flight director states resulting in four trials and twenty-four (or 4!) possible permutations of the four trials. Therefore, each participant was randomly assigned a number from 1 to 24, without replacement, in order to determine the order in which the scenario maneuvers were to be executed (Appendix C). The participant was not informed about the order of the maneuvers but was told what the two maneuvers entailed as well as the flight director conditions.

Subsequently, each participant was taken to the flight simulator and seated in the left pilot seat. The test administrator turned on the IFDs of the flight simulator in a simulation mode with the appropriate flight director state, established the flight emulation program in a paused state and activated the systems of the flight data recording and processing suite in a prestart state. The participant was given a headset that was coupled to the test administrator's headset using an inter-communication (intercom) system in the flight simulator. Once the volume levels on the intercom were adjusted for comfort by the participant and test administrator, the test administrator sat at a remote monitoring station and began to calibrate the eye-tracking camera system in the flight simulator with the participant sitting as if flying.

Over the headset, the test administrator explained, to the participant, brief descriptions of the maneuvers to be completed as well as the limitations of the flight simulator within the context of this study. The participant was given a chance to ask questions for clarity and receive answers. Following that, all systems of the flight data recording and processing suite were activated into an online state, ready to receive data. Once the systems were online, the test administrator issued a countdown then initialized the simulation with the autopilot engaged flying on an eastbound heading of 090^{o}, straight and level at 5,000 feet mean sea level (MSL).

The test administrator issued the command for the first maneuver and told the pilot “you have the flight controls” while disengaging the autopilot. When the maneuver was completed, the test administrator reengaged the autopilot while saying “I have the controls”.

Following each climbing turn maneuver, the flight data recording and processing suite was stopped and all data were saved to files. The test administrator reestablished the aircraft in the initial state and the systems of the flight data recording and processing suite ready to capture.

When the simulation was online, the test administrator used the autopilot to fly the aircraft eastbound, straight and level at 5,000 feet mean sea level (MSL). The test administrator and participant repeated this procedure until the straight and level maneuver was completed with and without the flight director, and the climbing turn maneuver was completed with and without the flight director. After the four trials, the participant was debriefed and thanked for contribution to the first experiment.

3.1.3 Experimental Trials

The experimenter used four trials to control the flight maneuver (straight and level or right climbing turn), and the flight director state (on or off). The order of the trials was determined and assigned to participants randomly.

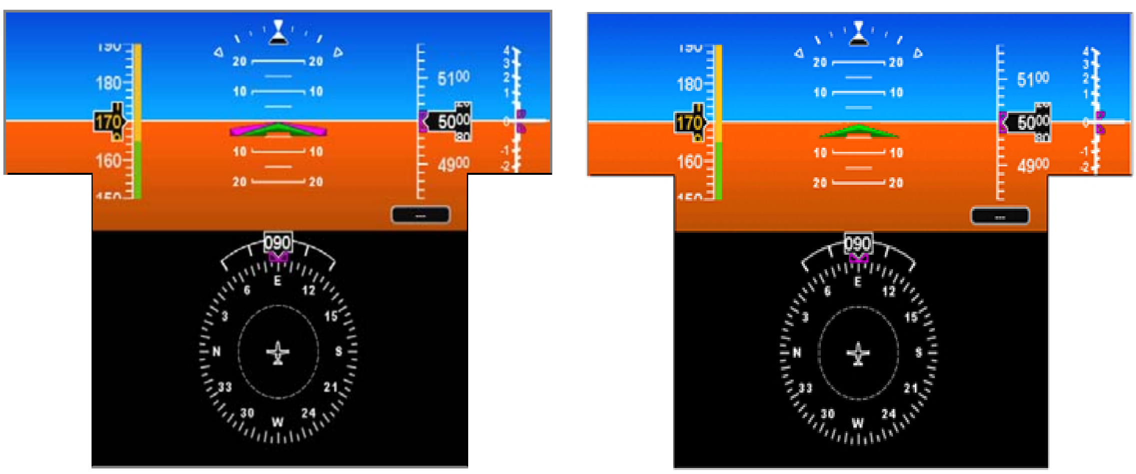


Figure 6. E2L display with (left) and without (right) the flight director for straight and level

During both trials with the straight and level maneuver, the participant was required to maintain an eastbound heading of 090^{o} at an altitude of 5,000 feet mean sea level (MSL) and airspeed of 170 knots for 5 minutes (see Figure 6). The airspeed was not evaluated but was required to be controlled. The tolerances under investigation are as follows: altitude, 5,000±100 feet; heading 090±10^{o} (FAA, 2002b).

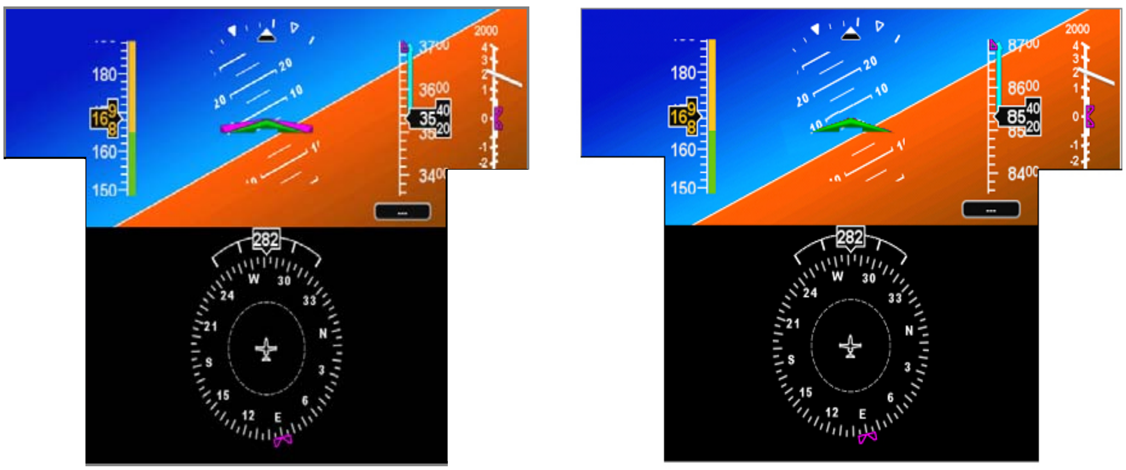


Figure 7. E2L display with (left) and without (right) the flight director for right climbing turn

During both trials with the climbing turn maneuver, the participant was required to begin from a eastbound heading of 090^{o} at an altitude of 5,000 feet mean sea level (MSL) and airspeed of 170 knots. Once established in initial configuration, the participant was required to initiate and complete a 3-minute constant rate climbing right turn with a vertical speed of 2,000 feet per minute (fpm) and a 25^{o} bank angle (see Figure 7). The airspeed was not measured but was required to be controlled within the normal operating range of the Cirrus SR22 (within the green arc of the airspeed indicator). The tolerances under investigation are as follows: vertical speed, 2,000±500 fpm; bank angle 25±5^{o} (FAA, 2002b).

3.1.4 Experimental Platform

The study was conducted in AviSAFE, which is described in Appendix A. Figure 8 shows AviSAFE with a participant and the test administrator in the simulation setting that was configured as the Cirrus SR22 aircraft.

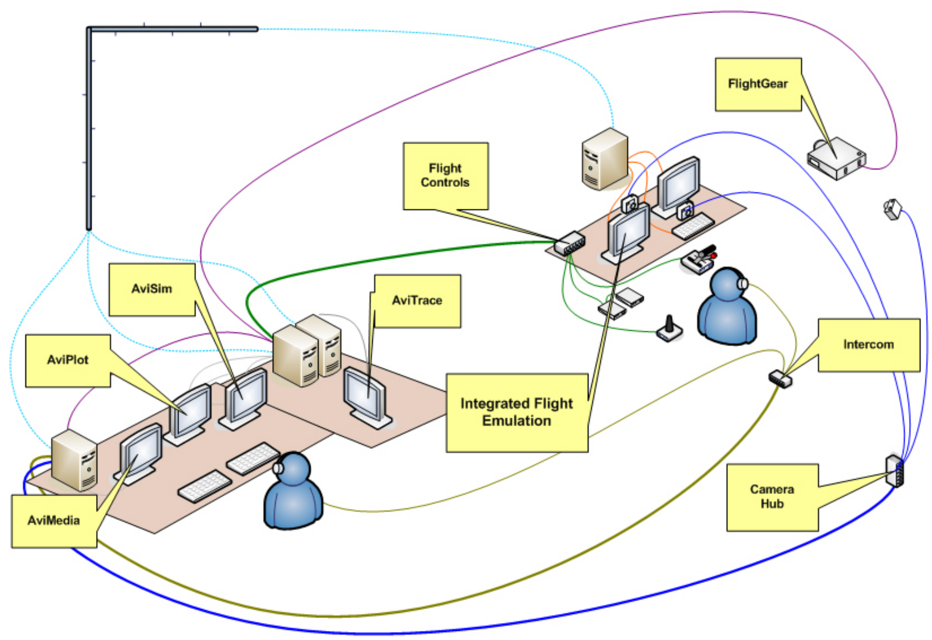


Figure 8. AviSAFE simulation configuration

3.1.5 Experimental Design

This study employed six doubly multivariate analyses of variance using a repeated measures design. For each analysis, the independent variable was flight director state with two levels, on and off. Four designs were used to analyze pilot performance.

Table 10. Within-subjects MANOVA design for altitude performance

Straight and Level – **Altitude**

Flight Director ON					Flight Director OFF				
SD	**RMSE**	**ND**	**TD**	**MTE**	**SD**	**RMSE**	**ND**	**TD**	**MTE**
s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1
s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2
s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3
…	…	…	…	…	…	…	…	…	…
s_{24}	s_4	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}

Table 11. Within-subjects MANOVA design for heading performance

Straight and Level – **Heading**

Flight Director ON					Flight Director OFF				
SD	**RMSE**	**ND**	**TD**	**MTE**	**SD**	**RMSE**	**ND**	**TD**	**MTE**
s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1
s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2
s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3
…	…	…	…	…	…	…	…	…	…
s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}

Table 12. Within-subjects MANOVA design for vertical speed performance

Climbing Turn – **Vertical Speed**

Flight Director ON					Flight Director OFF				
SD	**RMSE**	**ND**	**TD**	**MTE**	**SD**	**RMSE**	**ND**	**TD**	**MTE**
s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1
s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2
s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3
…	…	…	…	…	…	…	…	…	…
s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}

Table 13. Within-subjects MANOVA design for bank angle performance

Climbing Turn – **Bank Angle**

Flight Director ON					Flight Director OFF				
SD	**RMSE**	**ND**	**TD**	**MTE**	**SD**	**RMSE**	**ND**	**TD**	**MTE**
s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1	s_1
s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2	s_2
s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3	s_3
…	…	…	…	…	…	…	…	…	…
s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}	s_{24}

Table 10, Table 11, Table 12 and Table 13 show four MANOVA constructs with the five dependent variables (SD, RMSE, ND, TD and MTE) used to measure pilot performance on the four flight parameters (altitude, heading, vertical speed and bank angle).

Table 14. Within-subjects MANOVA design for dwell time performance during straight and level

Straight and Level – **Dwell Times**

Flight Director ON

DTP	DTS	DTC	DTO
s_1	s_1	s_1	s_1
s_2	s_2	s_2	s_2
s_3	s_3	s_3	s_3
…	…	…	…
s_{24}	s_{24}	s_{24}	s_{24}

Flight Director OFF

DTP	DTS	DTC	DTO
s_1	s_1	s_1	s_1
s_2	s_2	s_2	s_2
s_3	s_3	s_3	s_3
…	…	…	…
s_{24}	s_{24}	s_{24}	s_{24}

Table 15. Within-subjects MANOVA design for dwell time performance during the climbing turn

Climbing Turn – **Dwell Times**

Flight Director ON

DTP	DTS	DTC	DTO
s_1	s_1	s_1	s_1
s_2	s_2	s_2	s_2
s_3	s_3	s_3	s_3
…	…	…	…
s_{24}	s_{24}	s_{24}	s_{24}

Flight Director OFF

DTP	DTS	DTC	DTO
s_1	s_1	s_1	s_1
s_2	s_2	s_2	s_2
s_3	s_3	s_3	s_3
…	…	…	…
s_{24}	s_{24}	s_{24}	s_{24}

Two designs were used to analyze dwell time performance in the simulated flight deck environment. Table 14 and Table 15 show two MANOVA constructs with the four dependent variables (dwell time on the primary display - DTP, dwell time on the secondary display - DTS, dwell time on the center console - DTC & dwell time on the outside view - DTO) used to measure the dwell time performance in the four areas of interest.

With within-subjects (or repeated measures) MANOVA designs, practice effects and order effects that pose threats to internal validity were controlled by randomization the assignment of trials to subjects as well as by using the all possible presentation orders of the four trials that were specified by 24 permutations.

3.1.6 <u>Results</u>

For all of the trials in Experiment 1, participants were able to search, read and integrate the information presented in the graphical instruments of the integrated flight deck. The following results show the flight director effects on flight parameters such as altitude, heading, vertical speed and bank angle using derived measurements such as standard deviation (SD), root mean square error (RMSE), number of deviations (ND), time outside tolerance (TD) and mean time to exceed tolerance (MTE). The level of significance used to determine the hypotheses was signified by $\alpha = 0.05$. Flight director state was the single factor whose effect was controlled and measured using the within-subjects design.

Table 16. Means for altitude performance during straight and level

Straight and Level – **Altitude**

Flight Director ON					Flight Director OFF				
SD	**RMSE**	**ND**	**TD**	**MTE**	**SD**	**RMSE**	**ND**	**TD**	**MTE**
42.384	51.890	1.792	22.000	41.159	60.606	71.277	1.792	38.083	32.041

Table 16 shows the means of the performance measures derived from altitude during the straight and level maneuver for the two flight director states. There are noticeable differences for all measures across the flight director state except ND. The effect of the flight director on altitude performance is graphically shown in Figure 9.

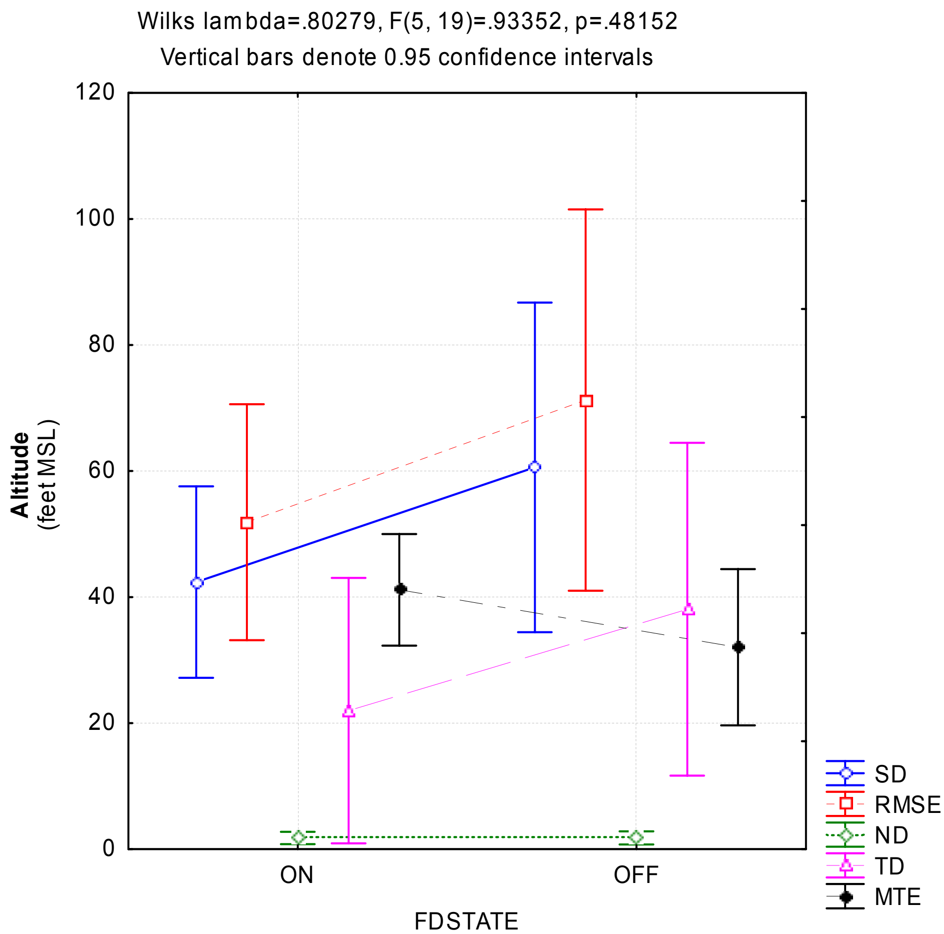


Figure 9. Graph of altitude performance during straight and level

The MANOVA for altitude performance revealed that Wilks' Lambda $\Lambda = .803$ and there was no statistical significance of the flight director's effect, $F(5, 19) = 0.93352$, $p = .482$. Therefore, there was insufficient evidence to conclude that while hand-flying, pilots experience better altitude performance on maintaining straight and level when using the flight director compared to when not using the flight director.

Table 17. Means for heading performance during straight and level

Straight and Level – **Heading**

Flight Director ON					Flight Director OFF				
SD	**RMSE**	**ND**	**TD**	**MTE**	**SD**	**RMSE**	**ND**	**TD**	**MTE**
2.218	2.288	3.167	9.542	32.679	2.337	2.510	2.458	19.667	31.768

Table 17 shows the means of the performance measures derived from heading during the straight and level maneuver for the two flight director states. There are noticeable differences for all measures across the flight director state. The effect of the flight director on heading performance is graphically shown in Figure 10.

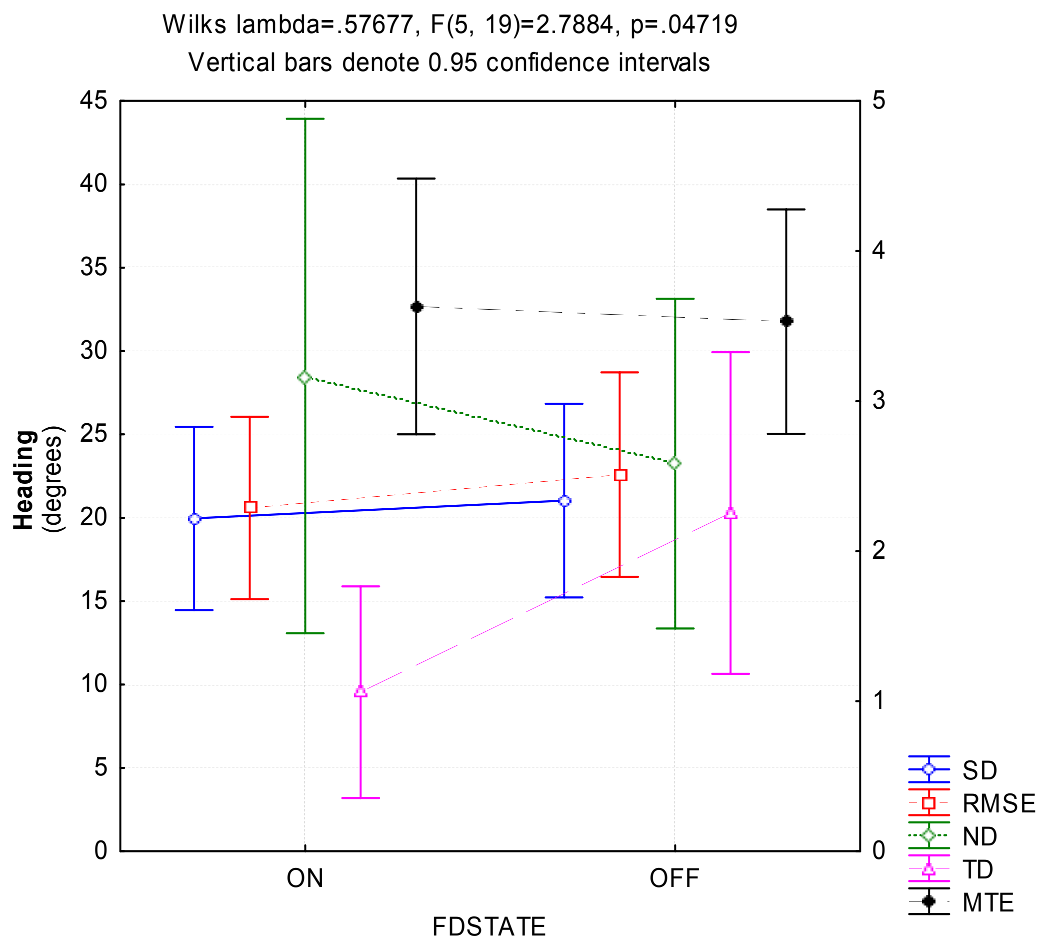


Figure 10. Graph of heading performance during straight and level

The MANOVA for heading performance revealed that Wilks' Lambda $\Lambda = .577$ and there was statistical significance of the flight director's effect on heading performance, $F(5, 19) = 2.788$, $p < .05$. Within-subjects contrasts of the flight director effect on the heading performance variables revealed that there was a significant effect on TD, $F(5, 19) = 8.201$, $p < .05$. There were statistically insignificant effects for standard deviation, $F(1, 23) = 0.132$, $p = 0.72$, root mean square error, $F(1, 23) = 0.429$, $p = 0.519$, number of deviations, $F(1, 23) = 0.717$, $p = 0.406$ and mean time to exceed tolerance, $F(1, 23) = 0.06$, $p = 0.809$. Therefore, while hand-flying, pilots experience better heading performance on maintaining straight and level when using the flight director compared to when not using the flight director.

Table 18. Means for dwell time performance during straight and level

Straight and Level – **Dwell Time**

Flight Director ON				Flight Director OFF			
DTP	**DTS**	**DTC**	**DTO**	**DTP**	**DTS**	**DTC**	**DTO**
298.375	0.083	0.042	1.417	295.917	0.500	0.083	3.583

Table 18 shows the means of the dwell time performance measures during the straight and level maneuver for the two flight director states. There are noticeable differences for all measures across the flight director state. Generally, the dwell time on the primary IFD decreased and dwell time on other areas of interest increased. The effect of the flight director on dwell time performance is shown in Figure 11.

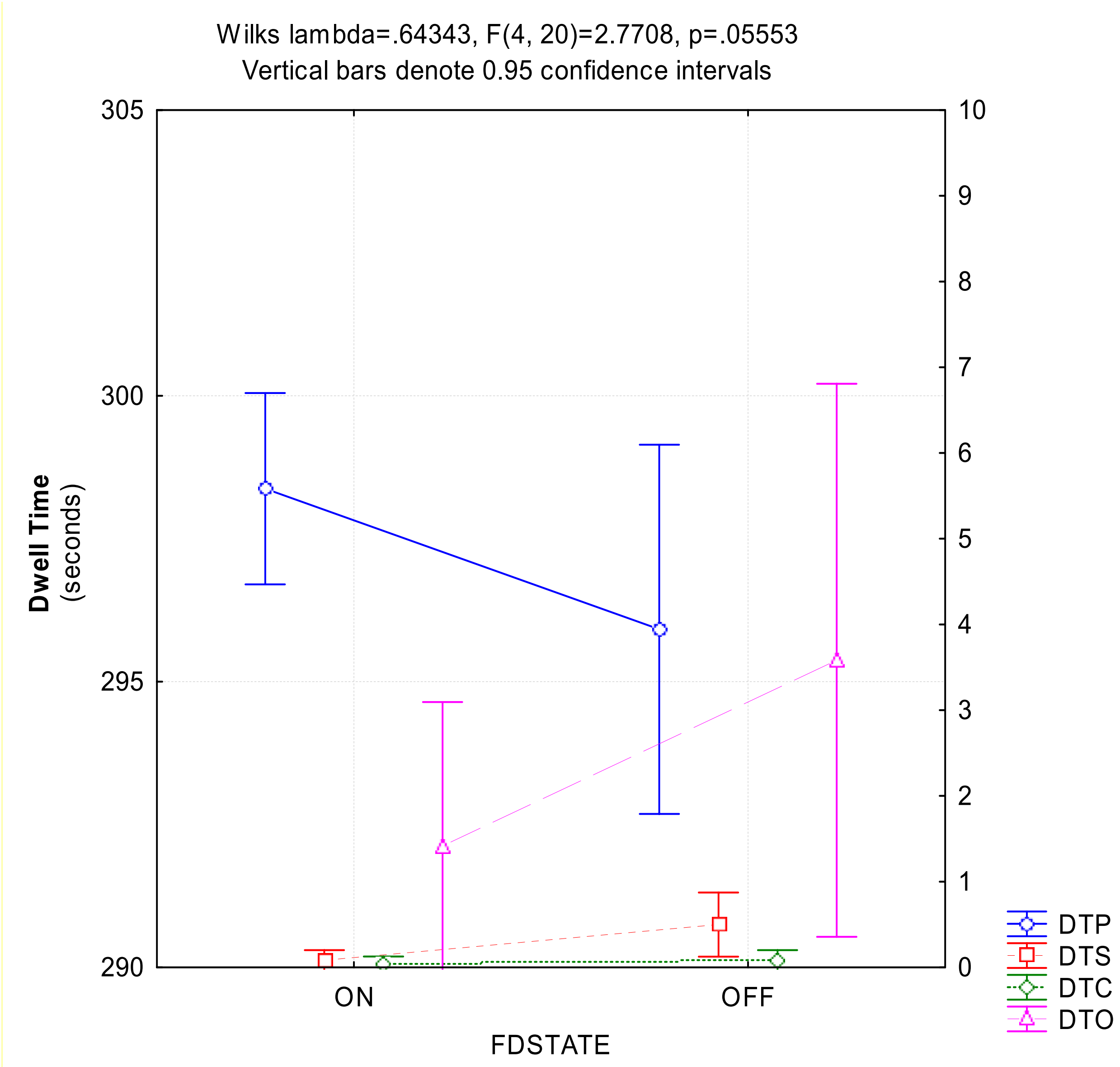


Figure 11. Graph of dwell time performance during straight and level

The MANOVA for dwell time performance revealed no statistical significance of the flight director's effect with Wilks' Lambda Λ = .643 and, $F(4, 20) = 2.771$,

$p = 0.056$. Therefore, there was insufficient evidence to conclude that flight director usage causes a different monitoring strategy during the straight and level maneuver.

Table 19. Means for vertical speed performance during the climbing turn

Climbing Turn – **Vertical Speed**

Flight Director ON					Flight Director OFF				
SD	**RMSE**	**ND**	**TD**	**MTE**	**SD**	**RMSE**	**ND**	**TD**	**MTE**
443.63	506.50	11.542	22.792	3.926	738.54	774.59	14.375	47.792	3.696

Table 19 shows the means of the performance measures derived from vertical speed during the climbing turn maneuver for the two flight director states. There are noticeable differences for all measures across the flight director state. The effect of the flight director on vertical speed performance is shown in Figure 12.

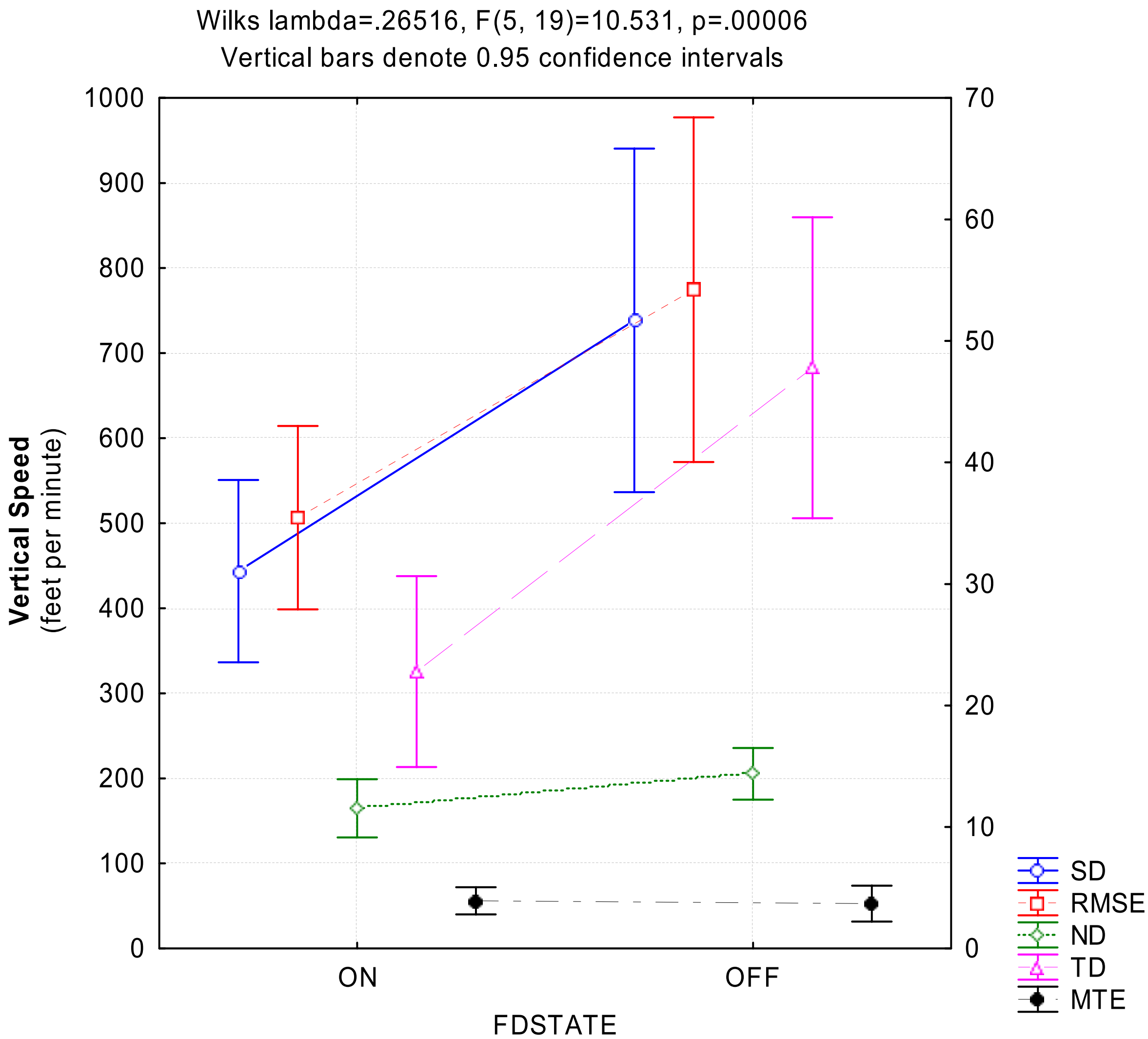


Figure 12. Graph of vertical speed performance during the climbing turn

The MANOVA for vertical speed performance revealed that Wilks' Lambda $\Lambda = .265$ and there was statistical significance of the flight director's effect on vertical speed performance, $F(5, 19) = 10.531$, $p < .001$. Within-subjects contrasts of the flight director effect on the vertical speed performance variables revealed the following statistical significant results: standard deviation, $F(1, 23) = 9.182$, $p < .05$, root mean square error, $F(1, 23) = 8.174$, $p < .05$ and time outside tolerance, $F(1, 3) = 29.818$, $p < .001$. There were statistically insignificant effects for number of deviations, $F(1, 23) = 2.530$, $p = .125$ and mean time to exceed tolerance, $F(1, 23) = .077$, $p = .783$. Therefore, while hand-flying, pilots experience better vertical speed performance on the climbing turn maneuver when using the flight director compared to when not using the flight director.

Table 20. Means for bank angle performance during the climbing turn

Climbing Turn – **Bank Angle**

Flight Director ON					Flight Director OFF				
SD	RMSE	ND	TD	MTE	SD	RMSE	ND	TD	MTE
2.912	3.245	1.917	7.000	113.39	3.509	4.358	2.667	26.833	90.239

Table 20 shows the means of the performance measures derived from bank angle during the climbing turn maneuver for the two flight director states. There are noticeable differences for all measures across the flight director state. The effect of the flight director on bank angle performance is graphically shown in Figure 13.

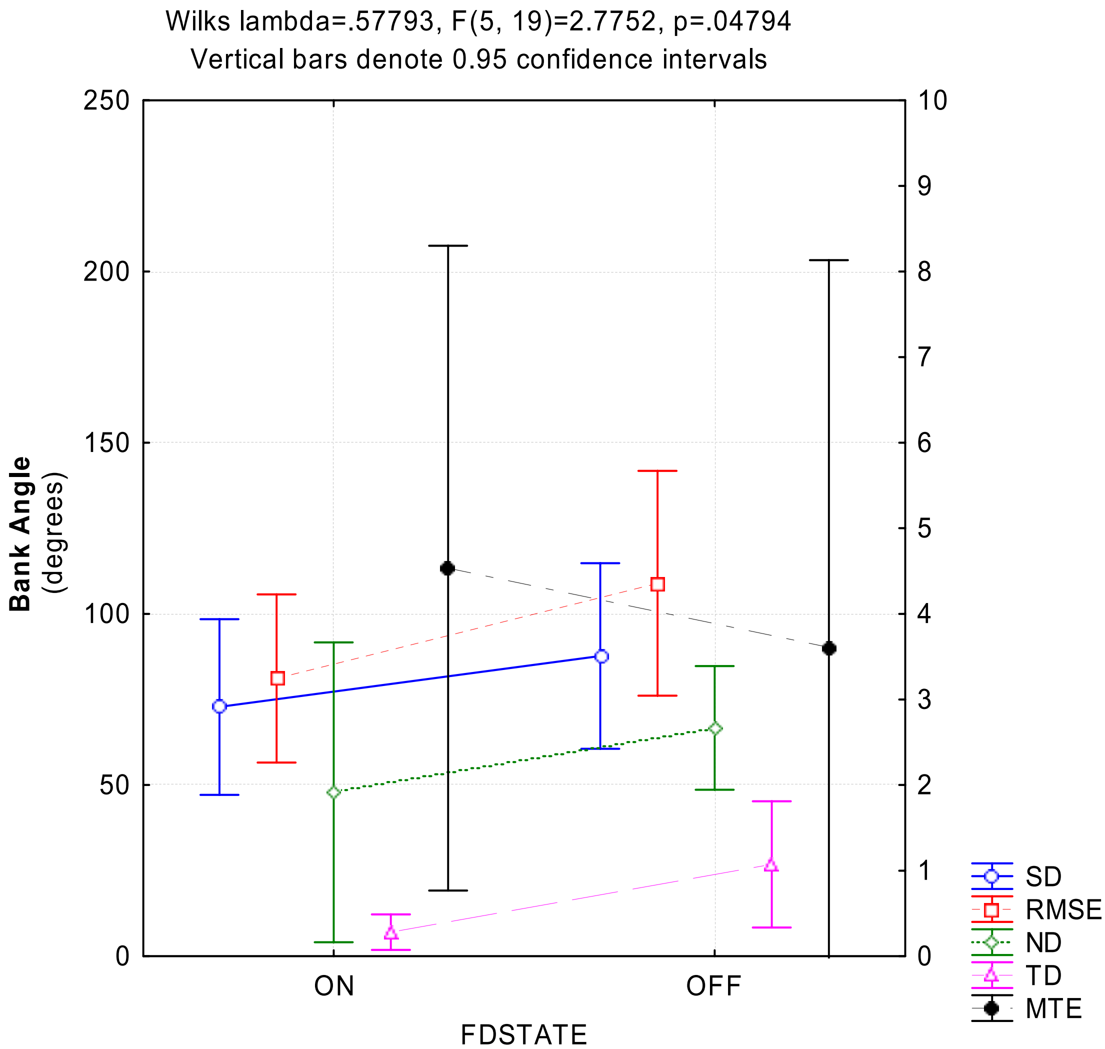


Figure 13. Graph of bank angle performance during the climbing turn

The MANOVA for bank angle performance revealed that Wilks' Lambda $\Lambda = .578$ and there was statistical significance of the flight director's effect on bank angle performance, $F(5, 19) = 2.775$, $p < 0.05$. Within-subjects contrasts of the flight director effect on the bank angle performance variables revealed the following statistical significant results: root mean square error, $F(1, 23) = 5.016$, $p < 0.05$ and time outside tolerance, $F(1, 23) = 4.497$, $p < 0.05$. There were statistically insignificant effects for standard deviation, $F(1, 23) = 4.234$, $p = 0.051$, number of deviations, $F(1, 23) = 0.695$, $p = 0.413$ and mean time to exceed tolerance, $F(1, 23) = 3.004$, $p = 0.096$. Therefore, while hand-flying, pilots experience better bank angle performance on the climbing turn maneuver when using the flight director compared to when not using the flight director.

Table 21. Means for dwell time performance during the climbing turn

Climbing Turn – **Dwell Time**

Flight Director ON				Flight Director OFF			
DTP	**DTS**	**DTC**	**DTO**	**DTP**	**DTS**	**DTC**	**DTO**
169.250	0.125	0.042	0.625	169.083	0.125	0.042	1.167

Table 21 shows the means of the dwell time performance measures during the climbing turn maneuver for the two flight director states. There are only noticeable differences for dwell times for the primary IFD and outside across the flight director states. Generally, the dwell time on the primary IFD slightly decreased and dwell time on the outside view increased. The effect of the flight director on dwell time performance is graphically shown in Figure 14.

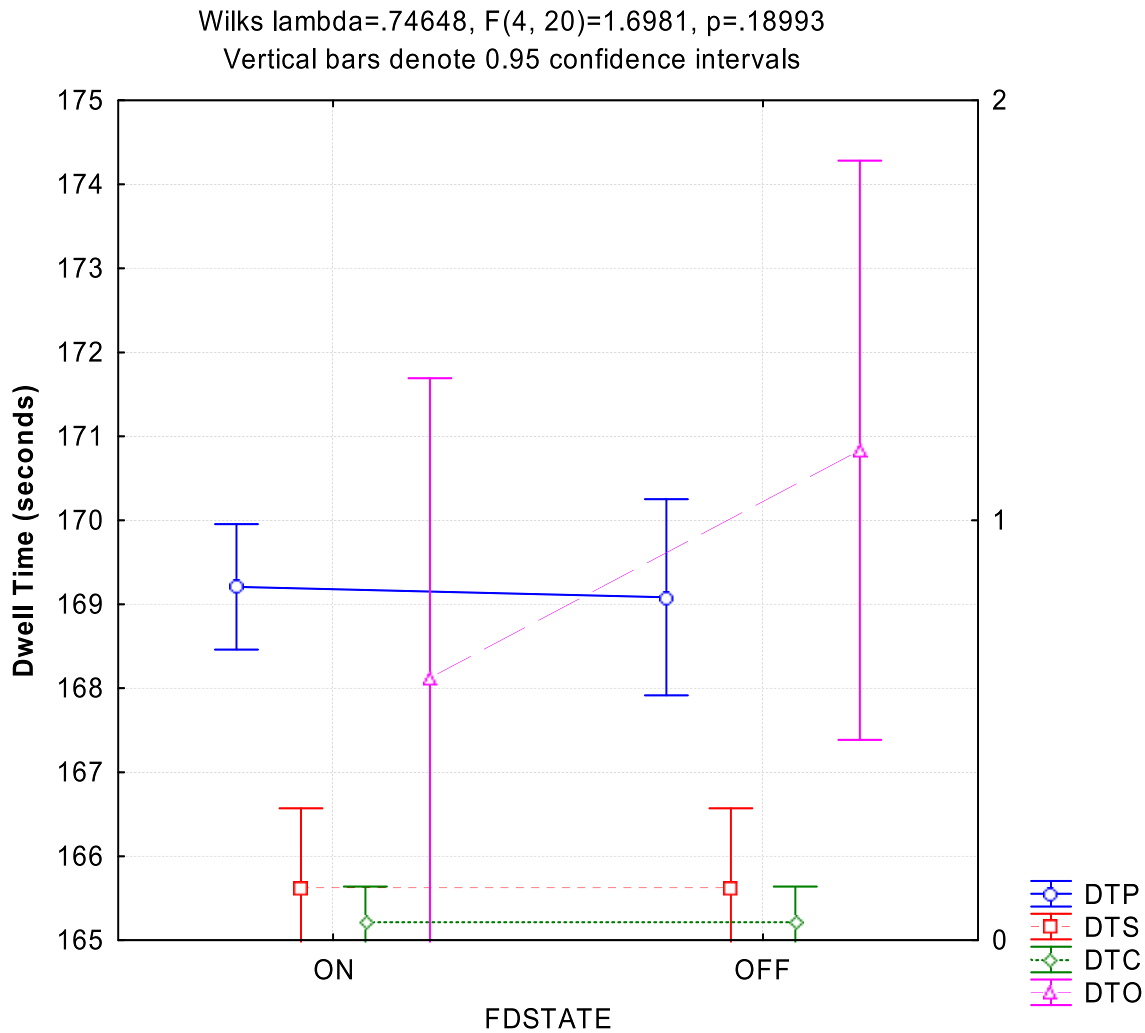


Figure 14. Graph of dwell time performance during the climbing turn

The MANOVA for dwell time performance revealed no statistical significance of the flight director's effect with Wilks' Lambda $\Lambda = .746$ and, $F(4, 20) = 1.698$, $p = 0.19$. Therefore, there was insufficient evidence to conclude that flight director usage causes a different monitoring strategy during the right climbing turn maneuver.

3.1.7 <u>Discussion</u>

Designers in various domains tend to use the presentation of visual cues to enhance tracking performance (e.g., the presentation of the magenta flight director command bars for tracking with the aircraft reference symbol). This design approach is widely adopted but not completely supported by recent findings in attention literature, which indicate that attentional problems can be explained by designers' increasing reliance on automation feedback that requires focal visual attention (Sarter, 2000; Wickens & Hollands, 2000). Experiment 1 presents two maneuvers that are similar in concept (i.e., tracking maneuvers). However, the two maneuvers may invoke different monitoring strategies (i.e., they follow the flight director as the primary feedback source or follow the prescribed instruments as the primary feedback source).

In the straight and level maneuver, subjective reports correlated with the objective measures and confirmed that the flight director was not very useful (see Table 16 & Table 17). There was not enough evidence for the straight and level maneuver to conclude that pilots experience better hand-flying performance when using the flight director. An improvement on this analysis could involve a subjective metric to determine the monitoring strategy employed, which can be correlated with the derived flight measures. Such analysis would provide information on the impact of the pilot's selected monitoring strategy for particular maneuvers.

During the straight and level maneuver, high frequency of the oscillations about the prescribed altitude and heading suggests that the pilots were generally focusing attention to "chase" the multi-dimensional cue of the flight director when it was presented. However, when the flight director was not presented, divided attention made it imperative for the pilot to "scan" multiple visual cues and extract important uni-dimensional information resulting in less frequent and smoother oscillations.

Pilots experienced better hand-flying performance on climbing turn maneuvers when using the flight director compared to when not using the flight director. Subjective accounts of the monitoring strategy tend to suggest that pilots were not necessarily playing closer attention to the flight director but were using it to orderly orient the aircraft reference symbol as opposed to randomly searching for an aircraft attitude that yielded the desired vertical speed and bank angle.

To increase the meaningfulness and strategic influence of usability data, a single dependent variable that does not sacrifice precision should be used to represent the entire construct of usability (Sauro & Kindlund, 2005). The experiments presented in this thesis aim to consolidate the various performance measures at the level of statistical analysis by utilizing the multivariate analysis of variance. Despite this attempt to consolidate information into one value, it is useful to have the performance measures as independent values since different types of inferences can be made about the pilot performance. For example, a pilot can make one deviation but spend the entire period outside the tolerance. In this case the pilot's performance would be poor if the average time of deviation is used to consolidate the two measures but the pilot would have good performance is the number of deviations is used alone.

Dwell times for the various areas of interest failed to demonstrate enough statistical evidence to conclude that flight director usage causes an unbalanced monitoring strategy during either maneuver as more time is spent focusing on the primary IFD with the flight director, and less time is spent focusing on the secondary IFD and outside view.

3.2 Experiment 2

In the second experiment, twenty-four pilots were selected to compare the usability of the flight management system (FMS) when the autopilot was engaged and when the autopilot was disengaged. Usability of the FMS was measured directly through the number of operations and task time required to complete the in-flight planning task. The usability of the FMS was also measured indirectly through the effect on pilot performance in the secondary hand-flying task. The following hypotheses were postulated for Experiment 2:

- On FMS tasks, pilots experience enhanced pilot performance when using the autopilot compared to when not using the autopilot.
- On FMS tasks, pilots operate the IFD more effectively and efficiently, during in-flight planning tasks, when using the autopilot compared to when not using the autopilot.
- On FMS tasks, commercial pilots experience the best pilot performance and student pilots experience the worst pilot performance.

- On FMS tasks, commercial pilots operate the IFD most effectively and efficiently, during in-flight planning tasks, and student pilots operate the IFD worst.
- There is a statistically significant interaction as pilot certification interacts with the autopilot state to affect pilot performance causing greater differences between pilot certification categories when the autopilot is off.
- There is a statistically significant interaction as pilot certification interacts with the autopilot state to affect IFD operation performance causing greater differences between pilot certification categories when the autopilot is off.

3.2.1 Participants

The original twenty-four pilots from FIT Aviation flight school that were involved in the first study were also the participants of this study. In this study, the participants were categorized according to their pilot certificate, forming the following three categories: student pilots, private pilots and commercial pilots. The in-flight planning tasks used involved some knowledge of instrument flight. However, the tasks were not too difficult so that the pilots without instrument-ratings could complete them.

3.2.2 Procedure

After a voluntary break for at least fifteen minutes, Experiment 2 began. The participant was then briefed on the purpose and structure of the study as well as the features and limitations of AviSAFE within the context of the second experiment. An introductory presentation was conducted to describe the flight planning features on the secondary IFD. Each participant was then randomly assigned to one of two trial sequences (Appendix D). The sequence trials either required the participant to fly with the autopilot first, then without, or to fly without the autopilot first, then with the autopilot. The participant was not informed about the order of the trials.

Subsequently, each participant was taken to the flight simulator and seated in the left pilot seat. The participant was fitted with a knee-board for recording information issued by the controller. The test administrator turned on the IFDs in the flight simulator in a simulation mode with the flight director on, established the flight emulation program in a paused state and activated the other systems of the flight data recording and processing suite in a prestart state.

The participant was then given a headset that was coupled to the test administrator's headset using an intercom system in the flight simulator. Similar to Experiment 1, the volume levels were adjusted for comfort then the test administrator sat at a remote monitoring station and began to calibrate the eye-tracking camera system in the flight simulator with the participant sitting as though he/she was flying. The participant was informed that a route would be programmed into the FMS on the secondary IFD. The prescribed cruise altitude was 10,000 feet.

The test administrator announced to the participant the limitations of the flight simulator within the context of Experiment 2. The participant was given a chance to ask questions for clarity and receive answers. Systems were brought online and the simulation was initialized with the autopilot engaged and executing a standard departure. The departure involved a standard rate right turn with a maximum power climb to the cruise altitude of 10,000 feet MSL. On climbing through 4,000 feet MSL, the test administrator issued the departure clearance from runway 29 at KORH (Worcester Regional Airport) to the GDM (Gardner VORTAC), then direct to the KPWM (Portland International Jetport), arriving on runway 36. Once the read-back was approved, the pilot was cleared to enter the plan in the FMS (see Figure 15).

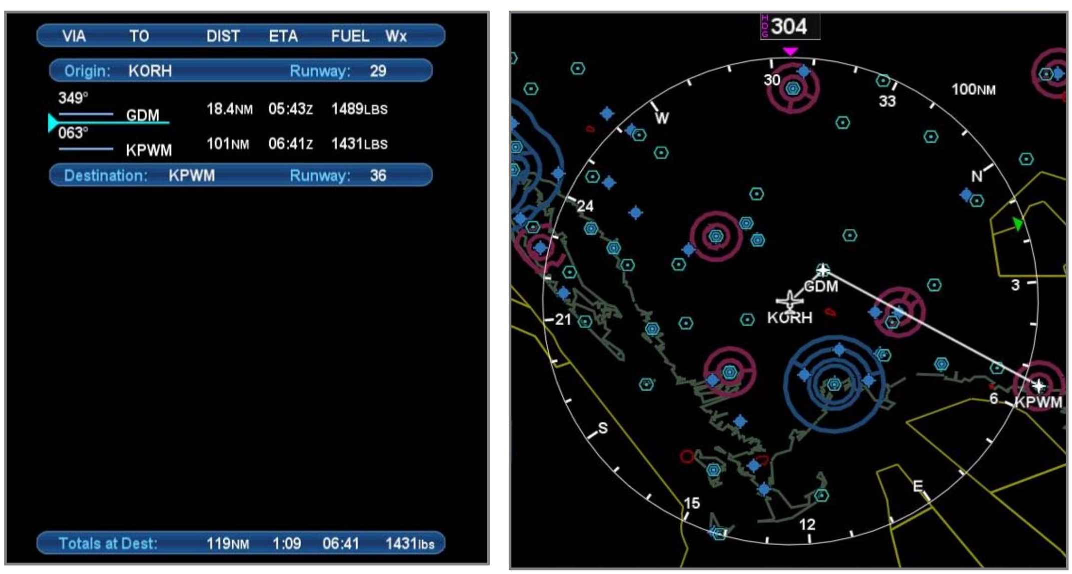


Figure 15. Flight plan (left) and MAP (right) for departure clearance

Upon reaching the cruise altitude, the test administrator issued the command for the pilot to maintain 180 knots indicated airspeed. At 5 nautical miles from GDM, the test administrator requested the pilot to use the MAP to determine if there were any airports directly under the route of flight. One of the airports expected to be reported was KMHT (Manchester Airport), an airport to be identified later in the scenario. On turning over GDM, the test administrator announced to the pilot whether or not the autopilot was going to be disengaged during the scenario. If the autopilot were to be disengaged, the test administrator positively transferred the controls at 95 nautical miles from KPWM. Immediately, after the pilot assumed flight controls, the test administrator commanded the pilot to maintain heading 065^{o}, altitude 10,000 feet and airspeed 180 knots.

Upon reaching 90 nautical miles from KPWM, the test administrator announced that "level 3 thunderstorms were reported in the vicinity of KMHT and that there was an amendment to the flight plan." Once the amended route was delivered and its read-back approved, the test administrator issued the command for the pilot to maintain a heading of 065^{o}, altitude of 10,000 feet and airspeed of 180 knots while reprogramming the given route into the flight management system. The pilot was told to maintain the current flight path until vectors were given to intercept the reprogrammed route shown in Figure 16.

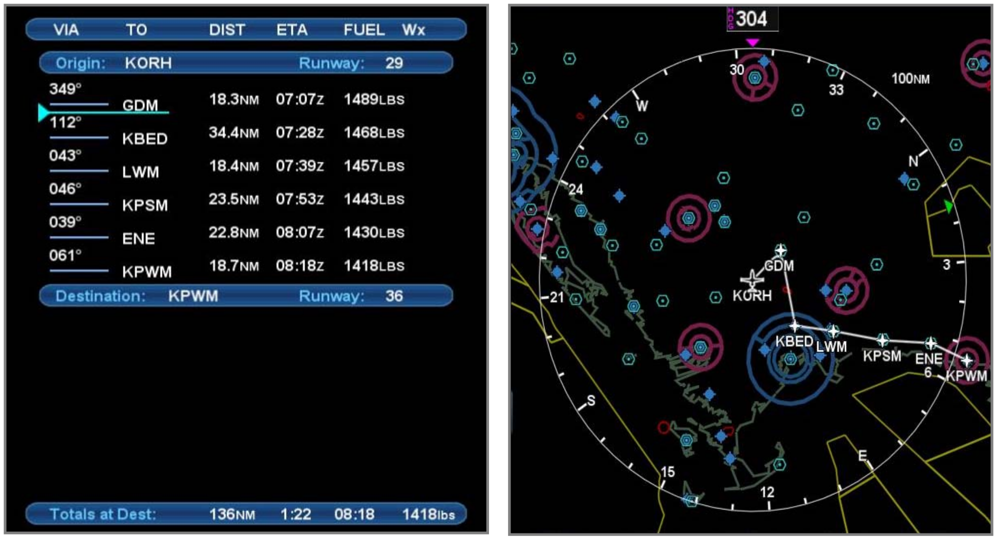


Figure 16. Flight plan (left) and MAP (right) showing amendments with the autopilot disengaged

Upon reaching 80 nautical miles from KPWM, the flight data recording and processing suite was stopped and all data were saved to files. The pilot was informed that the trial was over. The flight plan was cleared from the FMS and the flight data recording and processing suite was brought into the online state.

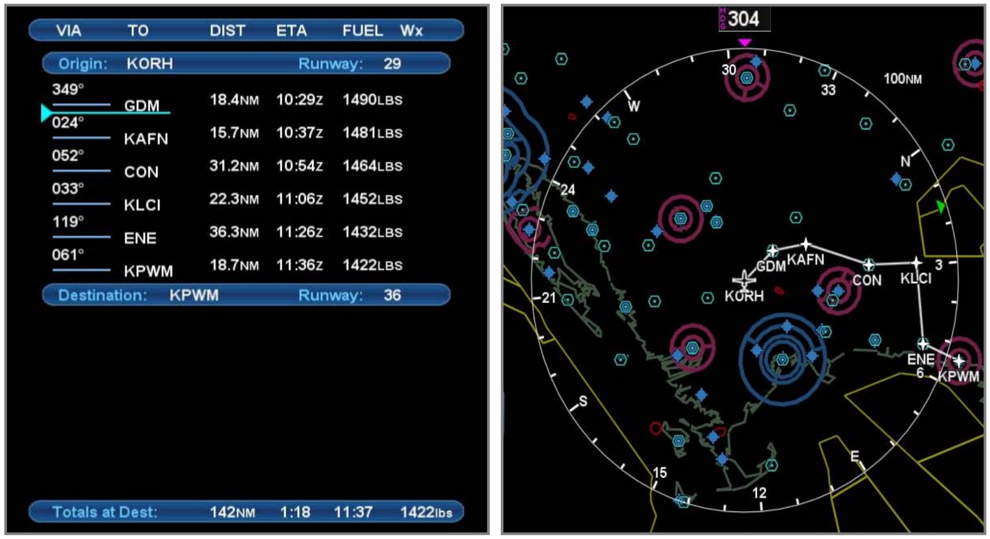


Figure 17. Flight plan (left) and MAP (right) showing amendments with the autopilot engaged

The steps of this procedure were repeated using the other autopilot condition and the other diversion route shown in Figure 17. When the data were saved and the flight data recorders stopped on the final trial, the pilot was vectored to intercept the reprogrammed route at 75 nautical miles from KPWM then allowed to fly the route to completion. Once done, the participant was debriefed and thanked for contribution to the second experiment.

3.2.3 <u>Experimental Platform</u>

The study was conducted in AviSAFE similarly to Experiment 1. The test administrator remained outside the simulator issued the commands over a headset intercom system.

3.2.4 <u>Experimental Design</u>

This study employed three two-factor multivariate mixed designs. For each design, the within-subjects independent variable was the autopilot state with two levels, on and off. For each design, the between-subjects independent variable was the pilot certificate with three levels, student, private and commercial. Table 22 and Table 23 show two MANOVA constructs with the five dependent variables (SD, RMSE, ND, TD and MTE) used to measure performance on the two flight parameters (altitude and heading, respectively).

Table 22. Two-factor mixed MANOVA design for altitude performance

Altitude

	Autopilot ON					Autopilot OFF				
	SD	RMSE	ND	TD	MTE	SD	RMSE	ND	TD	MTE
STUDENT	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8
PRIVATE	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}
COMMERCIAL	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}

Table 23. Two-factor mixed MANOVA design for heading performance

Heading

	Autopilot ON					Autopilot OFF				
	SD	RMSE	ND	TD	MTE	SD	RMSE	ND	TD	MTE
STUDENT	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8	s_1 s_2 s_3 … s_8
PRIVATE	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}	s_9 s_{10} s_{11} … s_{16}
COMMERCIAL	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}	s_{17} s_{18} s_{19} … s_{24}

One two-factor multivariate mixed design was used to analyze IFD operation performance in the simulated flight deck environment. Table 24 shows the MANOVA constructs with the three dependent variables (NOS – number of operations on the secondary IFD, NOK – number of operations on the keyboard & TT – task time) used to measure the IFD operation performance on two adaptable IFD components.

Table 24. Two-factor mixed MANOVA design for IFD operation performance

IFD Operations

	Autopilot ON			Autopilot OFF		
	NOS	NOK	TT	NOS	NOK	TT
STUDENT	S_1 S_2 S_3 … S_8	S_1 S_2 S_3 … S_8	S_1 S_2 S_3 … S_8	S_1 S_2 S_3 … S_8	S_1 S_2 S_3 … S_8	S_1 S_2 S_3 … S_8
PRIVATE	S_9 S_{10} S_{11} … S_{16}	S_9 S_{10} S_{11} … S_{16}	S_9 S_{10} S_{11} … S_{16}	S_9 S_{10} S_{11} … S_{16}	S_9 S_{10} S_{11} … S_{16}	S_9 S_{10} S_{11} … S_{16}
COMMERCIAL	S_{17} S_{18} S_{19} … S_{24}	S_{17} S_{18} S_{19} … S_{24}	S_{17} S_{18} S_{19} … S_{24}	S_{17} S_{18} S_{19} … S_{24}	S_{17} S_{18} S_{19} … S_{24}	S_{17} S_{18} S_{19} … S_{24}

With these repeated measures factor of the designs for Experiment 2, the practice effects, carry over and order effects that pose threats to internal validity were controlled by randomizing the trials using a counterbalancing technique. The counterbalancing required that for each level of pilot certification, the next subject received the opposite order that the previous subject received. Since there were eight pilots in each level of pilot certification, four pilots received “autopilot on then autopilot off”, the other four pilots received “autopilots off then autopilot on”.

3.2.5 Results

For both trials in Experiment 2 (with and without the autopilot), participants were able to search, read and integrate the information presented in the adaptable functions of the integrated flight deck. The following results show the autopilot and pilot certification effects on flight parameters such as altitude and heading using derived measurements such as standard deviation (SD), root mean square error (RMSE), number of deviations (ND), time outside tolerance (TD) and mean time to exceed tolerance (MTE). The level of significance used to determine the hypotheses was signified by $\alpha = 0.05$. Autopilot state was the within-subjects factor whose effect was controlled. Pilot certification was the between-subjects factor whose effect was observed even though not directly controlled by the experimenter.

Table 25. Means for altitude performance under different autopilot states

Altitude

	Autopilot ON					Autopilot OFF				
	SD	RMSE	ND	TD	MTE	SD	RMSE	ND	TD	MTE
STUDENT	0.010	0.143	0	0	81989.177	180.011	186.203	4.5	62.375	6.32
PRIVATE	0.012	0.202	0	0	178422.127	146.414	171.805	4	66.25	5.612
COMMERCIAL	0.025	0.136	0	0	153001.927	148.943	169.598	4.625	61	15.184

Table 25 shows the means of the performance measures derived from altitude during the two autopilot states. There are noticeable differences for all measures across the autopilot state. The effect of the autopilot state on the altitude performance dependent variables for each level of pilot certification are graphically shown in Figure 18, Figure 19, Figure 20, Figure 21 and Figure 30.

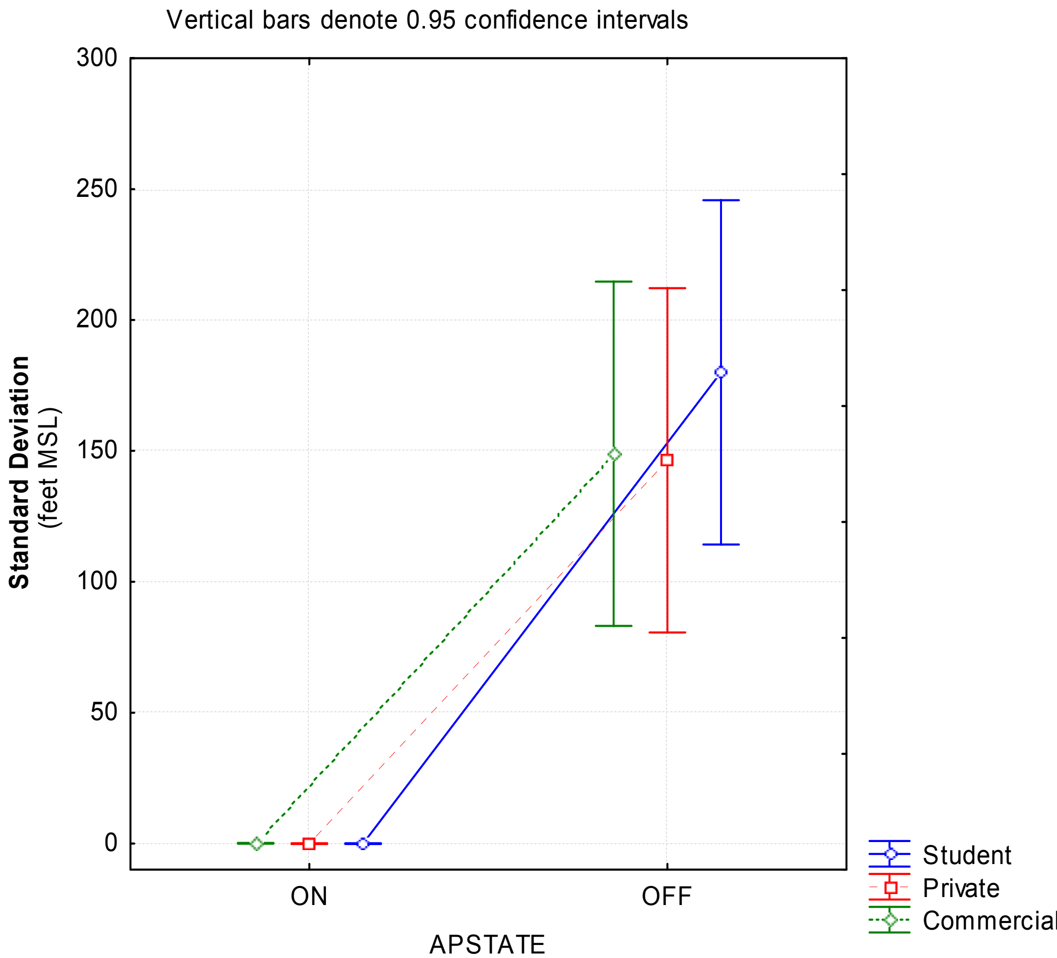


Figure 18. Graph of altitude SD performance under different autopilot conditions

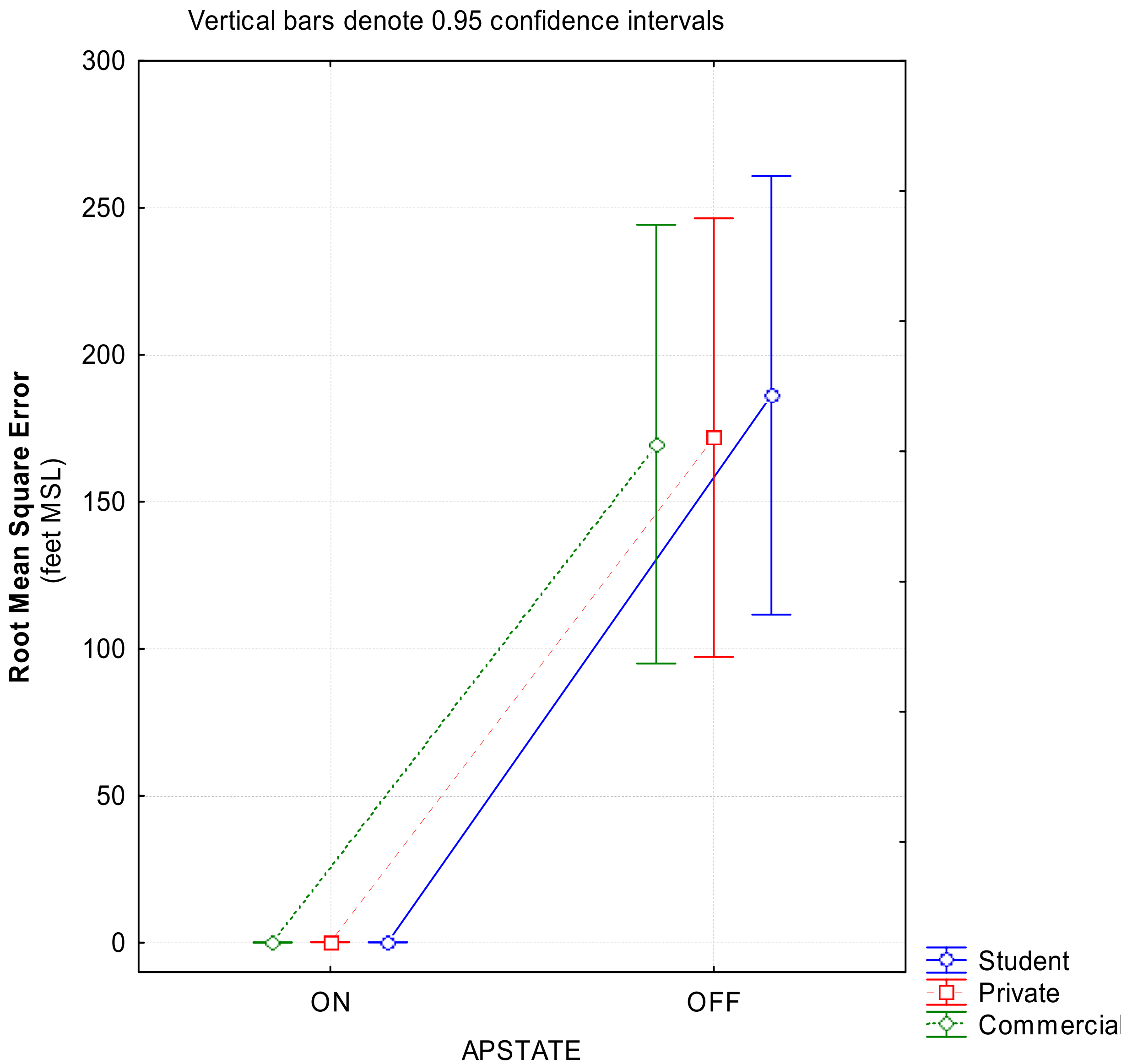


Figure 19. Graph of altitude RMSE performance under different autopilot conditions

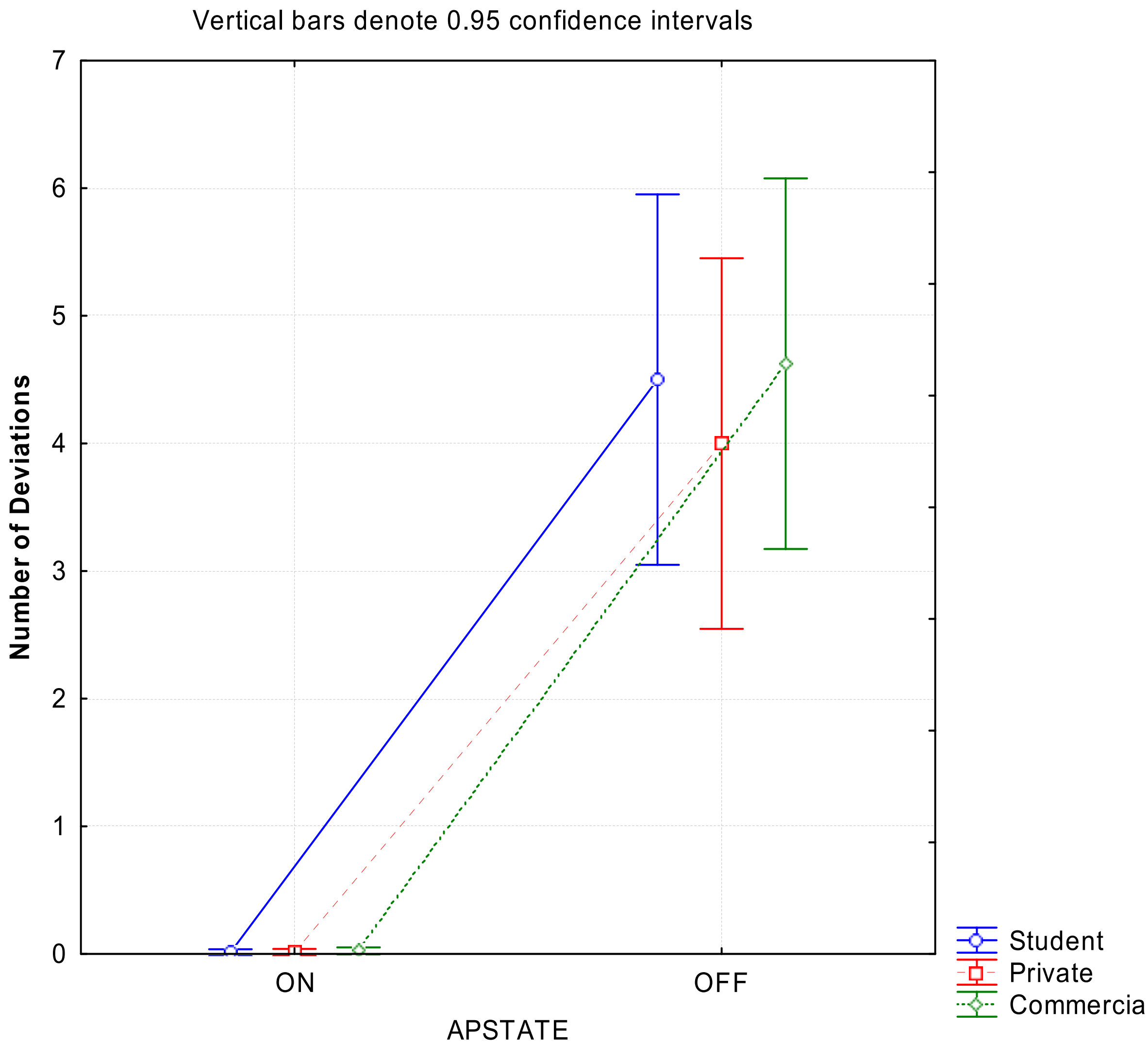


Figure 20. Graph of altitude ND performance under different autopilot conditions

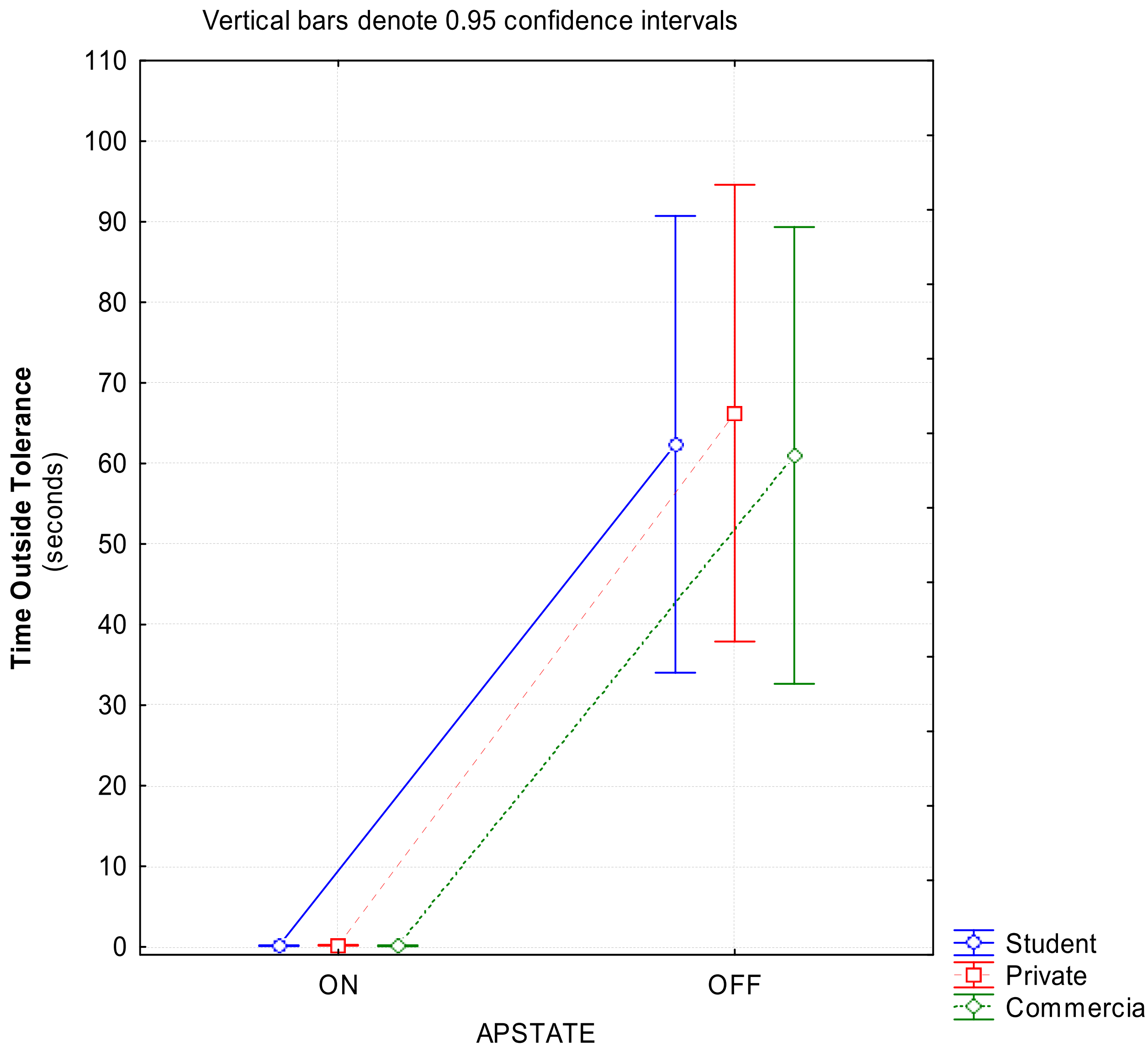


Figure 21. Graph of altitude TD performance under different autopilot conditions

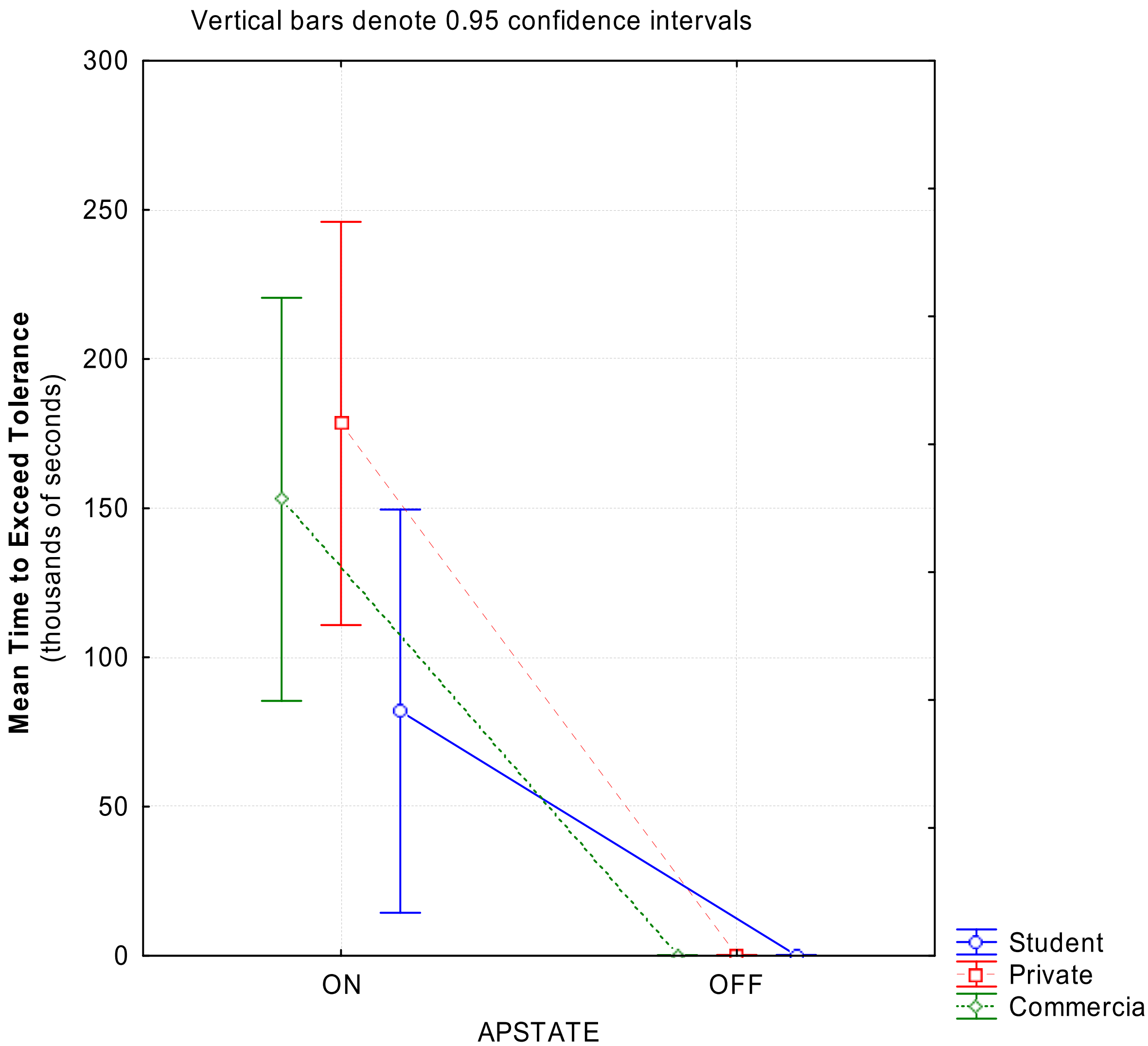


Figure 22. Graph of altitude MTE performance under different autopilot conditions

The MANOVA for altitude performance revealed that Wilks' Lambda $\Lambda = .113$ and there was statistical significance of the autopilot's effect on altitude performance, $F(5, 17) = 26.678$, $p < 0.001$.

Within-subjects contrasts of the autopilot effect on the altitude performance variables revealed the following statistical significant results: standard deviation, $F(1, 21) = 75.208$, $p < 0.001$, root mean square error, $F(1, 21) = 8.963$, $p < 0.05$, number of deviations, $F(1, 21) = 116.366$, $p < 0.001$, time outside tolerance, $F(1, 21) = 64.26$, $p < 0.001$ and mean time to exceed tolerance, $F(1, 21) = 53.882$, $p < 0.001$.

The main effect of pilot certification on altitude performance was not statistically significant, $F(10, 34) = 0.965$, $p = 0.491$.

The interaction effect between autopilot state and pilot certification was also not statistically significant, $F(10, 34) = 0.952$, $p = 0.501$.

Therefore, on FMS tasks, pilots experience enhanced altitude performance when using the autopilot compared to when not using the autopilot. However, there is insufficient evidence to conclude that pilot certification affects altitude performance or interacts with the autopilot state to affect altitude performance.

Table 26. Means for heading performance under different autopilot states

Heading

Autopilot ON

	SD	RMSE	ND	TD	MTE
STUDENT	*0.066*	*0.019*	*0.000*	*0.000*	*2777.509*
PRIVATE	*0.072*	*0.034*	*0.000*	*0.000*	*2017.466*
COMMERCIAL	*0.074*	*0.011*	*0.125*	*0.125*	*2134.050*

Autopilot OFF

	SD	RMSE	ND	TD	MTE
STUDENT	*5.596*	*5.679*	*6.375*	*53.625*	*18.602*
PRIVATE	*10.337*	*10.544*	*4.250*	*60.000*	*12.409*
COMMERCIAL	*7.625*	*7.877*	*3.875*	*46.625*	*23.833*

Table 26 shows the means of the performance measures derived from heading during the two autopilot states. There are noticeable differences for all measures across the autopilot state. The effect of the autopilot state on the heading performance dependent variables for each level of pilot certification are graphically shown in Figure 23, Figure 24, Figure 25, Figure 26 and Figure 27.

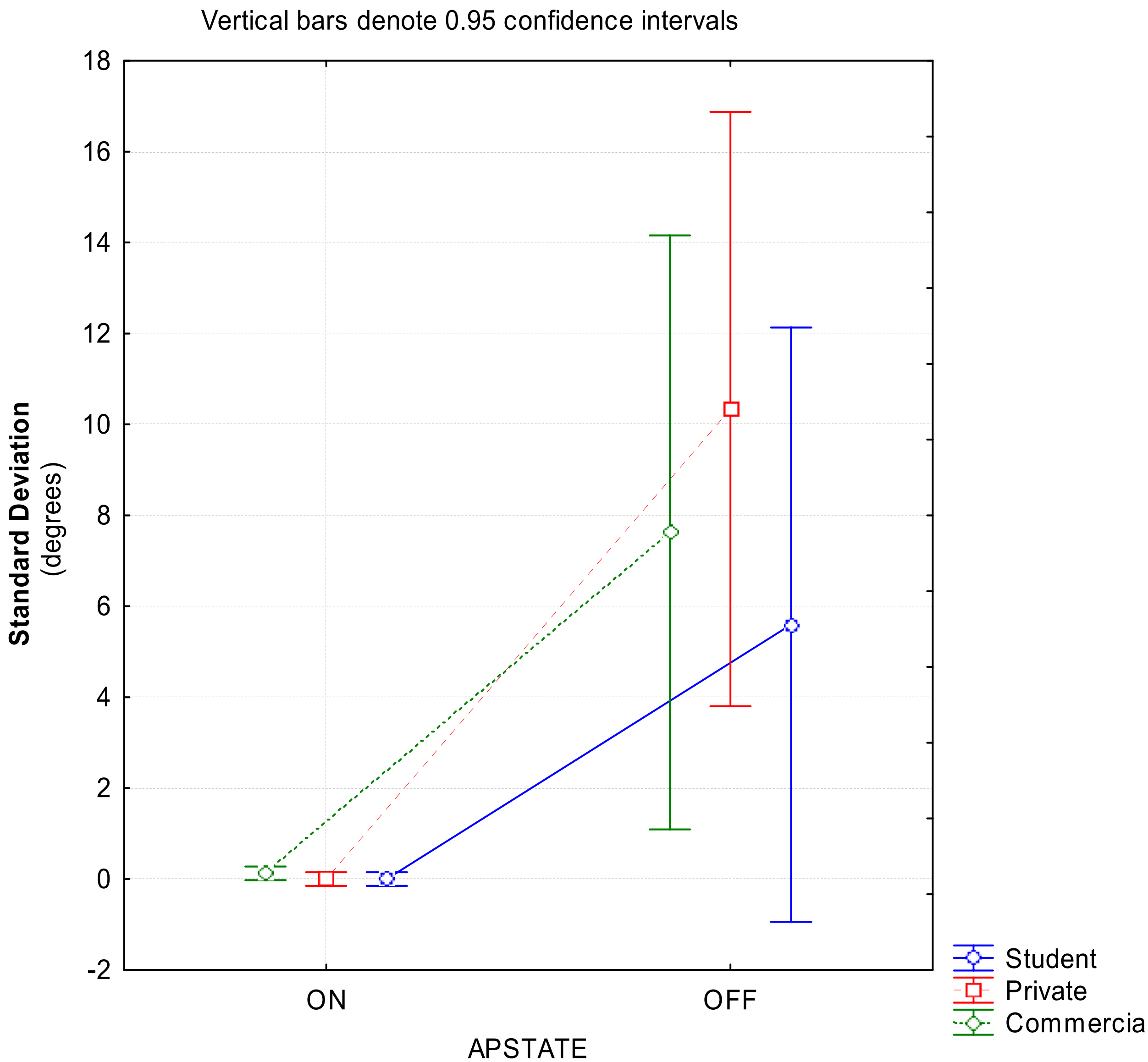


Figure 23 Graph of heading SD performance under autopilot conditions

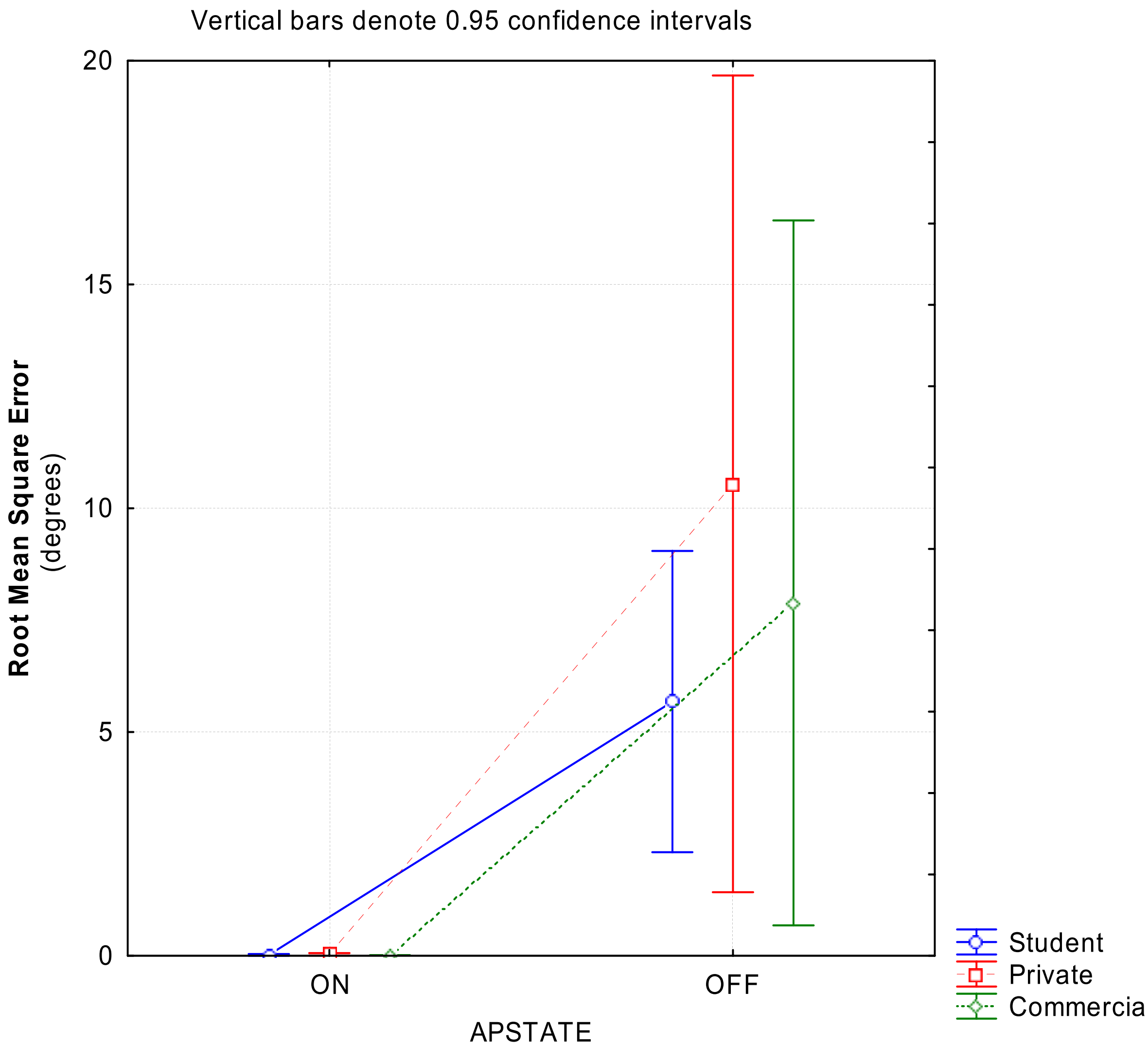


Figure 24. Graph of heading RMSE performance under different autopilot conditions

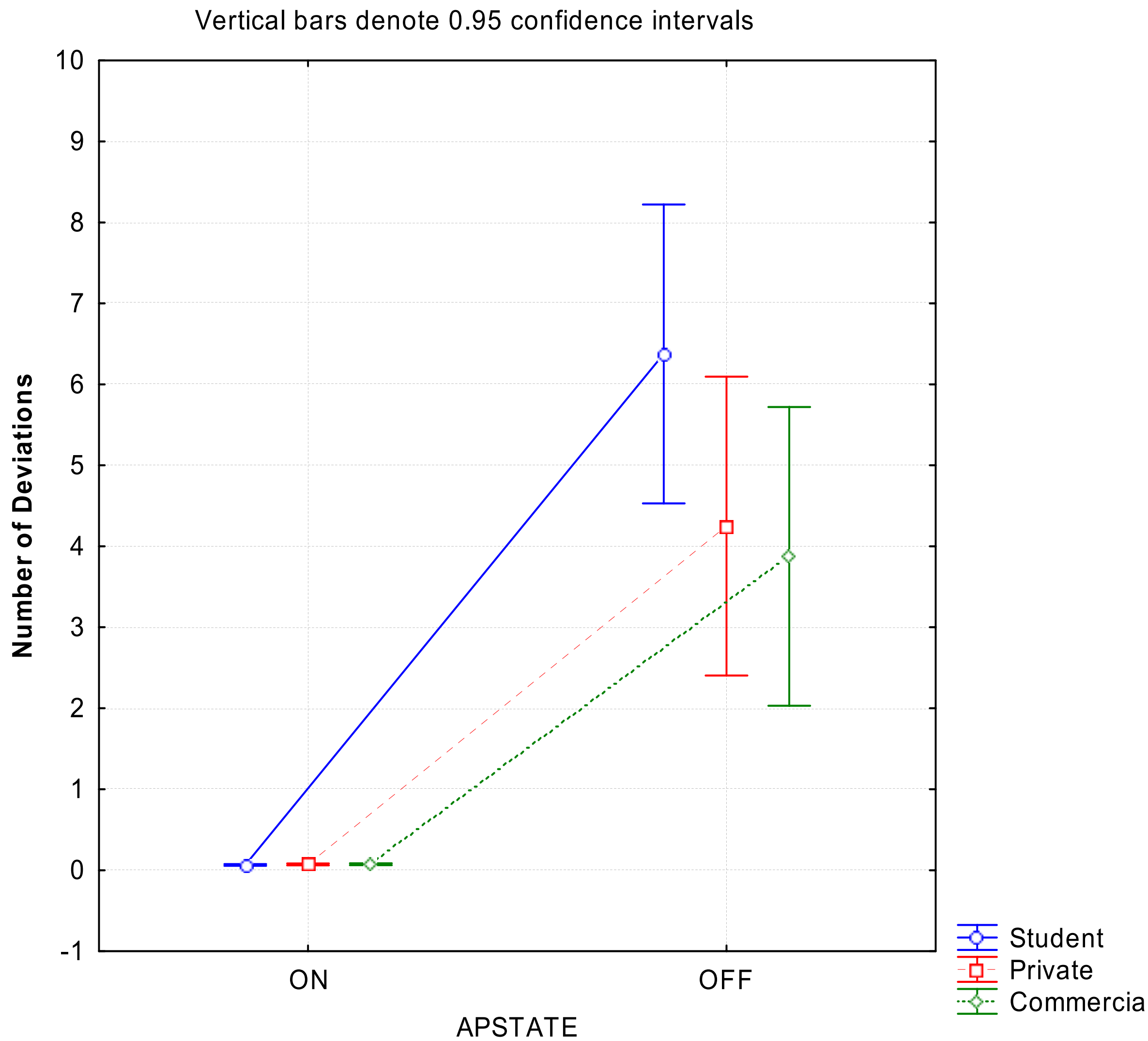


Figure 25. Graph of heading ND performance under different autopilot conditions

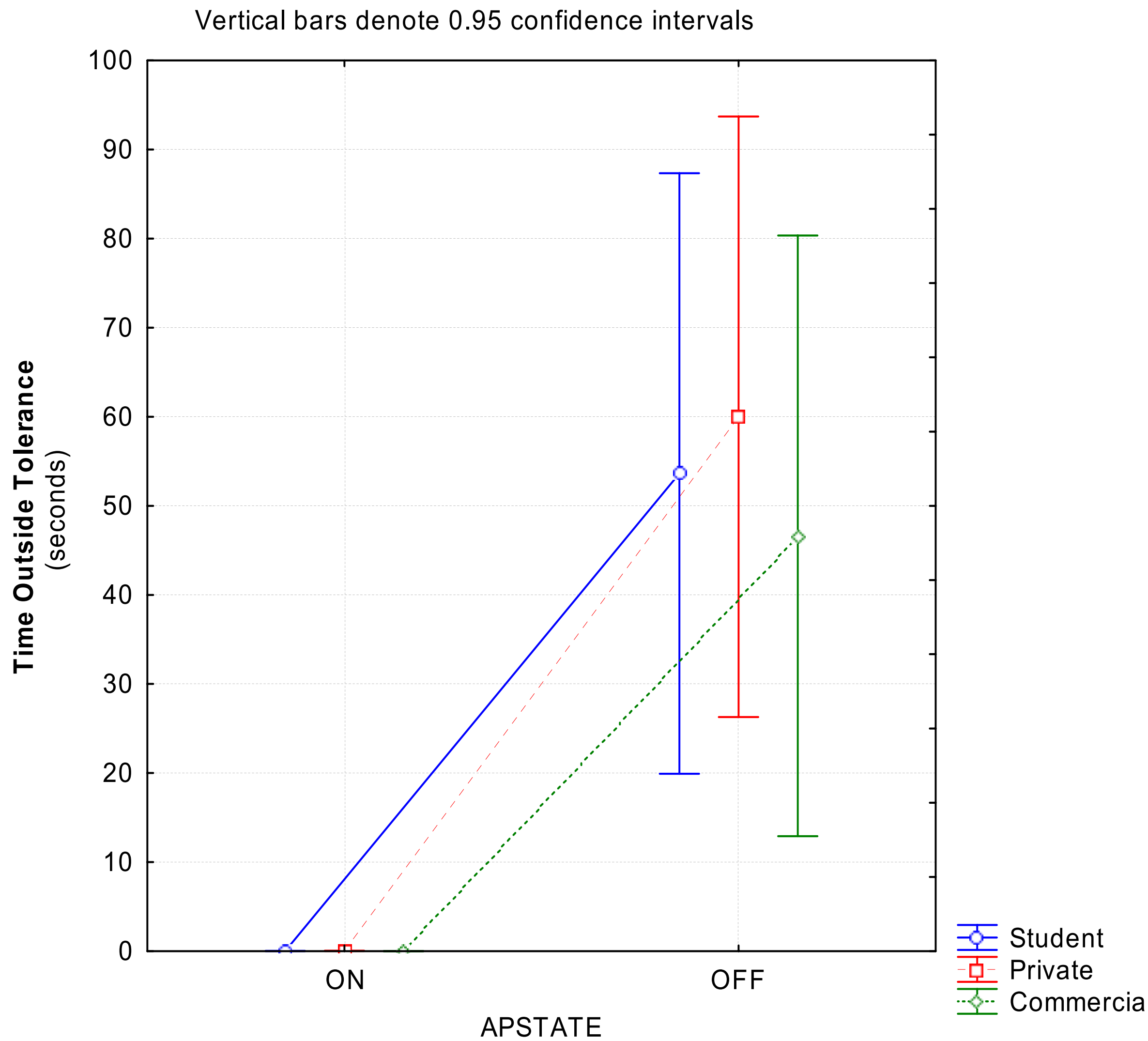


Figure 26. Graph of heading TD performance under different autopilot conditions

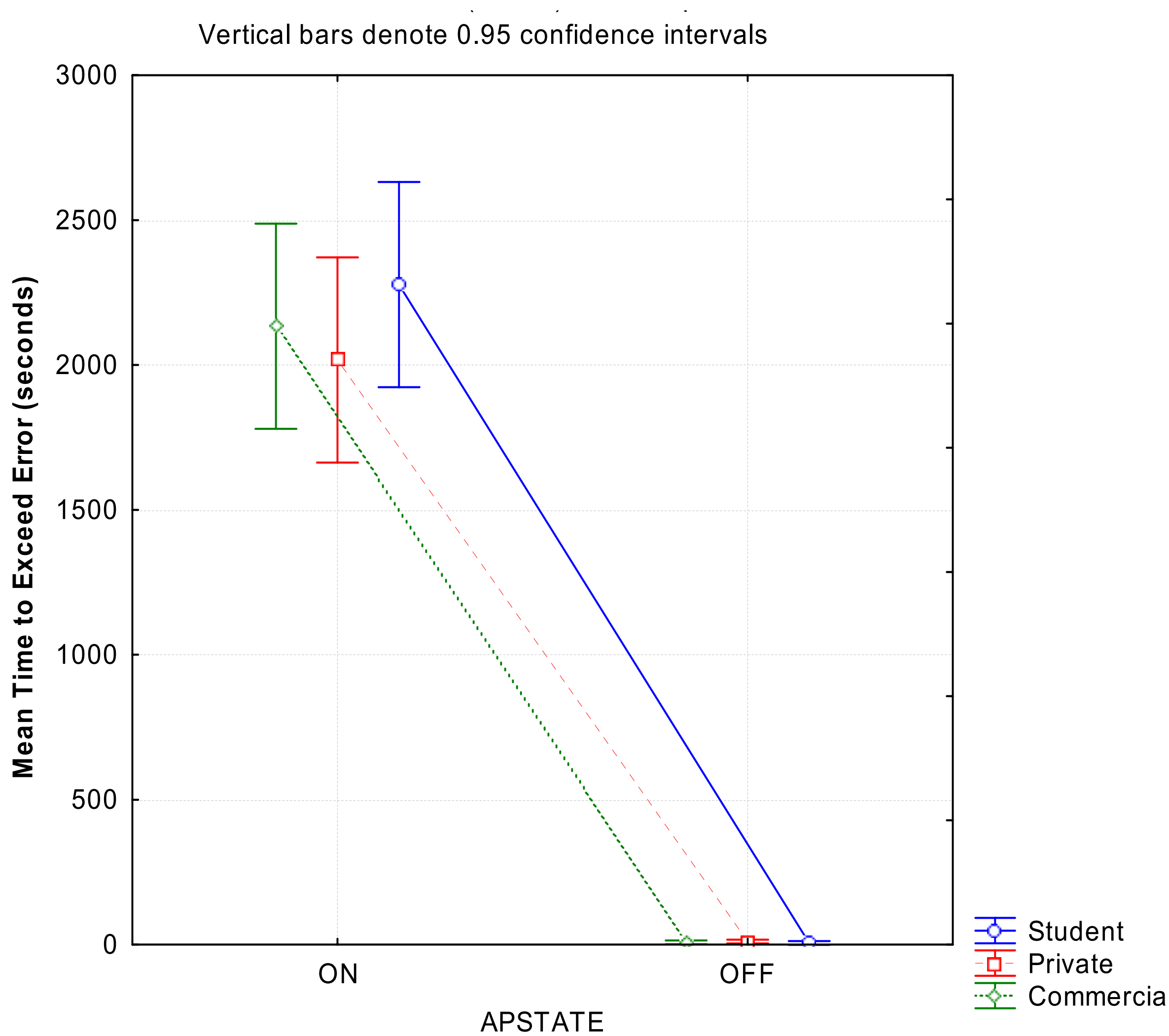


Figure 27. Graph of heading MTE performance under different autopilot conditions

The MANOVA for heading performance revealed that Wilks' Lambda $\Lambda = .036$ and there was statistical significance of the autopilot's effect on heading performance, $F(5, 17) = 90.751$, $p < 0.001$.

Within-subjects contrasts of the autopilot effect on heading performance variables revealed the following statistical significant results: standard deviation, $F(1, 21) = 18.431$, $p < 0.001$, root mean square error, $F(1, 21) = 19.467$, $p < 0.001$, number of deviations, $F(1, 21) = 86.298$, $p < 0.001$, time outside tolerance, $F(1, 21) = 32.557$, $p < 0.001$ and mean time to exceed tolerance, $F(1, 21) = 471.418$, $p < 0.001$.

The main effect of pilot certification on heading performance was not statistically significant, $F(10, 34) = 0.669$, $p = 0.745$.

The interaction effect between autopilot state and pilot certification was not statistically significant, $F(10, 34) = 0.651$, $p = 0.76$.

Therefore, on FMS tasks, pilots experience enhanced heading performance when using the autopilot compared to when not using the autopilot. However, there is insufficient evidence to conclude that pilot certification affects heading performance or interacts with autopilot state to affect heading performance.

Table 27. Means for IFD operation performance under different autopilot states

IFD Operations

Autopilot ON

	NOS	NOK	TT
STUDENT	*44.75*	*39.625*	*38.500*
PRIVATE	*50.375*	*41.625*	*32.875*
COMMERCIAL	*44.625*	*41.625*	*60.000*

Autopilot OFF

	NOS	NOK	TT
STUDENT	*59.25*	*49.75*	*78.125*
PRIVATE	*67.625*	*47.5*	*92.125*
COMMERCIAL	*66.75*	*55.625*	*112.625*

Table 27 shows the means of the performance measures derived from IFD operations during the two autopilot states. There are noticeable differences for all measures across the autopilot state. The effect of the autopilot state on the IFD operation performance dependent variables for each level of pilot certification are graphically shown in Figure 28, Figure 29 and Figure 30.

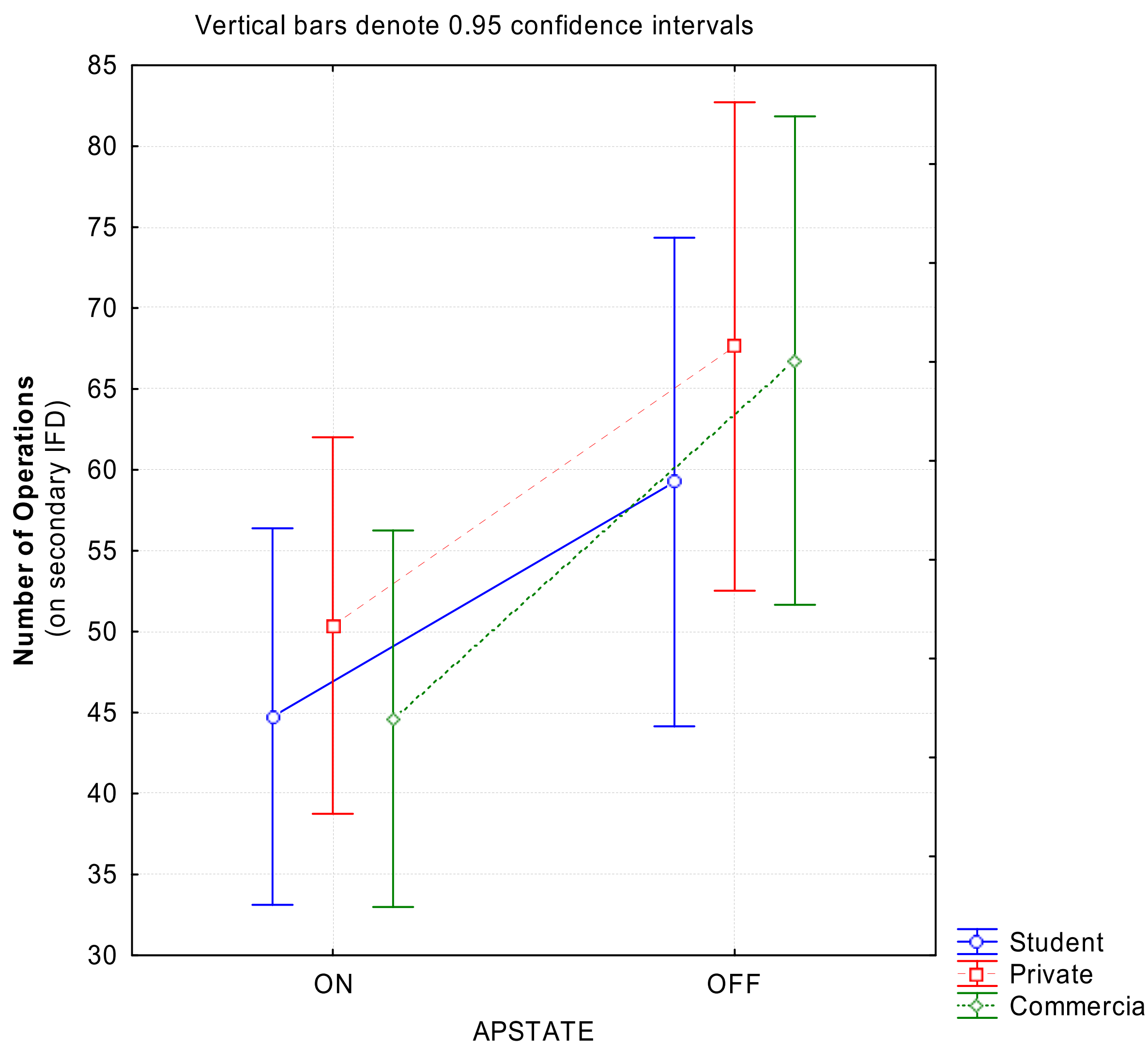


Figure 28. Graph of IFD operation NOS performance under different autopilot conditions

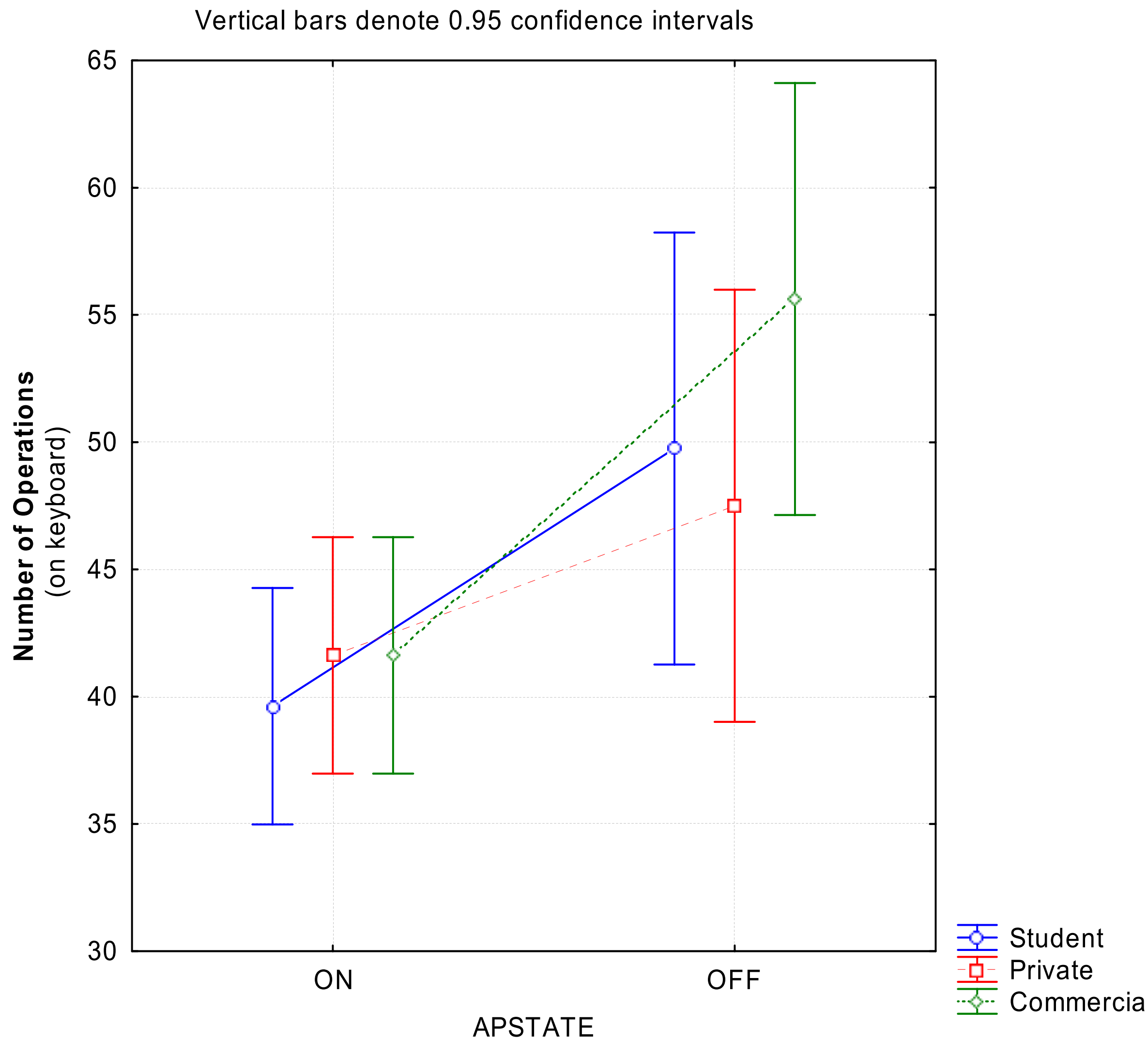


Figure 29. Graph of IFD operation NOK performance under different autopilot conditions

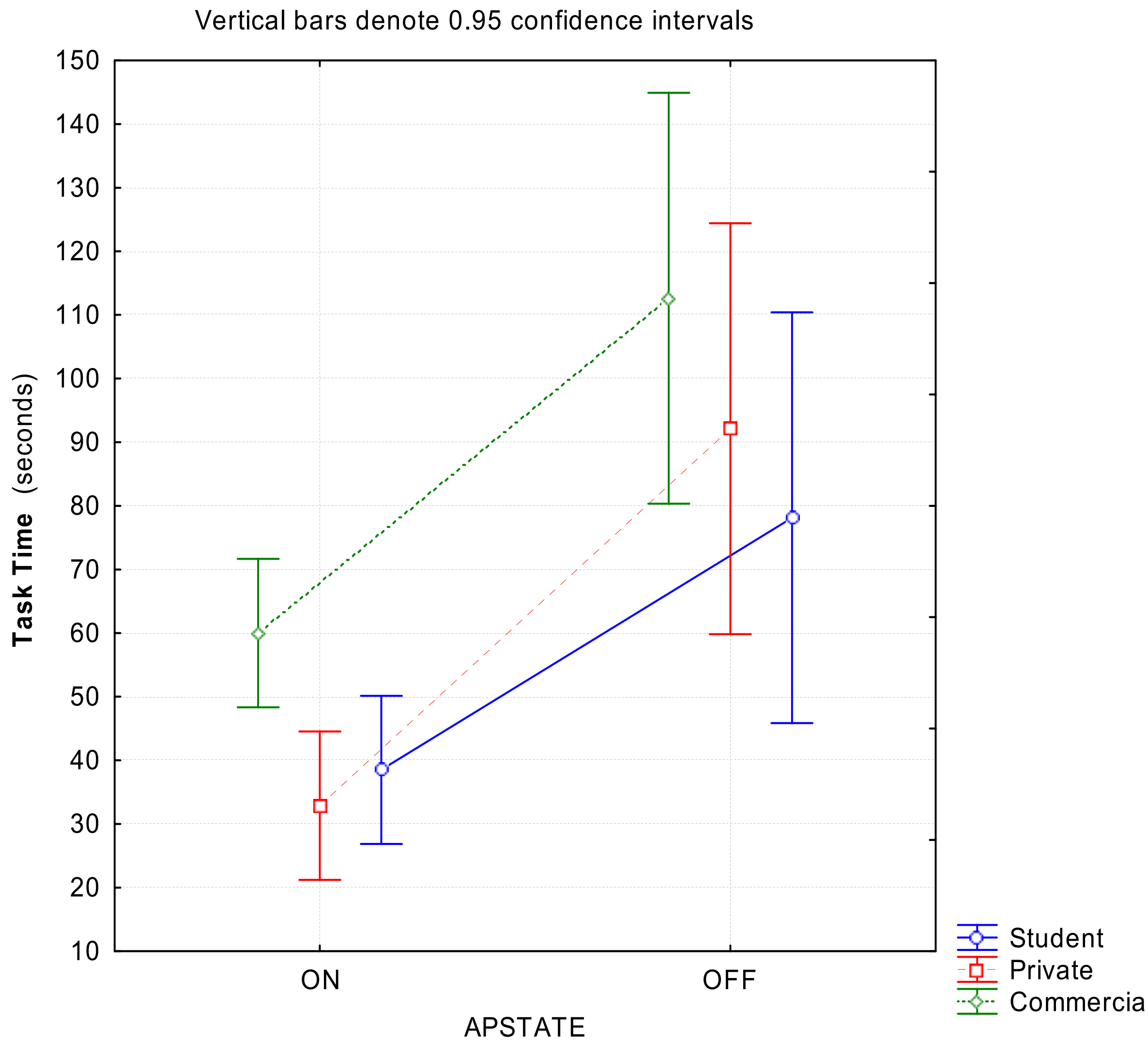


Figure 30. Graph of IFD operation TT performance under different autopilot conditions

The MANOVA for IFD operation performance revealed that Wilks' Lambda $\Lambda = .185$ and there was statistical significance of the autopilot's effect on heading performance, $F(3, 19) = 27.964$, $p < 0.001$.

Within-subjects contrasts of the autopilot effect on the IFD operation performance variables revealed the following statistical significant results: number of operations on the secondary IFD, $F(1, 21) = 22.174$, $p < 0.001$ and task time, $F(1, 21) = 34.47$, $p < 0.001$. The result for the number of operations on the keyboard was not statistically significant, $F(1, 21) = 1.041$, $p = 0.319$.

The main effect of pilot certification on IFD operation performance was not statistically significant, $F(6, 38) = 1.226$, $p = 0.315$.

The interaction effect between autopilot state and pilot certification was also not statistically significant, $F(6, 38) = 1.358$, $p = 0.256$.

Therefore, on FMS tasks, pilots operate the IFD, during in-flight planning tasks, more effectively and efficiently when using the autopilot compared to when not using the autopilot. Also, on FMS tasks, pilots complete flight plan amendments more quickly when using the autopilot compared to when not using the autopilot. However, there is insufficient evidence to conclude that pilot certification affects IFD operation performance or interacts with the autopilot state to affect IFD operation performance.

3.2.6 Discussion

The goal of Experiment 2 was to investigate the effects of autopilot usage and pilot certification on pilot performance and IFD operation performance. There were at least three critically important subtasks to be performed when the pilot used the navigation plan and moving MAP interfaces. The pilots were tasked to search, read, and integrate the information presented in aviation displays throughout the flight.

Three types of attention have been identified as applicable to these sorts of tasks: selective attention (search), focused attention on a located item (read), and divided attention between a pair of items (integrate) (Wickens, 2003; Wickens & Hollands, 2000). Under the autopilot off condition, especially, the pilot had to consistently employ selective attention in the dynamic flight environment. This required the pilot to use selective and sequential assessment of multiple information sources, in order to successfully and efficiently accomplish critical mission goal of avoiding the reported thunderstorms.

It was expected that the aircraft performance and task completion efficiency would be significantly influenced by the autopilot usage during a high-tempo scenario that involved an in-flight FMS amendment task. From the results, it can be inferred that while executing in-flight planning tasks, pilots maintain aircraft control better when using the autopilot compared to when not using the autopilot. Likewise, it can also be inferred that pilots complete flight plan amendments more effectively and efficiently when using the autopilot compared to when not using the autopilot.

Effectiveness was measured using the number of keyboard button presses (NOK), and the number of secondary IFD operations (NOS) required to amend the flight plan. These measures could indicate whether the pilot was making errors and repeating steps or whether the pilot was utilizing the auto-fill automated feature of the display. Efficiency was measured using the completion speed or time taken (TT) to successfully complete the task. Figure 28, Figure 29 and Figure 30 show that the effectiveness and efficiency measures decrease as the autopilot state changed from on to off.

It was expected that, when the autopilot is disengaged, there would be significant interactions between pilot performance and pilot certification on pilot performance and IFD operation. Commercial pilots were expected to have the best pilot performance and IFD operation while student pilots have the worst pilot performance and IFD operations. The lack of evidence to accept this hypothesis implies that the between-subjects effect of pilot certification is not statistically significant. Therefore, there was no statistical difference in the groups based on the pilot certificate which can be correlated to the pilot's flight experience. Further analysis of between-subjects factors should include rank of personal computer experience and typing experience (see Table 8 and Table 9). There are legitimate hypotheses involving these factors that may be tested. For example, a possible hypothesis could be pilot IFD operation performance is less efficient for pilots with low ranks of computer experience, and more efficient for pilots with higher ranks of computer experience.

4. Conclusion

Currently, there is a historic transformation taking place in the general aviation flight deck interface. The designers of general aviation avionics have the burden of proof to demonstrate the reliability and human centered efficacy of the systems. All practical and cost-effective techniques should be applied to reduce the impact of human factors issues and pilot interface considerations. The technique focused on throughout this thesis was evaluative-simulation, an automated usability testing technique that is applicable during the design-phase of product development.

The evaluative capability of the flight simulation environment required a dynamic closed loop system in which the pilot and simulation test administrator played active roles. Since there were several data models and flight parameters that needed to be computed, the primary limitations of the simulation environment were network bandwidth, external device bus speeds and internal computer processing power.

After specifying the tasks of the human factors evaluator and defining measurable targets for constraints, system level requirements of an integrated flight emulation and an associated flight data recording and processing suite were developed. The flight emulation was required to contain integrated devices that would allow the pilot and simulation test administrator to control different aspects of the flight model. The flight data recording and processing suite was required to contain automated functions that would facilitate the collection of data to produce objective pilot performance measurements. The high level requirements and low level requirements were derived. Once traceability verified that the defined systems would fulfill the user tasks while functioning within the fundamental constraints, the simulation environment was implemented.

To demonstrate how evaluative-simulation benefits flight deck designers, and to validate the functionality of the simulation environment that was developed, two experiment designs were postulated. The experiments incorporated the use of elemental features of TAA automation (i.e., flight director, autopilot and FMS) while utilizing majority of the real-time and playback functionality implemented in the flight data recording and processing suite. However, there were several interaction features of the simulator that were not applied in these two validation experiments.

The successful conduct of the experiments proved that the integrated flight emulation is capable of replicating trials of formal experiments. Even though none of the trials used exceeded thirty minutes, which was a design target before network lag might become noticeable, there were no symptoms of system failure or system performance degradation. The report of some statistically significant results proved that the simulator is capable of producing conclusions that can determine hypotheses during the decision-making process.

The real-time processing and analysis of pilot performance made it possible to issue immediate feedback to the participants. This improved the efficiency of the time spent in the simulation administration and data entry phases.

To fully validate the functionality the simulation environment, real flight data should be compared with the results of identical simulation runs. Once these results are correlated and the variance of data analyzed, it should be possible to determine how well the performance of the simulation matches that of the real aircraft. These analyses will indicate the level of the transfer of testing from the simulation environment to the real world and verify the performance of the flight simulator.

The rationale for HFT&E of IFDs is rooted in the fact that flight deck systems are developed to provide pilot of various levels of experience with the ability to aviate, navigate and communicate. The evaluative-simulation technique has been used to determine whether or not the flight director, autopilot, and FMS achieve their purposes, to augment pilot performance. Evaluative simulations can be introduced during the system design-phase, then implemented during the system development-phase when prototype technology becomes available.

5. List of References

American Psychological Association. (2001). *Publication manual of the American Psychological Association* (5th ed.). Washington, DC: American Psychological Association.

ANSI (2001). *Common industry format for usability test reports* (ANSI-NCITS 354-2001). Washington, DC: American National Standards Institute.

AOPA Air Safety Foundation. (2005). *Technically advanced aircraft: Safety and training* (Special Report). Frederick, MD: Author.

Besco, R., & Funk, K. (1999). Conceptual design guidelines to rediscover systems engineering for automated flight decks. *International Journal of Aviation Psychology, 9*(2), 189–198.

Clamann, M., & Kaber, D. (2004). Applicability of usability evaluation techniques to aviation systems. *International Journal of Aviation Psychology, 14*(4), 395–420.

Endsley, M. (1996). Automation and situation awareness. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 163–181). Mahwah, NJ: Lawrence Erlbaum.

FAA. (1996). *The interfaces between flightcrews and modern flight deck systems*. Washington, DC: Federal Aviation Administration.

FAA. (2002a). *Addressing human factors/pilot interface issues of complex, integrated avionics as part of the technical standard order process*. Washington, DC: Federal Aviation Administration.

FAA. (2002b). *Commercial pilot practical test standards for airplane*. Washington, DC: Federal Aviation Administration.

FAA. (2004). *Flight crew error/ flight crew performance considerations in the flight deck certification process* (HF-HWG Final report). Washington, DC: Federal Aviation Administration.

Funk, K., Lyall, B., Wilson, J., Vint, R., Niemczyk, M., Suroteguh, C., & Owen, G. (1999). Flight deck automation issues. *International Journal of Aviation Psychology, 9*(2), 109–123.

Hackos, J. & Redish, J. (1998). *User and task analysis for interface design.* NY: John Wiley and Sons, Inc.

ISO. (1998). *Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11: Guidance on usability* (ISO 9241-11:1998(E)). Geneva, Switzerland: Author.

Jentsch, F., & Bowers, C. (1998). Evidence for the validity of PC-based simulations in studying aircrew coordination. *International Journal of Aviation Psychology, 8*(3), 243–260.

Jeppesen Sanderson, Inc. (2004). *Guided flight discovery: Instrument commercial.* Englewood, CO: Author.

Johnson, N., & Rantanen, E. (2005). *Objective pilot performance measurement: A literature review and taxonomy of metrics*. Paper to be presented at the 13th International Symposium on Aviation Psychology.

Levy, B., Som, P. & Greenhaw, R. (2003). Analysis of flight technical error on straight final approach segments. *Institute of Navigation 59th Annual Proceedings*.

Lewis, J. (In Press). *Handbook of Human Factors and Ergonomics 3rd Edition*. Gavriel Salvendy Editor. John Wiley & Sons.

Longridge, T., Burki-Cohen, J., Go, T. & Kendra, A. (2001). Simulator fidelity considerations for training and evaluation of today's airline pilots. *Proceedings of the 11th International Symposium on Aviation Psychology.*

Mosier, K., Skitka, L., Heers, S., & Burdick, M. (1998). Automation bias: Decision-making and performance in high-tech cockpits. *International Journal of Aviation Psychology, 8*(1), 47–63.

Mumaw, R., Sarter, N., & Wickens (2001). *Analysis of pilots' monitoring and performance on an automated flight deck*. Paper presented at the 11th International Symposium on Aviation Psychology.

Mumaw, R., Nikolic, M., Sarter, & N., Wickens, C. (2001). A simulator study of pilots monitoring strategies and performance on modern . *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting.*

National Transportation Safety Board. (2000). *Santa Rosa, California 95448 Sonoma County Airport, 10/04/00* (NTSB Rep. No. LAX01FA003). Retreived July 22, 2005, from NTSB Online Database: http://www.ntsb.gov/ntsb/brief2.asp?ev_id=20001212X22176&ntsbno=LAX01FA003&akey=1.

National Transportation Safety Board. (2002a). Mary*land 21520 Garret County Airport 10/31/02* (NTSB Rep. No. NYC03FA015). Retreived July 22, 2005, from NTSB Online Database: http://www.ntsb.gov/ntsb/brief2.asp?ev_id= 20021112X05459&ntsbno=NYC03FA015&akey=1.

National Transportation Safety Board. (2002b). *Las\Vegas, New Mexico 87701, 11/03/02* (NTSB Rep. No. FTW03FA029). Retreived July 22, 2005, from NTSB Online Database: http://www.ntsb.gov/ntsb/brief2.asp?ev_id= 20021108X05449&ntsbno=FTW03FA029&akey=1.

National Transportation Safety Board. (2003). *San Jose, California, Reid-Hillview of Santa Clara Airport* (NTSB Rep. No. LAX03FA072). Retreived July 22, 2005, from NTSB Online Database: http://www.ntsb.gov/ntsb/brief.asp?ev_id= 20030206X00175&key=1.

Oman, C., Kendra, A., Hayashi, M., Stearns, M., & Burki-Cohen, J. (2001). Vertical navigation displays: Pilot performance and workload during simulated constant angle of descent GPS approaches. *International Journal of Aviation Psychology, 11*(1), 15–31.

Prinzel III, L., Comstock Jr., R., Glaab, L., Kramer, L., & Arthur, J. (2004). The efficacy of head-down and head-up synthetic vision display concepts for retro- and forward-fit of commercial aircraft. *International Journal of Aviation Psychology, 14*(1), 53–77.

Rantanen, E., & Talleur, D. (2001). Use of flight recorders for pilot performance measurement and training. *Proceedings of the 2001 International Aviation Training Symposium*. Oklahoma City, OK: FAA Academy.

Rantanen, E., Talleur, D., Taylor, H., Bradshaw, G., Emanuel, T., Lendrum, L., & Hulin, C. (2001). *Derivation of pilot performance measures from flight data recorder information*. Paper presented at the 11th International Symposium on Aviation Psychology.

Sarter, N. (2000). The need for multisensory interfaces in support of effective attention allocation in highly dynamic event-driven domains: the case of cockpit automation. *International Journal of Aviation Psychology, 10*(3), 231–245.

Sarter, N., Wickens, C., Mumaw, R., Kimball, S., Marsh, R., Nikolic, M., & Xu, W. (2003). *Modern flight deck automation: pilot mental model and monitoring patterns and performance*. Paper presented at the 12th International Symposium on Aviation Psychology.

Sarter, N., & Woods, D. (1995). How in the world did we ever get into that mode? Mode error and awareness in supervisory control. *Human Factors, 37*(1), 5–19.

Sarter, N., & Woods, D. (1994). Pilot interaction with cockpit automation. II. An experimental study of pilot's mental model and awareness of the Flight Management System. *International Journal of Aviation Psychology, 4*(1), 1–28.

Sarter, N., & Woods, D. (1992). Pilot interaction with cockpit automation: Operational experiences with the flight management system. *International Journal of Aviation Psychology, 2*(4), 303–321.

Sauro, J & Kindlund E. (2005). A method to standardize usability metrics into a single score. *Proceedings of the Conference on Human Factors in Computing Systems*, 401 - 409.

Schnell, T., Kwon, Y., & Merchant, S. (2004). Improved flight technical performance in flight decks equipped with synthetic vision information system displays. *International Journal of Aviation Psychology, 14*(1), 79–102.

Shneiderman, B. & Plaisant, C. (2005). *Designing the user interface: Strategies for effective human-computer interaction* (4th ed.). Boston: Addison Wesley.

TAA Safety Study Team. (2003). *General aviation technically advanced aircraft: FAA – Industry safety study*. Washington, DC: Federal Aviation Administration.

Tenny, Y., Rogers, W., & Pew, R. (1998). Pilot opinions on cockpit automation issues. *International Journal of Aviation Psychology, 8*(2), 103–120.

Wickens, C.D. (2003). Aviation displays. In P. Tsang & M. Vidulich (Eds.), *Principles and practices of aviation psychology* (pp. 147-199). Mahwah, NJ: Lawrence Erlbaum Publishers.

Wickens, C. D., & Hollands, J. G. (2000). *Engineering psychology and human performance* (3rd ed). Upper Saddle River, NJ: Prentice-Hall, Inc.

Wiegmann, D., & Shappell, S. (2001). *Applying the human factors analysis and classification system to the analysis of commercial aviation accident data*. Paper presented at the 11th International Symposium on Aviation Psychology.

Wiegmann, D., Shappell, S., Boquet, A., Detwiler, C., Holcomb, K., & Faaborg, T. (2005). *Human error and general aviation accidents: A comprehensive, fine-grained analysis using HFACS* (Final Technical Report AHFD-05-08/FAA-05-03). Urbana-Champaign: University of Illinois, Aviation Human Factors Division.

Williams, K. & Ball, J. (2004). *Usability and effectiveness of advanced general aviation cockpit displays for visual flight procedures*. U.S. Department of Transportation, Federal Aviation Administration, Office of Aerospace Medicine, Washington, DC. Publication No. DOT/FAA/AM-04/20.

6. Appendices

6.1 **Appendix A –** *AviSAFE Architecture*

Simulator evaluations use devices that present an integrated emulation (using flight hardware, simulated systems, or combinations of these) of the flight deck and the operational environment (FAA, 2004). AviSAFE contains an integrated flight emulation that can be flown with response characteristics that replicate, to some extent, the responses of the Cirrus SR22. However, some evaluations may be limited by the fidelity and realism of the simulation in its representation of the airplane, flight deck, external environment, and pilot operations. The certification teams of the FAA have noted that not all aspects of the simulation must have a high level of fidelity for any given compliance issue (FAA, 2004). Rather, the level of fidelity required should be determined in view of the issue being evaluated.

AviSAFE utilizes three essential computer programs that constitute its integrated flight emulation. The programs are coupled using the Transmission Control Protocol and Universal Datagram Protocol on the Internet Protocol (TCP/IP and UDP/IP). When data transfer must be reliable must be transferred critically, for example, the pilot command packets which are to be recorded twice per second, TCP/IP is chosen because errors and misplaced packets are automatically handled by the protocol. Whereas, when the reliability of data transfer is not as critical and reliability can be leveraged, for example, the outside display data packets which are to be updated ten to thirty times per second, UDP is chosen since the loss of two packets in a one second bin will have no perceptible effect.

The first program is AviSim, which is a six-degree of freedom simulation engine with a World Geodetic System (WGS-84) earth model, a configurable aircraft model that has been arranged to imitate the Cirrus SR22, and other physics models that emulate the atmospheric conditions, autopilot servo commands and several other environmental functions. AviSim accepts pilot commands via Universal Serial Bus (USB) devices such as a three-axis joystick with an autopilot engage/disengage button switch, a pair of rudder pedals, mockup autopilot, and power quadrant with throttle, mixture and propeller pitch.

Figure 31 shows the AviSim windows interface whose main component is an interactive tree structure that permits the operator to change the many parameters of the various models affecting the simulation. The parameters are changed by activating a radio-button to select the manual input option then a value is entered for the model parameter that is selected. Other radio-buttons allow the operator to select different forms of simulated input for parameters such as constant value, periodic sinusoidal changes or step-wise changes.

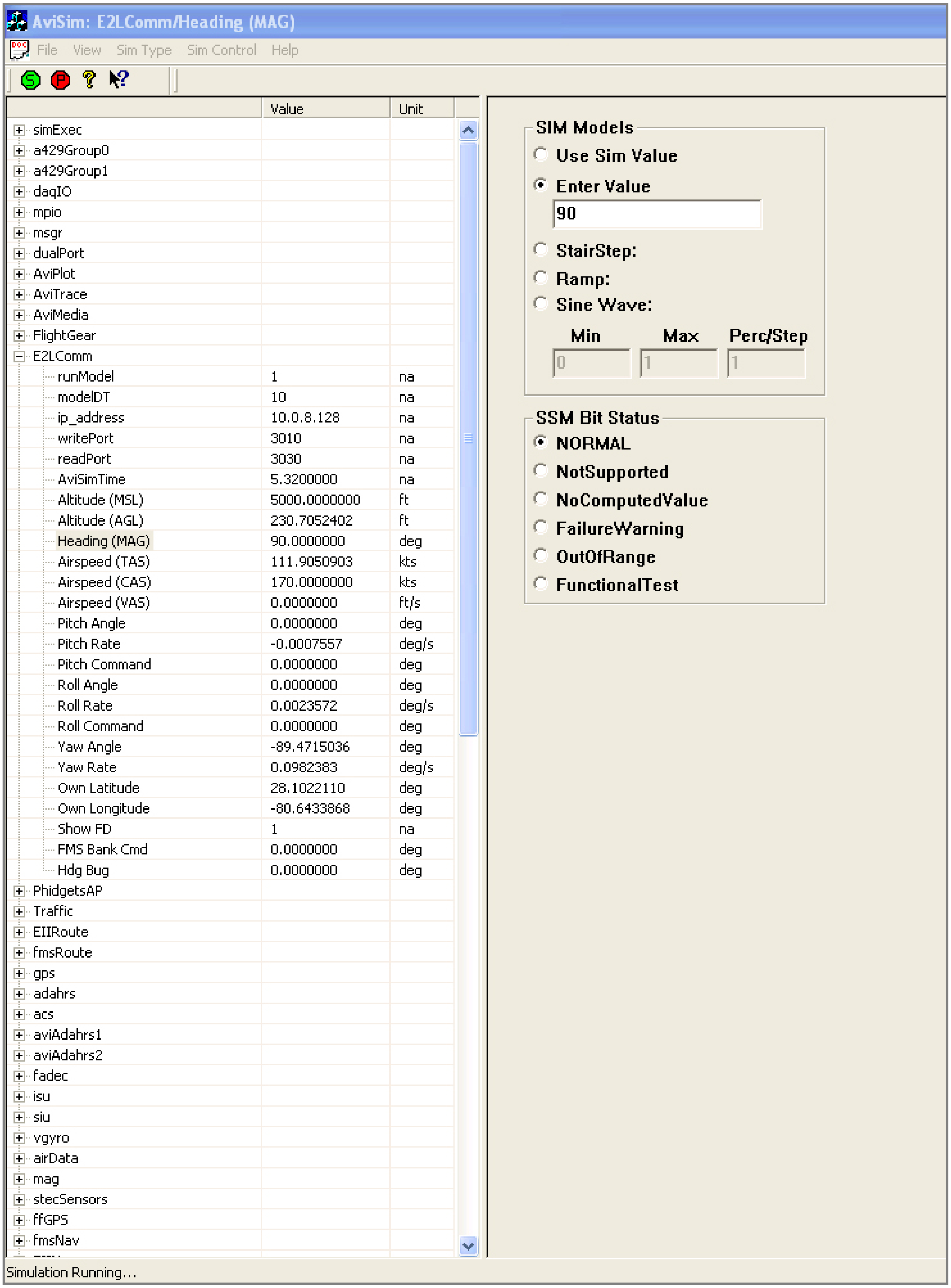


Figure 31. AviSim GUI with the E2L communications module expanded

AviSim is linked with all programs in the simulated environment. It sends air data, position data and flight director status to the cockpit flight deck program over UDP/IP, similar data are sent to the outside display program over UDP/IP, flight parameter data to the plotting program over TCP/IP, and unique time-stamps to the other flight data recording programs over TCP/IP. AviSim receives waypoint identifiers and position data from the cockpit flight deck program over UDP/IP.

The Entegra Loop (E2L) is a windows-based emulation of the Avidyne Entegra integrated flight deck with a simulation mode where the flight deck systems can be stimulated by external sources such as the models in AviSim. E2L accepts pilot commands via a USB keyboard and USB mouse.

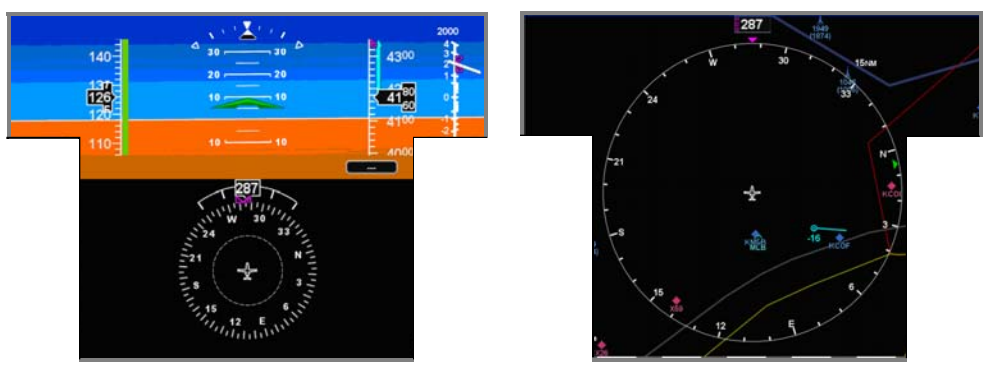


Figure 32. Dual E2L instances running with primary flight display (left) and moving MAP (right)

Figure 32 shows the E2L graphical user interface containing a replica of the integrated flight deck as it evolves in the design process. There are interactive buttons with function key shortcuts that are used to interact with E2L. The pilot operating

E2L uses the mouse and keyboard as input devices. Since E2L is a development version of the final product, not all modes have been implemented to date. However, the experiments in this thesis utilized only adequately functioning modes.

E2L uses UDP/IP connections to link with AviSim and one of the flight data recorders, AviTrace. It receives air data, position data and flight director status from AviSim at 10 Hz while returning waypoint identifiers and position data to AviSim when the user interacts with the FMS. E2L sends all simulated button presses and knob turns to AviTrace where they are time-stamped using AviSim time and stored for real-time analysis and/or playback.

FlightGear is an open-source flight simulator that can be modified so that it is driven by an external source such as AviSim. There are several menu options that can be used to configure FlightGear before and after startup. FlightGear has scenery databases that can be downloaded for most regions in the North American continent. In the AviSAFE configuration, FlightGear scenery is projected on a wall in front of the flight deck to simulate the outside view (see Figure 33). FlightGear uses a UDP/IP connection to link with AviSim so that it can receive air and position data. The networking configuration needed for FlightGear to communicate with AviSim was preset within AviSim.

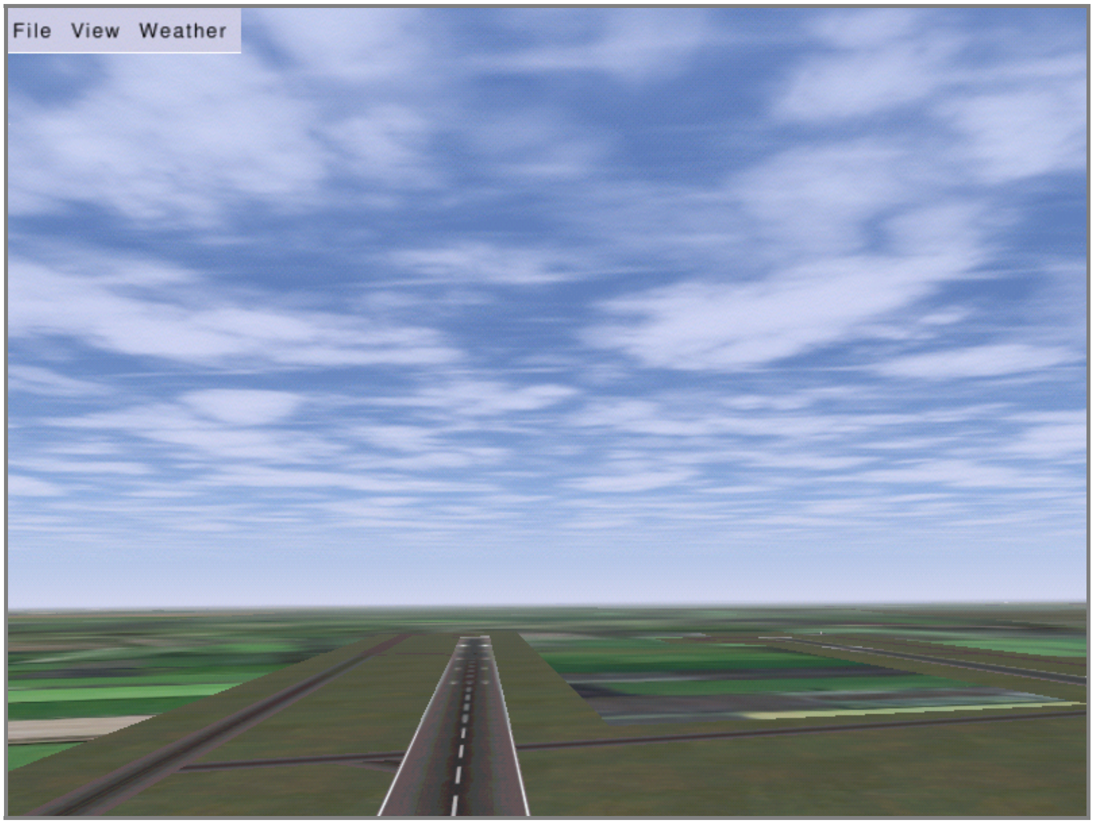


Figure 33. FlightGear scenery displaying outside view of the cockpit

AviSAFE has employed three seminal computer programs that constitute the FDR suite. Like the integrated flight emulation, the programs of the FDR suite are coupled using the Transmission Control Protocol and Universal Datagram Protocol on the Internet Protocol (TCP/IP and UDP/IP). The FDR suite is linked to the integrated flight emulation via TCP/IP and UDP/IP connections.

AviPlot is a plotting program that records and displays data for several flight parameters using a Cartesian coordinate system. The plots can be automatically preconfigured depending on the scenario to be evaluated thereby reducing the time taken to setup the system to record a scenario. AviPlot has features such as zooming and snapping to data that make real-time data processing a simple task. Of greater importance, error tolerances can be set up around a baseline plot and compared to the plot of the pilot that is performing real-time.

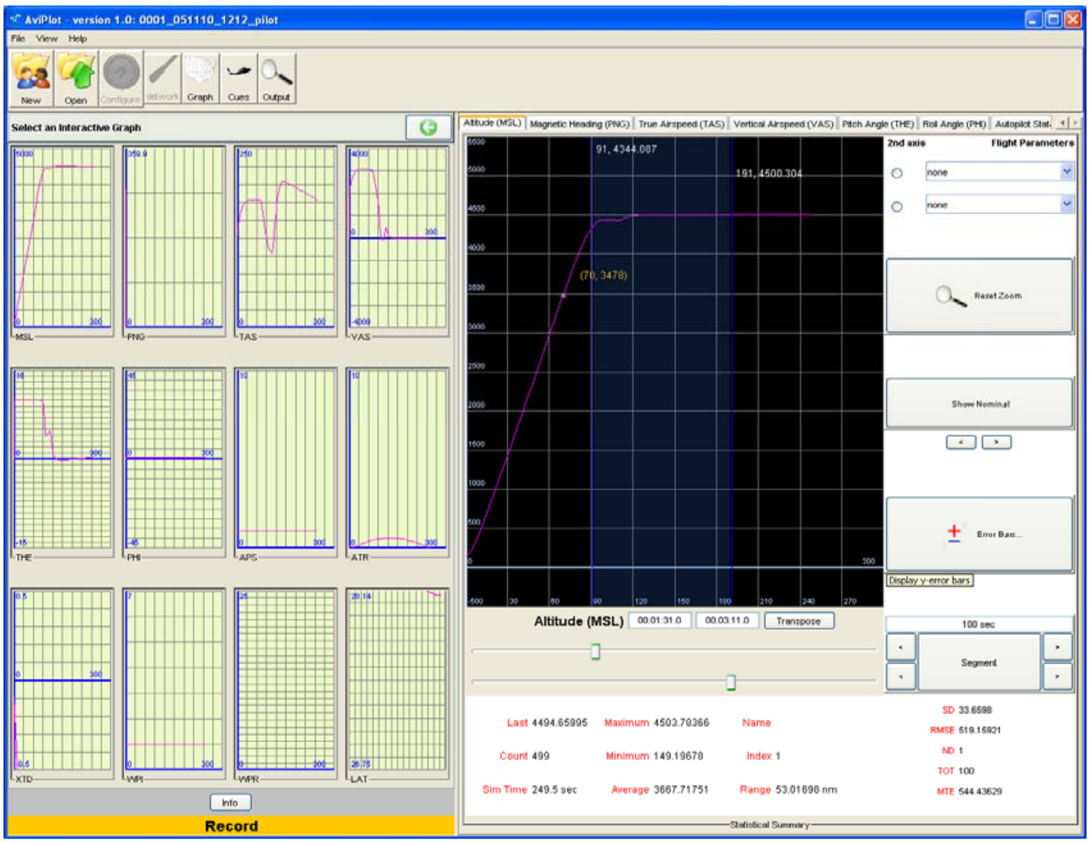


Figure 34. The AviPlot GUI in the multiple plot dialog configuration

Figure 34 shows the AviPlot windows interface with two main panels (dialog panel and active panel) that house several components. The dialogs associated with dialog panel are changed by activating a button above the dialog panel. Different plots are activated in the active panel by selecting the appropriate tab or selecting the appropriate plot from a multiple-plot control in the dialog panel. There are two slider bars with associated text boxes and buttons for controlling the location of the segmentation bars. The bottom pane of the active panel contains a statistical summary for the active plot including the performance measures for the defined segment. Special controls are activated by selecting the buttons on the right side of the active panel and by right-clicking to activate the context menu.

The error tolerances can be preset using the FAA's PTS or statistical criteria such as standard deviation or standard error of the mean. AviPlot also features segmentation sliders that allow the evaluator to analyze five performance measures for a defined flight segment. The performance measures that are automatically computed for the defined segment of the a particular flight parameter are standard deviation, root mean square error, number of deviations, time outside tolerance and the mean time to exceed tolerance.

AviPlot receives flight data (pitch, pitch rate, pitch command, roll, roll rate, roll command, yaw, yaw rate, yaw command, indicated airspeed, true airspeed, altitude above mean sea level, altitude above ground level, latitude, longitude, cross track error, waypoint range, waypoint index and autopilot status) from AviSim via TCP/IP at a rate of 2 Hz (Appendix E). AviPlot sends cue data via TCP/IP to AviTrace and AviMedia.

AviTrace is a trace program that records and displays data for E2L operations on the IFDs using the mouse and keyboard. The traces become visible on AviTrace's graphical user interface immediately as the E2L interaction occurs. AviTrace is capable of displaying interaction traces separate for the primary and secondary IFDs and the keyboard, or in a merged list where all interaction traces are displayed chronologically.

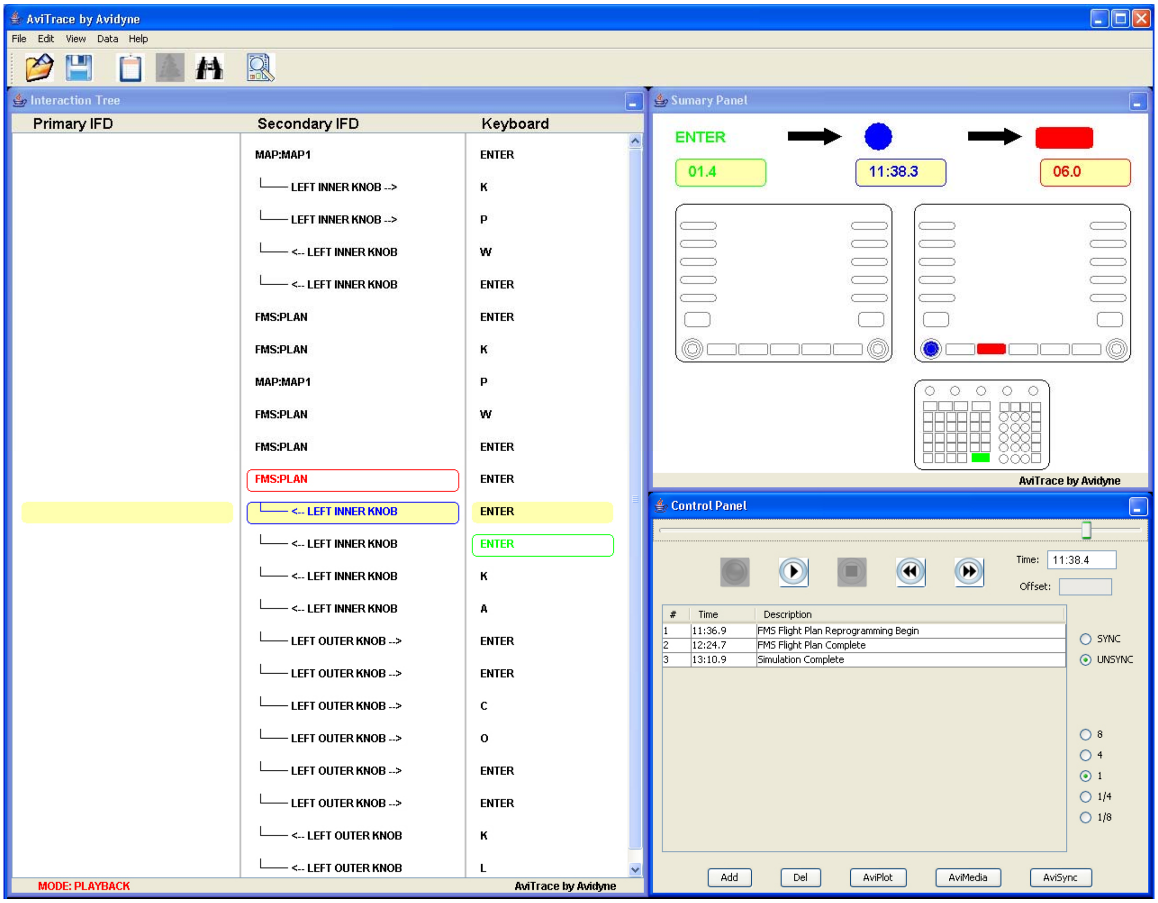


Figure 35. The AviTrace GUI in the tree configuration

Figure 35 shows the AviTrace windows interface with three main panels (interaction trace panel, graphical panel and playback control panel) that house several components. The interaction trace panel is capable of displaying merged traces or separate traces for the different interaction sources. The graphical panel contains summarized time information about the previous, current and next interactions. There are graphical representations of the interaction sources that highlight the location of the interactions real-time. The control panel contains playback controls, a timeline, time window and speed control buttons that are used to control the playback of data. There is a cue table and its associated buttons that are used to add, edit and delete cue points.

AviTrace receives interaction traces from E2L interaction sources via UDP/IP that simulate button presses and knob turns on the IFDs and keyboard. AviTrace also receives centralized time from AviSim. AviTrace sends cue data that is entered at a time of interest via TCP/IP to AviPlot and AviMedia.

AviMedia is a multimedia program that records and displays audio-video data streams. The data streams are saved directly to files for playback and analysis. AviMedia receives audio input from a pilot headset compatible intercom system and a Windows Media Player output source via stereo audio connections. The Windows Media Player application supplies a simulated air traffic control audio either from a live connection or a saved file playback. In the case of Experiment 2, two video files from high traffic periods are used during the experiment.

AviMedia receives video input from three web-video cameras via USB connections and from a computer's video card output via an S-Video connection. The web-cameras provide video data from three different locations in the flight simulator (between the primary IFD and the outside view; between the secondary IFD and the center console; and from overhead behind the pilot). The camera between the primary

IFD and the outside view captures the pilot looking directly ahead when the primary IFD is in the field of view (FOV). When the pilot looks over at the secondary IFD, the camera between the secondary IFD and the center console captures the pilot looking directly off-center with the secondary IFD in the FOV. The S-Video connection provides video data containing the FlightGear scenery in the outside view.

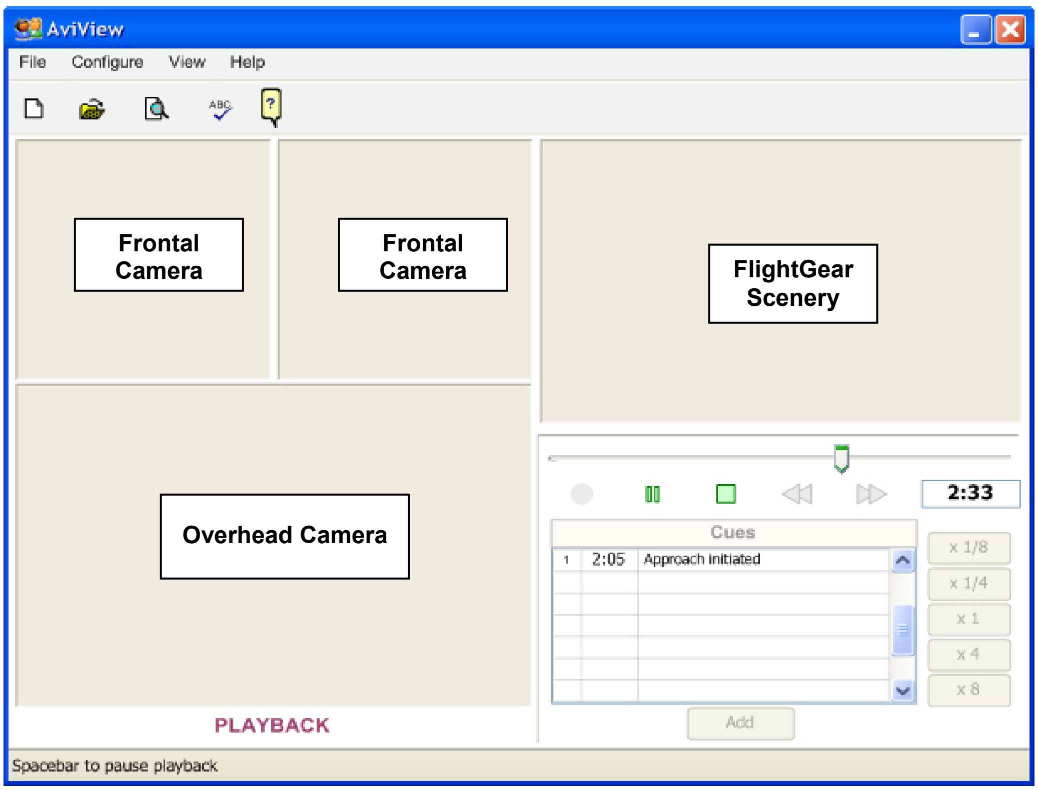


Figure 36. The AviMedia GUI in the playback mode

Figure 36 shows the AviMedia windows interface with five main panels (primary IFD frontal panel, secondary IFD frontal panel, overhead panel, FlightGear scenery panel and playback control panel) that house media components. Similar to

AviTrace, the control panel contains playback controls, a timeline, time window and speed control buttons that are used to control the playback of audio and video data streams. There is a cue table and its associated buttons that are used to add, edit and delete cue points.

AviMedia receives audio data from an intercom system and the Windows Media Player application. AviMedia receives video data from web-cameras and a computer video card output. AviMedia also receives centralized time from AviSim. AviMedia sends cue data that is entered at a time of interest via TCP/IP to AviPlot and AviTrace.

6.2 **Appendix B –** *Electronic Survey Instrument*

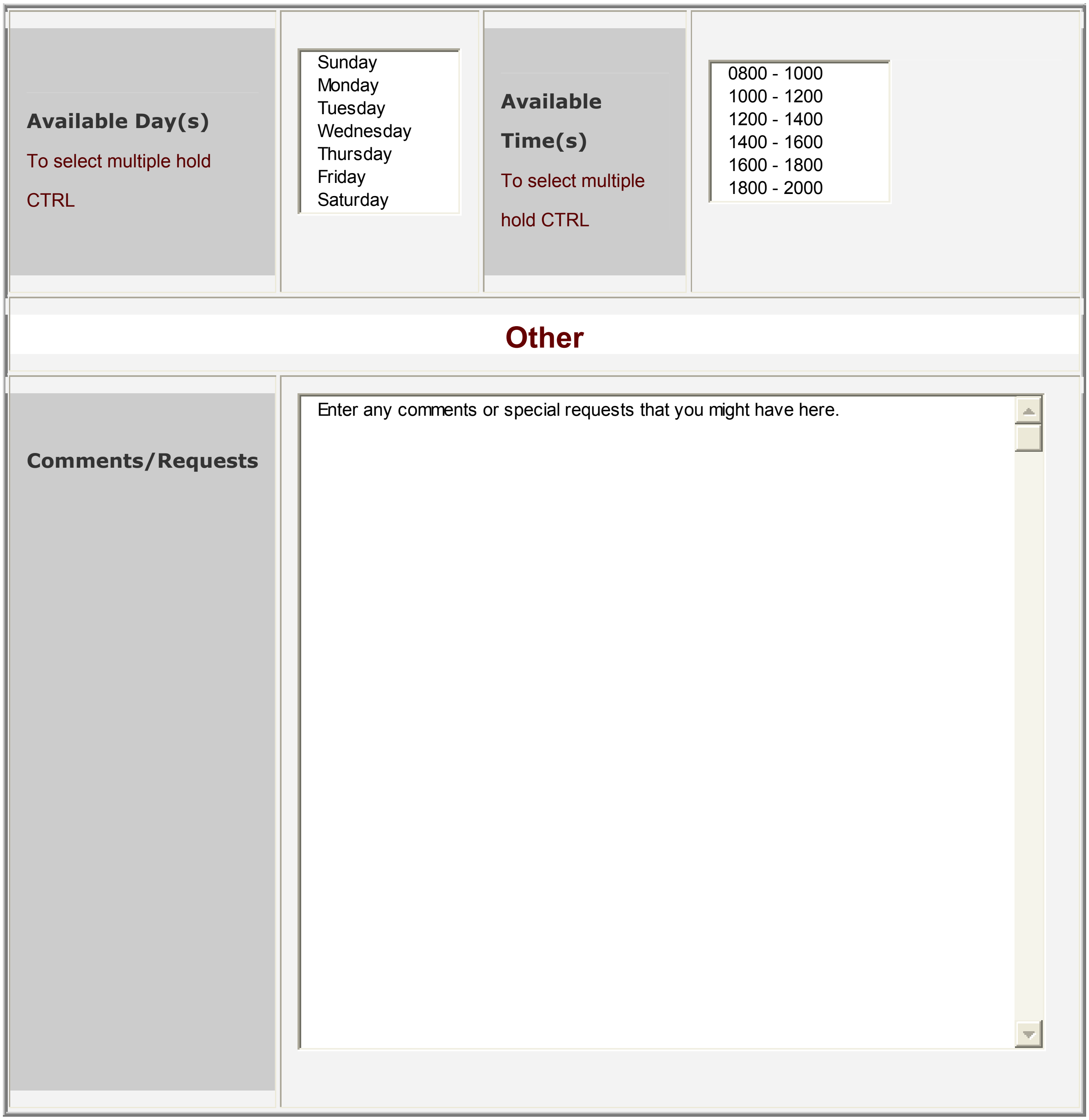
General Information	
Date	December 6, 2005
Full Name	
Sex	Male Female
Age	
Phone Number	
Email Address	
Flight Information	
Pilot License	Student Private Commercial
Flight Instructor	No Yes
Total Flight Hrs	0

Instrument Hrs Actual, simulated & simulator	0
Glass Cockpit Hrs Actual, simulated & simulator	0
Side-stick Hrs Actual, simulated & simulator	0

Computer Information

Computer Experience What level are you?	Basic Below average Average Above average Extensive
Typing Experience How well do you type?	Basic Below average Average Above average Extensive

Availability



Available Day(s)

To select multiple hold CTRL

Sunday
Monday
Tuesday
Wednesday
Thursday
Friday
Saturday

Available Time(s)

To select multiple hold CTRL

0800 - 1000
1000 - 1200
1200 - 1400
1400 - 1600
1600 - 1800
1800 - 2000

Other

Comments/Requests

Enter any comments or special requests that you might have here.

6.3 **Appendix C –** *Experiment 1 Trial Sequence Randomization*

Each participant randomly receives a trial sequence, *S*.

S	Trial 1	Trial 2	Trial 3	Trial 4
1	A	B	C	D
2	A	B	D	C
3	A	C	B	D
4	A	C	D	B
5	A	D	B	C
6	A	D	C	B
7	B	A	C	D
8	B	A	D	C
9	B	C	A	D
10	B	C	D	A
11	B	D	A	C
12	B	D	C	A
13	C	A	B	D
14	C	A	D	B
15	C	B	A	D
16	C	B	D	A
17	C	D	A	B
18	C	D	B	A
19	D	A	B	C
20	D	A	C	B
21	D	B	A	C
22	D	B	C	A
23	D	C	A	B
24	D	C	B	A

Code	FD STATE	Maneuver
A	FD ON	Straight and Level
B	FD OFF	Straight and Level
C	FD ON	Climbing Turn
D	FD OFF	Climbing Turn

6.4 **Appendix D –** *Experiment 2 Trial Sequence Randomization*

Each participant randomly receives a trial sequence, *S* within the participant's certificate category.

Pilot Certificate	*S*	Trial 1	Trial 2
Student	1	A	B
	2	B	A
	3	A	B
	4	B	A
	5	A	B
	6	B	A
	7	A	B
	8	B	A
Private	9	A	B
	10	B	A
	11	A	B
	12	B	A
	13	A	B
	14	B	A
	15	A	B
	16	B	A
Commercial	17	A	B
	18	B	A
	19	A	B
	20	B	A
	21	A	B
	22	B	A
	23	A	B
	24	B	A

Code	AP STATE
A	AP ON
B	AP OFF

6.5 **Appendix E –** *Flight Data Recorder Specifications*

The following table was adapted from Appendix M to Title 14 CFR Part 121.

Parameter	**Range**	**Sampling Interval** (sec)	**Resolution**
Time	0 to 4095 sec	4	1 sec
Pressure Altitude	-1000 to MCA+5,000 feet	1	5 to 35 feet
Indicated Airspeed	50 to Max V_{SO} KIAS	1	1 knot
Heading	0 to 360 degrees	1	0.5 degree
Pitch Attitude	± 75 degrees	1	0.5 degree
Roll Attitude	± 180 degrees	1	0.5 degree
Pitch Controls	Full range	0.5	0.2 percent of full range
Lateral Controls	Full range	0.5	0.2 percent of full range
Yaw Controls	Full range	0.5	0.2 percent of full range
Latitude	As installed	4	0.002 degree
Longitude	As installed	4	0.002 degree
Autopilot Status	0 – 1 Binary discrete	1	…