Summer 8-2017

Comparative Analysis of Conventional Electronic and OZ Concept Displays for Aircraft Energy Management

Erik Reese Baker
Embry-Riddle Aeronautical University

COMPARATIVE ANALYSIS OF CONVENTIONAL ELECTRONIC AND OZ CONCEPT DISPLAYS FOR AIRCRAFT ENERGY MANAGEMENT

By

Erik Reese Baker

A Dissertation Submitted to the College of Aviation
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Aviation

Embry-Riddle Aeronautical University
Daytona Beach, Florida
August 2017

COMPARATIVE ANALYSIS OF CONVENTIONAL ELECTRONIC AND OZ CONCEPT DISPLAYS FOR AIRCRAFT ENERGY MANAGEMENT

By

Erik Reese Baker

This Dissertation was prepared under the direction of the candidate's Dissertation Committee Chair, Dr. Haydee M. Cuevas, and has been approved by the members of the dissertation committee. It was submitted to the College of Aviation and was accepted in partial fulfillment of the requirements for the Degree of Doctor of Philosophy in Aviation

Haydee M. Cuevas, Ph.D.
Committee Chair

Andrew R. Dattel, Ph.D.
Committee Member

Antonio I. Cortés, Ph.D.
Associate Dean, School of Graduate Studies

Robert E. Joslin, Ph.D.
Committee Member

Alan J. Stolzer, Ph.D.
Dean, College of Aviation

David L. Still, Ph.D.
Committee Member

Michael P. Hickey, Ph.D.
Dean of Research and Graduate Studies

24 August 2017
Date

ABSTRACT

Researcher: Erik Reese Baker

Title: COMPARATIVE ANALYSIS OF CONVENTIONAL ELECTRONIC AND OZ CONCEPT DISPLAYS FOR AIRCRAFT ENERGY MANAGEMENT

Institution: Embry-Riddle Aeronautical University

Degree: Doctor of Philosophy in Aviation

Year: 2017

A repeated-measures, within-subjects design was conducted on 58 participant pilots to assess mean differences on energy management situation awareness response time and response accuracy between a conventional electronic aircraft display, a primary flight display (PFD), and an ecological interface design aircraft display, the OZ concept display. Participants were associated with a small Midwestern aviation university, including student pilots, flight instructors, and faculty with piloting experience. Testing consisted of observing 15 static screenshots of each cockpit display type and then selecting applicable responses from 27 standardized responses for each screen.

A paired samples *t*-test was computed comparing accuracy and response time for the two displays. There was no significant difference in means between PFD Response Time and OZ Response Time. On average, mean PFD Accuracy was significantly higher than mean OZ Accuracy ($M_{Diff} = 13.17$, $SD_{Diff} = 20.96$), $t(57) = 4.78$, $p < .001$, $d = 0.63$. This finding showed operational potential for the OZ display, since even without first training to proficiency on the previously unseen OZ display, participant performance differences were not operationally remarkable.

There was no significant correlation between PFD Response Time and PFD Accuracy, but there was a significant correlation between OZ Response Time and OZ

Accuracy, r (58) = .353, $p < .01$. These findings suggest the participant familiarity of the PFD resulted in accuracy scores unrelated to response time, compared to the participants unaccustomed with the OZ display where longer response times manifested in greater understanding of the OZ display.

PFD Response Time and PFD Accuracy were not correlated with pilot flight hours, which was not expected. It was thought that increased experience would translate into faster and more accurate assessment of the aircraft stimuli. OZ Response Time and OZ Accuracy were also not correlated with pilot flight hours, but this was expected. This was consistent with previous research that observed novice operators performing as well as experienced professional pilots on dynamic flight tasks with the OZ display. A demographic questionnaire and a feedback survey were included in the trial. An equivalent three-quarters majority of participants rated the PFD as “easy” and the OZ as “confusing”, yet performance accuracy and response times between the two displays were not operationally different.

DEDICATION

This dissertation is dedicated to my late father, John Austin Baker, Ed.D., who is my best friend, my life coach, my role model, and my inspiration. I only wish I could have completed this sooner so you could have seen me walk across that stage. You will always be with me, and following in your footsteps as I start my next chapter as university faculty, all that I do I dedicate to you. I love you, Dad.

ACKNOWLEDGEMENTS

This product was a culmination of over seven years of hard work. It obviously was not a solo effort. There are too many people to acknowledge here, but I will do my best to recognize the herculean efforts the following people have made to educate me, support me, encourage me, and love me.

My committee, especially my dissertation chair and friend, Dr. Haydee Cuevas, whose infectious positive attitude and limitless energy are a boon to all of her students. Dr. Andy Dattel, who offered advice and guidance on particulars of running an experiment. Dr. Robert Joslin, who dared me to think bigger and dig deeper on my research. Dr. Dave Still, who expertly supported and guided this research on his “baby”, the OZ display, and Dr. Tom Eskridge, who provided the OZ software and de-bugging, translating his advanced computer knowledge to a less technical newbie like me. Dr. Alan Stolzer, who kept his faith in me, offering me several deadline allowances. Dr. Mary Jo Smith, who offered statistical expertise and life tips, sometimes disguising them as each other. Dr. David Freiwald, who I not only credit with my current employment, but offered collegial support and superior academic guidance throughout. Dr. Kevin O’Leary, who avoided Patriot gloating to the best of his ability and kept calling and texting me during the dark ages of the dissertation phase where you are all but forgotten. Dr. Randy DeMik, the best mentor and friend I could ever ask for, who patiently protected me during my first year as an academic, and offered expert guidance throughout the dissertation process. Kathy Mazies, the warm friend who took superior care of me during my first year at Lewis and tirelessly became the public face of my experiment. Kyle Meyer, the bright and industrious graduate assistant, that worked with me on my pilot study and effortlessly produced participants for my experiment. The Aviation & Transportation Department at Lewis University, for giving me the "keys to the city" this year and allowing me to produce my best work and thrive as a new teacher and ABD. The amazing programmers at Millisecond Software who answered all my questions about Inquisit 5 Lab. Panera, Down to Earth, and Starbucks, who provided the internet, the coffee, and the health food "haven" when I spent countless days working on this product.

My mom, Leila Ann Baker, whose support, love, and daily inquisitive concern kept me honest about my progress, my motivation, my dedication, and the final goal.

My children, Lena, Ingrid, Henry, and Martin, whose entire lifetime has been burdened by their father's never-ending academic pursuits. Thank you for your patience, love, kindness, and suggestion to "just write 'I love airplanes' 500 times and call it good!" was considered more than once or twice. Let the promise of finally acquiring a family dog commence!

My wife, Amy Elizabeth Baker, whose entire 14-year history with me has been framed by some sort of intense co-lateral graduate school. You have never wavered in your support, your approval, your sacrifice for this goal. What will life be like post-doc? I am ready to find out. This next 14 years is dedicated to YOU!

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

This study seeks to determine whether an ecological interface design (EID) such as the OZ concept cockpit display can improve a pilot's energy management situation awareness (EM SA) over a conventional electronic flight display when observing display screenshots of an airplane operating at high altitude. The two aerodynamic manifestations of concern during climb and subsequent level cruise are: (a) when airplanes fly near their absolute aerodynamic ceiling—all available thrust is required to maintain balanced level flight; and (b) when airplanes decelerate into the region of reverse command—increased thrust is required for slower flight. If cockpit displays provided situation awareness (SA) specifically focused on energy management (EM) (i.e., thrust required with respect to thrust available or lift over drag (L/D)), then pilots would have the opportunity to recognize EM deviations early and perhaps correct from a developing loss of control inflight (LOC-I) incident. In this chapter, the following relevant concepts will be briefly discussed to provide the context for this study:

- Loss of Control Inflight
- Airplane Upsets
- Reverse Command and Coffin Corner
- Energy Management Situation Awareness
- Information Requirements
- Ecological Interface Design
- OZ Concept Display

Loss of Control Inflight

LOC-I is defined by the Federal Aviation Administration (FAA) as a category of airplane accident or incident resulting from a deviation from intended flight path (FAA, 2015a). LOC-I was the leading cause of fatalities in commercial aviation between 2001 and 2011 (ICAO, 2014b), involved in 22 commercial jet accidents and over 40% of all fatalities between 1999 and 2008 (Belcastro & Jacobson, 2010).

Airplane Upsets

Brooks, Ransbury, and Stowell (2014) claim that airplane upsets often precede fatal LOC-I events. An airplane upset is defined by both the FAA and International Civil Aviation Organization (ICAO) as "an unintentional exceedance of flight parameters in normal line operations or training—greater than 25° nose up, greater than 10° nose down, greater than 45° bank, or flying within the above parameters but at airspeeds inappropriate for the conditions" (FAA, 2015a, p. 2; ICAO, 2014a, p. *x*). Unintended stalls are a subset of airplane upset, since during a stall the airplane is at an inappropriate airspeed for the conditions (ICAO, 2014a). Aerodynamic stalls are responsible for nearly half of all fatal LOC-I accidents (Brooks et al., 2014).

Reverse Command and Coffin Corner

Two complex aerodynamic concepts are the region of reverse command and the coffin corner. The region of reverse command can occur at any altitude and is explained when an airplane is flown at speeds slower than the maximum lift over drag ratio (L/D_{max}), commonly known as best endurance airspeed. While maintaining a level altitude in reverse command, any decrease in desired airspeed requires a counter-intuitive increase in thrust due to the accumulation of induced drag as a function of lift (Carbaugh, Rockliff, & Vandel, 2008). When an airplane flies faster than L/D_{max} airspeed, the relationships are normal in so much as an increase in desired speed necessitates a logical increase in thrust.

LOC-I incidents sometimes occur when commercial and business jets fly at the edge of their flight envelope—the *coffin-corner*—which is the confluence of the airplane's altitude ceiling (the highest it can possibly fly based on engine performance) and the narrow margin between structural overspeed (too fast) and aerodynamic stall (too slow) (FAA, 2013). Airplanes are purposefully flown in that confluence in order to reap the operational cost benefits of fuel efficiency and faster ground speeds due to the less dense air at altitude, thereby placing airplanes at risk for high altitude upset when cruise true airspeed (TAS) is so near to the stall margin. These emergent operational preferences add complexity to high altitude decision making and overall SA, increasing the risk for LOC-I.

Energy Management Situation Awareness (EM SA)

Poor energy management has been a causal factor in airplane upset LOC-I accidents (Belcastro & Jacobson, 2010; ICAO, 2014a). EM is the process by which pilots safely and efficiently convert thermal energy (e.g., fuel) into potential energy (e.g., altitude) and kinetic energy (e.g., airspeed) (Merkt, 2013). SA is "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley, 1995, p. 36). Therefore, EM SA is the degree to which a pilot has an accurate awareness of the energy state of their airplane.

EM SA data. Although the two main devices manipulated on an airplane for EM are engine throttles and primary flight controls, a pilot may draw upon other available information in the cockpit to make EM decisions. Cockpit designers have been able to provide a plethora of data crucial to proper energy management to pilots, such as atmospheric conditions and gross weight of the airplane. However, the pilot is still left responsible for the interpretation and application of the information to determine the airplane's energy state. Furthermore, EM information is scattered throughout the cockpit displays and in deep layers of flight management system (FMS) pages. Aircraft parameter data "come from different sources, differing in importance, and using different scales and frames of reference… [and do not] have any explicit relationship with other [data] streams–except in the operator's head" (Eskridge, Still, & Hoffman, 2014, p. 91), meaning pilots spend considerable cognitive capital processing raw data in order to build EM SA comprehension.

Information Requirements

Any new primary flight display (PFD) design must first ascertain the information requirements of the pilots, that is, the types of flight performance information that would aid in making safe, timely, and accurate decisions during the conduct of a flight. Airplane cockpits require pilots to scan a multitude of different displays to find critical flight performance information (Temme, Still, & Acromite, 2003). The FAA published Advisory Circular 25-11B (2014) mandates that critical flight information, called primary flight information (PFI), be displayed on the PFD located in the primary field of view (PFOV). Important issues to consider are whether different combinations of information sources should be displayed on the PFD during different flight regimes and how information should be displayed on the PFD in order to maintain the airplane in a safe operating flight envelope with respect to altitude, airspeed, attitude, angle of attack, and thrust.

Ecological Interface Design

An EID is a human-computer interface for a complex sociotechnical system, such as found in medicine, aviation, and nuclear power (Vicente, 2002). While PFDs are data driven, full of various aeronautical data, push cognitive limits, and offer limited support to SA, the philosophy behind EID is to provide operators with specific work domain constraints, aiding their SA, so they can make informed decisions during complex scenarios. EIDs present both physical and functional information through intuitive graphical features promoting Endsley's levels of SA: perception, comprehension, and projection of future states (2012; 1995).

OZ Concept Display

Toward the goal of "reducing the cognitive effort required to maintain operator SA" (IMHC, 2015, para. 3), the Florida Institute of Human & Machine Cognition (IHMC) developed a new type of airplane instrument display called the OZ concept display. By providing all three levels of SA information as described by Endsley (1995): perception (Level 1), comprehension (Level 2), and projection (Level 3), IMHC asserts that the OZ display "presents data in a manner that allows the operator to understand the situation effortlessly" (IMHC, 2015, para. 2).

Significance of this Research

LOC-I and airplane upsets during the enroute phase of flight are the leading cause of accidents and fatalities worldwide (Carbaugh, Rockliff, & Vandel, 2008). Emphasis on preventing LOC-I is a priority for the ICAO commercial aviation field as well as the U.S. general aviation (GA) community (ICAO, 2014b, para. 1; NTSB, 2015).

Smith, Fadden, and Boehm-Davis (2005) examined pilot performance on altitude, airspeed, and heading control using the OZ display in comparison to an analog, round-dial, Cessna cockpit. Their study recommended further investigation of the OZ display in comparison to modern electronic flight information system (EFIS) technology, with respect to its ability to integrate multiple types of information. The functionality of the OZ display incorporates EM SA data, such as L/D max speed, into the typical information delivered to the pilot by a PFD such as altitude, heading, and airspeed. Angle of attack (AOA) information is intended for future OZ designs, pending software modifications (D. L. Still, personal communication, September 16, 2015). Smith et al. (2005) recognized the necessity to identify technology that could improve the next generation of flight displays. Notably, the OZ concept display, built for a light single-engine Cessna 172, has not been evaluated in the high altitude aerodynamic regime where commercial aviation commonly experiences LOC-I (Gerold, 2003). The proposed research aims to address this gap.

Statement of the Problem

Current cockpit PFDs may not provide pilots with adequate energy management SA to prevent LOC-I during high altitude cruise flight. Modern flight operations, intent on minimizing operating costs, seek flight paths at higher altitudes due to reduced fuel burn and favorable upper-level winds on certain routes in cruise, but often inadvertently place turbojet or turboprop airplanes near aerodynamic stall with minimal excess thrust or power, respectively. Thrust, normally associated with turbojet engines, is defined as work over distance (Dole & Lewis, 2000). Power, normally associated with turboprop or turboshaft engines, is defined as work over time (Dole & Lewis, 2000). Any mention of thrust or power in this document may be interchanged, respective to the type of engine in question. A lack of availability of information about the thrust curve and area of reverse command can make EM decisions critical at high altitude and slow cruise speeds since the airplane has no EM margin for error.

Current flight displays may not provide pilot SA in terms of energy management during high altitude regimes because these displays may not contain sufficient Level 2 SA (comprehension) or Level 3 SA (projection) information. The typical PFD and engine indicating and crew alerting system (EICAS) of modern airplanes offers Level 1 SA (perception) information but does not seem to support comprehension (Level 2 SA), and projection of future states is limited (Level 3 SA). Even with modern airplane stall alerting systems such as stick shaker, stick pusher, and AOA gauges, conditions leading to aerodynamic stalls might be avoided by proper reference to an intuitive, Level 3 SA display of EM information, such as depicted on an OZ concept display.

If the off-nominal conditions surrounding the occurrence of high altitude upsets are becoming more likely, pilots should be provided with all available technology to properly analyze and take corrective action on complex energy management situations. In particular, providing PFDs with thrust required / thrust available curve information could reduce and possibly prevent the occurrence of aerodynamic stalls.

Purpose Statement

This study will evaluate a PFD and the OZ concept display in terms of how effectively each display provides pilots with EM SA when observing display screenshots of an airplane operating at high altitude. This study will also examine the potential utility of those displays in providing specific forms of EM SA such as thrust required / thrust available curve information. An empirical study utilizing instrument panel screen shots requiring participants to make energy management decisions will comparatively measure the EM SA demonstrated from a PFD and the OZ concept display. Pilot EM SA will be assessed using the validated SA measure Situation Present Assessment Method (SPAM) to evaluate differences when using either a PFD or an OZ concept display.

Research Question

Will an ecological interface design, such as the OZ concept cockpit display, provide increased EM SA compared to a PFD, in terms of greater response accuracy and quicker response times, when instrument-rated pilots are presented with high altitude EM decisions?

Hypotheses

Research 1) Participants presented with high altitude EM decisions will exhibit a difference in response accuracy and response times when using the OZ concept cockpit display compared to when using a conventional electronic flight display.

Null 1) Participants presented with high altitude EM decisions will exhibit no difference in response accuracy and response time when using the OZ concept cockpit display compared to when using a conventional electronic flight display.

Research 2) Pilot flight hours will be correlated with speed and accuracy of EM decisions when presented with a conventional electronic flight display.

Null 2) Pilot flight hours will not be correlated with speed and accuracy of EM decisions when presented with a conventional electronic flight display.

Research 3) Pilot flight hours will not be correlated with speed and accuracy of EM decisions when presented with the OZ concept cockpit display.

Null 3) Pilot flight hours will be correlated with speed and accuracy of EM decisions when presented with the OZ concept cockpit display.

Delimitations

This study recruited FAA instrument-rated aviation students, faculty, and staff from a small private Midwest university, for a within-subjects design to evaluate the EM SA differences between the OZ concept display and a conventional electronic display. This study will compare the OZ display for a Cessna-172 to a regional jet PFD for a CRJ-700; both aircraft are modeled in Microsoft Flight Simulator X. The experiment is a simulation of instrument flying with no incidental out-the-window cues, as may be present when flying in IMC.

Limitations and Assumptions

The use of the term "conventional" in this report will refer to analog and electronic versions of cockpit displays presently used in GA, commercial, and military aircraft. The use of the term "conventional analog" will refer to round dial (RD) gauges, and "conventional electronic" will refer to digital PFDs.

The participants, recruited from the single, Midwest aviation university, will signify a convenience sample but will generalize to the larger population of FAA instrument-rated pilots. The within-subjects, repeated-measures design reduces the effect of individual differences. It is assumed that any measured EM SA differences between the displays in the proposed study will yield similar differences in other FAA instrument-rated pilots under the same scenarios.

While it is possible that mere perception, comprehension, and projection of EM SA in certain participants will not elicit the additional physical motor skills required during corrective action in an airplane, it is assumed that increased EM SA indicates an opportunity for pilot intervention and corrective action. Additionally, it is assumed that measured increases in EM SA through an improved cockpit display design would translate into fewer EM incidents in practice.

Proprioceptive environmental cues to include visual, aural, tactile / haptic cues, both from within and outside the cockpit, would equally affect a pilot's EM SA for either type of display. Although most air carriers have two pilots who influence each other's EM SA through their crew resource management (CRM), the experiment assumes that the individual EM SA effect would be the same for either type of display. Pilots would physically apply the thrust level and pitch control inputs consistent with what they indicated in their responses to interpreting the displays.

Definition of Terms

Aerodynamic Ceiling	Altitude where high speed MACH pre-stall buffet and low speed IAS pre-stall buffet meet, commonly referred to as the *coffin corner* (FAA, 2013).
Airplane Upset	An airplane in flight unintentionally exceeding the parameters normally experienced in line operations or training: pitch attitude >25° nose up or >10° nose down, bank angle >45°, or within the above parameters but flying at airspeeds inappropriate for the conditions (FAA, 2015a).

Angle of Attack (AOA)	Angle between the relative wind and the chord of the airfoil (FAA, 2008b).
Coffin Corner or Q-Corner	Term used to describe operations at high altitudes where low indicated airspeed (IAS) yield high true airspeed (TAS) (as indicated by Mach number) at high angles of attack. The high AOA results in flow separation that causes buffet. Turning maneuvers at these altitudes increase the AOA and result in stability deterioration with a decrease in control effectiveness. The relationship of stall speed to the critical Mach (Mcr) narrows to a point where sudden increases in AOA, roll rates, and / or disturbances (e.g., clear air turbulence) cause the limits of the flight envelope to be exceeded. Coffin corner exists in the upper portion of the flight envelope for a given gross weight and G-force (FAA, 2013).
Electronic Flight Display	The modern cockpit display featuring cathode ray tube (CRT), liquid crystal display (LCD), or light-emitting diode (LED) electronic digital flight instruments, a technological advance past mechanical gauges (Garland et al., 1999).

Energy Management (EM)	Process where pilots safely and efficiently manipulate thermal energy (fuel) into potential energy (altitude) and kinetic energy (airspeed) (Merkt, 2013).
Energy Management Situation Awareness (EM SA)	Degree to which a pilot has an accurate awareness on the energy state of their airplane.
Flight Envelope	The region of flight parameters surrounded by load factor ("G") limit and airspeed structural limit. If an aircraft is pushed outside of this region then it is considered to be operating "outside the envelope" (Dole & Lewis, 2000).
High Altitude	Operations above flight level 250 (FL250) or 25,000 feet mean sea level (MSL) (Carbaugh, Rockliff, & Vandel, 2008).
Indicated Airspeed (IAS)	Direct instrument reading of airspeed uncorrected for atmospheric density, installation error, or instrument error (FAA, 2008b).
Induced Drag	Portion of drag that is the inevitable consequence of the production of lift; varies indirectly with airspeed (FAA, 2008b).
L/D max	Lowest point on the total drag curve. The speed slower than L/D max is known as slow flight, "back

	side of thrust / power curve", or "region of reverse command". Speed faster is known as normal flight or "front side of thrust / power curve" (Carbaugh, Rockliff, & Vandel, 2008).
MACH number	Ratio of the true airspeed to the speed of sound (FAA, 2008b).
Maximum Altitude	Altitude that the airplane is either thrust-limited: no longer able to provide any rate of climb; buffet-limited: where 1.3 g loading due to turning, maneuvering, or turbulence would result in pre-stall buffet; or structural-limited: pressurization load limits on airframe (FAA, 2008b).
Optimum Altitude	Altitude that a given thrust setting results in the corresponding maximum range / minimum fuel burn speed. This altitude is not constant but increases with a decrease in temperature, a reduction in weight of the airplane, or a reduction in speed of the airplane (FAA, 2008a).
Parasitic Drag	Portion of drag created by the shape of the airplane and is not associated with the production of lift; varies directly with airspeed (FAA, 2008b).
Power Available (PA)	Maximum power an engine of a propeller / rotor aircraft can produce (Dole & Lewis, 2000).

Power Curve	Depiction of power and airspeed, normally a "J" shaped curve, applicable to propeller or rotor driven aircraft (Dole & Lewis, 2000).
Power Required (PR)	Amount of power equal to total parasitic and induced drag (Dole & Lewis, 2000).
Service Ceiling	Altitude that produces a rate of climb of 100 feet per minute (fpm; Hurt, 1965).
Situation Awareness (SA)	The perception of elements in the environment within a volume of time and space (Level 1 SA), the comprehension of their meaning (Level 2 SA), and the projection of their status in the near future (Level 3 SA) (Endsley, 1995, p. 36).
"Six-Pack" Gauges	The six flight instruments chosen by the British Royal Air Force in 1937 to be the standard cockpit set-up: altimeter, airspeed indicator, attitude indicator, turn and bank indicator, vertical speed indicator, and directional gyro (Williamson, 1937).
Stall	Sudden reduction in lift occurring at the critical angle of attack when airflow separates from wing surface (FAA, 2008b).
Thrust Available (TA)	Maximum thrust an engine on a turbojet aircraft can produce (Dole & Lewis, 2000).

Thrust Curve	Depiction of thrust and airspeed, normally a "U" shaped curve, applicable to thrust producing turbojet aircraft (Carbaugh, Rockliff, & Vandel, 2008).
Thrust Required (TR)	Amount of thrust equal to total parasitic and induced drag on a turbojet aircraft (Dole & Lewis, 2000).
True Airspeed (TAS)	Value for indicated airspeed when corrected for air compressibility, air density, and position error (FAA, 2008b).

List of Acronyms

ANOVA	Analysis of Variance
AH	Abstraction Hierarchy
AOA	Angle of Attack
ATC	Air Traffic Control
ATP	Airline Transport Pilot
CAST	Commercial Aviation Safety Team
CICTT	CAST / ICAO Common Taxonomy Team
CFII	Certified Flight Instructor Instrument
CFIS	Controlled Flight into Stall
CRT	Cathode Ray Tube
DOT	Department of Transportation
EFIS	Electronic Flight Information System

EGPWS	Enhanced Ground Proximity Warning System
EICAS	Engine Indicating and Crew Alerting System
EID	Ecological Interface Design
EM	Energy Management
EM SA	Energy Management Situation Awareness
EPR	Engine Pressure Ratio
ERAU	Embry-Riddle Aeronautical University
FAA	Federal Aviation Administration
FFS	Full Flight Simulator
FLXX0	Flight Level XX Thousand Feet
FMS	Flight Management System
g	Gravitational Force or Load
GA	General Aviation
IAS	Indicated Airspeed
ICAO	International Civil Aviation Organization
IHMC	Florida Institute of Human & Machine Cognition
IRB	Institutional Review Board
ISA	International Standard Atmosphere
LCD	Liquid Crystal Display
L/D	Lift over Drag
LED	Light-emitting Diode
LOC-I	Loss of Control Inflight
LOFT	Line Oriented Flight Training

Mcr	Critical Mach
MSFS	Microsoft Flight Simulator
MSL	Mean Sea Level
NTSB	National Transportation Safety Board
OA	Optimum Altitude
PA	Power Available
PFD	Primary Flight Display
PFI	Primary Flight Information
PFOV	Primary Field of View
PR	Power Required
RD	Round Dial Display
RMS	Root Mean Square
SA	Situation Awareness
SME	Subject Matter Expert
SPAM	Situation Present Assessment Method
SRK	Skill-Rule-Knowledge
SVS	Synthetic Vision System
TA	Thrust Available
TAS	True Airspeed
TR	Thrust Required
UCD	User Centered Design

CHAPTER II

REVIEW OF THE RELEVANT LITERATURE

This literature review details industry collective knowledge on the aeronautical hazard of airplane upset and loss of control inflight, and through the examination of National Transportation Safety Board (NTSB) reports, FAA advisory circulars, ICAO training manuals, and scholarly journals, directs the reparative focus to the potential contributions of improved energy management situation awareness. In addition to strengthening aerodynamic education, pilot muscle memory, and stall recovery procedural training, the prominence of EM on the OZ display provides an opportunity for research on its proposed benefits to pre-upset recognition. While the literature and industry knowledge of this topic are based on conventional displays, visual stall aids, and angle of attack indications, the contributions of a thrust curve energy management depiction on the OZ display have been briefly yet successfully studied. This initial OZ simulator testing proved the utility of the system to provide improved adherence to flight performance parameters during elementary flight maneuvers, despite distractions of mental tasks and simulated turbulence, on a light, single engine, low altitude plane. The gap uncovered is what OZ and its EM functional interface can contribute to a high altitude airplane upset scenario.

Airplane Upset Recovery Training Aid

An international industry working group with members from Boeing, Airbus, Flight Safety Foundation, U.S. Department of Transportation (DOT), FAA, and NTSB was established in 1998 to address an increasing rate of high altitude upsets per number of flight operations (Carbaugh et al., 2008; FAA, 2008). High altitude upsets are defined as those occurring at or above 25,000 feet mean sea level (MSL). The Training Aid was updated in 2004 and again in 2008 with the most recent information and recommendations.

> The goal of this group was to educate pilots so they have the knowledge and skill to adequately operate their airplanes and prevent upsets in a high altitude environment. This should include the ability to recognize and prevent an impending high altitude problem and increase the likelihood of a successful recovery from a high altitude upset situation, should it occur. (FAA, 2008a, p. 1)

The training aid supplement informed pilots of high altitude aerodynamic concepts such as optimum altitude (OA), the cruise altitude for minimum fuel consumption (minimum cost), and the fact that increased international standard atmosphere (ISA) temperatures will lower OA.

The report also cited the incomplete scenarios that, until as recently as 2009, were used for the training of aerodynamic stall recognition and recovery. Most flight simulator training programs practice stalls at mid altitudes like 15,000 feet MSL, conveniently weaving them into the rest of their simulator training profiles. The shortfall of this training method is that pilots would not have received practical flight training for stalls in the other regimes where commercial airplanes spend most of their flight time, particularly at typical cruising altitudes ranging from 35,000 to 39,000 feet MSL, with much different ISA temperatures resulting in different engine performance.

Precursors to Loss of Control. High altitude aircraft upset and aerodynamic stalls are major precursors to loss of control inflight (LOC-I) accidents (Brooks et al., 2014), which remain a major proportion of all transport airplane accidents. Lambregts, Nesemeier, Wilborn, and Newman (2008) examined 74 upset accidents between 1993 and 2007 from worldwide scheduled transport and commuter category airplanes and reported that 27 stalls (36%) were responsible for 848 fatalities (26%). Six stalls were induced by the autopilot's attempt to maintain a vertical speed or a selected altitude in a thrust-limited scenario, three of which occurred above FL250. Notably, only seven stalls (16%) occurred from a grouping of 43 LOC upset accidents between 1958 and 1993 found during their research, as compared with the 27 stalls (36%) in the 15 years studied between 1993 and 2007 (Lambregts et al., 2008).

Loss of Control Inflight

At the 2014 LOC-I Symposium, ICAO stated reducing LOC-I was a "global aviation safety priority" (ICAO, 2014b, para. 1), as evidenced in Figure 1. In addition, the NTSB made preventing LOC-I part of its Top Ten Most Wanted list for 2015, claiming over 40% of fixed wing general aviation (GA) fatal accidents between 2001 and 2011 were attributed to LOC-I (NTSB, 2015). LOC-I remained on the NTSB Most Wanted List for 2016, which reported it as increasing to 47% of fatal fixed-wing GA accidents between 2008-2014 in the U.S., resulting in 1,210 fatalities (NTSB, 2016).

According to the U.S. GA Joint Steering Committee, during the period 2001–2010, LOC-I was the top category for fatal accidents in GA (Brooks et al., 2014), resulting in a greater number of fatalities than the next five categories combined.

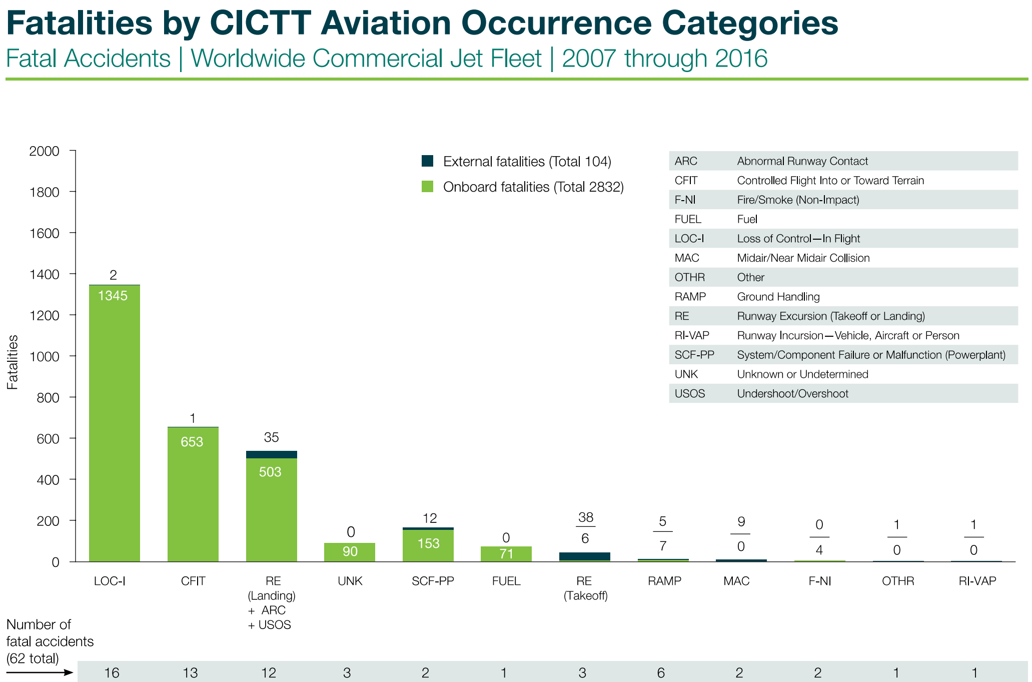

Figure 1. Fatalities by CICTT Aviation Occurrence Categories, Worldwide Commercial Jet Fleet, 2007-2016. Adapted from "Statistical Summary of Commercial Jet Airplane Accidents, Worldwide Operations, 1959-2016" by Aviation Safety, Boeing Commercial Airplanes, 2017, www.boeing.com/news/techissues/pdf/statsum.pdf, p. 22. Copyright 2017 by Boeing.

Lambregts et al. (2008) reviewed 74 LOC-I accidents from 1993 to 2007 involving 42 hull losses and 3,241 fatalities and categorized them five ways: aerodynamic stall, flight control system malfunctions, spatial disorientation (confusion) of the crew, contaminated airfoil (icing), and atmospheric disturbance (e.g., weather, turbulence, wind, temperature). Belcastro and Jacobson (2010) referred to LOC-I *off-nominal conditions*, categorized as adverse conditions occurring onboard the vehicle, external hazards and disturbances, and abnormal flight conditions.

Belcastro and Foster (2010) reviewed 126 Part-121 LOC-I accident reports from 1979 to 2009 using data from the Aviation Safety Network and NTSB websites.

Inappropriate crew response occurred in 42.8% of accidents, and stall / departure from controlled flight occurred in 38.9% of accidents. A three-dimensional scatterplot of worst-case combinations of LOC causal factors, shown in Figure 2, revealed stall / departure from controlled flight in conjunction with inappropriate crew response to be the most frequent and most dangerous combination, totaling 778 fatalities.

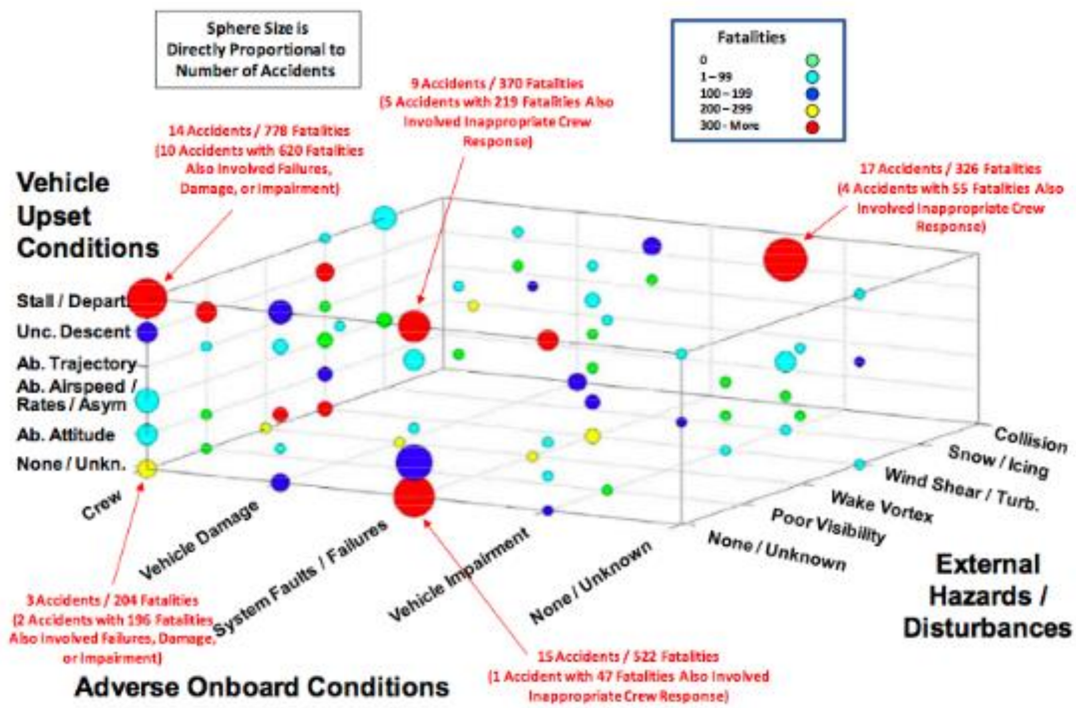

Figure 2. Identification of Overlap in LOC Causal and Contributing Factor Combinations, 1979-2009. Adapted from "Aircraft Loss-of-Control Accident Analysis" by C. Belcastro & J. Foster, 2010, p. 5.

Examples of LOC-I Accidents

The following LOC-I accident summaries were extracted from the NTSB Aviation Accident Database, all involving aircraft speed decay at altitude followed by aerodynamic stall. Any non-reported near-miss incident—where corrective pilot intervention prevents an unsafe event—would most likely not be quantified or captured in any safety databases. Some additional accidents that occurred between 1994 and 2013 where LOC-I was the causal factor are listed in Table 1 (ASN, 2015; NTSB, 2015).

- On April 29, 1993, a Continental Express Embraer-120 turboprop (FTW93MA143) experienced an aerodynamic stall and subsequent loss of control around 17,400 feet when the pilots selected an improper vertical mode on the autopilot (to climb at an angle too high and airspeed too slow) outside of normal parameters for the conditions, resulting in left engine propeller and cowling damage as well as substantial damage to the airplane fuselage during a forced landing.
- On June 4, 2002, a Spirit Airlines McDonnell Douglas MD-82 turbojet (CHI02IA151) experienced an aerodynamic stall and subsequent loss of control around flight level FL330 when the engine inlet probes became blocked with ice, resulting in a false engine pressure ratio (EPR) and subsequent retardation of the throttles by the auto-throttle system. As airspeed decayed, pitch attitude increased (corresponding AOA), and the airplane eventually stalled. No injuries or airplane damage occurred.
- On October 14, 2004, a Northwest Airlink Bombardier CL-600 (DCA05MA003) experienced an aerodynamic stall and subsequent loss of control around FL410

when both engines flamed out, resulting in a fatal crash of both pilots and hull loss of the airplane.

- On June 14, 2008, a FedEx Douglas MD-10 (DCA08FA075) experienced an aerodynamic stall and subsequent loss of control around FL330 when the airplane could not maintain pilot selected airspeed in a turn and later when buffeting occurred during an airframe slat overspeed at high altitude. Substantial damage to the elevators and stabilizer occurred.
- On February 12, 2009, a Colgan Air Dash-8 (DCA09MA027) experienced an aerodynamic stall during approach due to inappropriate pilot response after airspeed decay led to a stick shaker warning. The airplane crashed, killing 49 people onboard and one on the ground. This incident in particular led to Congressional legislation (PL 111-216, the Airline Safety and FAA Extension Act of 2010) and FAA regulation changes requiring, in particular, a minimum of 1,500 flight hours prior to joining the crew of a commercial airliner.
- On June 1, 2009, an Air France A-330 (DCA09RA052) experienced an aerodynamic stall when at high altitude (at FL350) their pitot tubes became blocked with ice and their automatic flight control systems disconnected due to the disagreement in air data. The aircrew failed to properly diagnose the attitude / thrust situation and made inappropriate flight control responses resulting in the stall and eventual crash, killing 228 people over the Atlantic Ocean.
- On July 6, 2013, an Asiana B-777 (DCA13MA120) flight crew failed to properly manage approach glidepath and inadequately monitored approach airspeed leading to the crash, total hull loss, and 3 ground fatalities.

Table 1

Selection of Additional LOC-I Accidents

Date	Aircraft identification with NTSB Ref. No.	Summary
1/7/94	United Express (Jetstream 4101 twin turboprop) DCA94MA027	Airspeed decayed to stall on approach, improper pilot response to stall warning, raised flaps
12/22/96	Airborne Express (DC-8) DCA97MA016	Improper pilot response during stall recovery at altitude
2/16/98	China Airlines (A300) DCA98WA044	Airspeed decay to stall during go-around
12/11/98	Thai Airways (A310) DCA99W021	Airspeed decay to stall during go-around
8/16/05	West Caribbean (MD-82) DCA05RA093	Improper pilot response during airspeed decay to stall at high altitude
8/22/06	Pulkovo (TU-154) N/A	High altitude turbulence caused AOA increase and airspeed decrease to stall
12/29/12	Lancair IV-P (single piston prop) WPR13FA076	Inadvertent aerodynamic stall / spin
8/9/13	Rockwell Commander 690B (twin turboprop) ERA13FA358	Inadvertent aerodynamic stall / spin
12/12/13	Piper PA-24 Commanche (single piston prop) CEN14FA084	Inadvertent aerodynamic stall / spin

Note. Adapted from NTSB Aviation Accident Database & Synopses query page, accessed on April 24, 2015, and ASN Aviation Safety Database query page, accessed on April 24, 2015.

Controlled flight into stall. A subset of these LOC-I accidents, called controlled flight into stall (CFIS), is when fully functional airplanes unintentionally decelerate through stall airspeed due to improper use or confusion with automated systems by the pilot. Sherry and Mauro (2014, p. D1-1) examined 19 LOC-I accidents all caused by a "complex sequence of automation behaviors" followed by the inability of the flight crew to intervene properly. Sherry and Mauro asserted that although many of the historic accident causes and safety issues have already been addressed through advances in technology, training procedures, and maintenance practices, still many LOC-I accidents occur from a range of various factors, which no common intervention strategy could universally mitigate.

Sherry and Mauro (2014) concluded that in each of these 19 CFIS accidents, the pilot decision-making required to properly identify a low airspeed condition and respond correctly was not supported by the available automation cues. The degree to which the automation was controlling the airplane in a manner not intended by the crew, the status of the sensors, and the degree to which aircraft systems were degraded was not obvious to the flight crew. Sociologist Charles B. Perrow (1984) accurately describes these types of automation complexity failures as situations that "both start a fire and simultaneously deactivate the fire alarm" (Sherry & Mauro, 2014, p. D1-8).

Stall corrective action. The FAA recommended corrective procedures for high altitude aerodynamic stall emphasize a reduction in AOA by establishing a nose down attitude and accelerating by descending, since an increase in thrust while maintaining altitude is not always an available option and could actually aggravate the stall (Carbaugh et al., 2008). Unfortunately, prior to the original FAA AC 120-109 published in 2012, industry training programs emphasized stall recovery standards to a zero loss of altitude, since that was the FAA Practical Test Standard. High altitude line oriented flight training

(LOFT), a type of simulator training that mimics typical “line” operations in a chronological flow of takeoff, departure, cruise, approach, and landing instead of artificial simulator sequences of emergency after compound emergency, was a change recommended by the industry to provide more effective, real-world, applied training, and included familiarization with high altitude slowdowns and approaches to stalls (Carbaugh et al., 2008). The FAA also suggested incorporating real-world startle and surprise elements into full-motion flight simulator scenarios, since a quick and correct response to stall conditions is paramount to successful recovery, even during potentially confusing circumstances (FAA, 2015b, p. 16).

Round Dial Cockpit Display

The round dial “six-pack” cockpit instrumentation containing airspeed indicator, attitude gyro, altimeter, turn and bank indicator, heading indicator, and vertical speed indicator, shown in Figure 3, and its modern replacement the PFD, shown in Figure 4, may not be the best way for pilots to make informed decisions about energy management (Temme et al., 2003). The analog cockpit display shown in Figure 3 has six or more independent gauges, and it would take considerable time to perceive and comprehend all the information contained on it (Hamilton, 2001).

Figure 3. An emulation of a round dial cockpit display. Adapted from “Principles for Human-Centered Interaction Design, Part 1: Performative Systems” by T. Eskridge et al., 2014, p. 89. Copyright 2014 by T. Eskridge et al.

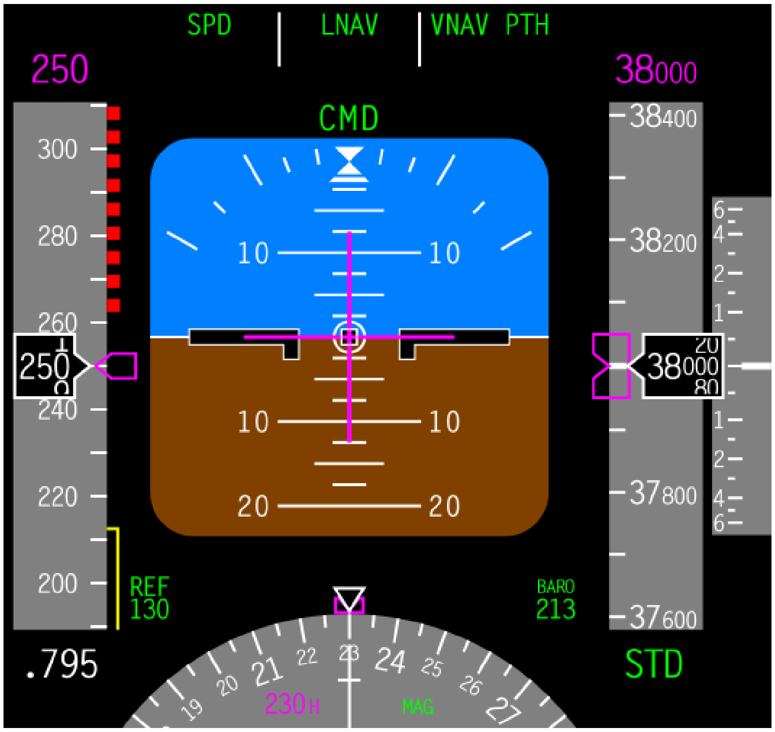

Figure 4. Modern electronic flight information system panel primary flight display. Adapted from Primary Flight Display, by Denelson83, retrieved from https://commons.wikimedia.org/wiki/File:Primary_Flight_Display.svg

Visual stall aids. Pilots are tasked with monitoring the automated flight control systems, but often their effectiveness is hampered by the complex logic embedded inside of the aircraft's computerized systems, through mode selections based on algorithms sometimes not fully understood. Instead of "a specific fix for only one of many potential one-of-a-kind problems," Sherry and Mauro recommend, "improving the detection and intervention strategies" (2014, p. D1-9). The FAA has recommended yellow bands placed on the airspeed tape indicator to depict the 1.3V_{stall} range in addition to the red bands already in wide usage on modern PFD airspeed tapes to depict the V_{stall} range itself (FAA Safety Recommendation A-10-011, 2013; Sherry & Mauro, 2014). The FAA also recommends redundant aural and visual warnings of impending low airspeed conditions (FAA A-10-012, 2015). Both of these FAA safety recommendations followed the 2009

Colgan Air crash in Buffalo, NY, in which inappropriate pilot response to stall warnings was deemed the cause (NTSB Aviation Accident Report 10/01, 2010).

Modern PFDs often contain colored speed bars on the airspeed tape indicator corresponding to the stall speed and / or stall buffer (normally 130% of the stall speed or 1.3*Vs), but these values change often due to their dependence on aircraft configuration, bank angle, g-loading, and ice accretion. The amber or yellow speed bands on a PFD, indicating the speed range margin for aerodynamic buffet, do not supply any indication of excess thrust (available versus required) to maintain current altitude and airspeed. Even armed with these visual color-coded airspeed indicators with aural alerts, accidents continue to occur through numerous singular and often difficult to duplicate faults that manifest between pilots and complex autopilot systems during particular situations (Sherry & Mauro, 2014, D1-2).

Current cockpit instrumentation displays airspeed, engine thrust, and altitude, as well as important environmental information such as air temperature, air density, ISA values, and true airspeed (TAS). Derived flight information such as optimum altitude and maximum or ceiling altitude can be crucial to decisions about energy management but may be hidden deep within layered digital display menus available in a flight management systems (FMS).

Thrust curve. In normal flight operations, pilots fly the aircraft inside the region of normal command, or the *front side* of the thrust / drag curve. The front side is defined by increases in thrust resulting in increases in aircraft airspeed. After any disturbance in this region, the aircraft airspeed will return to its original airspeed as long as thrust does not change (Carbaugh et al., 2008). The *back side* of the thrust / drag curve, or region of reverse command, is defined when a reduction in airspeed counter-intuitively produces an increase in induced drag that necessarily requires an increase of thrust in order to maintain level flight. Flight in this region is normally only purposely flown during final approach and landing and during slow cruise flight in training scenarios.

Stalls can occur at any altitude, any airspeed if the critical angle of attack is exceeded and even in a descent. Weather has considerable effects on the performance of an airplane. The jet stream upper air currents can submit decreasing velocity wind shear to an airplane, actually pushing it over to the backside of the thrust curve. Ice accumulation on the airframe that changes the airfoil shape may also lead to stalls at higher airspeed and lower angles of attack than expected. A visual indicator of an aircraft's current state on the thrust / drag curve would provide immediate SA for proper control of thrust and attitude. An improvement to the current PFD paradigm could provide pilots with critical energy management SA information to curb high altitude aircraft upsets. Other possible additions to the PFD that may benefit SA are angle of attack, a readout for optimum airspeed and altitude based on aircraft weight and outside air temperature, and any attention-grabbing, visually coded textual performance warnings.

Ecological Interface Design (EID)

Ecological interface design is a framework focused on designing human-computer interfaces for complex sociotechnical systems (Vicente, 2002, p. 63). Born out of research at the Riso National Laboratory in Denmark in the 1960s, the philosophy behind EID is to provide operators with specific work domain constraints, aiding some aspect of their SA, so that they can make informed decisions during complex scenarios (Vicente, 2002, p. 63). The functional strength of an EID lies as a monitor of complicated systems and relationships, and as such, are found in the medical and nuclear power industries where monitoring operators are faced with important and time-critical decisions, cannot afford to make errors, and require SA-enhancing equipment (Burns et al., 2008).

Sharp and Helmicki (1998) describe an EID interface to aid neonatal intensive care providers with pediatric tissue oxygenation assessment, while Burns et al. (2008) report that EIDs function as the display and interface for the turbine, condenser, and feedwater systems in nuclear power plants. Sharp and Helmicki (1998) found 15 of 16 physicians made more accurate diagnoses with the EID as compared to the traditional interface, and Burns et al. (2008) observed an improvement in SA using EID as compared to traditional displays used in nuclear power plant control rooms.

One theory supporting EID is the skill-rule-knowledge (SRK) taxonomy created by Rasmussen (1983): skill-based behavior is considered the most basic, muscle-memory type of automation for simple, learned tasks; rule-based behavior is an intermediate level where experience links signal cues to appropriate responses; and knowledge-based behavior is reserved for tasks requiring analysis, troubleshooting, and problem resolution.

The EID framework adheres to three design principles: operators act directly on the interface, the interface matches domain constraints to perceptual information, and the interface provides a mental model for problem solving (Vicente, 2002, p, 64). Modern human-machine systems aim to automate routine skill-based or rule-based activities, so EID is well suited for job environs where operators occasionally encounter non-routine situations that require difficult decision-making (Rasmussen, 1983). Sharp and Helmicki (1998, p. 354) describe an EID as providing “abstract meaning from the many raw streams of data" traditionally presented individually to the operator, often innocuous when comprehended singularly, but catastrophic in certain juxtapositions when considered holistically.

Knowledge-based behavior can be more successful when operators are equipped with an EID versus a conventional display (Vicente, 2002, p. 62). Through a specially designed graphical / conceptual interface, EID continues to promote skill- and rule-based behavior, conserving cognitive resources, but also enhances analytical and problem solving support through its perceptual and intuitive displays designed to tackle more challenging and error-prone knowledge-based behavior. Interestingly, Burns et al. (2008) consider that EID may not be as effective as traditional interfaces at supporting rule- and skill-based behavior since the higher-level presentations mask the information required for a known or learned situation.

A second theory supporting EID is abstraction hierarchy (AH), also developed by Rasmussen (1985). The AH consists of five levels listed from high to low: (a) functional purpose, (b) abstract function, (c) generalized function, (d) physical function, and (e) physical form. The AH describes the relationships between the system, operator objectives, and the methods available to achieve these objectives. In contrast to task analysis normally used for system design, EID, through the AH, employs work domain analysis. Whereas task analysis is limited to addressing only what is predicted or anticipated, work domain analysis is robust enough to address unforeseen events and non-routine situations. Vincente showed that EIDs support problem-solving in unanticipated situations (as cited in Burns et al., 2008) and suggests EIDs can improve SA when implicit procedures are not available or do not exist.

OZ Cockpit Concept Display

OZ is a novel cockpit display using basic principles in vision science and aerodynamics (Eskridge et al., 2014). With a goal of reducing or simplifying the instrument scan, OZ replaces the PFD with a graphic depiction of aircraft performance, as shown in Figure 5, from which the pilot would otherwise have to construct a mental model from multiple instrument gauges (Eskridge et al., 2014, p. 91). OZ reduces the traditionally detailed visual scan workload in flight to a nearly instantaneous perception of the entire flight system (Eskridge et al., 2014, p. 91).

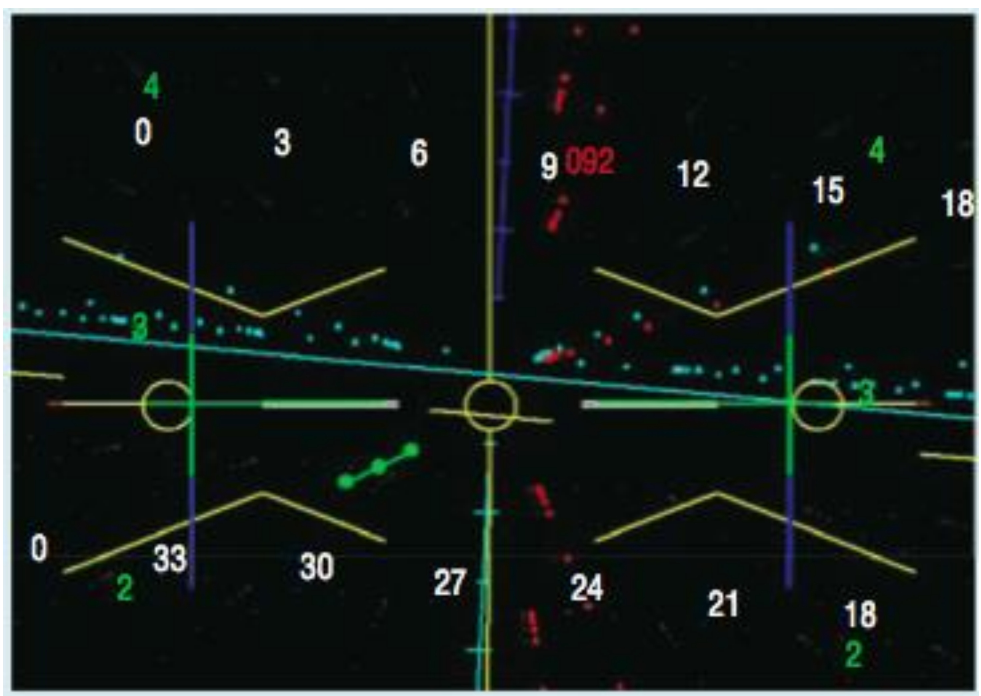

Figure 5. OZ display for fixed wing turboprop. Adapted from “Principles for Human-Centered Interaction Design, Part 1: Performative Systems,” by T. Eskridge et al., 2014, p. 90. Copyrght 2014 by T. Eskridge et al.

The OZ concept display provides intuitive energy management information and may mitigate LOC-I by displaying an airplane's current power setting in relation to its minimum allowable speed (stall), its maximum lift (lift / drag), and its maximum allowable speed (structural limits), as shown in Figure 6 (Gerold, 2003). An easily visible, readily apparent, and highly intuitive graphical power / drag curve provides EM SA to pilots.

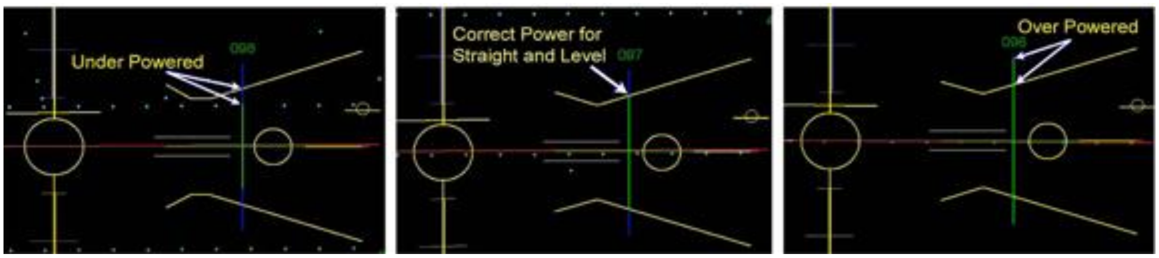

Figure 6. Turboprop power on an OZ concept display. Adapted from "OZ Human Centered Flight Displays," by D. Still & T. Eskridge, PPT slide 11, retrieved January 25, 2015 from http://www.imhc.us/groups/oz/.

The OZ display also uses a green line to depict the power applied and a blue line to depict power available. A vertical line on the horizontal axis represents the current airspeed. This depiction is designed to allow a pilot to quickly perceive the necessary power required to achieve a desired airspeed. The OZ concept display has the potential to reduce pilot workload and improve situational awareness, particularly in energy management (Albery, 2007, p. B189). The desired result of the OZ principle is that the pilot can simply look at the display to determine system status, required actions, and amount of correction, rather than having to update a mental model containing low-level data referenced from multiple gauges dispersed throughout the cockpit (Eskridge et al., 2014, p. 91).

Temme, Still, and Acromite (2003) initially conceived the OZ concept display from the standpoint of reducing the time it took for pilots to consult their instrument gauges. In addition, they tried to solve the dual quandaries that current display technology did not effectively convey aircraft parameter deviations to the pilot and that these same displays were confusing to interpret due to the multitude of different frames of reference and units of measure displayed. OZ, the name given during prototype-testing when a participant pilot was told to pay no attention to the programmer behind the rear projection screen, uses the additive effect of both vision channels, focal (central) and ambient (peripheral), to counteract the excessive time required to linearly perceive and comprehend aviation flight information from multiple displays.

OZ also employs a user-centered design (UCD) approach by attempting to reduce human informational processing and the overall cognitive workload of flight (Temme, Still, & Acromite, 2003). The aim is to produce a cockpit display that can be quickly comprehended by pilots in the complex aviation environment, increasing overall SA. Specifically, OZ combines and reduces the multitude of contrasting airframe and environmental data so that the pilot can directly perceive the aircraft's performance capabilities.

The structure of the OZ concept display consists of both a *starfield* and aircraft metaphor that act similar to what is known as synthetic vision, a digital graphical depiction in the cockpit of real-world topography and an almost bird's-eye view of the airplane and its surroundings, as opposed to the nose-of-the-airplane perspective provided in conventional displays (Temme, Still, & Acromite, 2003). The starfield metaphor, shown in Figure 7, provides aircraft attitude and geo-spatial information in the form of a vertical and horizontal grid representing the physical space around the aircraft composed of dots spaced every 10° of heading and every 500 feet of altitude.

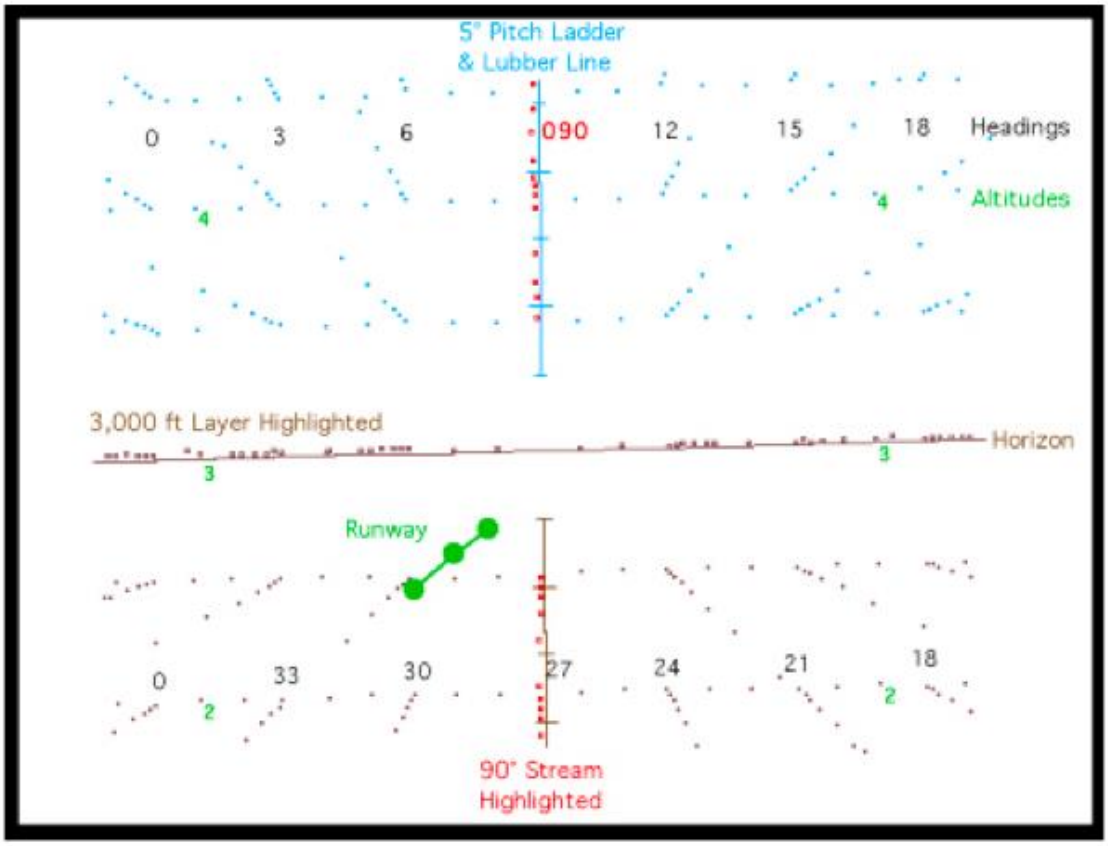

Figure 7. Starfield metaphor. Adapted from "OZ: A human-centered computing cockpit display," by L. A. Temme, D. L. Still, and M. Acromite, 2003, In *45th Annual Conference of the International Military Testing Association*, Pensacola, FL (pp. 70-90).

The aircraft metaphor, shown in Figure 8, provides performance and maneuvering capability information, as well as attitude and configuration, through lines and circles typeset in certain arrangements. The upper and lower bent wings provide the airspeed limits between overspeed (where the line ends on the outer edge) and stall (inner edge). The angle bend or *wing pinch* in each of the four wings corresponds to L/D max, or the maximum ratio of lift to drag speed. The angled wing bends also provide a useful tool for setting a perfect standard rate turn when rolling the airplane to align the angled line onto the horizon. The horizontal position of the vertical speed struts indicates airspeed and moves outboard toward the overspeed limit of the bent wings when airspeed increases. The green section of the vertical speed struts shows power applied, the total green and blue length of the vertical struts indicates power available, and un-accelerated flight power is achieved when the green section just touches the upper and lower bent wings. The two outboard circles indicate desired airspeed selected by the pilot.

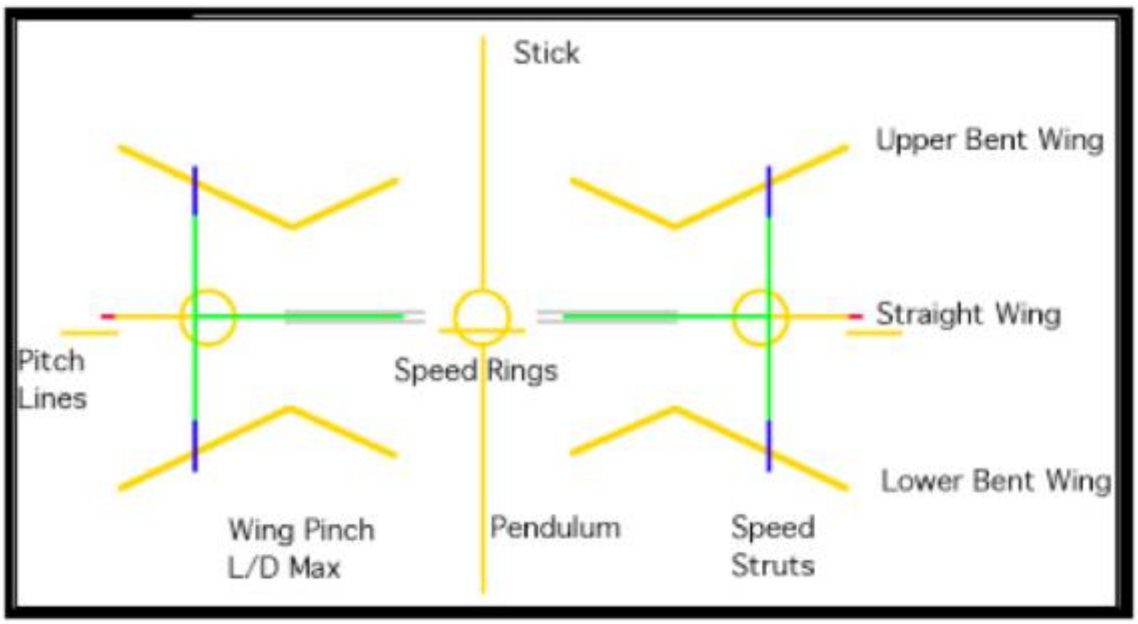

Figure 8. Aircraft metaphor. Adapted from "OZ: A human-centered computing cockpit display," by L. A. Temme, D. L. Still, and M. Acromite, 2003, In *45th Annual Conference of the International Military Testing Association*, Pensacola, FL (pp. 70-90).

The result of this radical design shift is an intuitive and direct visual representation on a single display of all the complex mental models of aerodynamic relationships and aircraft performance the pilot would otherwise have to construct and continuously update from various displays around the cockpit (Temme, Still, & Acromite, 2003). The operator of an OZ display is able to simultaneously acquire several different sources of cockpit information from only one display, thus greatly reducing the time required to scan multiple instruments and saving precious central focal vision for more important and immediate tasks such as forward visual scans or processing emergency procedures. Temme, Still, and Acromite (2003) claim the processed data provided by OZ does not provide misleading information or make decisions for the pilot. OZ reduces the human cognitive workload of piloting by instantaneously providing Levels 1 and 2 SA information (perception and comprehension).

In order to determine that an aircraft is approaching stall speed on a PFD airspeed indicator, a pilot must first remember the actual value of the aircraft's stall speed, next reference the airspeed indicator, and then integrate that data into his mental model in the context of the aircraft configuration (e.g., weight, landing gear position, flap position) and aircraft state (e.g., attitude, climb, descent) to determine the consequence of that relationship. In the OZ display, the perception and comprehension and projection of an aircraft's proximity to stall speed is as simple as visually referencing the power bars in relation to the inboard ends of the wings, as shown in Figure 6. No specific airspeed stall values need to be remembered or recalled for the given configuration, weight, or g-loading (Eskridge et al., 2014, p. 92).

Some Level 1 SA (perception) energy management information is present on most glass cockpit PFDs today, such as upper and lower airspeed limits as shown in Figure 4. Acquiring all facets of relevant EM information in a commercial transport or business jet requires a time-consuming scan among multiple gauges located in different places around the cockpit (e.g., instrument panel PFD and multi-function display, center console / pedestal control display unit and FMS). This increases the chance for errors during perception or comprehension. With the OZ concept display, different types of information are merged so they can be processed simultaneously rather than serially during an instrument scan. The guiding principle behind the OZ concept display is simply to reduce cognitive workload (Eskridge et al., 2014, p. 92).

While PFDs provide only digitalized versions of the same legacy instruments aviation has continued to use, EID displays such as the OZ concept display, present both physical and functional information through intuitive graphical features promoting Endsley's (1995) triad of SA: perception, comprehension, and projection of future states. Similarly, work from Borst et al. (2008, p. 159) analyzed the Synthetic Vision System (SVS) and the Enhanced Ground Proximity Warning System (EGPWS), two advanced cockpit display technologies meant to enhance pilot terrain and obstacle SA. Borst et al. claim that SVS supports perception (Level 1 SA) but fails to provide comprehension or projection (Levels 2 & 3 SA). EGPWS conversely provides pilots with terrain and obstacle information but without "underlying data and rationale" (Borst et al., 2008, p. 159). Borst et al. found in their research that comprehension and perception could be improved by additionally providing pilots with aircraft performance characteristics and flight conditions, effectively establishing constraints for their forthcoming complex decision-making. Borst's team recommended an EID display for terrain awareness during total engine failure of a GA aircraft since the display supported all three levels of SA. Their results showed better terrain awareness with the EID in terms of achieving pilot goals within aircraft performance and environmental constraints, but also reported that the pilots flew riskier yet still successful flight profiles with this enhanced information, negatively affecting their margin of safety.

OZ experimentation. Temme, Still, and Acromite (2003) performed a series of exploratory experiments in a desktop flight simulator to analyze the differences in capabilities of the OZ concept display compared to independent, analog, round-dial displays. This initial OZ testing consisted of five incremental experiments to determine if pilot performance, mainly altitude and heading adherence, was improved by the OZ concept display as compared with an analog round dial display found in many GA airplanes. Training before the trials was minimal and restricted to answering questions about the instruments but did not explain how to use the instrumentation to minimize heading and altitude errors (Temme, Still, & Acromite, 2003, p. 86). Performance scores were referenced to root mean square (RMS) values obtained from altitude and heading deviations from the desired target values. Temme, Still, and Acromite (2003) chose RMS because it incorporates variability of the individual score and also its position relative to a target value instead of relative to an average score.

Study 1 compared two non-pilot participants on both an OZ and an analog round dial display (RD) flying a Cessna 172 at constant airspeed, heading, and altitude for about three minutes per trial. The counterbalanced trials consisted of four levels of turbulence, from nil to severe. Participants would switch from about three minutes of one display type to about three minutes of the other display type for about one to two hours. Results indicated that the participants using the OZ display performed better in terms of altitude and heading adherence than when they used the RD. Also, no considerable differences in performance were found using the OZ display for the varying degrees of turbulence, while in the RD trials, heading and altitude adherence became progressively worse when turbulence increased. The non-flight experienced and novice participants learned the OZ display quickly since their performance scores did not improve much during the course of the trials (like the RD performance improved over time) but remained considerably better than the RD scores from the beginning. These results suggest the OZ display is easy to learn (performance remained consistently superior) as compared to the RD (where performance continued to improve over time) and also enhances superior pilot performance compared to analog gauges, when operated in varying degrees of turbulence (Temme, Still, & Acromite, 2003, p. 90).

Study 2 used the same two non-pilot participants maintaining the same altitude and heading assignment using both the OZ and RD displays but also included an OZ display trial where words appeared in the center of the display at a rate of one word per second, and participants would have to read them aloud. This experiment was to test the ability of OZ to allow for focal tasks irrespective of the ambient channel control dedicated to flying the aircraft using the OZ symbology. The results indicated the participants performed equally as superior on the unaltered OZ display as they did on the word-task OZ display, when compared to the RD.

Study 3 used four different non-pilot participants to analyze the performance differences when obscuring varying degrees of the OZ display from 0% to 80% in 20% increments. Again, these new participants performed considerably better on altitude and heading adherence using OZ as compared to the RD, regardless of turbulence intensity. In addition, the unaltered OZ and the obscured OZ both fared equally well in terms of heading adherence in all levels of turbulence, indicating large portions of the OZ display are not required to maintain flight performance within set tolerances. The amount of obscuration and turbulence level did affect performance in the altitude maintaining task, prompting a design change of inserting altitude pitch cues on the left and right outer borders of the screen instead of just on the center vertical axis, where the obscuration blocked them from the participants.

Study 4 compared the performance of four trained pilots to that of three non-pilot volunteers who had not already been in any of the experiments. The results indicated that the trained pilots performed equally well using OZ or the RD, regardless of the turbulence level. The non-pilots performed better using the OZ over the RD, and the effect of turbulence on their performance was less with OZ than with RD. Overall, pilots performed better than non-pilots, turbulence impacted non-pilots more than pilots, and the difference in performance between OZ and RD was greatest for the non-pilots. The performance of the non-pilots was equivalent to the previous three experiments, which all used different groups of non-pilots.

Study 5 used the same four pilots who were in Experiment 4 but made the task considerably more difficult. The participants now had to maintain a specified airspeed with active throttle control as well as maintain specified headings and altitudes, while flying slower than the minimum drag speed (L/D max) in the area of reverse command. The results showed that OZ performance was better than RD in both altitude and heading adherence, in all levels of turbulence. Mean altitude error remained consistent on the OZ display, even as turbulence increased. Temme, Still, and Acromite (2003) expressed their view that since the experienced pilots who have flown RD in the real airplane performed better with OZ during the experiment, the OZ is not only easier to operate but also easier to learn.

The stated primary goal of Temme, Still, and Acromite (2003) was to determine if the OZ concept display could reduce the time and cognitive effort required to maintain an accurate and timely instrument scan since an OZ display could provide the pilot with the same information normally retrieved from multiple gauges throughout the cockpit. Temme, Still, and Acromite (2003) noted that it is faster for pilots to read and interpret information from the OZ graphical display versus a RD, which presents information primarily with alphanumerics. The human visual system takes longer to process and integrate information from the dials, gauges, arrows, and pointers of a RD (Temme, Still & Acromite, 2003, p. 75). Temme, Still, and Acromite (2003) were hopeful that their display may eventually integrate every piece of information a pilot needs to properly fly an aircraft, noting that it supports situation awareness and spatial orientation particularly well.

Further experimentation with OZ. Recognizing the capability of a functional display to improve pilot task accomplishment and aircraft state awareness through a presentation of complex, high level information, Smith (2008) performed a series of experiments exploring the benefits of the OZ display over RDs. In particular, OZ uses graphics to provide a visual representation of the functional relationships between thrust, drag, and airspeed on the lift / drag curve.

In three incremental studies, Smith et al. (2004), Smith et al. (2005), and Smith and Boehm-Davis (2005) compared both novice and experienced pilot performance between the OZ display and an analog RD Cessna 172 display when pilots were tasked to perform flight maneuvers, secondary tasks, and turbulent conditions. The results of these studies suggested that a visual graphic depiction of the power curve, like on the OZ display, improves the ability of pilots to directly manage energy (energy management or EM) in the airplane, rather than relying on analog RD gauges. RDs alternatively require pilots to extract information from various separate instruments, and then process, comprehend, and finally act on that information to make the proper power level and flight control manipulations. Smith (2008) makes two assertions from this research: OZ improves and standardizes energy management, and OZ improves the understanding and knowledge of energy management. This research was approached from the perspective of improving pilot training in order to reduce the number of hours required for proficiency. Smith (2008) purports that flight training using an OZ display can teach energy management more effectively, and in addition, positively transfer that increased knowledge to a RD display. He viewed the OZ as a training device only and recognized it is not the standard PFD pilots find on commercial, military, and GA aircraft.

Using a repeated-measures, mixed design study, Smith (2008) reported that after a comparison of novice flight participants randomly assigned to four treatments of flight training followed by flight performance (OZ-OZ, OZ-RD, RD-OZ, RD-RD), those trained on the OZ display showed considerably more accurate knowledge and performance for EM than those trained on RDs. Participants using OZ exclusively for both training and performance sections showed considerably more accurate flight performance. Smith (2008) concluded that EM knowledge transfer did occur from OZ to RD, and that performance improvements (OZ better than RD) were consistent in power, airspeed, and pitch. Smith's (2008) approach was to identify display technology that will improve cockpit design in the future.

The first of Smith et al.'s (2004) three experiments involved experienced regional airline pilots who performed a series of 11 maneuvers (climbing / descending, banking turns, airspeed acceleration / deceleration) on the OZ display and a RD for a Cessna 172 PC emulator. Smith et al. (2004) found that power settings during maneuvers were managed more consistently with OZ, and there were considerably fewer differences in parameter deviations between participants. Pilots using a RD seemed to rely on *heuristics*, a way of managing an analytical task without analytical behavior, in that they used a rudimentary, binary, one-or-the-other approach to energy management and instead controlled airspeed or altitude with pitch alone, thus devolving the flight performance problem into something considerably easier. Overall, pilots using a RD applied more power than necessary for the maneuvers. His explanation was that the functional OZ display gave pilots perceptual cues for power settings, changing this task from an analytical problem-solving one to a substantially easier skill-based one. Smith et al. (2004) references the Skill-Rule-Knowledge framework of behaviors Rasmussen (1983) developed, listed from most basic to most advanced: fastest, skill-based perceptual-motor behavior using environmental cues; rule-based behavior associating perceptual signs with appropriate goals; and slowest, knowledge-based behavior requiring analytical problem-solving.

In a second experiment, Smith et al. (2005) used general aviation pilot participants on similar maneuvers but also incorporated turbulence and a secondary "distracting" task of locating an object's position in relation to the aircraft's position using another display in the cockpit (e.g., a radar screen or TCAS display). This experiment included slow flight (near Vs) maneuvers in order to capture more complex energy management scenarios. The results showed a decrease in accuracy of RD flight parameter performance with a corresponding increase in difficulty on the secondary tasks of turbulence mitigation or object location, while OZ participants showed no decrease in flight performance with added task workload. Also, Smith et al. (2005) reported that by simplifying the tasks of proper energy management, pilots had available cognitive capacity to perform the additional scenario tasks of locating the object and mitigating the turbulence. Smith et al. (2005) recommended further research comparing the OZ concept display to modern glass cockpit displays since EFIS cockpits are more effective at integrating flight information into the digital PFD versus RD, which have multiple gauges providing individual presentations of information.

Smith and Boehm-Davis (2005) performed a third experiment to evaluate the ability of functional OZ displays to better train pilots. The method followed the same 11 maneuvers as the first experiment, except with novice participants. Results showed novices using OZ were more accurate and consistent in their flight performance than those using RD. Also, the performance of the novices was similar to the first experiment when experienced regional pilots flew the same maneuvers. Since flight experience seemed not to be a factor, Smith and Boehm-Davis proposed that the direct perceptual graphics of OZ do make energy management easier as compared with the cognitive requirements using a RD.

Expanding on Previous OZ Studies

While the studies by Smith (2008), Smith and Boehm-Davis (2005), Smith et al.

(2005), and Smith et al. (2004) focused on measuring manual pilot flight performance using OZ versus RD in a light single engine aircraft, this proposed research study would examine the ability of pilots to extract energy management information from the displays themselves. Whereas Smith et al. (2005) studied if pilots flew more precisely in terms of altitude, airspeed, and heading control with OZ opposed to RD, this research studies whether pilots have enhanced overall EM SA. In addition, this research will examine pilots EM SA when flying a twin-engine jet aircraft at high altitudes at or above Flight Level 250, as opposed to the earlier studies which used a Cessna 172 operating at low altitudes. The different performance characteristics of the dissimilar aircraft, as well as the very different flight atmospheric environment, will expand the body of knowledge in regards to functional displays, human cognition, and EM SA.

Situation Awareness (SA)

Aviation performance of pilots can be evaluated through assessment of their situation awareness: the degree to which a pilot is cognizant of what is going on, in, and around the aircraft. In her seminal paper, Endsley (1995, p. 36) defined SA as “the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future.” This definition implies three levels of SA: perception (Level 1 SA), comprehension (Level 2 SA), and projection (Level 3 SA); see Figure 9. Endsley, Bolstad, Jones, and Riley (2003, p. 268) further define SA as the “internalized mental model of the current state of the operator’s environment.”

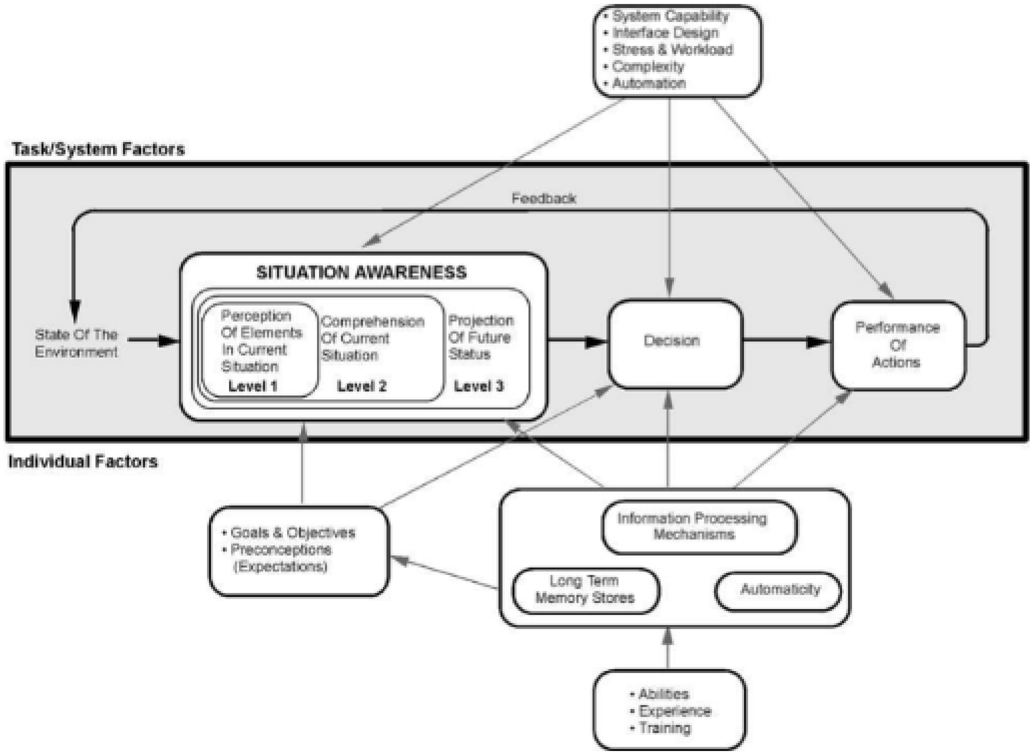

Figure 9. Model of situation awareness. Adapted from "Toward a Theory of Situation Awareness in Dynamic Systems," by M. R. Endsley, 1995, *Human Factors 37,* 32-64. Copyright 1995 by M. R. Endsley.

Carol (1992) describes pilot SA as continuous perception, the ability to forecast, and the execution of tasks based on those details. SA can also be thought of as three-dimensional problem solving, complicated by time compression, encompassing an individual's experience and capabilities. Tenney, Adams, Pew, Huggins, and Rogers (1992) explain that SA factors heavily in overall piloting performance with the increasing automation in flight cockpits and subsequent increased dependence on human cognition over traditional piloting stick and rudder physical motor skills.

Since at present analog and electronic cockpit instruments typically present only Level 1 SA (perception) information, the pilot is left to rely on working memory, mental models, and experience in order to tackle many complex goals involved with modern-day aviation. Endsley and Jones (2012) describe Level 2 SA (comprehension) information, as the “synthesis of disjointed Level 1 elements and a comparison of that to one’s goals” (p. 16). By placing Level 2 SA elements onto the PFD directly, which an EID like the OZ display does, the pilots would gain increased SA almost immediately and without a corresponding increase in cognitive workload. This heightened SA would give pilots an understanding of what the data and instrumentation mean in regards to the goal of maintaining the aircraft within safe and stable performance parameters, instead of relying on the pilot to deduce a correspondingly corrective flight control response. This design gives the pilots the information they need to make better decisions about an aircraft flight profile and aims to reduce the human errors that do exist.

Level 3 SA (projection) information is more critical to this system design than Level 2 SA since high altitude upsets usually result from a pilot not knowing that the current (benign) configuration will result in an impending undesirable aerodynamic event. Alternatively, the presentation of Level 2 SA information with trend data might suffice for Level 3 SA information. Airplane upsets might occur less frequently if pilots are provided greater comprehension of current performance parameters, but more importantly, projection of future states of such parameters can be more beneficial to pilots than even outside corrective action.

Goal versus data processing. When new display systems are designed and installed into aircraft, there are limitations to how many instrument indications a pilot can mentally integrate and also to the physical space remaining in the cockpit. Due to the physical limitations of the PFD, the information presented must be prioritized, but there are obvious disadvantages to only displaying data and thus encouraging data-driving processing, which is highly taxing cognitively (Endsley & Jones, 2012). Pilots must mentally process the data and then make their own Level 2 and 3 SA decisions. In goal-driven processing, typically more intuitive and less cognitively taxing, pilots can use the Level 1, 2, and 3 SA information and spend most of their cognitive resources on higher-order goals instead. However, displaying only goal-driven processes would deprive the pilot of the "raw data" required to make complex decisions. Pilots would have to trust the aircraft computers if they were not able to see the actual parameters it uses to formulate an automated decision. A decision between goal- and data-driven processing thereby occurs. The optimal solution from a display perspective is a compromise--displayed information among data and goals, potentially customized by the pilot depending on flight regime and individual preference (Endsley & Jones, 2012).

The quality of perception, recognition, and future projection of aircraft states are essential to energy management, since aircraft are complicated machines in dynamic environments, and often pilot inputs and their consequent implications take time to manifest. If EM is not precise, aircraft upset can lead to aerodynamic stall, sometimes unrecoverable. Adequate EM SA can safely navigate an aircraft through the environmental hazards of high altitude flight. This research will examine the ability of the OZ display to enhance EM SA in high altitude flight.

Mode awareness. Modern cockpits containing complex FMS, autopilots, and auto-throttles, designed to relieve pilot workload, might actually increase cognitive demand (Endsley & Jones, 2012). Greater SA has an unintended casualty, as the complex cockpit display designs that provide it might actually hinder pilots from proper perception and comprehension of flight information. The phenomenon known as *mode awareness* is the degree to which a pilot is aware of how the computer logic is controlling the aircraft (Sarter & Woods, 1995). Pilot error due to a lack of awareness or confusion related to the functions of the automated systems is often cited in accident reports (Sherry & Mauro, 2014). An EID display, such as the OZ display, using situation awareness oriented design principals, may allow for greater comprehension and computer logic transparency by providing improved perception of relevant flight information.

Modern aircraft operations have seen pilot duties evolve from a very hands-on and active participant requiring frequent stick and rudder manipulation to a detached supervisory monitor of advanced flight computers capable of controlling aircraft from takeoff to landing (Sarter & Woods, 1995). The cognitive demands required of pilots increased greatly with the introduction of automated flight controls and computers. Not only were pilots responsible for knowing how to fly their airplane properly, but now in the modern age they also have to master the rules, algorithms, and operating systems of the flight computers onboard.

In order to operate these advanced systems properly, pilots must first integrate large quantities of data from various sources throughout the cockpit and next maintain an accurate mental model of what the aircraft is doing in its airspace and what the automation is doing with the input it has received, both from the pilot and from pre-selected internal algorithms (Sarter & Woods, 1995). Lastly, pilots need to understand the complex rule structures of the autopilot, FMS, and flight director included, so they can correctly predict future states of the aircraft and future behavior of its automation. One of the hazards of technological advancements in the cockpit has been automated autopilot systems that contain a plethora of different ways (modes) to accomplish the same aviation task, often providing pilots with confusing automation choices, each with its own rule set or algorithm. Pilots unaware of the selected (active) mode and what rules pertain to that particular mode may inadvertently be the root for errors of commission or omission due to what is called a lack of *mode awareness* and a general gap in the mental model.

Sarter and Woods (1995, p. 122) describe the phenomena of mode awareness as a special form of SA for “human supervisory controllers” (e.g., pilots monitoring an airplane autopilot’s various autonomous modes) and their need to properly understand what the airplane is doing, how it is doing it, and why it is doing it. Mode awareness problems originate from “technology-centered automation”—those systems without a visible or obvious trail of status and behavior, evidently designed without the operator (pilot) in mind (Sarter & Woods, 1995, p. 115).

Sarter and Woods (1995) imply that cognitive demand required for proper mode awareness is poorly supported by displays currently available. The penalty for low mode awareness is “new human-machine error forms and new paths toward accidents” (Sarter & Woods, 1995, p. 111) as evidenced when confused pilots incorrectly command an FMS or autopilot with respect to altitude, airspeed, or lateral or vertical navigation. Pilots, either through a lack of full understanding of the system or a misinterpretation of the current mode, can unwittingly place an aircraft in hazardous situations. Examples include the accidents of Air France 447 in June 2009 (NTSB Report DCA09RA052), where the pilots were confused about autopilot functionality and capabilities and failed to properly diagnose an approach to stall situation, or Asiana 214 in July 2013 (NTSB

Report DCA13MA120), where the pilots' insufficient monitoring of airspeed resulted from automation misunderstanding of the complexities of the auto-throttle and autopilot systems. The recommendation made by Sarter and Woods (1995) is to take a "process-oriented [vice] product-oriented approach … to design tools that support monitoring, assessment, and awareness demands on [pilots]" (p. 122). Mode awareness confusion can inhibit a pilot's ability to discern how an aircraft is trending from an EM perspective.

Increased air traffic density brought along improved technologies and revised regulations and procedures. On account of this, a flight deck automation working group was established to update a 1996 FAA report indicating vulnerabilities in pilot management of automation and SA (Abbott, McKenney, & Railsback, 2013, p. 1). The working group, consisting of key members from industry as well as aviation researchers, addressed the safety and efficiency of modern flight deck systems for flight path management and energy-state management (Abbott, et al., 2013, p. 1). Finding 2 of 28 was devoted to "vulnerabilities in pilot knowledge and skills for manual flight operations", specifically inadequate energy management (Abbott et al., 2013, p. 2). According to Abbot et al. (2013, p. 73), ubiquitous and powerful cockpit technology led to a degradation in manual piloting skills, decreased recognition of an energy deviation, and more difficult recovery from an upset condition. The working group recognized that as modern piloting tasks change from manual manipulation of flight controls to passive monitoring of complex computer systems, the skills required to safely fly must be updated and changed in professional training settings as well as primary flight training centers (Abbott et al., 2013, p. 121).

Situation Awareness Measurement

The effectiveness of any human-factors derived design is determined by how well SA is supported throughout the design. Since SA is “an internalized mental construct” (Endsley & Jones, 2012, p. 259), accurate measurement is difficult to define or obtain. It is rare that individual participants show consistency between behaviors and outcomes. Furthermore, many human factors either combine to influence SA or completely mask it altogether. Several methods are available to measure SA. Endsley and Jones (2012) assert that the most complete analysis utilizes a combination of methods.

Breton, Tremblay, and Banbury (2007) list 28 different tools and use two distinctions to classify SA measurement tools: on- or off-line and direct or indirect. Whereas on-line methods tend to include the SA queries concurrently during the scenario, off-line methods typically pause the testing scenario to cognitively assess SA through interrogation of environmental and operational knowledge of the test scenario. Off-line methods may even assess SA retrospectively after the scenario is complete. Artificiality occurs using either on- or off-line methods that interrupt a scenario with external stimuli, but some positive results have been seen when the queries are organic to the scenario, for example, if air traffic control (ATC) is asking the pilots for information similarly as would happen in the real world.

Indirect methods, sometimes referred to as inferred or process-oriented methods, measure process, behavior, and performance. These indirect methods analyze communication and psychomotor mannerisms such as eye tracking or heart rate to gain insight toward processes or mechanisms used to obtain SA. Inferred methods also may examine behavior traits, such as reaction time or decision-making, or performance results on applicable tasks within a scenario. Inferred methods are usually subjective by nature, based upon a subject matter expert (SME) assessment of a participant's SA relative to the known truths of the scenario, but even so it has been shown that performance, behavior, and process (e.g., what a participant is looking at through eye-trackers) do not always correlate directly to SA. High SA does not always result in high performance, and high performance does not always stem from high SA (Endsley, 1995). Furthermore, methods that assess the participant's process of acquiring SA cannot analyze what the participant will actually do with the information that was gathered. Memory, attention span, and cognition all combine during the perception / comprehension / projection phases of SA.

Direct methods rely on the subjectivity of SMEs and raters to decipher the complex synthesis of the internal human cognition that is SA. The Situation Present Assessment Method (SPAM), described as a real-time probe technique, is an example of a direct method. SPAM examines participant's Level 1 and Levels 2/3 SA by measuring the reaction time to probes (Breton et al., 2007). Unlike other methods, SPAM interrogations do not artificially interrupt or stop the scenario but incorporate these queries into its natural flow of task performance. In an experiment, Endsley, Sollenberg, and Stein (2000) found a slight correlation between SPAM response times and mental workload.

Summary

Aircraft upsets are a hazard in modern aviation and a safety concern throughout the industry as the frequency of accidents in GA, corporate, military, and commercial aviation increases. Knowledge-based actions required to manage airplane energy exist during all phases of flight. However, the high altitude environment complicates this relationship between thrust, lift, and drag since aircraft are often at cruise with a narrow margin between overspeed and stall. PFDs and analog six-pack RD cockpits display individual pieces of critical information in series, whereas the OZ concept display merges compound data into a comprehensive presentation of aircraft state in a parallel fashion (Eskridge et al., 2014, p. 92). Research with the OZ concept display has reported advantages of a functional display for delivering valid and reliable perceptual cues to pilots. The OZ display's predominant graphical feature of the thrust curve provides thrust required, thrust available, overspeed, and stall limits directly to the pilot. Smith (2008), Smith and Boehm-Davis (2005), Smith et al. (2004), Smith et al. (2005), and Temme et al. (2003) have all reported improvements in flight performance over RDs when both novices and experienced pilots use the OZ display. EM SA could be enhanced through the use of the OZ display to reduce aircraft upsets. The proposed study will utilize a within-subjects, counter-balanced experimental design to compare the OZ graphical EID display and conventional electronic flight displays for providing EM SA.

CHAPTER III

METHODOLOGY

The study protocol was submitted for review and approval by the Institutional Review Board (IRB) for Embry-Riddle Aeronautical University and Lewis University (see Appendix A). Prior to participation in the study, participants were asked to read and complete an informed consent form (see Appendix B). Treatment of all participants was in accordance with the ethical standards of the American Psychological Association.

Research Approach

This empirical study evaluated how energy management situation awareness (EM SA) was supported by both a PFD, similar to what is used in modern transport aircraft, and the OZ Concept ecological interface design (EID) display, designed by the Institute for Human Machine Cognition (IHMC). This study ascertained how effective each display was at providing pilots with EM SA, operationally defined as speed and accuracy of on-screen responses to EM questions when observing display screenshots of an airplane operating at high altitude. Participants were seated at a personal computer terminal displaying an EM situation depicted on a static cockpit display screenshot, they interpreted the aircraft's energy state and acknowledged the requirement, if any, for aircraft flight control inputs by selecting multiple options from a list of possible on-screen responses. The two cockpit display designs were evaluated in terms of each condition's utility for supporting pilot EM SA. The degree of EM SA supported was determined by accuracy and response time to the on-screen queries.

An experiment utilizing aircraft instrument panel screenshots requiring participants to make energy management decisions comparatively measured the EM SA demonstrated on both a PFD and an OZ concept display. The purpose of this experiment was to examine the overall EM SA of each display screenshot. Rather than running participants through full motion flight simulators, this screenshot technique was faster and less costly to implement, reflected speed of understanding more directly, and provided a similar level of fidelity for both displays. Admittedly, proprioceptive, aural, and haptic cues important to developing and maintaining SA, are lost when just using a static, artificial, out-of-context screenshot method. Pilot EM SA was assessed during the experimental trials using the validated SA measure SPAM to evaluate differences in pilot analysis of each aircraft energy state condition when presented with either a PFD or an OZ concept display. SPAM measures reaction time (response latency) (Durso et al., 1997) to real-time probes in action scenarios in order to assess SA (Breton et al., 2007).

Design and procedures. A within-participants, repeated-measures experiment exposed each participant to a series of static cockpit screenshots, 15 from each of two display types. Each participant sat in front of a personal computer terminal. Initially, participants were presented with a demographic questionnaire ascertaining their age, gender, race / ethnicity, and flight experience / qualifications (see Appendix C). Participants were required to qualify for an FAA Third Class medical certificate, effectively screening for vision issues (color, depth, etc.) and any other physical or psychological impairments that would affect their ability to participate in the study. Participants were also required to hold a current FAA instrument rating, since knowledge and understanding of controlling an aircraft using only an instrument gauge scan was necessary for the trials. The demographic survey also ascertained flight experience particular to instrument time, inferring instrument scanning skills and electronic flight instrument system (EFIS) experience, as well as high altitude background, inferring training, education, and flying qualities. The participants were next presented with a short tutorial explaining each of the two cockpit display types, the symbology presented, and the complete taxonomy of choices for the on-screen responses.

Once the experiment began, participants were presented with a series of 15 cockpit display screenshots from each display type. Each participant received the screenshot series in the same order but in a counterbalanced design, so that each participant viewed all of one cockpit type followed by all of the other type. For each cockpit screenshot, the participant viewed the cockpit display screenshot photo and assessed the current and near future state of the aircraft's energy, decided what is currently happening to the aircraft with respect to energy state, and what, if any, pilot action is required in regards to energy management in terms of airspeed, aircraft pitch attitude, or altitude adjustment.

The EM situations chosen were sufficiently complex to result in measurable difficulty with conventional flight displays. This particular study did not consider aircraft configuration, such as speed brakes or lateral or directional control inputs, but those could be evaluated for future research. The addition of AOA to a PFD provides very useful EM information but was not included on either of these two display conditions and was not evaluated in this study but may provide fruitful future research.

When each new screenshot was presented, the participant acknowledged the aircraft's current energy state and any input intentions (thrust levers, pitch adjustments) that may be required by selecting from the appropriate on-screen response options. The list of standardized responses, shown in Table 2, was on-screen for the participants to choose, so as to decrease the variation in responses (see also Appendix D). This list was also included in the tutorial each participant received prior to the experiment.

Table 2

Selection of Standardized Participant Responses to Screenshot Stimuli

Attitude	Airspeed	Energy State	Thrust Lever Input
Climbing	Accelerating	Power avail > Power req	Add Thrust
Descending	Decelerating	Power req > Power avail	Reduce thrust
Level	Constant	Power avail = Power req	Pitch Up
Left Bank	Near Overspeed		Pitch Down
Right Bank	Near Stall		Roll Wings Level
			No Input Required

Following the experiment, participants were presented with an on-screen computerized survey requesting open-ended, qualitative, free-form, typed feedback regarding their attitudes and impressions of the experiment structure, the two display types, the difficulty or ease of display interpretation; their level of understanding about energy management, and their corresponding situation awareness of it (see Appendix E). This provided insight for further design modifications and future studies.

Apparatus and materials. Fifteen screenshots of the electronic flight display PFD (see Figure 10) were taken during a desktop simulator (*Flight Sim X*) flight of a medium sized regional jet, a CRJ-700. Fifteen screenshots of the EID OZ Concept flight display (see Figure 11) were taken during a desktop simulator (*Flight Sim X*) flight of a single-engine Cessna 172. The OZ emulator is currently only programmed with C-172 performance data. The series of display screenshots detailed the aircraft's instrumentation at high altitude cruise flight through various degrees of energy states: balanced flight (thrust required < thrust available), the region of reverse command (thrust required > thrust available), and energy states leading up to aerodynamic stall. The simulator scenario placed both aircraft at its respective optimum altitude, or highest sustainable altitude for that specific gross weight and atmospheric temperature, and included various combinations of thrust level position, bank angle, and pitch angle.

The screenshot stimuli and on-screen responses were shown to the participants in a slideshow format, as shown in Figures 10, 11, and 12. A precision computer-based psychological testing software, "Inquisit 5 Lab" by Millisecond, was used to present the slideshow and record elapsed time for each screenshot and subsequent response selection. Participants were seated at a desktop personal computer terminal in a secluded, private office away from visual and aural distraction, they viewed the slideshow on a personal monitor, and they used a pointing device "mouse" to select "click" on-screen responses and progress to the next screenshot stimuli. The 27 possible responses were shown on-screen directly below the display graphics for both conditions, in a checkbox fashion, where participants were instructed to mark all that apply using their computer mouse.

Figure 10. Display screenshot testing stimuli for PFD display condition.

Indicate the condition. Mark all that apply.

❑ Climbing ❑ Descending ❑ Level ❑ Left bank ❑ Right bank ❑ Other

❑ Accelerating ❑ Decelerating ❑ Constant ❑ Near overspeed ❑ Near stall ❑ Other

❑ Power avail > power req ❑ Power avail = power req ❑ Power avail < power req ❑ Other

Indicate the response. Mark all that apply.

❑ Add thrust ❑ Reduce thrust ❑ Pitch up ❑ Pitch down ❑ Roll wings level ❑ No input required ❑ Other

Figure 11. Standard list of responses to high altitude energy management decisions.

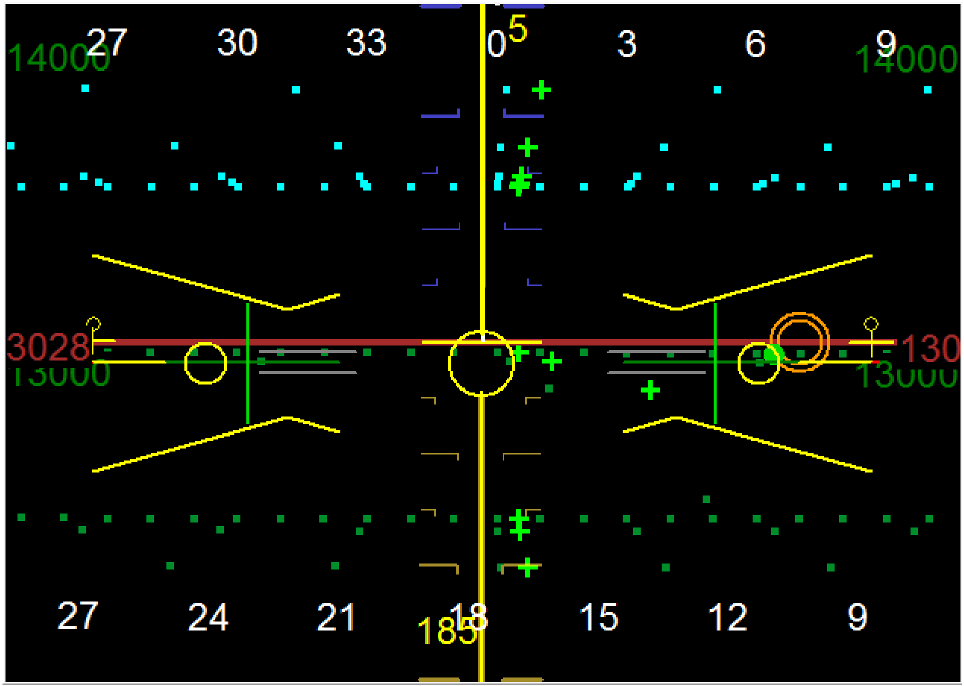

Figure 12. Display screenshot testing stimuli for OZ display condition.

Population/Sample

This study recruited a volunteer sample of Lewis University undergraduate and graduate students, staff, faculty, and flight instructors who held at least an FAA instrument rating and qualified for an FAA third class medical certificate, by advertising the study through mass email to the aviation department, posted flyers in the aviation building, and word of mouth through the flight school instructors and aviation faculty. There was no benefit offered to the participants in this study. This sample had varied aviation backgrounds in operational experience, licenses, type ratings, flight hours, instructor time, instrument time, and high altitude exposure. The evaluation of the experimental results determined the utility of both display types in regards to EM SA. The repeated-measures, within-participants design de-emphasized individual differences in aeronautical skill, experience, and cognitive ability, and mitigated potential variance in sample selection.

Testing was conducted within the Lewis University Aviation & Transportation Department in Romeoville, IL, over the course of 20 days during the summer term. Ninety (90) volunteer participants initiated the study, returning 85 valid and complete responses for both sets of stimuli. Data from five participants were excluded from the analysis due to technical / procedural problems. The valid cases were further reduced to N=58 after removing those participants without a medical clearance and an instrument license. Data from student pilots currently pursuing an instrument rating were retained.

A power analysis using G*Power 3.1 software determined a minimum sample size to be between 54 participants, based on a *t*-test for two paired dependent means, a two-tailed alpha between .05, and an effect size of .5. This was confirmed with Faul et al. (2009) and Cohen (1992).

Sources of the Data

Data was collected for each participant linking display screenshot, participant on-screen question response, and elapsed response time. A response answer key was created by a subject matter expert (SME), a professional pilot with an Airline Transport Pilot (ATP) rating, Certified Flight Instructor Instrument (CFII) certificate, a first class medical certificate, and over 4,800 flight hours in fixed and rotary wing aircraft. The answer key was validated by a second SME, a professional pilot with an ATP, a Gold Seal CFII, a first class medical, and over 12,700 flight hours in fixed wing aircraft. The accuracy of responses and elapsed response times were both considered ratio data. Display type (PFD vs. OZ display) and participant demographic data (gender, age, flight experience, recency, etc.) was considered nominal data and stratified according to the sample composition.

The participants' EM SA during these scenarios was measured with the Situation Present Assessment Method (SPAM) in real-time. Breton et al. (2007) endorse the use of SPAM in flight simulators where the inclusion of SPAM queries does not require suspension (stopping or pausing) of the scenario. SPAM also distinguishes workload from SA by only measuring the elapsed response time once the participant begins processing the query. The assumption was that if workload is low, the participant would be able to answer the query immediately, whereas if workload is high, the participant may be delayed answering the query. In this design, the query shows up simultaneously with the display condition, so workload was not measured or considered. SPAM differs from other common SA measures in that SPAM requires only that a participant know where to find targeted information, rather than having to recall from memory that piece of information (Durso et al., 1997).

SPAM in this experiment queried participants on the nature of flight control response required, thrust lever and / or pitch control, if any, with regards to the participant's interpretation of aircraft EM, current and near future aircraft energy states. SPAM measured response time and accuracy of the participants' survey query answers, as shown in Table 2. The measured data was the participant's on-screen responses to each static screenshot of cockpit gauges, either PFD or OZ display, portraying a particular aircraft energy state. The screenshot and the survey question both appeared on the screenshot. Elapsed time required to make the response after seeing the screenshot condition was also measured. The graded response data was paired with the elapsed time data as the dependent variables. The independent variable was the display type: PFD or OZ display.

Each participant used a mouse to click on 27 possible responses for each screen; they were instructed to check any / all statements that applied to the aircraft scenario depicted. The stimuli response data were scored in the following manner: the participant's responses were compared to the answer key, summing matches of either a correct response or a correct blank or omission, depending on the question and screen. Either a checked box where the answer key indicated a blank or a blank where the answer key indicated a checked box, would both be considered a "miss" and not counted in the match totals. These 27 responses for each of the 15 screens in that particular display type corresponded to four constructs: airspeed, attitude, energy, and corrective response. However, the PFD Accuracy and OZ Accuracy variables were aggregates for the sum of all questions, respective to display type. The maximum possible score for each display was 405.

Data Collection Device

Instrument reliability. The stability and consistency of the SA measure queries in this design were ensured since each given stimulus situation only had one correct answer. Any individual should be expected to repeat their same answers during identical trials performed at a later time. Assuming learning effects stemming from repeated testing, this trial only included one pass through each display type. Any learning effects gained from the display design tutorial prior to beginning the experiment was equal across the participants. The given responses for each specific screenshot scenario should remain the same in future trials.

Reliability was improved by utilizing a computer to collect data for the dependent variables of response accuracy and response time. In an effort to minimize the learned and expected behavior during the trial for individual participants, the order of displayed screenshots was randomized using a random number generator to order the screenshots, and their occurrence within the trial was not necessarily chronological with respect to when they occurred in the previously flown simulator scenario. However, the experiment was counter-balanced so that half of the participants received the random order of display "A" screenshots followed by the random order of display "B" screenshots, while the other half of participants received the opposite order. The same randomly chosen order of screenshots within each display set was consistent for each participant, ensuring reliability between participant trials.

This experiment employed the same simulator scenario parameters flown in both the PFD and the OZ display and then used these screenshots for the participants who experienced both in a repeated measures design. Reliability was strengthened by running each participant through an identical single series consisting of both display sets during the same flight scenario, essentially providing each participant with two attempts at each energy state screenshot. Cronbach's coefficient alpha was computed to test construct question internal consistency for the PFD Accuracy and OZ Accuracy variables.

Instrument validity. Durso et al. (2006) support the incremental validity of SA queries using SPAM when used with several cognitive tests, since SA is a construct above and beyond these underlying cognitive mechanisms. The screenshot used in this experiment and the SPAM questionnaire was rated by two SMEs, satisfying face validity prior to implementation with the participants. The two SMEs also ensured that the questions asked were clearly stated and that the participants understood the question the way it was intended.

Treatment of the Data

Descriptive statistics. Data were collected for each participant linking display type, screenshot scenario energy state, accuracy of response, and elapsed response time. Descriptive statistics included measures of central tendency, dispersion, distribution, and percentiles. The data were sorted, aggregated and cleaned using Microsoft Excel, and the various statistical analyses were run using IBM SPSS Statistics Version 23.

Hypotheses testing. Research Hypothesis 1 (Participants presented with high altitude EM decisions will exhibit a difference in response accuracy and response time when using the OZ concept display compared to when using the conventional electronic flight display) was tested using a paired samples *t*-test, comparing the differences of EM SA performance means in accuracy and time within individual participants on both the PFD and the OZ display.

Null Hypothesis 1 (Participants presented with high altitude EM decisions will exhibit no difference in response accuracy and response time when using the OZ concept cockpit display compared to when using a conventional electronic flight display) was tested using a paired samples *t*-test, comparing the differences of EM SA performance means in accuracy and time within individual participants on both the PFD and the OZ display.

Research Hypothesis 2 (Pilot flight hours in airplanes will be correlated with speed and accuracy of EM decisions when presented with the conventional electronic flight display) and Null Hypothesis 2 (Pilot flight hours in airplanes will not be correlated with speed and accuracy of EM decisions when presented with the conventional electronic flight display) were tested using Spearman's correlation statistics to measure strengths of relationships.

Research Hypothesis 3 (Pilot flight hours in airplanes will not be correlated with speed and accuracy of EM decisions when presented with the OZ concept EID display) and Null Hypothesis 3 (Pilot flight hours in airplanes will be correlated with speed and accuracy of EM decisions when presented with the OZ concept EID display) were tested using Spearman's correlation statistics to measure strengths of relationships.

Assumptions for the *t*-tests are: independent observations, randomly sampled data, normal distribution of all dependent variables, and homogeneity of covariance (Field, 2009). These assumptions were met since the independent performance scores were generated by different individuals, and participants served as their own controls (repeated measures design). Assumptions for a non-parametric Spearman's correlation test are that the data be ordinal (Field, 2009).

Qualitative data. The post-experiment written survey administered immediately following the experiment while the participant remained seated at the computer terminal provided qualitative data in the form of open-ended, free-form, typed feedback about the experiment structure, the two display types, the difficulty or ease of display interpretation, the level of understanding about energy management, and corresponding situation awareness of it.

Pilot Study

A pilot study was performed with N=14 participants to gain some initial analysis for the hypotheses and relationships between the PFD and the OZ display, as well as test the software stimulus delivery and overall flow of the experiment. Results showed non-significant differences in means between PFD and OZ on both accuracy and response time, yet the OZ display did fare better: 3.13 seconds faster and a slightly more accurate aggregate result (1.64 to 1.51 out of 4.00). The accuracy grading rubric awarded one point for correct answers to each of four constructs: airspeed, attitude, energy, and corrective response. A perfect accuracy score was a 4.00 for each of the ten screens on that display condition. Participants were more accurate using the PFD for attitude and corrective response individually, but more accurate using the OZ for airspeed and energy management individually.

The stimuli screenshots, originally taken from low altitude scenarios, were replaced with more complex aerodynamic scenarios at maximum operational altitude, near coffin corner for both aircraft types. Some of the demographic survey answer choices were adjusted based on SME and participant feedback. A screen counter (*n* of 15) was added when number of total stimuli went from 20 to 30.

CHAPTER IV

RESULTS

Descriptive Statistics

Participant Demographics. The 58 valid cases consisted of 54 males (93.1%) and 3 females (5.2%) with 1 preferring not to answer. There were 35 participants (60.3%) in the 18-24 year old age group, 15 (25.9%) in the 25-34 year old age group, and 8 (13.8%) spread among the other decade groups from 35-74 years old. Forty-three participants (74.1%) identified as being not Hispanic or Latino, 11 (19.0%) as Hispanic or Latino, and 4 (6.9%) preferred not to answer. Fifty-one participants (87.9%) identified as being White, 3 (5.2%) as Black or African American, 2 (3.4%) as Asian, 1 (1.7%) as American Indian or Native Alaskan, and 1 (1.7%) preferred not to answer.

Twenty-five participants (43.1%) held a bachelor's degree, 17 (29.3%) completed some college but no degree, 6 (10.3%) held a high school diploma, 6 (10.3%) held a master's degree, 2 (3.4%) held an associate's degree, and 2 (3.4%) held a doctoral degree. Twenty-nine participants (50.0%) were university students in aviation fields, 15 (25.9%) flight instructors, 6 (10.3%) commercial pilots, 5 (8.6%) chief or assistant chief pilots, 2 (3.4%) aviation faculty, and 1 preferred not to answer.

Thirty-five participants (60.3%) held a current first class medical certificate, 11 (19.0%) a second class, and 12 (20.7%) a third class. Eighteen participants (31.0%) held a private pilot certificate and were working toward their instrument rating, 9 (15.5%) an instrument rating, 1 (1.7%) a commercial rating, 16 (27.6%) a certified flight instructor instrument rating, 7 (12.1%) a multi-engine rating, 6 (10.3%) an airline transport pilot rating, and 1 (1.7%) a remote pilot operator license. Five participants (8.6%) had between 51 and 100 flight hours in airplanes; 27 (46.6%) between 101 and 500; 10 (17.2%) had between 501 and 1,000; 4 (6.9%) between 1,001 and 1,500; 9 (15.5%) between 1,501 and 5,000; 2 (3.4%) between 5,001 and 10,000; and 1 (1.7%) more than 10,001.

Dependent Variables. Each participant viewed 15 screens of each display type for a total of 30 screens. On each screen, participants indicated a response to 27 different checkboxes for current airspeed (7 boxes), aircraft attitude (7 boxes), current energy (5 boxes), and pilot corrective response (8 boxes). Correct responses matching the key were summed for each display type for a total of two performance scores: both PFD Accuracy and OZ Accuracy had a maximum perfect score of 405: 27 responses multiplied by 15 screens. Elapsed time spent viewing and responding to each of the 15 screens for each display type was recorded in milliseconds and then averaged for a total of two time scores: PFD Time and OZ Time. Response times and response accuracy were analyzed independently and not according to the level of accuracy. Response times associated with low accuracy were not discounted in the analysis. Descriptive statistics are shown in Table 3. Histograms are shown in Figures 13 through 18. Response time for PFD and OZ both had positive skew with most participants registering quicker response times, while both PFD and OZ accuracy scores displayed a more normal curve appearance centered on their mean. Response time variables were analyzed after a square root transformation reduced the skew and improved normality, as shown in Figures 15 and 16, but no remarkable results differing from the original analysis emerged with the transformed data.

Table 3

Dependent Variables

Variable	*M* (*SD*)	Min.	Max.	Skewness	Kurtosis
PFD Time	39.88 (14.10)	20.70	86.70	1.132	1.144
OZ Time	43.35 (19.46)	15.86	114.85	1.270	2.019
PFD Accuracy	348.84 (12.95)	321.00	377.00	.304	-.272
OZ Accuracy	335.67 (13.22)	304.00	361.00	-.131	-.441

Note. N = 58. Times are displayed in seconds. Min. = Minimum; Max. = Maximum.

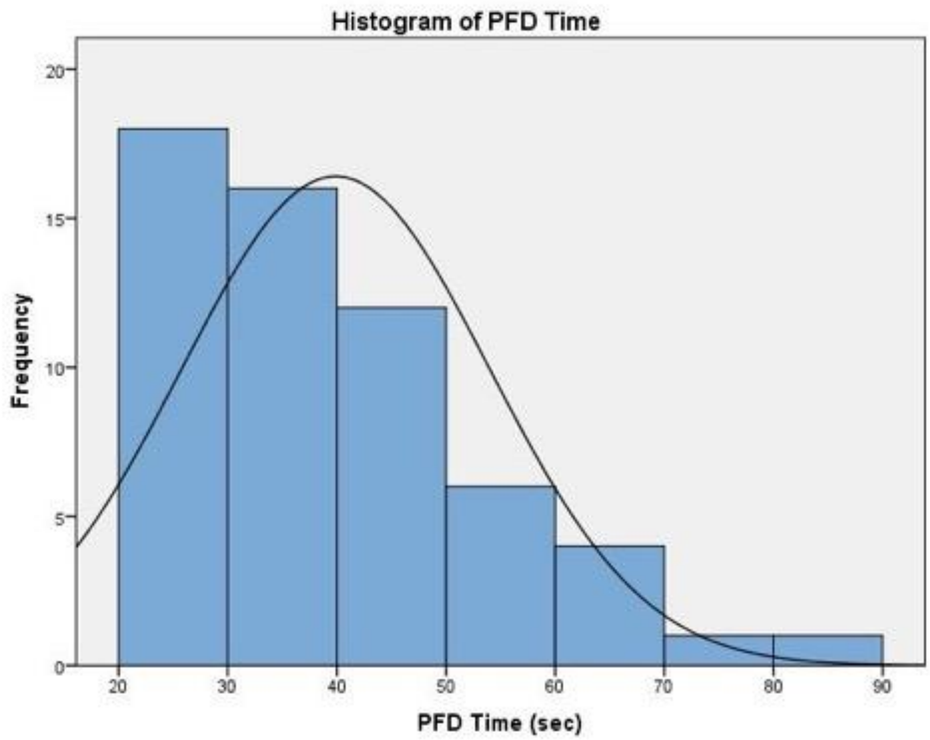

Figure 13. Total elapsed time participants spent viewing and responding to high altitude energy management decisions presented on a PFD.

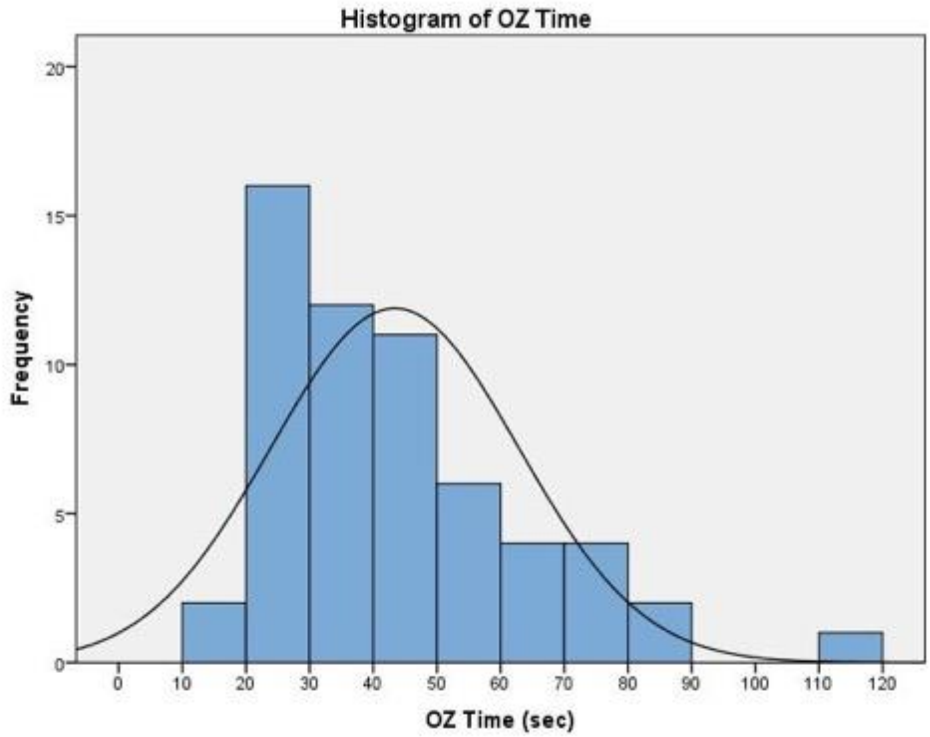

Figure 14. Total elapsed time participants spent viewing and responding to high altitude energy management decisions presented on an OZ display.

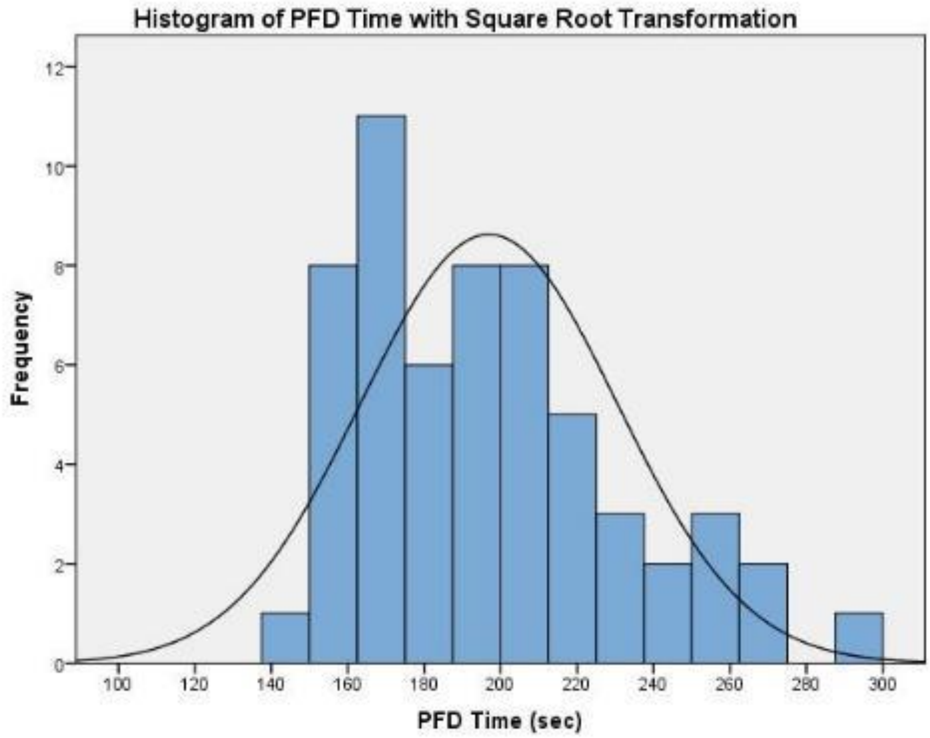

Figure 15. Total elapsed time participants spent viewing and responding to high altitude energy management decisions presented on a PFD with transformation.

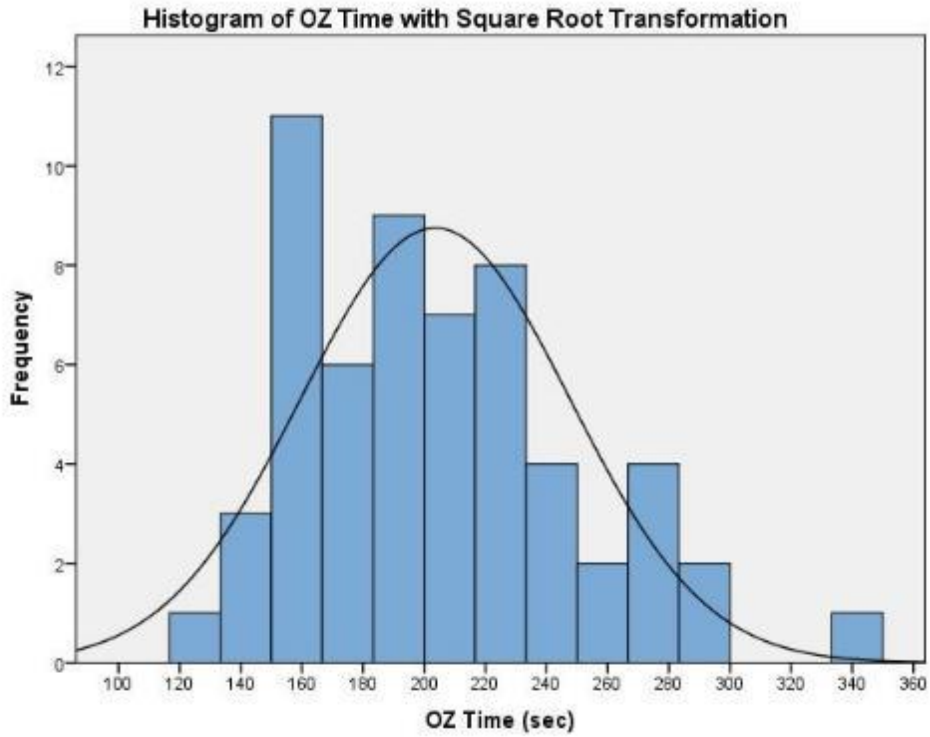

Figure 16. Total elapsed time participants spent viewing and responding to high altitude energy management decisions presented on an OZ display with transformation.

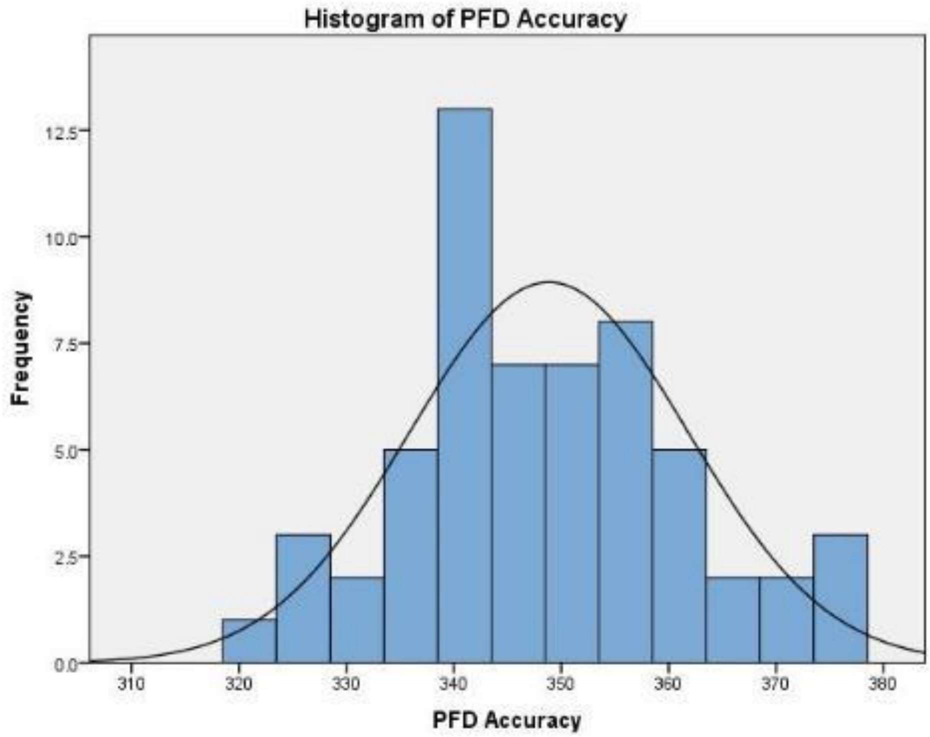

Figure 17. Total correct responses to high altitude energy management decisions presented on a PFD.

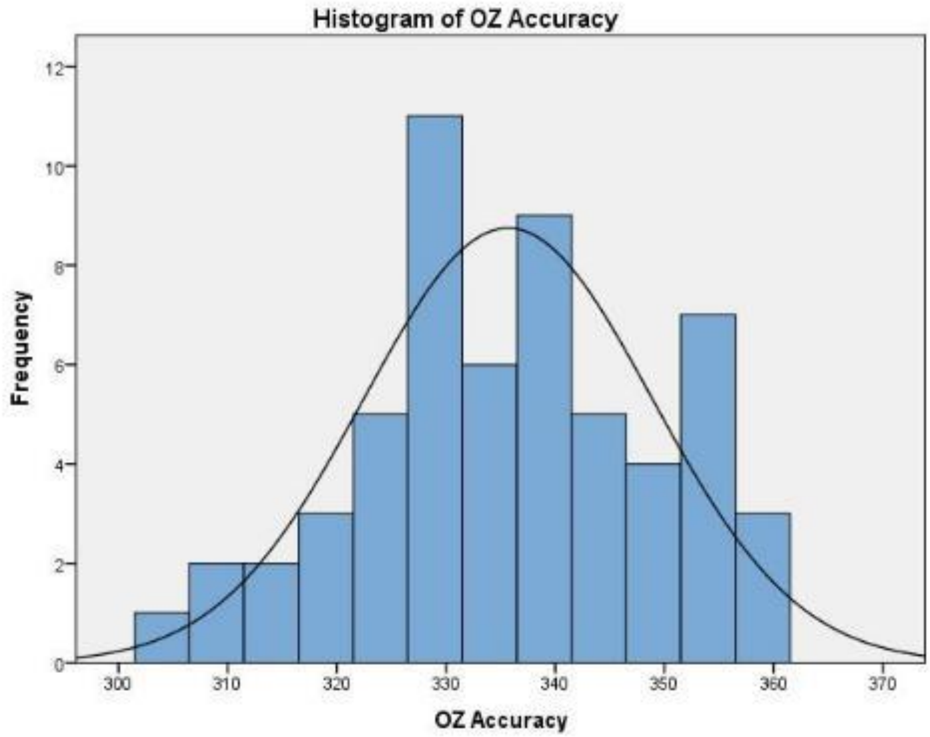

Figure 18. Total correct responses to high altitude energy management decisions presented on an OZ display.

Reliability Testing

Cronbach's coefficient alpha was computed to assess whether the items aggregated into the accuracy score had acceptable internal consistency. The alpha for the aggregated items relating to the PFD display was .13, and the alpha for the aggregated items relating to the OZ display was .63. The difference in these coefficient alpha values was unexpected since the questions were the same for both conditions.

Hypothesis Testing

Hypothesis 1. Null Hypothesis 1 (Participants presented with high altitude EM decisions will not exhibit any differences in response accuracy and response times when using the OZ concept display compared to when using the conventional electronic flight display) was tested using a paired samples *t*-test, comparing the differences of EM SA performance means in accuracy and time within individual participants on both the PFD and the OZ display. There was not a statistically significant difference in means between PFD Time and OZ Time. On average, mean PFD Accuracy was significantly higher than mean OZ Accuracy scores (M_{Diff} = 13.17, SD_{Diff} = 20.96), $t(57) = 4.78$, $p < .001$, $d = 0.63$, as shown in Table 4. Thus, Null Hypothesis 1 was retained for response times but rejected for response accuracy.

Table 4

Paired Samples Test

Variable	M_{Diff} (SD_{Diff})	*t*(57)	*p*	LL	UL	Cohen's *d*
PFD-OZ Time	-3.47 (17.83)	-1.48	.143	-8.16	1.21	0.19
PFD-OZ Accuracy	13.17 (20.96)	4.78	<.001	7.65	18.68	0.63

Note. Times are displayed in seconds. LL = 95% Confidence Interval Lower Limit; UL = 95% Confidence Interval Upper Limit.

To further explore this finding, a Pearson bivariate correlation was run between PFD Accuracy and PFD Time as well as between OZ Accuracy and OZ Time. Results showed no significant correlation between PFD Accuracy and PFD Time, r (58) = .011, $p > .05$. However, results showed a significant correlation between OZ Accuracy and OZ Time, r (58) = .353, $p < .01$. For the OZ display condition, better accuracy was associated with slower response times. Scatterplots for accuracy versus time for both PFD and OZ are shown in Figures 19 and 20, respectively.

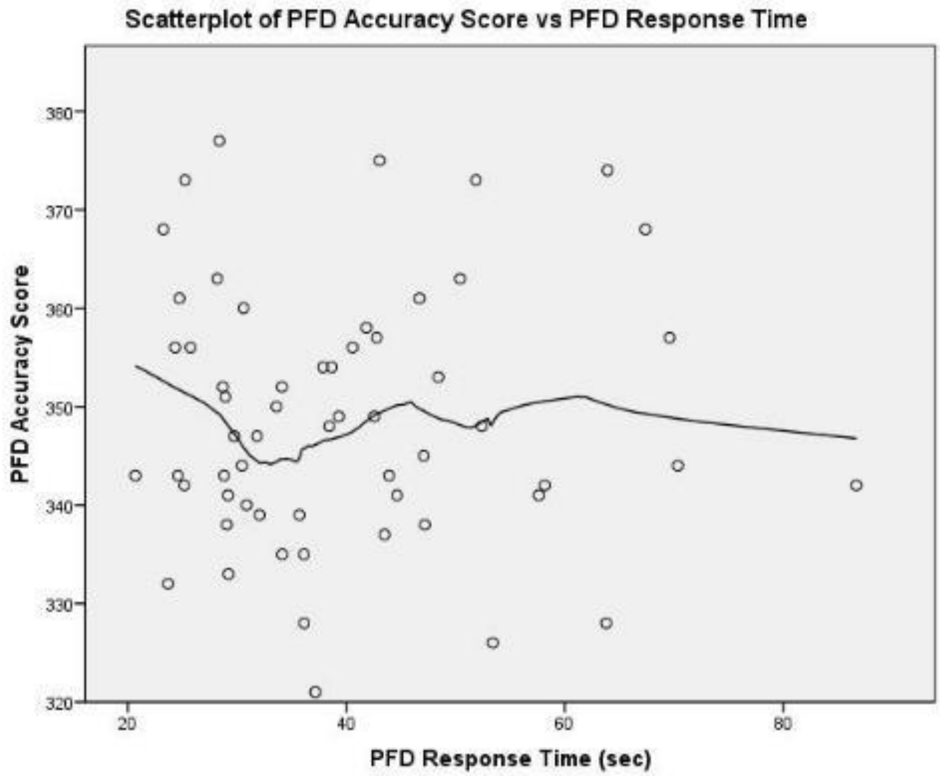

Figure 19. Scatterplot of elapsed time participants spent viewing and responding to high altitude energy management decisions presented on the PFD plotted against total correct responses on the PFD.

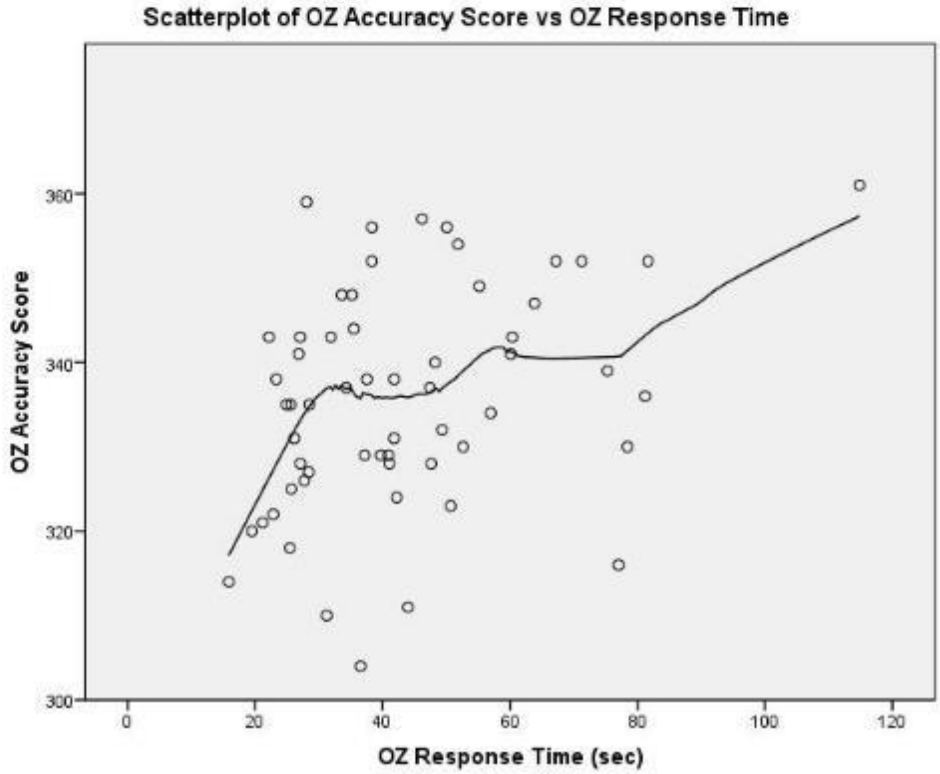

Figure 20. Scatterplot of elapsed time participants spent viewing and responding to high altitude energy management decisions presented on the OZ plotted against total correct responses on the OZ.

Hypothesis 2. To evaluate Null Hypothesis 2 (Pilot flight hours in airplanes will not be correlated with speed and accuracy of EM decisions when presented with the conventional electronic flight display), a non-parametric Spearman's rho correlation was run between pilot flight hours in airplanes (categorical data) and PFD Time and PFD Accuracy. Neither of these two relationships were significant, as shown in Table 5. Thus, Null Hypothesis 2 was retained.

Table 5

Correlation of Accuracy and Time with Flight Hours

Variable	*N*	Correlation coefficient	*p*
PFD Time	58	.055	.680
PFD Accuracy	58	.228	.085
OZ Time	58	.024	.860
OZ Accuracy	58	.149	.264

A one-way analysis of variance (ANOVA) revealed that PFD Accuracy score means were not significantly different across all seven flight time categories, $F(6,51) = 1.11$, $p > .05$, $\omega = .11$, as shown in Figure 21, but that PFD RT means were significantly different across all seven flight time categories, $F(6,51) = 2.50$, $p < .05$, $\omega = .37$, as shown in Figure 22. Further exploring this result, by removing the three cases above 5,001 flight hours, PFD Response Time means were not significantly different across the remaining five flight time categories, $F(4,50) = .751$, $p > .05$, $\omega = .14$, as shown in Figure 23.

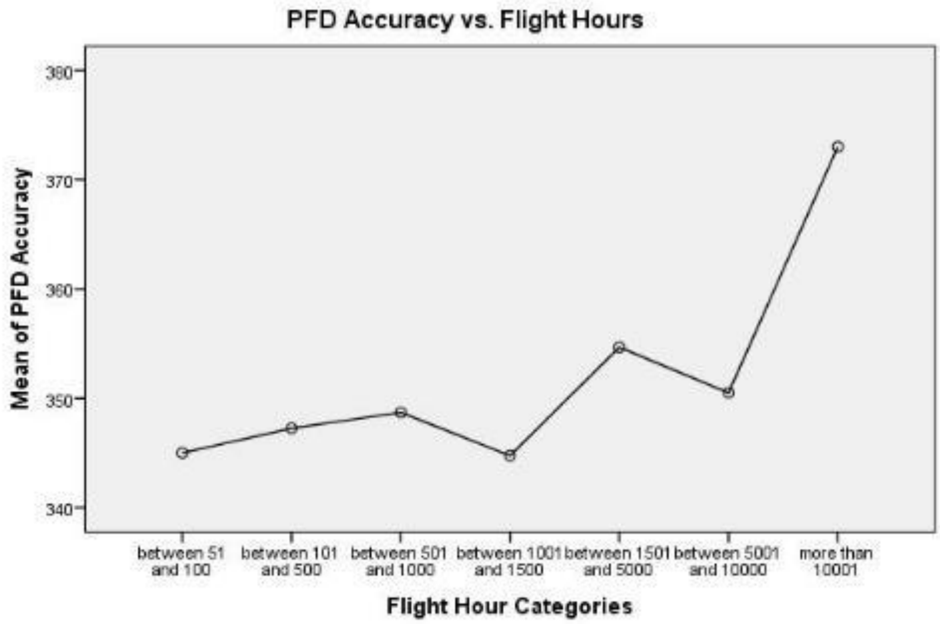

Figure 21. Plot of total correct response means to high altitude energy management decisions presented on a PFD against all seven categories of participant flight hour demographics.

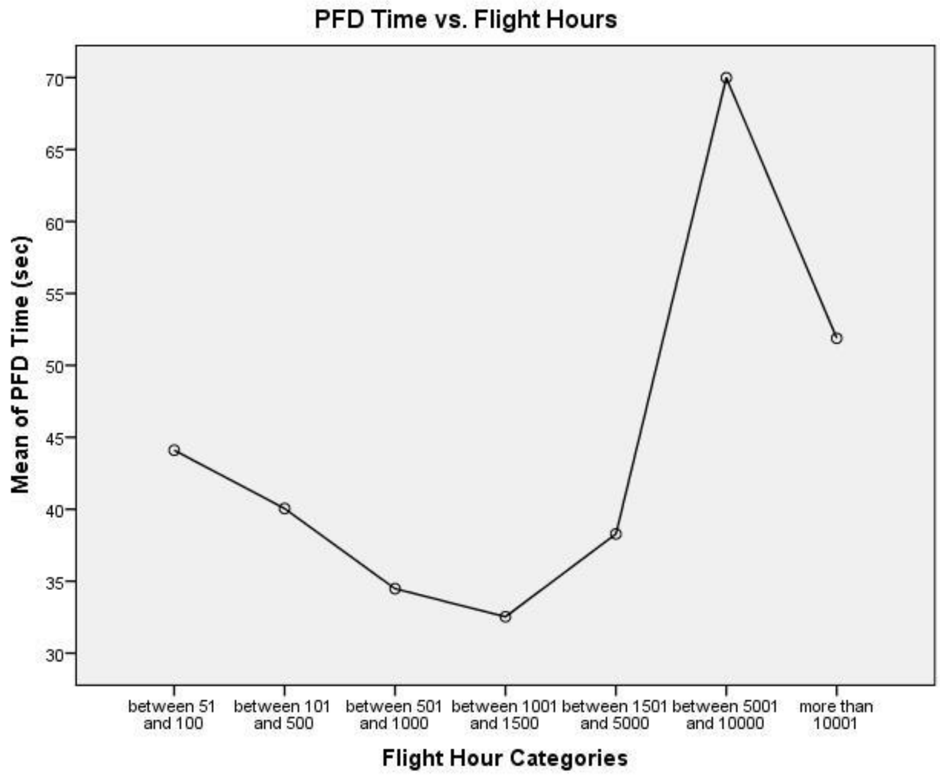

Figure 22. Plot of elapsed time means participants spent viewing and responding to high altitude energy management decisions presented on a PFD against seven selected participant flight hour demographic categories

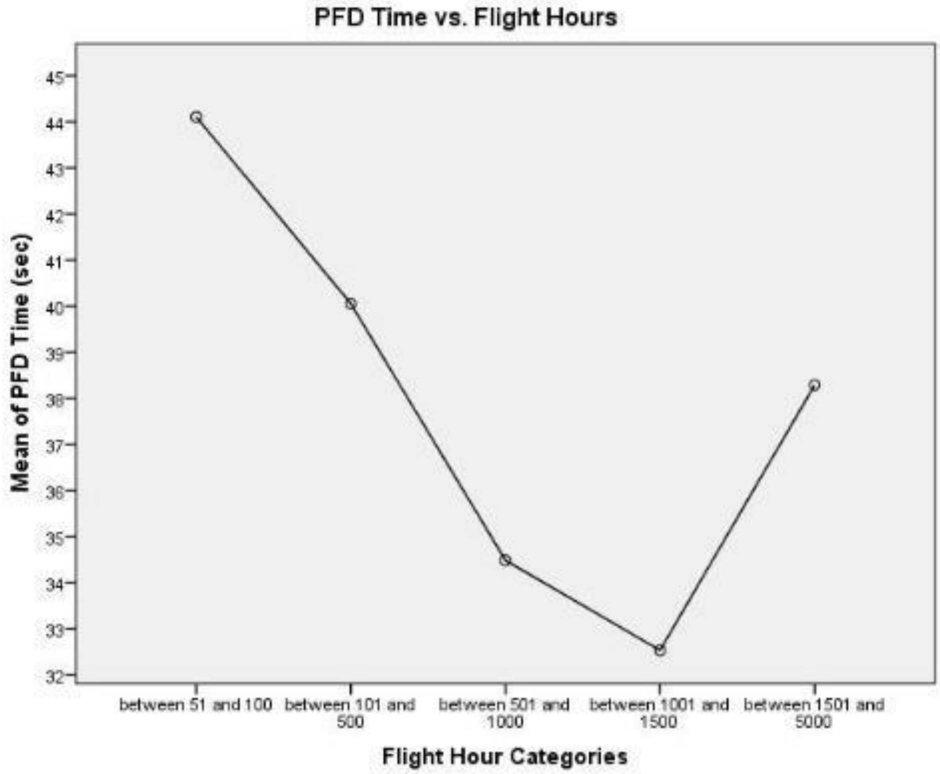

Figure 23. Plot of elapsed time means participants spent viewing and responding to high altitude energy management decisions presented on a PFD against five selected participant flight hour demographic categories.

Hypothesis 3. To evaluate Null Hypothesis 3 (Pilot flight hours in airplanes will be correlated with speed and accuracy of EM decisions when presented with the OZ concept EID display), a non-parametric Spearman's rho correlation was run between pilot flight hours in airplanes (categorical data) and OZ Time and OZ Accuracy. Neither of these two relationships were significant, as shown in Table 5. Thus, Null Hypothesis 3 was rejected.

A one-way ANOVA revealed that neither OZ Accuracy score means $F(6,51) = 1.24$, $p > .05$, $\omega = .15$ or OZ Response Time means $F(6,51) = .83$, $p > .05$, $\omega = .14$ were significantly different across all seven flight time categories, as shown in Figures 24 and 25.

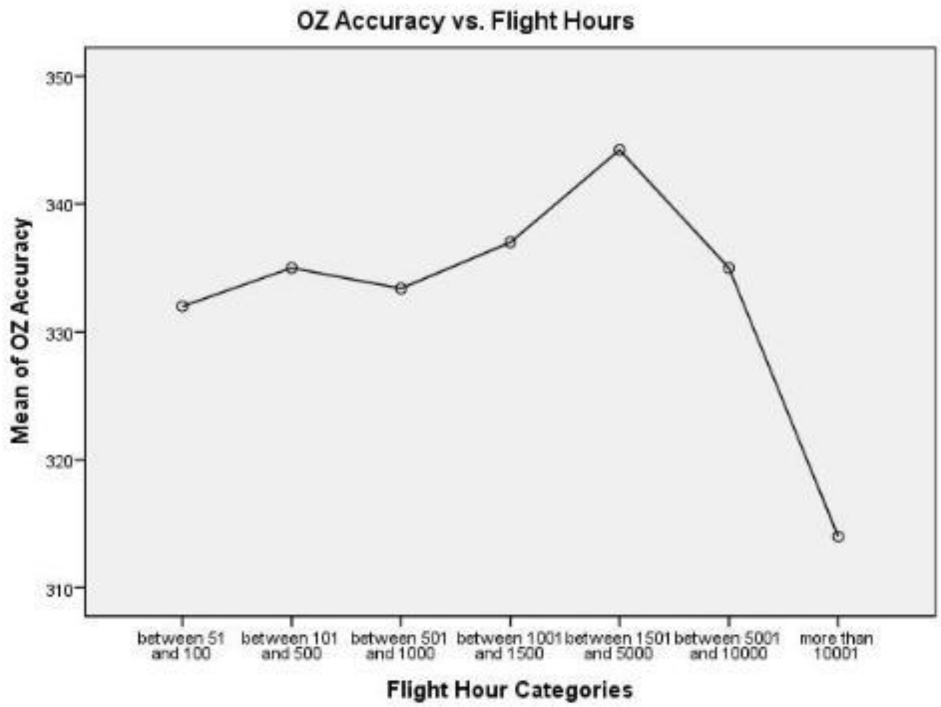

Figure 24. Plot of total correct response means to high altitude energy management decisions presented on an OZ display against all seven categories of participant flight hour demographics.

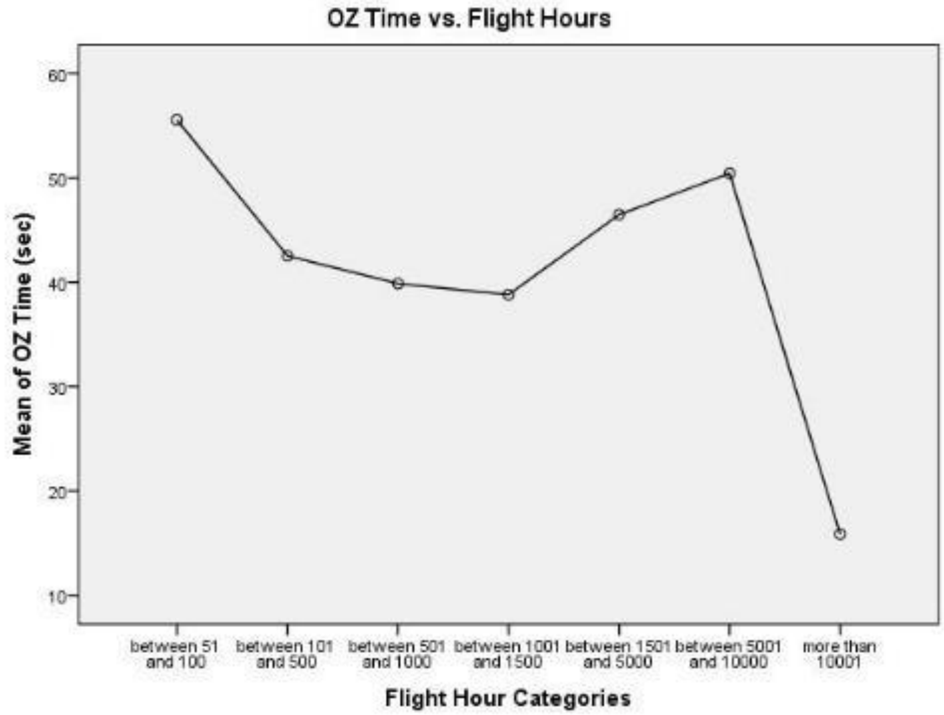

Figure 25. Plot of elapsed time means participants spent viewing and responding to high altitude energy management decisions presented on an OZ display against all seven categories of participant flight hour demographic.

Qualitative Feedback Data

The post-experiment feedback survey question responses are presented in Appendix F. Questions #1-3 allowed participants to indicate multiple responses, while questions #4-6 permitted mutually exclusive responses.

When asked about the overall study difficulty, a majority indicated it was "interesting" (58.6%), "confusing" (56.9%), and "challenging" (50.0%). Participants regarded the difficulty of the PFD display as "easy" (69.0%), "interesting" (25.9%), "challenging" (15.5%), and "fun" (15.5%). Participants regarded the difficulty of the OZ display as "confusing" (72.4%), "interesting" (53.4%), "challenging" (44.8%), "difficult" (43.1%), and "annoying" (13.8%). On a question forcing only one exclusive answer, 42

(72.5%) participants indicated they would “definitely not”, “no”, or “maybe” desire to fly an aircraft with the OZ display. Consistent to the previous question, another forced exclusive answer to the usefulness priority of an OZ display compared with a PFD display or round dial analog gauges resulted in 28 (48.3%) indicating they would prefer a PFD first, analog second, and OZ third. Nineteen (32.8%) indicated they would prefer a PFD first, followed by OZ, with analog last. Seven (12.1%) indicated they would prefer analog primarily, followed by a PFD, and lastly OZ. Only 3 (5.2%) indicated they would prefer flying an OZ aircraft ahead of either a PFD or analog cockpit.

CHAPTER V

DISCUSSION, CONCLUSIONS, AND RECOMMENDATIONS

Discussion

Hypotheses Findings.

Hypothesis 1. Mean PFD Accuracy scores were significantly higher (13 points) than mean OZ Accuracy scores, and mean PFD Response Times were 3.5 seconds faster than mean OZ Response Times (not statistically significant). This finding showed operational potential for the OZ display since even without first training to proficiency on the previously unseen OZ display, participant performance differences were not operationally remarkable from the PFD, a display with which 82.75% of participants stated a “decent amount”, “extensive”, or “expert” experience.

While Temme, Still, and Acromite (2003) and Smith (2008) studied OZ after participants were given ample time to train to a level of proficiency where their performance could not noticeably improve further, this study only provided two static pictures describing what participants would see on the OZ display. Perhaps with increased training, a more detailed tutorial prior to the experiment phase, and a check for adequate understanding prior to the stimuli, results from this study could be consistent with the superior pilot performance of OZ compared with analog gauges noticed in the earlier studies. That this study showed non-significant differences in response time and an operationally non-remarkable difference in accuracy score, provides encouragement for future OZ research in different aerodynamic regimes, different aircraft, and different pilot demographic groups.

A perfect accuracy score was 405 correct responses, indicating the means of 348 and 335 correctly matching responses, PFD and OZ respectively, equated to 85.93% and 82.72% accuracy, respectively. The high accuracy scores were undoubtedly affected by the large percentage of correct "blank" responses and suggest a different query response instrument should be used in future testing. OZ response times were only 3.5 seconds slower (approximately 8.75% of the faster PFD time), and OZ accuracy was only 13 points lower (approximately 3.21% of a perfect score of 405) than PFD accuracy, so with proper training and practice, OZ performance could potentially improve.

The positive skew of response times on both displays suggests that values were not normalized around a mean but concentrated near shorter response times and decreased in frequency as response time increased. Reasons include participants may have been rushing or guessing, leading to data contamination. Response time histograms were transformed using a square root transformation, but no remarkable results emerged from analyzing the transformed variables, probably due to the robustness of these tests. Field (2009, p. 155) quotes Glass, Peckham, and Sanders (1972) in that "the payoff of normalizing transformation in terms of more valid probability statements is low."

The significant positive correlation between OZ Time and OZ Accuracy was surprising since accuracy plotted against response time was expected to resolve similar to a normal curve with the highest accuracy scores reserved for a mean response time with decreased accuracy correlated to both tails of the curve representing very fast and very slow response times. In other words, participants responding too quickly would be expected to display poor accuracy, as well as those responding too slowly, perhaps if they had difficulty interpreting the displays and were taking a long time to respond, would also be expected to display decreased accuracy. However, this was not the case, since the correlation was moderately strong and showed that participants responding more slowly also scored more accurately. It would be interesting to study this same relationship in a new experiment that included greater training on the OZ display prior to testing. Perhaps more interesting was that PFD Accuracy and PFD Times were not correlated, meaning that there was greater variability and perhaps more randomness in response times and accuracy scores on the PFD display. Accuracy did not seem to change as response time varied suggesting that the PFD provides aircraft performance information in a meaningful and discernable manner. These findings suggest the participant familiarity of the PFD resulted in accuracy scores unrelated to response time, compared to the participants' lack of training, exposure, and experience with the OZ display where longer response times manifested in greater understanding of the OZ display.

Hypothesis 2. Hypothesis 2 was rejected as speed and accuracy of EM decisions on the PFD display were not correlated with pilot flight hours. It was thought that increased experience would translate into a faster and more accurate assessment of the aircraft stimuli; perhaps the stimuli were either too easy or too difficult to decipher, no matter the flight experience of the participant.

Hypothesis 3. Hypothesis 3 was not rejected as speed and accuracy of EM decision on the OZ display were not correlated with pilot flight hours. Participants, regardless of flight experience, fared equally well on the OZ display, yet the aggregate accuracy score was slightly worse on the OZ display than on the PFD display. This was consistent with previous research from Smith (2008) that observed novice operators performed as well as experienced professional pilots on dynamic flight tasks with the OZ display.

Qualitative Feedback. The qualitative post-experiment feedback survey revealed a comfort with the PFD and analog round dial gauges, consistent with the types of aircraft cockpit displays familiar to modern aviators like these participants and an uneasiness with the OZ display consistent since this was the first time these participants saw the OZ display and they were afforded little training or explanation of the symbology or logic behind the display. Interestingly, a similar majority of participants rated the PFD display as "easy" and the OZ display as "confusing," yet as discussed earlier, performance accuracy scores and response times between the two display types were not correspondingly different.

Limitations

Categorical Data. Demographic information was collected at the beginning of each trial as categorical data. While information such as age and flight hours could have been collected as continuous ratio data allowing for more powerful analysis using different types of procedures, the decision was made to use standard categories to make participant data entry simple. While these categories hindered the variance of the sample data since raw values were unavailable, the range of flight hours categories provided was considered adequate to sort participants into groups of flight experience for the analyses planned.

Construct Validity. This research discovered the difficulties of measuring the transient and ephemeral qualities of SA. Situation awareness is a complex construct, and there should be increased effort in attempting to align measures of energy management

SA to analogous elements such as airspeed recognition, attitude recognition, energy recognition, and corrective response selection. While response time and accuracy attempted to infer EM SA, there is no assurance that pilots demonstrating high SA will always make correct cockpit decisions and vice versa with pilots exhibiting low SA who do not always make poor choices in the cockpit, a paradox further highlighted by Endsley (1995).

Sample Size. The small sample (N=58) after filtering out those without an instrument rating was disappointing after more than 90 participants volunteered for the trial. Participants were not discouraged from taking the trial without an instrument rating. Notably, results revealed that none of the four dependent variables were significantly different across the seven flight hour categories, suggesting no contamination occurred from the 19 participants who did not hold an instrument rating but were working toward one. Nonetheless, more effort should be made on future studies to screen for the instrument rating prior to initiating the experiment to increase the generalizability of results.

Stimuli. The query response instrument could be improved in a future study. Some of the post-test feedback comments revealed that the lack of previous aircraft state context surrounding the aircraft scenario created an artificial difficulty in deciphering the current energy state and recommended corrective actions. Moreover, the PFD stimuli were difficult to analyze due to the absence of trend information on the airspeed and attitude indicators normally present on genuine PFD displays and the absence of energy information on any PFD display.

Accuracy Scoring. The method chosen to grade the response items was that a correct match to the answer key (either checked or not checked) was equally awarded one point, while disagreements (either checked when key was blank or blank when key was checked) were awarded zero points. Un-checked responses were considered correct since they matched the un-checked blank on the answer key. Since there were 640 unmarked responses (79% of the 810 total), participants could earn a correct answer simply by leaving the question blank or skipping it altogether, so accuracy scores were weighted toward those participants who left blanks due to lack of knowledge, lack of time, or other reasons. Other scoring methods were evaluated (awarding one point for correct matches but awarding zero points for correct blank matches and subtracting one point for disagreements), but the original method was retained for its tendency to accurately reward participants for correctly interpreting the displays and responding to the correct options. The inherent difficulties in scoring a multiple response query highlights the need for a better designed query response instrument to control for this phenomenon.

Response Time. Response time data were not evaluated in accordance with associated accuracy scores, so, in effect, a response time from a low accuracy screen was not discounted as invalid. For example, if a participant quickly completes a screen, recording a fast response time but a very inaccurate score, their response time data should be considered invalid. In future studies, there should be an algorithm developed that can appropriately discern a valid response time value from an invalid one.

Also, since each screen had different sets of possible answers, there is the possibility that more difficult screens would take longer to answer and vice versa. This then would invalidate the average response time data since they would not be standardized across screens or across display conditions. The within-subjects design does attempt to mitigate this by aggregating all response times for a certain display, but there is no assurance that total aggregate for each screen is equivalent in difficulty. This then should be corrected for future studies, equivalent difficulty in screens and displays, so that the participant mean difference in time and accuracy could be validly compared.

Recommendations

While this study did not find significant differences in EM SA provided by the OZ display in comparison with the PFD display, the complexities and nuances of these cognitive constructs are admittedly difficult to capture. The experience and comfort level using the two conditions (PFD and OZ) were not equal, could have affected the results, and therefore should be controlled in a future study.

This study employed static screenshots without any context or trend information, so pilot participants were at a disadvantage when they attempted to decipher the scenarios. A future study could employ dynamic simulator scenarios in that the pilot participant is fully engaged in the moments prior to the queries, providing greater SA to the pilot before they are asked to ascertain the current SA or EM situation.

While Smith (2008) studied normal flight tasks, a new area for EID display research could involve unusual attitude recovery performance using existing OZ software and Cessna flight simulators compared with traditional electronic PFD displays, as well as analog round dials, which are still predominantly in use in U.S. flight schools. An unusual attitude is akin to a flight upset, when the aircraft is off its normally expected flight trajectory, typically experienced prior to LOC-I.

A consideration for future research is in the accuracy and fidelity of the scenario stimuli themselves. The OZ software is currently only programmed to emulate a single-engine Cessna 172 in Microsoft Flight Simulator (MSFS). It would be beneficial to study pilot performance with an OZ display programmed for a common commercial airliner to analyze the effectiveness of that display to relevant professional aircraft platforms.

This study relied on the complexity of aerodynamic relationships of coffin corner flight operations where aircraft were limited by altitude, airspeed, and engine thrust. This was achievable on the PFD displays taken from a CRJ-700 on MSFS flown near operational ceiling, but accuracy of engine parameters is uncertain in that software. A more realistic depiction of aircraft operational parameters would be found in an FAA approved "Level-D" full flight simulator (FFS) under a complex aerodynamic high altitude energy scenario. A future study which employed the OZ software programmed for a twin-engine airliner and the fidelity and accuracy of a dynamic FFS should provide more clarity on the hypothesis that the OZ display could enhance EM SA.

The sample population centered on an aviation university with only slight variation in pilot experience and demographics, limiting the generalizability of this study. A larger multiple site study could investigate display differences with a more varied pilot base. Future studies could enhance reliability by running the same participants through multiple trials to ensure the reliability of the data obtained from just one run through the trial.

Increased EM SA was inferred by two factors: greater accuracy of the on-screen responses and faster response times. Subsequent research could develop a better scoring method to synthesize response time and accuracy so that a metric could account and weigh the distinction between fast / inaccurate and slow / accurate responses. A quick, yet inaccurate response is not operationally similar to a slow and accurate response.

Conclusions

Overwhelmingly, the feedback from the participants after leaving the testing center was that they were intrigued by the OZ display and wanted more time to fly and learn it. This study was created to evaluate the OZ display for its unique EID capabilities, specifically in the EM realm. While the almost 100-year-old instrument layout in cockpits worldwide will most likely not undergo a radical modification, a potential utilization of this OZ display could be realized in the flexible and customized digital real estate of advanced electronic cockpit displays. An initial hypothesis was that the OZ display, as compared with a PFD, would provide greater EM information more directly to the pilot thus improving his SA. This hypothesis was directed toward the operational theory that increased EM SA would decrease a flight crew's potential for LOC-I by avoiding aircraft upsets or correcting from them more quickly. This compound goal requires extensive research, and this study hoped to start this process by identifying potential benefits and drawbacks to certain cockpit instrumentation.

REFERENCES

Abbott, K., McKenney, D., & Railsback, P. (2013). *Operational use of flight path management systems: Final Report Recommendations of the Flight Deck Automation Working Group: Performance-based operation Aviation Rulemaking Committee and Commercial Aviation Safety Team.* Washington D.C.: FAA. Retrieved from https://www.faa.gov/about/office_org/headquarters_offices/avs/offices/afs/afs400/parc/parc_reco/media/2013/130908_parc_fltdawg_final_report_recommendations.pdf on 30 March, 2017.

Albery, W. B. (2007). Multisensory cueing for enhancing orientation information during flight. *Aviation, Space, and Environmental Medicine*, *78*(Supplement 1), B186-B190.

ASN. (2015). *ASN aviation safety database*. Retrieved from https://aviation-safety.net/databse

Belcastro, C. M., & Foster, J. V. (2010). Aircraft loss-of-control accident analysis. In *Proceedings of AIAA Guidance, Navigation and Control Conference, Toronto, Canada, Paper No. AIAA-2010-8004*.

Belcastro, C. M., & Jacobson, S. R. (2010). Future integrated systems concept for preventing aircraft loss-of-control accidents. In *AIAA Guidance, Navigation, and Control Conference*, pp. (2-5).

Boeing. (2017). *Statistical summary of commercial jet airplane accidents, worldwide operations, 1959-2016*. Aviation Safety, Boeing Commercial Airplanes. Retrieved from www.boeing.com/news/techissues/pdf/statsum.pdf.

Borst, C., Sjer, F. A., Mulder, M., Van Paassen, M. M., & Mulder, J. A. (2008). Ecological approach to support pilot terrain awareness after total engine failure. *Journal of Aircraft, 45*(1), 159-171. doi:10.2514/1.30214

Breton, R., Tremblay, S., & Banbury, S. (2007). Measurement of individual and team situation awareness: A critical evaluation of the available metrics and tools and their applicability to command and control environments (Technical Report DRDC Valcartier TR 2003-283). Valcartier, Quebec, Canada: Defence Research & Development Canada – Valcartier. Retrieved from http://pubs.drdc-rddc.gc.ca/BASIS/pcandid/www/engpub/DDW?W%3DSYSNUM=527639&r=0

Brooks, R., Ransbury, P., & Stowell, R. (2014). *Addressing on-aeroplane upset prevention & recovery training: Primary considerations for the safe and effective delivery of UPRT*. Retrieved from http://www.UPRTA.org.

Burns, C. M., Skraaning Jr, G., Jamieson, G. A., Lau, N., Kwok, J., Welch, R., & Andresen, G. (2008). Evaluation of ecological interface design for nuclear process control: situation awareness effects. *Human Factors, 50*(4), 663-679.

Carbaugh, D., Rockliff, L., & Vandel, B. (2008). *Airplane upset recovery training aid revision 2*. Retrieved from https://www.faa.gov/other_visit/aviation_industry/airline_operators/training/media/AP_UpsetRecovery_Book.pdf.

Carol, L. A. (1992). Desperately seeking SA. *TAC Attack (TAC SP 127-1), 32*(3), 5-6.

Dole, C. E., & Lewis, J. E. (2000). *Flight theory and aerodynamics* (2nd ed.). New York, NY: Wiley & Sons.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32-64.

Endsley, M. R., Bolstad, C. A., Jones, D. G., & Riley, J. M. (2003, October). Situation awareness oriented design: From user's cognitive requirements to creating effective supporting technologies. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 47, No. 3, pp. 268-272).

Endsley, M. R., & Jones, D. G. (2012). *Designing for situation awareness: An approach to user-centered design* (2nd ed.). Boca Raton, FL: CRC Press - Taylor & Francis.

Endsley, M. R., Sollenberger, R., & Stein, E. (2000). Situation awareness: A comparison of measures. *Proceedings of the First Human Performance, Situation Awareness, and Automation Conference*. Savannah, Georgia.

Eskridge, T. C., Still, D., & Hoffman, R. R. (2014). Principles for human-centered interaction design, Part 1: Performative systems. *Intelligent Systems, IEEE*, *29*(4), 88-94.

FAA. (2008a). *Airplane upset recovery training aid revision 2, supplement #1: High altitude operations*. Retrieved from http://www.faa.gov/other_visit/aviation_industry/airline_operators/training/media/Appendix_3-E_HighAltOperations.pdf

FAA. (2008b). *Pilot's handbook of aeronautical knowledge: FAA-H-8083-25A.* Oklahoma City, OK: U.S. Department of Transportation, Federal Aviation Administration, Airman Testing Standards Branch.

FAA. (2013). *Aircraft operations at altitudes above 25,000 feet mean sea level or mach numbers greater than .75, Ch. 1*. AC 61-107B.

FAA. (2015b). *Stall prevention and recovery training*. AC 120-109A.

FAA. (2015a). *Upset Prevention and Recovery Training.* AC 120-111.

Gerold, A. (2003). OZ- A revolution in cockpit display. *Avionics Magazine, 27*(10), 27-30.

Hamilton, S. (2001). Thinking outside the box at the IHMC. *Computer*, *34*(1), 61-71.

Hurt, H. H. (1965). *Aerodynamics for naval aviators* (revised ed.). (NAVWEPS 00-80T-80). Patuxent River, MD: Naval Air Systems Command.

ICAO. (2014a). *Manual on aeroplane upset prevention and recovery training*. (Doc 10011, AN/506).

ICAO. (2014b). *Loss of control in-flight symposium*. Retrieved from http://www.icao.int/Meetings/LOCI/Pages/default.aspx

Lambregts, A. A., Nesemeier, G., Wilborn, J. E., & Newman, R. L. (2008). *Airplane upsets: Old problem, new issues*. AIAA Modeling and Simulation Technologies Conference and Exhibit, AIAA 2008-6867.

Merkt, J. R. (2013). Flight energy management training: Promoting safety and efficiency. *Journal of Aviation Technology and Engineering, 3*(1), 24-36.

NTSB. (2015). *Aviation accident database & synopses*. Retrieved from http://www.ntsb.gov/_layouts/ntsb.aviation/index.aspx on April 24, 2015.

Sarter, N. B., & Woods, D. 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.

Sharp, T. D., & Helmicki, A. J. (1998). The application of the ecological interface design approach to neonatal intensive care medicine. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (Vol. 42, No. 3, p. 350-354). Los Angeles, CA: SAGE Publications.

Sherry, L., & Mauro, R. (2014). Controlled Flight into Stall (CFIS): Functional complexity failures and automation surprises. In *Integrated Communications, Navigation and Surveillance Conference (ICNS), 2014* (pp. D1-1).

Smith, C. F. (2008). *The effect of functional display information on the acquisition and transfer of novice piloting knowledge* (Doctoral dissertation). Retrieved from ProQuest Dissertations & Theses Global. (Order No. 3321174)

Smith, C. F., & Boehm-Davis, D. A. (2005). Improving novice flight performance using a functional display. In *Proceedings of the International Society of Aviation Psychology 13th Annual Meeting*, *51*(25), 1583-1587, Oklahoma City, OK.

Smith, C. F., Boehm-Davis, D. A., & Chong, R. (2004). Improving piloting performance through the use of a functional aviation display. In *Proceedings of the International Society of Aviation Psychology 13th Annual Meeting*, Oklahoma City, OK.

Smith, C. F., Fadden, S., & Boehm-Davis, D. A. (2005). Use of a functional aviation display under varying workload conditions. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*, *49*(1), 59-63. Santa Monica, CA: HFES.

Temme, L. A., Still, D. L., & Acromite, M. (2003). OZ: A human-centered computing cockpit display. In *45th Annual Conference of The International Military Testing Association*, Pensacola, FL, p. 70-90.

Tenney, Y. J., Adams, M. J., Pew, R. W., Huggins, A. W., & Rogers, W. H. (1992.) *A principled approach to the measurement of situation awareness in commercial aviation*. (NASA contractor report 4451). Langley Research Center: NASA.

Vicente, K. J. (2002). Ecological interface design: Progress and challenges. *Human Factors, 44*(1), 62-78. doi:10.1518/0018720024494829

Williamson, G. W. (1937). The new service blind flying panel described: Six essential instruments correctly grouped. *Flight, 1495*(XXXII), 193-195.

APPENDIX A

Permission to Conduct Research

Embry-Riddle Aeronautical University
Application for IRB Approval
Expedited Determination

Principle Investigator: Erik Baker **Other Investigators:** Dr. Haydee Cuevas **Role:** Student **Campus:** Daytona Beach **College:** COA

Project Title: ***Comparative Analysis of Conventional Electronic and OZ Concept Display for Aircraft Energy Management***

Submission Date: 5/1/2017 **Determination Date:** 5/12/2017

Review Board Use Only

Exempt: No

Approved:

Mike Wiggins	M.B. McLatchey	May 23, 2017 Expires: May 22, 2018
Pre-Reviewer Signature	Chair of the IRB Signature	Date of Approval / Expiration Date

Brief Description: The purpose of this study is to examine the effectiveness of two different aircraft cockpit displays in promoting energy management situation awareness. Using computer images and a questionnaire.

This research falls under the **expedited** category as per 45 CFR 46.110 (b) because one or both of the following apply:

(1) ☒ some or all of the research appearing on the list below are found by the reviewer(s) to involve no more than minimal risk,

(2) ☐ minor changes in previously approved research during the period (of one year or less) for which approval is authorized.

Research activities that (1) present no more than minimal risk to human subjects, and (2) involve only procedures listed in one or more of the following categories. The activities listed should not be deemed to be of minimal risk simply because they are included on this list. Inclusion on this list merely means that the activity is eligible for review through the expedited review procedure when the specific circumstances of the proposed research involve no more than minimal risk to human subjects. (Bankert & Amdur 2006)

Bankert, E. A., Amdur, R. J., (2006) *Institutional Review Board Management and Function, Second Edition*, pp. 517-518.

1. ☐ Prospective collection of biological specimens for research purposes by noninvasive means.

2. ☐ Collection of data from voice, video, digital, or image recordings made for research purposes.

3. ☒ Research on individual or group characteristics or behavior (including, but not limited to, research on perception, cognition, motivation, identity, language, communication, cultural beliefs or practices, and social behavior) or research employing survey, interview, oral history, focus group, program evaluation, human factors evaluation, or quality assurance methodologies. (NOTE: Some research in this category may be exempt from the HHS regulations for the protection of human subjects 45 CFR 46.101(b)(2) and (b)(3). This listing refers only to research that is not exempt.) [This means research that presents *more than minimal risk to human subjects*.]

Bankert, E. A., Amdur, R. J., (2006) *Institutional Review Board Management and Function, Second Edition*, pp. 517-518.

LEWIS UNIVERSITY
A Catholic and Lasallian University

January 30, 2017

MEMORANDUM

TO: Erik Baker,
Department of Aviation and Transportation
College of Arts and Sciences

FR: Erin Zimmer, PhD
Institutional Review Board, Chair

RE: Research proposal involving human subjects for the project entitled "Analyzing Energy Management Situation Awareness in OZ Concept Display"

This is to inform you that the above-named application for human subjects research has been approved by Institutional Review Board (IRB) at Lewis Univeristy. Although you may begin data collection immediately, please be advised that federal regulations require the IRB be made aware of all research activities that place human subjects at maximum or minimum risk.

The approval is effective for one year from the date of this letter. If your project will continue beyond that date, or if you intend to make modifications to this study, you will need additional approval and should contact IRB for assistance. You will need to continue to get annual approval and supply progress updates on the project until you no longer retain any identifiers that could link the subjects to the data.

It is important for you to note as the primary investigator involved with human subject research, you are responsible for ensuring that this project has current IRB approval at all times. You are also responsible for retaining the signed consent forms obtained from your subjects for a minimum of three years after the study is complete. In addition, you are required to report to the IRB any injuries, unanticipated problems or risks to the subjects as well as any changes to the protocol as outlined in your proposal.

Best wishes for success in your research projects.

EZ

cc: Office of the Provost

One University Parkway • Romeoville, IL 60446-2200
(800) 837-9000 • (815) 838-0500 • *lewisu.edu*

CAMPUS SITES: ROMEOVILLE • CHICAGO • HICKORY HILLS • OAK BROOK • TINLEY PARK • ALBUQUERQUE, NM

APPENDIX B

Informed Consent Form

AGREEMENT TO PARTICIPATE IN

Comparative Analysis of Conventional Electronic and OZ Concept Display for Aircraft Energy Management

STUDY LEADERSHIP. We are asking you to take part in a research project that is led by Erik Baker, Ph.D. Candidate, Embry-Riddle Aeronautical University.

PURPOSE. The purpose of this study is to examine the effectiveness of two different aircraft cockpit displays in promoting energy management situation awareness. We hope to use what we learn from this study to enhance flight safety by quantifying effectiveness of energy management situation awareness in general aviation operations.

ELIGIBILITY. To be in this study, you must be 18 years or older.

PARTICIPATION. During the study, you will be asked to take part in an experimental session where you will be seated at a computer terminal.

1. You will observe a computer monitor, respond to an on-screen question, and then repeat this procedure about 30 times. The experimental session will take about 20 minutes of your time.
2. This computer-based evaluation may include: (a) determining flight information from the on-screen display, and (b) responding to on-screen questions with a pointer device (mouse).
3. We will ask you to respond to a demographic questionnaire prior to starting the experimental session. After the experiment has finished, we will also ask you to respond to a satisfaction survey regarding the nature of the cockpit displays used in the trial.

RISKS OF PARTICIPATION. The risks of participating in this study are minimal, no more than in everyday life.

BENEFITS OF PARTICIPATION. We do not expect the study to benefit you personally; however, the research may help us learn how to improve aviation safety for the general aviation community.

VOLUNTARY PARTICIPATION. Your participation in this study is completely voluntary. You may stop or withdraw from the study at any time or refuse to answer any particular question without it being held against you. Your decision whether or not to participate will have no effect on your current or future connection with anyone at Embry-Riddle Aeronautical University or Lewis University. Any data collected from participants who “opt-out” before or during the experiment will be considered incomplete but will remain CONFIDENTIAL and be kept in the same manner as complete participant data, further explained below.

RESPONDENT PRIVACY. Your individual information will be protected in all data resulting from this study. Your responses to this survey will be CONFIDENTIAL. In order to protect the confidentiality of your responses, any information that is obtained in connection with this study and that can be identified with you will remain strictly confidential and will be disclosed only with your permission or as required by law. Confidentiality will be maintained by means of assigning a code number for all data collected from each participant. We will use only the code number, and therefore no names or identifying information will be used in this study or in any of the research reports. Results will be reported at the group level and will not identify you individually.

***Confidentiality** means that you will know or can readily learn the participant's identity, but you will not disclose or make it possible for anyone outside of the research team to learn it.*

FURTHER INFORMATION. If you have any questions or would like additional information about this study, please contact Erik Baker, baker7fa@my.erau.edu, 815-836-5936.

The ERAU Institutional Review Board (IRB) has approved this project. You may contact the ERAU IRB with any questions or issues at (386) 226-7179 or teri.gabriel@erau.edu. ERAU's IRB is registered with the Department of Health & Human Services – Number – IORG0004370.

CONSENT. Your selection of the on screen "AGREE" radio button will replace your signature, indicate that you understand the information on the form, that someone has answered any and all questions you may have about this study, that you are 18 years of age or older, and that you voluntarily agree to participate in it. You may print a copy of this form for your records. A copy of this form can also be requested from Erik Baker, baker7fa@my.erau.edu, 815-836-5936.

I DO NOT agree to participate

I agree to participate

APPENDIX C

Demographic Questionnaire

Demographic Questionnaire

1) "What is your gender?"

"Female",
"Male",
"I prefer not to answer this question",
"Other (please specify):"

2) "What is your age?"

"18-24 years old",
"25-34 years old",
"35-44 years old",
"45-54 years old",
"55-64 years old",
"65-74 years old",
"75 years or older",
"I prefer not to answer this question"

3) "What is your ethnicity?"

"Hispanic or Latino (a person of Cuban, Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race)",
"Not Hispanic or Latino",
"I prefer not to answer this question"

4) "What is your race? Mark one or more races to indicate what you consider yourself to be."

"American Indian or Alaska Native (a person having origins in any of the original peoples of North and South America (including Central America) who maintains cultural identification through tribal affiliation or community attachment)",
"Asian (a person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian Subcontinent, including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam)",
"Black or African American (a person having origins in any of the black racial groups of Africa)",
"Native Hawaiian or Other Pacific Islander (a person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands)",
"White (a person having origins in any of the original peoples of Europe, the Middle East, or North Africa)",
"I prefer not to answer this question",
"Info you would like to add:"

5) "What is the highest degree or level of schooling you have completed? If currently enrolled, highest degree received so far."

"Doctoral or professional degree",
"Master's degree",
"Bachelor's degree",
"Associate's degree",
"Postsecondary non-degree award",
"Some college, no degree",
"High school diploma or equivalent",
"Less than high school",
"I prefer not to answer this question"

6) "What is your current employment status? Mark all that apply."

"University student (non-aviation)",
"University student (aviation)",
"Flight Instructor",
"Chief or Asst. Chief Pilot",
"Commercial pilot",
"Corporate pilot",
“Cargo pilot”,
"General Aviation pilot",
"Faculty (aviation)",
"Faculty (non-aviation)",
"I prefer not to answer this question",
"Other (Please specify):"

7.) “What FAA medical certificate do you currently hold?”

“None”,
“Third class medical”,
“Second class medical”,
“First class medical”,
“I prefer not to answer this question”

8) "What pilot ratings do you currently hold? Mark the highest or all that apply."

“None”,
"Student",
"Recreational",
"Sport",
"Private",
"Instrument",
"Commercial",
"Certified Flight Instructor",

"Certified Flight Instructor Instrument",
"Multi-Engine Instructor",
"Airline Transport Pilot",
"Remote Pilot",
"Dispatcher",
"I prefer not to answer this question",
"Other (please specify):"

9) "How many flight hours do you have?"

"none",
"between 1 and 50",
"between 51 and 100",
"between 101 and 500",
"between 501 and 1000",
"between 1001 and 1500",
"between 1501 and 5000",
"between 5001 and 10000",
"more than 10001",
"I prefer not to answer this question"

10) "How much experience do you have with electronic flight displays (EFIS, PFD, glass, etc.)?"

"None",
"A little",
"Decent amount",
"Extensive",
"Expert",
"I prefer not to answer this question"

11) "How much experience do you have with high altitude flight (above FL250)?"

"None",
"A little",
"Decent amount",
"Extensive",
"Expert",
"I prefer not to answer this question"

12) "How much experience do you have with video gaming?"

"None",
"A little",
"Decent amount",
"Extensive",

"Expert",
"I prefer not to answer this question"

13) "How much experience do you have with the OZ display?"

"Never heard of it",
"A little",
"Decent amount",
"Extensive",
"I have previously participated in this trial",
"I prefer not to answer this question"

APPENDIX D

Data Collection Device

Energy Management Trials

1) "Indicate the condition. Mark all that apply."

"Climbing",
"Descending",
"Level",
"Left bank",
"Right bank",
"Other (Please specify):"

"Accelerating",
"Decelerating",
"Constant",
"Near overspeed",
"Near stall",
"Other (Please specify):"

"Power avail > power req",
"Power avail = power req",
"Power avail < power req",
"Other (Please specify):"

2) "Indicate the response. Mark all that apply."

"Add thrust",
"Reduce thrust",
"Pitch up",
"Pitch down",
"Roll wings level",
"No input required",
"Other (Please specify):"

APPENDIX E

Post-experiment Feedback Survey

Feedback Survey

1) "I found this study to be: __________. Mark all that apply."

"Easy",
"Difficult",
"Confusing",
"Interesting",
"Challenging",
"Exciting",
"Fun",
"Annoying",
"I prefer not to answer this question",
"Other (please specify):"

2) "I found the PFD display to be: __________. Mark all that apply."

"Easy",
"Difficult",
"Confusing",
"Interesting",
"Challenging",
"Exciting",
"Fun",
"Annoying",
"I prefer not to answer this question",
"Other (please specify):"

3) "I found the OZ display to be: __________. Mark all that apply."

"Easy",
"Difficult",
"Confusing",
"Interesting",
"Challenging",
"Exciting",
"Fun",
"Annoying",
"I prefer not to answer this question",
"Other (please specify):"

4) "I would like to fly an aircraft with the OZ display."

"Definitely not",
"No",
"Maybe",
"Yes",
"Definitely yes",
"I prefer not to answer this question",
"Other (please specify):"

5) "How do you compare the usefulness of the OZ display to a PFD or traditional gauges?"

"OZ best, PFD next, round dials last",
"OZ best, round dials next, PFD last",
"PFD best, OZ next, round dials last",
"PFD best, round dials next, OZ last",
"Round dials best, OZ next, PFD last",
"Round dials best, PFD next, OZ last",
"I prefer not to answer this question",
"Other (please specify):"

6) "How do you rate the usefulness of the OZ display for energy management (EM)?"

"Not useful at all",
"Useful for EM, but not for anything else",
"Useful for EM and some other aviation tasks",
"Useful for EM and most aviation tasks",
"Very useful for all aviation tasks",
"Essential",
"I prefer not to answer this question",
"Other (please specify):"

APPENDIX F

Post-experiment Feedback Results

Q1) Response to Study Difficulty

Response	*N*	Percent
Easy	3	5.2
Difficult	9	15.5
Confusing	33	56.9
Interesting	34	58.6
Challenging	29	50.0
Exciting	11	19.0
Fun	12	20.7
Annoying	2	3.4
Needed more info	1	1.7

Q2) Response to PFD Difficulty

Response	*N*	Percent
Easy	40	69.0
Difficult	2	3.4
Confusing	3	5.2
Interesting	15	25.9
Challenging	9	15.5
Exciting	4	6.9
Fun	9	15.5
Annoying	2	3.4
Became easier	1	1.7
Easier to understand	2	3.4
Missing trend info	1	1.7

Q3) Response to OZ Difficulty

Response	*N*	Percent
Easy	7	12.1
Difficult	25	43.1
Confusing	42	72.4
Interesting	31	53.4
Challenging	26	44.8
Exciting	8	13.8
Fun	5	8.6
Annoying	8	13.8
Great concept	1	1.7
Liked target A/S	1	1.7
Display sweet	1	1.7
Easier with practice	1	1.7

Q4) Response to Desire to Fly an Aircraft with OZ

Response	*N*	Percent
Definitely not	12	20.7
No	11	19.0
Maybe	19	32.8
Yes	9	15.5
Definitely yes	5	8.6
Yes with more practice	2	3.4

Q5) Response to Priority of Usefulness

Response	*N*	Percent
PFD, Round, OZ	28	48.3
PFD, OZ, Round	19	32.8
Round, PFD, OZ	7	12.1
OZ, PFD, Round	3	5.2
Prefer not to answer	1	1.7

Q6) Response to Usefulness of OZ for EM

Response	*N*	Percent
Not useful	13	22.4
Useful for EM	12	20.7
Useful for some	19	32.8
Useful for most	5	8.6
Very useful for all	4	6.9
Essential	1	1.7
Confusing	1	1.7
Need to understand better	1	1.7
Useful for military	1	1.7
Need more practice	1	1.7