2017

A Study of Human-Machine Interface (HMI) Learnability for Unmanned Aircraft Systems Command and Control

Tom Haritos
Nova Southeastern University, tharitos@k-state.edu

NSUWorks Citation

Tom Haritos. 2017. *A Study of Human-Machine Interface (HMI) Learnability for Unmanned Aircraft Systems Command and Control.* Doctoral dissertation. Nova Southeastern University. Retrieved from NSUWorks, College of Engineering and Computing. (1018)
https://nsuworks.nova.edu/gscis_etd/1018.

A Study of Human-Machine Interface (HMI) Learnability for Unmanned Aircraft Systems Command and Control

by

Tom Haritos

A dissertation submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy
in
Computing Technology in Education

College of Engineering and Computing
Nova Southeastern University
2017

We hereby certify that this dissertation, submitted by Tom Haritos, conforms to acceptable standards and is fully adequate in scope and quality to fulfill the dissertation requirements for the degree of Doctor of Philosophy.

Laurie Dringus — 11/07/2017

Laurie Dringus, Ph.D. — Date

Chairperson of Dissertation Committee

Martha M. Snyder — 11/07/2017

Martha M. Snyder, Ph.D. — Date

Dissertation Committee Member

11/7/2017

Ling Wang, Ph.D. — Date

Dissertation Committee Member

Approved:

11/07/2017

Yong X. Tao, Ph.D., P.E., FASME — Date

Dean, College of Engineering and Computing

College of Engineering and Computing
Nova Southeastern University

2017

An Abstract of a Dissertation Submitted to Nova Southeastern University in Partial
Fulfillment of the Requirements for the Degree of Doctor of Philosophy

A Study of Human-Machine Interface (HMI) Learnability for Unmanned Aircraft Systems Command and Control

by

Tom Haritos

October 2017

The operation of sophisticated unmanned aircraft systems (UAS) involves complex interactions between human and machine. Unlike other areas of aviation where technological advancement has flourished to accommodate the modernization of the National Airspace System (NAS), the scientific paradigm of UAS and UAS user interface design has received little research attention and minimal effort has been made to aggregate accurate data to assess the effectiveness of current UAS human-machine interface (HMI) representations for command and control. UAS HMI usability is a primary human factors concern as the Federal Aviation Administration (FAA) moves forward with the full-scale integration of UAS in the NAS by 2025.

This study examined system learnability of an industry standard UAS HMI as minimal usability data exists to support the state-of-the art for new and innovative command and control user interface designs. This study collected data as it pertained to the three classes of objective usability measures as prescribed by the ISO 9241-11. The three classes included: (1) effectiveness, (2) efficiency, and (3) satisfaction. Data collected for the dependent variables incorporated methods of video and audio recordings, a time stamped simulator data log, and the SUS survey instrument on forty-five participants with none to varying levels of conventional flight experience (i.e., private pilot and commercial pilot).

The results of the study suggested that those individuals with a high level of conventional flight experience (i.e., commercial pilot certificate) performed most effectively when compared to participants with low pilot or no pilot experience. The one-way analysis of variance (ANOVA) computations for completion rates revealed statistical significance for trial three between subjects [F (2, 42) = 3.98, p = 0.02]. Post hoc t-test using a Bonferroni correction revealed statistical significance in completion rates [*t* (28) = -2.92, p<0.01] between the low pilot experience group (M = 40%, SD =. 50) and high experience group (M = 86%, SD = .39). An evaluation of error rates in parallel with the completion rates for trial three also indicated that the high pilot experience group committed less errors (M = 2.44, SD = 3.9) during their third iteration when compared to the low pilot experience group (M = 9.53, SD = 12.63) for the same trial iteration.

Overall, the high pilot experience group (M = 86%, SD = .39) performed better than both the no pilot experience group (M = 66%, SD = .48) and low pilot experience group (M = 40%, SD =.50) with regard to task success and the number of errors committed. Data collected using the SUS measured an overall composite SUS score (M = 67.3, SD = 21.0) for the representative HMI. The subscale scores for usability and learnability were 69.0 and 60.8, respectively.

This study addressed a critical need for future research in the domain of UAS user interface designs and operator requirements as the industry is experiencing revolutionary growth at a very rapid rate. The deficiency in legislation to guide the scientific paradigm of UAS has generated significant discord within the industry leaving many facets associated with the teleportation of these systems in dire need of research attention.

Recommendations for future work included a need to: (1) establish comprehensive guidelines and standards for airworthiness certification for the design and development of UAS and UAS HMI for command and control, (2) establish comprehensive guidelines to classify the complexity associated with UAS systems design, (3) investigate mechanisms to develop comprehensive guidelines and regulations to guide UAS operator training, (4) develop methods to optimize UAS interface design through automation integration and adaptive display technologies, and (5) adopt methods and metrics to evaluate human-machine interface related to UAS applications for system usability and system learnability.

Acknowledgments

There are many individuals that I would like to acknowledge for providing support throughout this dissertation process. First, I would like to extend sincere gratitude to my dissertation advisor Dr. Laurie Dringus for working with me on this project. Thank you for the continued encouragement and guidance as I completed this dissertation. I wish to also extend great thanks to my dissertation committee members Dr. Ling Wang and Dr. Martha Snyder for agreeing to work with me on this research project.

To my wife Krista, thank you for supporting me as I completed my doctoral work. The love and support you have provided for our family and children, Victoria and Costa, as I completed this degree did not go unnoticed. Thank you.

Victoria, you were two years old when I ventured into this doctoral program. You have blossomed into a beautiful young lady. Thank you for your patience and understanding. I hope I have been a good role model. Costa, you were born in the midst of this dissertation and you have grown before my eyes and faster than I could have even imagined. Costa is three years old. Costa, thank you for being you. Last, I would like to thank my parents for their continued support and encouragement. Thank you all.

Table of Contents

Chapters

Appendices

List of Tables

Tables

List of Figures

Figures

Chapter 1

Introduction

Background

In the last two decades unmanned aircraft systems (UAS) have served an important role to leverage United States (U.S.) military efforts in the Middle East and other parts of the world. Unmanned aircraft systems have been instrumental for the Department of Defense (DOD) in intelligence, surveillance, and reconnaissance information gathering (ISR). More recently, UAS have transitioned from military to civilian and a number of viable public and commercial applications have emerged.

According to the U.S. Government Accountability Office (2012; 2015) practical and sustainable UAS applications for public use include: (1) law enforcement surveillance and intelligence gathering, (2) search and rescue, (3) wildfire monitoring and tracking, (4) ecosystem conservation, (5) power-line inspection (6) weather research and (7) remote sensing for geospatial sciences. Practical applications for monetary gain are often described as commercial applications and include: (1) videography and photography, (2) agriculture, (3) maritime (4) land cover mapping (5) surveying, (6) engineering and (7) education.

Unfortunately, the UAS industry is at an infant state regarding policy and legislation leaving many critical issues that influence the safe and effective integration of UAS into the National Airspace System (NAS) unresolved. At present, the NAS is undergoing a significant process of evolutionary change to maintain stride with the development of new aircraft technologies and the increased air traffic demands projected to occur by the year 2025 (Federal Aviation Administration, 2009). The expansion and modernization of the NAS requires new methods to manage and monitor the increased air traffic, the impact on airport capacity, the increased workload associated with air traffic controllers, and the potential for full-scale commercial UAS operations (U.S. GAO, 2012).

As prescribed in U.S. Code: Title 49 subpart 106, the Federal Aviation Administration (FAA) is the governing agency allocated by the Department of Transportation (DOT) to enforce the rules regarding the safe operation of aircraft in the NAS. Under this prescribed criteria, the FAA requires all aircraft operations in the NAS to be conducted by a licensed pilot in a registered and certified aircraft to ensure airworthiness prior to attaining operational approval for flight in the NAS. As regards UAS, the unmanned aircraft (UA) and the corresponding elements required for command and control (i.e., control station, communications and navigation equipment, data links, telemetry, and other associated support equipment) must also adhere to Title 49 U.S. Code § 44103 titled: registration of aircraft. Therefore, UAS operations in the U.S. for large category UAS is heavily restricted by the FAA for safety-related concerns of the public and users of the NAS (GAO, 2012).

Safety-related issues stem from the absence of common policies and regulations and promulgate to design, manufacturing, and operating inadequacies, training inefficiencies, inconsistencies in physical and logical control orientation, and irregularities in display design technology, terminology, and symbology (Terwilliger, Ison, Vincenzi & Liu, 2014; GAO, 2012; Maybury, 2012; Cooke, 2008; Tvaryanas, Thompson, & Constable, 2005). Despite the inadequacies, the FAA has carved a path for federal, state, and local government agencies to obtain a Certificate of Authorization or Waiver (COA) to operate UAS in the NAS for research, training, or both under explicit provisions described by the COA agreement. A COA is an authorization issued by the FAA Air Traffic Organization (ATO) to a public operator for a specific UAS activity within a predefined sector of airspace using a specific air vehicle (GAO, 2012). At present, the Department of Defense (DoD) is the largest public use UAS operator by means of the COA. The COA agreement serves as temporary solution for UAS integration into the NAS as the FAA continues to explore the safest and most appropriate measures (i.e., equipment requirements, policy, regulations, and prescribed standardized training criteria; GAO, 2012) for the full-scale integration of UAS into the NAS.

In 2015, the FAA released a Notice for Proposed Rulemaking (NPRM) for non-government applications (i.e., civilian operators) and as regards the operation of UAS that weigh below 55 pounds (i.e., small UAS; sUAS) to perform activities for monetary incentive by applying to the FAA for an exemption to section 333 of Public Law 112-95. Public Law 112-95 is often cited and referred to as the FAA Modernization and Reform act of 2012. Public Law 112-95 prescribes legislation to improve the safety of aviation by manipulating the capacity of the NAS, to provide a framework for integrating new technologies such as air-borne based navigation, to accommodate increased traffic demands and the integration of UAS by advancing the technology requirements for certain segments of the NAS.

According to Vincenzi, et al. (2015), requests for Section 333 exemptions were reviewed by the FAA on a case by case basis for low-risk UAS operations and were accompanied by a blanket COA limiting altitude to 200 feet above the ground, five miles away from an airport with a control tower, and away from populated areas. Although restrictive, the Section 333 served as a mechanism to accommodate the growing consumers of small UAS (sUAS) who wish to implement commercial off the shelf UAS technologies for a variety of professional job tasks.

More recently in 2016, the FAA amended its regulations to Title 14: Aeronautics and Space and added Part 107: Small Unmanned Aircraft Systems. As defined in Part 107, a sUAS is a remotely operated aircraft that weighs less than 55 pounds. The new rule provides a small segment of the UAS industry with initial operating procedures, safety rules for certain operations in the NAS, and addresses topics such as airspace restrictions, certification requirements for sUAS operators, and operational limits as a means to promote the safety of the NAS (U.S. Government Publishing office, 2017).

Unfortunately, 14 CFR Part 107 does not pertain to systems in excess of 55 pounds and specifically, does not pertain to medium altitude long endurance (MALE) UAS which is the primary focus of this research. Another limitation to Part 107 as noted by Jimenez, et al. (2016) is the fact the regulations do not address interface design nor do the regulations provide recommendations for the necessary aspect of information required for sUAS operators. Jimenez, et al. (2016) suggested that identifying the types of information required for sUAS operators across various flight regimes is a critical element to designing optimal HMI for sUAS applications. They recommend the implementation of standardized displays similar to that of conventional aviation where standardized displays are mandated to optimize the safety of sUAS.

Although a noteworthy effort on the part of the FAA to implement sUAS regulations, further research is deemed necessary to address issues that remain unresolved. For instance, access to airspace for large scale UAS and UAS operations is still prohibited due to the potential safety challenges and associated risks larger systems pose when sharing airspace with conventional aircraft (Tvaryanas, Thompson, & Constable, 2005; GAO, 2008). Issues as regards the curbed integration of larger systems as identified by the GAO (2012) include: (1) the inability for UAS to detect, sense, and avoid other aircraft and airborne obstacles, (2) vulnerabilities in the command and control paradigm, (3) limited human factors engineering incorporated in UAS technologies, (4) lack of standards to guide the safe integration of UAS, and (5) the lack of capability to transition UAS into the Next Generation Airspace System (NextGen). Interestingly, these issues are of pertinent matter to all UAS and UAS operations as solutions to the aforementioned problems hold the key to the safe and effective full-scale integration of UAS into the NAS regardless of vehicle size and weight.

The unresolved problem areas associated with UAS must be considered early in the integration process as there are certainly several operational advantages from removing the human from the aircraft; however, the operational benefits also come with a surplus of risk (Reynolds, et al. 2011) As an example, human-machine interaction (HMI) has been identified as one of the primary human factors concerns with UAS applications as the operator and vehicle are segregated. Modes of information presentation and information exchange have been described as non-effective and non-efficient (Vincenzi, et al., 2015; Terwilliger, et al., 2014; Maybury, 2012; Cooke, 2008; Williams, 2004). In fact, current HMI representations are a direct reflection of early development and rapid prototyping where manufacturers concentrated solely on the air vehicle and sensor payload with little emphasis placed on the HMI for command and control (Vincenzi, et al., 2015). Since the Department of Defense (DoD) has been the largest consumer of tactical-close range and medium altitude long endurance UAS, the designs have been greatly influenced by DoD requirements. To make matters worse, Vincenzi, et al., (2015) ascertained the designs for UAS have been service-branch specific and numerous systems have been developed in the last two decades without any coordination among UAS manufacturers or across the military branches of the DoD. The resultant outcome as often reported is a high variability in systems design and a leading human factors concern regarding the command and control paradigm for UAS operations in the NAS. According to Cooke (2008) human performance degradation in terms of erroneous control inputs, over correction in direct control, and overall flight degradation due to a lack of situational awareness was a leading cause of procedural errors in part of the crew.

Vincenzi, et al. (2015) suggested the trickling design paradigm of early HMI development has not only stifled initiatives for research to develop more modern and usable UAS HMI but a multitude of antiquated design representations for command and control continue to appear on the marketplace. Most often, these designs fail to meet consistent design standards limiting usability and interoperability across platforms. In fact, legacy interface as regards the command and control loop for UAS has been unchanged for decades leaving sophisticated air vehicles to be operated with non-sophisticated user interface technologies (Vincenzi, et al., 2015) and more recently with the advent of sUAS, with technologies better suited for recreational use (e.g., tablets and smartphones).

Holden, et al. (2013) defined human-computer interaction (HCI) as the methods by which humans and computer-based applications communicate, share, and transfer information to accomplish and fulfill operational objectives and job tasks. Alternately, information architecture (IA) is the classification, categorization, and presentation of system information on the HMI. They suggested that HCI and IA designs must support the human operator and crew combined to perform tasks effectively as poorly designed systems often lead to a wide range of user inefficiencies that include difficulties in entering, navigating, accessing, and understanding system state information.

Sub-optimal HCI and IA can lead to a host of HMI discrepancies that hinder human performance as a system which is difficult to use and difficult to learn impacts a user’s ability to interact effectively with the system. The goal for any designer is to achieve a user interface design that is comprehensible, intuitive, and one that provides a usable format to support the human in performing mission critical tasks while correspondingly reducing overhead tasking (Holden, et al., 2013).

According to Vincenzi, et al. (2015) UAS operators are not only subject to poor display design and poor information presentation which drastically increases overhead tasking and significantly impacts effective performance, sensory deprivation in the vestibular, peripheral visual, kinesthetic and tactile domain coupled with poor display design exacerbates an environment for errors to occur and for safety to become compromised.

Preece, Rogers and Sharp (2015) described the term effectiveness as to how good a product is at doing what it is supposed to do. For instance, does a product allow users to access the appropriate information they need and when they need it to carry-out their job tasks efficiently? As regards efficiency, they described the term as one used consistently with the term effectiveness in usability evaluations and often refers to the way a product supports the user in carrying out their tasks.

Congruently, Holden, et al. (2013) suggested negligible HCI and IA affect overall system effectiveness and efficiency by imposing substantial overhead tasking that expend the user's cognitive resources in other areas not related to the primary operational task. The resultant outcome is human performance degradation resultant from a significant reduction in a user's situational awareness for the current system state, current system health, and the surrounding operational environment. From an aerospace lens, the aforementioned imposes significant human factors concerns as poor system interaction often translates into an increased number of operator procedural errors on the job. In an aviation paradigm, this could have dire consequences as a high severity procedural error could lead to a UAS mishap. In fact, UAS have been subject to high mishap rates as reported by the DoD (U.S. GAO, 2012).

Damilano, Guglieri, Quagliotti and Sale (2012) suggested the term human factor encompasses many avenues in aviation such as: (1) display design, (2) automation interaction, (3) task performance errors (4) procedural errors (5) lack of operator training and many more. The physical separation of human and machine imposes several human factors complexities when negotiating UAS HMI designs. Therefore, manufacturers of UAS HMI should consider a means to mitigate design inadequacies by examining industry trends related to the flight deck design of conventional aircraft. As an example, the researchers compare the UAS HMI to the Flight Management System (FMS) presently available to pilots of commercial jetliners and ascertained the implementation of a consistent design similar to the FMS for jetliners may help to improve UAS operator performance by incorporating an appropriate level of automation to mitigate overhead tasks that may otherwise substantively increase cognitive demands that lead to high workload (Damilano, Guglieri, Quagliotti & Sale, 2012).

Congruently, Vincenzi, et al. (2015) postulated that varying degrees of automation and autonomy may serve fruitful in combating the effects of sensory deprivation and degraded situational awareness due to the lack of somatosensory cues. However, they urge caution as automation and autonomy may open the path for other unanticipated side effects such as increased mental workload as the pilot transitions from an active participant of the system to a passive monitor. Similar to conventional aviation and the use of automation on the flight deck which is indeed invaluable, they suggested vigilance and complacency issues should be anticipated with UAS due to the high mental workload associated with highly automated systems (Vincenzi, et al., 2015).

As regards performance and performance errors, Sauro and Lewis (2012) defined errors as unintended actions, slips, mistakes, or omissions a user makes while attempting a task. They suggested errors that can be linked to the HMI are defined as user interface (UI) problems and serve as excellent diagnostic information to investigate the usability of a system interface. Vincenzi, et al. (2015) ascertained the potential for UAS will become apparent and leveraged through research, testing, and the assimilation of advanced HMI for UAS command and control. They suggested that under the current and tight regulatory framework in the United States, it is a challenge and nearly impossible to conduct research and test systems in the live operating environment which greatly restricts the prospect for immediate and dramatic improvements to UAS HMI systems design. Therefore, by attaining a better understanding of current HMI designs through research, simulation and usability testing, future design standards could be achieved incrementally and over time. The goal is to identify problematic areas and to better understand aspects of current design modalities that impede human performance in an effort to design systems with attributes that maximize operator performance through user-centered design epistemology.

According to Preece, et al. (2015) usability refers to an interactive system that is easy to learn, effective to use, and enjoyable from the user's perspective. As defined by Nielsen (2012), usability refers to a quality attribute that evaluates ease of use and directly relates to user efficiency and effectiveness in carrying out system specific tasks. Nielsen characterizes usability through five quality attributes: (1) learnability, (2) efficiency, (3) memorability, (4) errors, and (5) satisfaction. The learnability of a system refers to ease of use for a first time user to perform basic tasks while efficiency denotes the task speed once a user becomes familiar with the design. Memorability is related to proficiency and aims to determine how easily users recover when separated from a system after a specified duration. Errors are typically associated with mistakes, lapses, and slips and are determined by the number of errors users make when performing a specific task, the severity of the error, and the proficiency in recovery time from the error. Last, satisfaction refers to the user's perception on how enjoyable the computer-application is for the specific job task (Nielsen, 2012).

Of the five quality attributes, the learnability of UAS HMI for the command and control of medium-altitude long-endurance UAS is of particular interest as these systems impose a substantive human factors concern from an HCI lens. The element of sensory deprivation coupled with poor HMI and IA has led to suboptimal human systems integration (HIS) resulting in operator error as indicated by a significant number of UAS accidents in the last two decades (Tvaryanas, Thompson, & Constable, 2005; Maybury, 2012). Cooke (2008) suggested the most prevalent cause of crew procedural errors in UAS accidents is a result of poor interaction with the HMI.

Similarly, Nielsen (2012), Sauro and Lewis (2012) and Chimbo, Gelderblom, and Villiers (2011) postulated the theory of learnability is a sub-principle of usability and relates to improving operator effectiveness and efficiency through human centered designs. Grossman, Fitzmaurice, and Attar (2009) ascertained the notion of learnability is an important and well-accepted aspect of usability, yet there is little consensus across the HCI community on how learnability should be defined, evaluated, and measured as learnability does not solely lie within the HCI community. It is important to note that the attributes of learnability are dispersed across many fields to include psychology, technical communication, artificial intelligence, and many more (Grossmann et al., 2009).

In the interim, Preece, Rogers and Sharp (2015) suggested learnability refers to a novice user’s initial understanding of a system as well as the attainment of maximum performance over time after a user becomes familiar with the system. Sauro and Lewis (2012) suggested learnability signifies and can be evaluated by how quickly a new user can become efficient with a system (i.e., error free interaction) as a more learnable system reduces the time it takes to complete tasks as the user spends more time interacting with the system. Other definitions postulated that learnability is related to the time it takes users to learn to use the commands for a task or the effort of a typical user in performing a set of tasks on interactive systems but no clear agreement on a definition has been reached (Chimbo, Gelderblom & Villiers, 2011).

In a meta-analysis consisting of 88 research articles from Computer Human-Interaction (CHI) and Transactions on Computer Human Interaction (TOCHI) dating from 1982 to 2008, Grossman et al. (2009) consolidated and organized the usage of the term learnability into eight categories. Interestingly, they found forty-five articles used the concept of learnability without a clear definition, five defined learnability as “easy to learn,” three as “easy to use,” seventeen as “first time user performance,” four as “first time performance after instruction,” eight as “a change in performance over time,” four as the “ability to master a system” and two as the “ability to remember skills over time.”

Although inconsistencies are evident across HCI literature for defining, evaluating, and measuring learnability, Chimbo et al. (2011) ascertained that it is more important to understand the characteristics of a system that produce good learnability than to define the term itself. Rafique, Weng, Wang, Abbasi and Lew (2012a) suggested good learnability in systems design often leads to reasonable learning times, adequate productivity during the learning phase, and high satisfaction in new users. Chimbo et al. (2011) determined that learnability can be measured through user performance elements geared to reveal learning related issues and gains associated with computer-based applications. For instance, error rates, time to complete tasks, and percentage of total functionality serve as viable sources of user data to investigate learnability for computer-based applications.

Similarly, Lazar, Feng, and Hochheiser (2010) and Sauro and Lewis (2012) ascertained that completion rates, task performance, task time, errors, and satisfaction ratings are fundamental usability metrics often used to evaluate HMI for UI problems. Lazar, Feng, and Hochheiser (2010) suggested other metrics useful for usability testing include the average time it takes a user to recover from an error, the time spent using the search or help functionality, and key logging to record user inputs when interacting with a system.

As regards the scope of this research, learnability was defined as a user's initial performance with a system after instruction (i.e., initial learnability) and performance gains on a moderate complexity task after a user becomes familiar with the basic functionality of the system (i.e., extended learnability). This definition was formulated as an amalgamation of captured characteristics or elements from the definitions presented earlier in this chapter. The refined definition serves best to describe the fundamental attributes of learnability within the context of this research. Learnability was manifested in terms of effectiveness, efficiency, and satisfaction and measured using the fundamental metrics of completion rates, task time, errors, and satisfaction as suggested by Sauro and Lewis (2012) and Lazar, Feng, and Hochheiser (2010) and the ISO 9241-11 (1998).

Problem Statement and Goal

According to Holden, et al. (2013) operational spaceflight has rarely been examined from an HCI perspective. Similarly, Damliano, Guglieri, Quagliotti, and Sale (2011) and Terwilliger et al. (2014) suggested that considerable research investment has been vested in the development of new HMIs for modern conventional aircraft in an effort to improve operator performance through user-centered designs but little research attention has been expended to investigate the potential problems associated with current UAS user interface for command and control. Terwilliger et al. (2014) suggested that this problem is significant as the UAS industry and in particular, the DoD, is highly dependent upon these HMI designs to command and control UAS from the ground.

Damliano et al. (2011) suggested that new HMI designs for modern manned aircraft collect, process, analyze, and present relevant flight information to the flight crew when the information is necessary and required (i.e., adaptive display technologies). This form of information presentation often referred to as information automation aids pilots to perform their functional job task with significantly low overhead and was identified as an element to enhance UAS HMI designs for operational effectiveness. It is also important to note that UAS have very low reliability when compared to conventional aircraft.

Similarly, Damliano et al. (2011) and Terwilliger et al. (2014) ascertained that system reliability is low as efforts to collect scientific data to determine system capabilities based on user requirements for new and innovative UAS HMI with advanced automation, procedures, and concepts is nearly absent across the aviation community. Congruently, Maybury (2012) suggested the current state of UAS is in dire need of attention as current HMI representations have been scrutinized for poor human factors and ergonomics (HFE) that stem from poor HCI due to the absence of a regulatory framework to guide industry standards for systems design.

According to Chimbo et al., (2011) the aim of design guidelines, standards and principles is to help designers improve the usability of their products by designing in accordance to rules that aid developers of these designs to make successful design decisions. Design guidelines are set in place to restrict the range of design decisions that may negatively impact a product’s overall usability (Chimbo, et al., 2011). Maybury (2012) suggested the absence of a regulatory framework has omitted practical HF and HCI usability testing in current UAS HMI leaving the industry saturated with non-intuitive HMIs that have misplaced, hard to find, and sometimes erroneous impracticable features of the software making the system very difficult to learn and difficult to use.

Therefore, by focusing on the learnability of current UAS HMI representations, a small part of the complexity associated with the usability of current designs can be examined. By studying user interactions, Su and Liu (2012) ascertained interface designers gain a better understanding of the users, the users various roles, capabilities, and expectations. Su and Liu (2012) and Shamsuddi, Sulaiman, Syed-Mohamad, and Zamli, (2011) suggested that cognition, navigation, appearance and usability testing serve as critical mechanisms to the design of any complex man-machine system as the success of these outcomes typically denote the level of understandability by the designer of the relationship and interaction between human and machine.

The addressable problem of the proposed study evaluated the learnability of an industry-standard UAS HMI as system usability is often poor and attaining the knowledge and skills required to reach a level of proficiency to perform near error free interaction is often substantive. The absence of human factors design principles for UAS HMI has impeded human performance and exacerbated the potential for inefficiencies and human error in UAS operators caused by design. In Tvaryanas (2006), the researcher warned aeromedical examiners to anticipate degraded performance when consulting or investigating future UAS mishaps as human performance error is a significant concern.

Williams (2004) suggested that improved design practices established on aviation display concepts and development focused around the tasks of the user may aid to reduce human error by improving overall system usability. At present, the majority of UAS designs do not capture aviation human factors epistemology as their designers and manufacturers are not aircraft manufacturers. The outcome is a high variability in UAS and UAS HMI representations (Williams, 2004).

Maybury (2012) claimed the advent of modern technologies such as UAS escalates the complexity for aviation usability testing and places UAS on the "avant-guard" of usability (Maybury, 2012, p. 2). Maybury suggested UAS most often represent a rich, difficult, and critical environment to accurately assess usability as they are complex, interrelated systems of systems operated in diverse and dynamic environments. Similarly, Terwilliger et al. (2014) and Maybury (2012) agreed the implementation and use of a iterative task analysis throughout the design process is necessary to better understand key task components, user needs, the user mental representation for displayed information, and other resources required to better optimize UAS HMI designs. Congruent with Williams (2004), the notion on the importance of established design principles for command and control user interface in UAS applications has been well articulated but over a decade later, minimal progress has been made.

Guided by the principles of usability research, this study investigated learnability in UAS interface design as the current state of HCI for UAS applications has received limited attention. In fact, usability research as it pertains to UAS HMI is absent from HCI literature. This research provided the UAS industry with baseline usability data on the system learnability associated with current HMI representations used in the command

and control loop for medium altitude long endurance UAS. The study furnished important information to the fields of aviation on the significance of sound HCI principles as considerations for future UAS HMI designs and introduced the HCI community to usability testing in complex UAS applications.

Research Questions

The following research questions were investigated as part of this study:

1. How accurately did task completion rates such as task completion time, time until failure, total time on task, and errors (Sauro & Lewis, 2012) serve to measure the learnability of the UAS HMI representation?
2. Were participants satisfied with the level of interaction to perform the specific set of operational UAS tasks as regards the System Usability Scale (Brooks, 1996)?
3. Based on the System Usability Scale as scored by Sauro and Lewis (2012), did participants find the UAS HMI usable and learnable?
4. Was incremental learning exhibited as participants become more familiar with the HMI (i.e., reduction in terms of task completion rates and errors)?
5. To what degree did the level of conventional flight experience (i.e., subsequent learning) impact system learnability as regards the dependent variables and perceived satisfaction when compared to those without any conventional flight experience?

Relevance and Significance

MacDonald and Atwood (2013) advocated a need for the HCI community to extend their vision for evaluating system usability in modern and future man-machine applications. They suggested the process of usability evaluation for interactive applications must be adapted to follow the process of technological trends and societal change. The adaptation of evaluation methods is a dominant theme within the HCI community and how the HCI community maintains pace with technological trends in an effort to advance the state-of-the-art for man-machine systems (MacDonald & Atwood, 2013). With that stated, realizing the vast potential and economic benefits with the integration of UAS in the NAS, Congress mandated the FAA Modernization and Reform Act of 2012. The legislation mandates the FAA to advance the state-of-the-art for UAS and accelerate the enactment of these systems as part of the Next Generation Air Transportation System (NextGen) transformation scheduled to be complete by the year 2025.

As a significant component addressed in the FAA Modernization and Reform Act of 2012, the legislation described the vast potential for civil and public UAS applications in the domestic United States and the prospective economic impact these systems will have once integrated with conventional aircraft in the NAS. According to the Association for Unmanned Vehicle Systems International (AUVSI; 2013) the domestic economic forecast of UAS integration as projected by Congress is said to surpass $13.6B between the years of 2015 and 2017. Impending forecasts by AUVSI (2013) suggested the industry to reach $82.1B dollars by the year 2025.

Terwilliger et al. (2014) stated that recent congressional mandates have readily forced the FAA to move forward with provisional legislation which has expanded opportunities available for civil and commercial use of UAS in the NAS. Terwilliger, et al. (2014) described the UAS market as one that has experienced significant growth with the advent of increased computing technologies, expanded UAS application opportunities (e.g., agriculture, linear infrastructure inspection, horizontal infrastructure inspection, thermography and mapping) and the availability of materials and related technologies for the construction and design of these systems. Alternately, the current state of HMI technology and display design for UAS is in dire need of attention as current UAS HMI representations contain many issues and challenges related with ineffective and inefficient information presentation and information exchange (Terwilliger et al, 2015; Maybury, 2012).

Terwilliger, et al. (2014) identified four primary HMI inadequacies that have significantly curbed the state-of-the-art for HMI advancement in UAS command and control. First and congruent with many others described the absence of a regulatory structure to guide the UAS community as the primary cause of the troubles that currently loom the industry. Terwilliger et al. (2014) suggested this is a consistent theme across literature and appears to be the systemic problem when discussing UAS systems design.

The absence of design guidelines has led to HMI that are not optimized for use with teleoperation as information presentation and information exchange has yet to be closely examined. At present, UAS HMI designs present an overabundance of visual information to the user during all stages of operation and lack information in other human sensory channels (e.g., peripheral visual, proprioceptive and vestibular). Industry HMI designs for command and control often include multiple software applications and multiple displays of information presented to the user (e.g., moving map, primary flight display, payload imagery, health and aircraft system state information), hierarchical menu structures to alter system state parameters, QWERTY keyboards, joysticks, and trackballs to input commands for vehicle control and navigation. These elements combined generate a significant level of overhead which directly impacts and allocates the user's attentional resources to secondary and tertiary tasks; not the primary task of operating the air vehicle.

Holden, et al. (2013) ascertained that if displays are not designed with a fully developed operations concept in the planning phase, fine-grained iterative task analysis, and knowledge of human information processing capabilities and limitations, the layout, format, and mode of information presentation on the HMI may be sub-optimal to support operator task performance. Typical repercussions include users misinterpreting, overlooking, or ignoring the original intent of information as the information presented does not accommodate the human mental representation of the current system state. Additional problems may arise when there is improper function allocation or an improper level of automation between the human and the system, or when interaction with the system is confusing, inefficient, or difficult to learn. According to Holden, et al. (2013) poorly designed user interface and displays for command and control and navigation negatively impact user performance by extending task completion times and elevating task execution errors by significantly increasing overhead tasking due to poor information presentation.

The third delinquency in UAS HMI was the absence of adaptive and flexible automation to ease operator workload. Terwilliger et al., (2014) suggested that UAS HMIs must be adaptive and flexible as regards automation so that the operator may adjust the level of automation required to suit their present workload. In the future, adaptive automation is believed to serve as a mechanism to reduce operator workload by alleviating aspects of overhead tasking associated with secondary and tertiary task. Adaptive automation closely resembles the concept of information automation currently present on HMI representations found in today's modern jet aircraft (Damilano et al., 2012). Similarly, Vincenzi, et al. (2015) and Piuzzi, Cont and Balerna (2014) recommended the use of adaptive displays as the functionality greatly optimizes user performance as an acceptable level of workload is constantly maintained by the system.

Last, recommendations to circumvent the element of sensory deprivation in the vestibular and peripheral-visual sensory regions through the HMI are highly desired. Similarly, Tvaryanas, Thompson, and Constable (2005) and Hopcroft, Burchart and Vince (2006) ascertained sensory deprivation is degraded in the visual, kinesthetic, vestibular, and auditory domains, In manned aviation, pilots use cueing information in addition to information presented on the HMI to distinguish orientation and other performance related variables such as the distance to the runway (Terwilliger et al., 2012).

Future UAS HMI should incorporate elements of human sensory cueing to enhance situational awareness by providing a "seat in the pants" feel of what the air vehicle is doing. Designed properly, human sensory cuing could correspond to the flight information displayed on the HMI enabling the opportunity for the operator to gain a clear understanding of the system state as they operate remotely from the ground (Terwilliger et al., 2014). Combined with adaptive displays, this may greatly benefit to improve UAS reliability as user-centered assistive technologies will reduce operator workload by minimizing overhead tasks, regulating the level of workload and by providing relevant flight information through multiple human sensory modalities.

Barriers and Issues

The following barriers and issues were identified in this research: First, employers with the number of UAS operators required for this research are large corporations, U.S. Government agencies, and U.S. military services. Direct access to survey these UAS operators from an expert perspective was not feasible for this study. Second, the researcher used a purposive sample of 45 participants. Students from the Embry-Riddle Aeronautical University (ERAU) comprised the sample population. These students occupied the desired FAA flight certificates and the number of flight hours required for both the low and high experience desired of the sample population. Low experience was defined as a private pilot with an instrument rating and less than 250 flight hours. High experience was defined as a commercial pilot with an instrument rating and between 250-1000 flight hours.

Limitations and Delimitations

Limitations

One limitation of the proposed study was the convenience sample of the participants and the absence of random sampling. Combined with elements of effect size, this impacted the study's overall generalizability and therefore the researcher believes it was necessary to address. As regards effect size, Sauro and Lewis (2012) defined usability testing as either summative or formative. Summative evaluations focus on measures of performance related to the successful completion of specific task goals whereas formative evaluations are designed to detect and reveal general usability problems. This study was designed as a summative study and focused on measures of performance related to the successful completion of task goals as participants interacted with the representative UAS HMI.

Sauro and Lewis (2012) suggested that summative evaluations closely resemble the mechanics of traditional experiments and therefore may require a larger sample population when compared to formative usability evaluation. This research implemented a convenience sample of forty-five participants. Sauro and Lewis (2012) suggested that larger sample populations are useful in some settings but are rare in usability testing. Alternately, if viewing this study through the lens of a true experiment, the sample size population may be considered modest, limiting again, the generalizability of the results. However, as a causal-comparative design, the sample population of 45 was adequate for this usability test.

As regards random sampling, random sampling in this case did not suffice and the researcher implemented a purposeful sampling method to determine the sample population based on predefined criteria. In this particular case, experience was a component of FAA level of certification and the number of flight hours. Combined, these criteria afforded the researcher to attain a more accurate and well defined representative sample. A sample population defined by participants that occupied similar attributes and experience within their respective groups was desired to minimize extraneous variables.

Delimitations

The study was limited to compare one UAS HMI presented as a high fidelity simulation of a medium-altitude long endurance UAS and not the actual equipment. The simulator used in this study offered a high fidelity mockup of the Predator MQ-1/ MQ-9 user command and control interface. The HMI symbology, control layout, menu structure, and terminology offered a spatial layout of information presentation and information exchange highly representative of these systems but were not the same. Nonetheless, operational flight techniques and skills are trained, applied, and evaluated using these devices. Second, since the simulator was representative of a medium-altitude long endurance UAS, the researcher suggested that other categories and class of UAS may not hold similar design characteristics and therefore, user performance aspects related to the findings of this study should correspond solely to HMI designs for medium-altitude long endurance UAS.

Definitions of Terms

The following section provided a list of key terms, their associated definitions and acronyms as used throughout this research. These definitions represent commonly accepted usages in the field of HCI, usability evaluation, and aeronautical science. The terms defined here are significant enough to mention or are not readily known

Certification of Authorization or Waiver (COA): A Certification of Authorization or Waiver (COA) is an authorization issued by the Federal Aviation Administration (FAA) to a public operator for a specific unmanned aircraft activity (Austin, 2010).

Commercial Pilot Certificate: A commercial pilot is one that may be compensated for flying. Training for the commercial certificate places more emphasis on a better understanding of aircraft systems and a higher standard of airmanship. The commercial certificate itself does not allow a pilot to fly in instrument meteorological conditions (IMC). Therefore, commercial pilots without an instrument rating are restricted to daytime flight within 50 nautical miles (NM) when flying for hire (Department of Transportation, 2008).

Effectiveness: Refers to how good a product is at doing what it is supposed to do (Preece, Rogers & Sharp, 2015).

Efficiency: Refers to the way a product supports the user in carrying out their tasks (Preece, Rogers & Sharp, 2015).

Errors: Any unintended actions, slips, mistakes, or omissions a user makes while attempting a task (Sauro & Lewis, 2012).

Federal Aviation Administration (FAA): The national agency allocated to the Department of Transportation (DOT) with the responsibility to implement and regulate standards for the air-worthiness of all civilian aircraft and to ensure safe operation in the national airspace system (NAS; Austin, 2010).

Federal Aviation Regulations (FAR): Directives issued by the FAA to govern flight operations, construction of aircraft, and the training requirements to obtain pilot certificates and ratings. Federal aviation regulations are identified specifically by Title number 14(i.e., Aeronautics and space) and fall within the larger group of rules obtained in the Code of Federal Regulations (CFR; Willits, Abbott, & Kailey, 2004).

Ground control Station (GCS): A GCS is a system that facilitates remote operation of unmanned aircraft by a pilot and in some configurations singular or multiple sensor or payload operators, when the portability of the system is determined by the size and complexity of the system (Austin, 2010).

Human-computer interaction (HCI): a discipline that studies and describes how humans and computer-based systems communicate, share information, and accomplish tasks (Holden, Vos & Martin, 2013).

Instrument Rating: The instrument rating is an option that allows a private pilot and/or commercial certificated pilot to fly in a wider range of weather conditions (i.e., limited out-of-the-window visibility). Aircraft control is maintained solely by reference to the cockpit instruments and not by reference to the ground or horizon (Willits, Abbott, & Kailey, 2004).

Learnability: Is a novice user's initial performance with a system after instruction (i.e., initial learnability) and user performance gains on a specific set of tasks after a user becomes familiar with the basic functionality of the system (i.e., extended learnability; Sauro & Lewis, 2012; Nielsen, 2012; Rogers, Sharp & Preece, 2011; Grossman et al., 2009).

Medium Altitude Long Endurance (MALE): A long endurance unmanned aerial vehicle controlled remotely by an internal pilot from a ground control station (GCS). The MALE UAS typically flies at an altitude window between 10,000 to 30,000 feet (Austin, 2010).

Private Pilot Certificate: The majority of active pilots typically hold a Private Pilot Certificate. This certificate permits command of any aircraft appropriate for this specified rating. The aircraft flown with noncommercial purpose and provides almost unlimited authority to fly under visual flight regulations (VFR). Passengers may be carried; however, a private pilot may not be compensated in any way for services as a pilot (Department of Transportation, 2008).

Tactical UAS: A medium range aerial vehicle with a range of 60 to 185 miles. These systems are typically smaller than MALE UASs with limited endurance operated remotely from a GCS (Austin, 2010).

Task Time: The duration a user spends on an activity (Sauro & Lewis, 2012).

Task Completion Time: Time of users who completed the task successfully (Sauro & Lewis, 2012).

Time until Failure: Time on task until users give up or complete the task incorrectly (Sauro & Lewis, 2012).

Total Time on Task: The total duration of time users spend on a task (Sauro & Lewis, 2012).

Unmanned Aircraft System (UAS): The term UAS refers to the complete complex unmanned system composed of a control element, a data and voice communication element, and an air vehicle element required for mission operation (U.S. GAO, 2008).

Unmanned Aircraft Vehicle (UAS): A UAS is the airframe component of the UAS that does not contain a pilot onboard and is either directly operated from a ground control station or autonomously (U.S. GAO, 2008).

Usability: Usability is a quality attribute that assesses how easy user interfaces are to use. The word "usability" also refers to methods for improving ease-of-use during the design process (Nielsen, 2012).

Chapter Summary

Chapter 1 introduced the background, identified the research problem, described the goal, and identified the research questions. The problem was learnability of an industry-standard UAS HMI as system usability is often poor and attaining the knowledge and skills required to reach a level of proficiency to perform near error free interaction may be substantive. The goal was to provide the UAS industry with baseline usability data on the learnability of current HMI representations for command and control. The study furnished important information to the fields of aviation of the significance for sound HCI principles in future UAS HMI designs and introduce the HCI community to usability testing in complex UAS applications.

Chapter 2

Review of the Literature

Introduction

Shneiderman and Plaisant (2010) defined teleoperation as the direct manipulation of a computer application and its processes by human operators to control physical aspects of the design in complex environments or settings. The term is often used interchangeably with “remote control” as operators interact with a computer application to perform the task from a distance (Shneiderman & Plaisant, 2010). The concept of teleoperation best describes the complex process associated with the command and control paradigm of unmanned aircraft system (UAS) as the operator input commands and controls the air vehicle from a ground control station (GCS).

According to Shneiderman and Plaisant (2010) in traditional direct-manipulation interfaces, objects and actions are depicted continuously and users typically point, click, and drag to change parameters and receive feedback efficiently unlike interfaces that may require a user to type specific parameters on a keyboard. The UAS is a direct-manipulation interface that requires operators to point, click, drag, and type information into a keyboard to manipulate variables and adjust parameters that affect vehicle state (i.e., pitch, roll, yaw, altitude, heading, and airspeed).

Shneiderman and Plaisant (2010) suggested designers should expend additional design efforts in direct-manipulation applications to aid users with potential lag and latency in system response loops, incomplete feedback, and the increased likelihood of breakdowns. They suggested more complex error recovery procedures are necessitated to circumvent potential limitations in direct-manipulation systems as these problems are directly related to hardware and software limitations, constraints in the physical environment, network design issues, and the complexity of the dynamic operational task environment. The architecture of remote environments specifically for UAS introduces several complicating factors such as time delays, incomplete feedback, and unanticipated interference that may negatively impact the effectiveness, efficiency, and safety of teleoperation. Similarly, Tvaryanas (2006) and Maybury (2012) suggested the lack of peripheral visual and vestibular sensory input in unmanned flight exacerbates the problem by introducing considerable perceptual delays between manual control input and system state feedback. The latency coupled with soda straw views of the real world greatly impairs an operator’s situation awareness by increasing workload which could easily lead to overall flight degradation (Cooke & Pedersen, 2010).

Human-Computer Interaction

According to Lazar, Feng, and Hochheiser (2010) the field of HCI is interdisciplinary and an amalgamation of computer-science, human factors engineering, psychology, sociology and many others. Holden, et al. (2013) suggested HCI and information architecture (IA) designs must support the human operator and crew tasks as poorly designed systems could lead to a wide range of potential user problems such as having difficulties in entering, navigating, accessing, and interpreting system state information. Further, Holden, et al. (2013) ascertained ineffective HCI and IA also imposes substantial overhead tasking that expend the user’s cognitive resources in other areas not related to the primary operational task and this imposes a significant safety concern as it pertains to operating air vehicles thereby, having a negative impact on operator effectiveness. UAS operators often fall victim to overhead tasking as they must filter through large volumes of system state and payload (i.e., visible camera, infrared camera) information typically presented by multiple resources and on multiple displays that require attention resource allocation. This often leads to reduced situational awareness of the operational environment and in some instances, systems state, as operators struggle to allocate attention between visual and auditory stimuli in the GCS (Terwilliger et al., 2015). These problems often become intensified when environments are dynamic, unpredictable, and when operating procedures are sub-par. Further, cognitive overload is directly related to overhead tasking as the quantity of information presented to the user from multiple visual streams may tax cognitive and mental resources leading to situations where users do not have a clear understanding of the spatial and temporal state of the system (Holden, et al., 2013).

Zhang and Walji (2011) suggested the goal with HCI epistemology is to reduce overhead tasks by improving the transparency and predictability of system processes and responses associated with system state information. Consequences associated with overhead tasking is potentially far more critical in flight applications as the environment is dynamic and minimal time exists to identify mistakes and correct slips (Holden, et al., 2013). Yin, Wickens, Helander, and Laberge (2014) ascertained that higher levels of predictive information particularly in aviation systems may lead to improved operator performance by enhancing the transparency associated with the uncertainty of prediction. Yin et al. (2014) postulated an important element of predicative displays is the interface design (i.e., HMI) and its effect on operator control and performance.

Terwilliger et al. (2014) suggested the most prevalent deficiency presented in UAS HMI design is the lack of basic sensory cues related to the forces of flight. These cues are typically used by conventional pilots to fly the aircraft and serve to confirm system state information, aid in decision making, and enhance operator situation awareness. They suggested designers of GCS and HMI for UAS should consider these cues as integral aspects of the design as a means to enhance operational effectiveness. Designing systems with the user in mind means designing interface that are functional, intuitive, easy to learn, and easy to understand so that operators may extract pertinent information when that information is needed (Terwilliger et al., 2015).

Learnability

Learnability is characterized in many ways. Some describe learnability as the ease of use on initial user performance and improvements in performance when interacting with a system over-time (Grossman et al., 2009; Chimbo et al., 2011). Others suggested learnability is the capability of a software product to enable the user to learn how to use it effectively within a reasonable amount of time (Shamsuddin, Sulaiman, & Zamli, 2011). Nielsen (1994) suggested a highly learnable system is one that allows users to reach a reasonable level of proficiency in a short span of time. Similarly, Rafique, Weng, Wang, Abbasi, and Lew (2012b) ascertained that good learnability often leads to adequate learning times, ample productivity during the learning stages, and overall better satisfaction in new learners. Correspondingly, Shamsuddin, Sulaiman, and Zamli (2011) suggested the quality of user interactions on a system directly relates and reflects the user’s knowledge and understanding of the system. However, when system features are difficult to initially learn, a significant amount of time and investment is required to reach an adequate level of proficiency to perform system related tasks in an effective and efficient manner.

According to Madni and Sievers (2014) today's modern systems are very complex and dynamically changing requiring new elements and new techniques for interface testing, validation, and verification. Rafique et al. (2012) suggested this has certainly been evident in many fields such as Geographic Information Systems (GIS) and web applications as many new features have been introduced to the users in an effort to maintain pace with technological trends associated with hardware and software advancement. Learning and understanding the new features have been described as challenging and frustrating. According to Madni and Sievers (2014) the overall goal with usability and systems integration is to assure interfaces can be adapted in a well-understood manner and with relatively modest effort.

Correspondingly in an exploratory investigation of workplace user frustration with computers, Lazar, Jones, and Shneiderman (2006) found that employees spent nearly 45% of their time dealing with frustrating experiences related to error messages, missing, hard to find, and impracticable systems features. Combined, these elements were attributed as leading factors to poor system learnability. Similarly, Maybury's (2012) assessment of UAS applications corresponded directly to the learnability issues identified by Lazar, Jones, and Shneiderman (2006). The only disparity was that these system inadequacies are unacceptable in UAS applications as the flight environment is dynamic and only a fraction of time exists to detect and correct performance errors (Holden, et al., 2013). Therefore, unmanned systems must incorporate designs in which the physical orientation of controls is logical and consistent across systems as well as provide uniformity in display symbology, terminology, and functionality (Maybury, 2012). Human-centered design considerations may aid to reduce training times and improve operator performance by considering and designing systems around the human element. Guidelines provided by the ISO such as ISO/IEC 25010:2010 and similar should be considered when defining UAS HMI characteristics related to human interaction as poor characteristics significantly impact system usability in terms of effectiveness, efficiency, satisfaction and safety of the intended system use. Limitations described as the aforementioned directly impact the overall quality of the system and currently impact systems in use today.

System quality as defined by ISO/IEC 25010 (2011) is the degree to which the system satisfies the intended needs of its stakeholders in a valuable manner. In accordance to the ISO guidelines, learnability can be measured as the extent to which a product or system can be used with effectiveness, efficiency, and satisfaction by specified users to achieve specific learning objectives within a specific context of use. This dimension of learnability was formally defined as the quality-in-use (QinU; Rafique et al., 2012b). In this study, learnability is defined as a user's initial performance with a system after instruction (i.e., initial learnability) and performance gains on a specific set of tasks after a user becomes familiar with the basic functionality of the system (i.e., extended learnability).

The QinU directly relates to the capability of a software product to influence user effectiveness, productivity, safety, and satisfaction when using the software product to achieve specific goals within a predetermined context of use. The QinU as described by ISO 25010 (2011) holds five attributes. The five attributes are: (1) effectiveness, (2) efficiency, (3) freedom of risk, (4) satisfaction, and (5) context coverage. Rafique et al. (2012b) ascertained that quality is achieved when users meet their target operational goals with effectiveness, efficiency, satisfaction, and safety. The researcher suggested that characteristics of utility such as functionality, suitability, reliability, and performance efficiency are the characteristics that define the quality of an overall system.

According to Chimbo et al. (2011) learnability is comprised of specific attributes that can be measured to reveal learning related issues and gains associated with computer-based applications. Similarly, MacDanold and Atwood (2013) described the five most common user performance metrics for measuring learnability as: (1) time to complete tasks, (2) error rate, (3) accuracy, (4) task completion rate, and (5) satisfaction. Congruently, Grossman, Fitzmaurice, and Attar (2009) suggested that error rates, time to complete tasks, and percentage of total system functionality are highly desirable performance attributes when investigating learnability using summative evaluation methods. Further, user experience may play a significant role in determining the learnability of systems as novice users may interact differently when compared to experienced users. Keyboard inputs, number of clicks, mouse movements and scrolling indicate a user’s knowledge on the sequence of actions to perform and complete a task (Shamsuddin, Sulaiman, & Zamli, 2011).

Similarly, Grossman, Fitzmaurice, and Attar (2009) described learning as a function of experience and suggested the type of user is inherently important when measuring learnability in computer-based applications. They suggested a prevalent trend in defining learnability is to also address the type of user for which learning is geared in an effort to delineate between novice, intermediate, and expert users. Davis and

Wiedenbeck (as cited in Grossman, Fitzmaurice, and Attar, 2009) defined subsequent learning as a user who has no specific system experience but experience with a similar type of system and in a similar domain. For this study, 15 participants had no previous conventional flight experience, 15 held at least a private pilot, and 15 held a commercial pilot certificate. None of the participants had any experience with UAS applications and in particular, had no interaction with this system prior to this experiment.

Unmanned Aircraft Systems

The term unmanned aircraft system describes a complex system of systems with many interrelated elements other than the air vehicle. A UAS includes the UAS (i.e., airframe), ground control station (GCS), data and voice communication infrastructure, and other related support subsystems to permit remote operations from the ground (Austin, 2010; GAO, 2008). Unmanned aircraft systems are typically described by their size, weight, vehicle design and their capability to achieve a specific mission set; yet, classifying these vehicles in a systematic fashion for commercial applications is challenging as new technologies are rapidly introduced (Goldberg, 2010). As existing technologies become enhanced and new technologies are introduced, smaller systems with much higher capability will replace the roles of their predecessors. However, without regulatory guidance for systems design, a lack of standardization will continue to exist across platforms and UAS control interface. The high level of variability in current system design yields an imperfect grouping of UAS systems and introduces many concerns from a human factors perspective (Simmons, Liu, & Vincenzi, 2008).

The most observable difference between manned and unmanned flight is the lack of an onboard pilot and essentially the method in which an operator controls the UAS (e.g., external versus internal command and control; Gundlach, 2012). Currently, there are three primary command and control interface modalities used for UAS operations. The command and control modalities are: (1) external piloting using a hand held radio controlled interface and visual line of sight, (2) internal piloting where the system is remotely operated from a GCS using communication data links and on-board sensors to acquire telemetry data and spatial information for navigation and control, and (3) full autonomous flight (Austin, 2010; Goldberg, 2012; Gundlach, 2012). Interestingly, each mode of operation requires a unique set of skills to ensure effective, efficient and safe systems integration (Goldberg, 2012).

Human Systems Integration

Madni and Sievers (2014) suggested systems integration (SI) is concerned with establishing an intelligible whole from component subsystems that offers the required functionality for the human operator to perform specific tasks to achieve the goals for a specific mission set. According to Tvaryanas (2006) human systems integration (HSI) is based on the premise that humans are critical elements within a system and describes a process model for optimizing human performance. Tvaryanas (2006) defined human performance as a quality function within the seven domains of HIS and includes (1) human factors engineering, (2) personnel, (3) training, (4) manpower, (5) environment, safety, and occupational health, (6) habitability, and (7) survivability. However, Madni and Sievers (2014) also suggested the concept of SI becomes far more difficult as system complexity and a need for adaptability increases. Tvarynas (2006) postulated systems complexity and design limitations are often circumvented by augmenting training or by simply selecting personnel with more experience and training in a similar domain. Dalamagkidis, Valavanis, and Piegl (2012) ascertained the differences in control interface between manned and unmanned vehicle system operations may certainly necessitate specialized training for UAS operators as remote interaction (i.e., internal versus external) from a GCS offers a new mode of air vehicle control. Williams (2012) examined the level of manned flight experience and the effect of sensory information on airmen reactions to system failures. Using a between subjects design, Williams (2012) manipulated two levels of sensory information (i.e., visual versus visual and auditory) and two levels of experience (i.e., non-airmen versus airmen). Sixteen out of 32 participants were certificated at least as private pilots while the remaining 16 had no previous flight experience. It is also important to note that participants had no previous experience flying a UAS.

Williams (2012) reported two significant findings. First, the proportion of participants that failed to respond to an engine failure within five seconds was greater in the visual domain only condition regardless of previous airmen experience. Second, the proportion of participants responding to a heading control deviation prior to a visual indication on the graphical user interface (GUI) was significantly greater with rated airmen than non-airmen. Participants certificated as rated airmen applied corrective action to heading deviations whereas non-airmen did not input any corrective action to compensate for the deviations in heading. Further, rated airmen also flew significantly closer to the flight path than non-rated airmen. Under this condition the findings suggested that technical flight error correction in the manual control condition was much higher for rated airmen than non-airmen.

The results of this experiment ascertained significant differences may exist between participants with previous airmen experience than those without. Williams (2012) postulated that airmen training and previous airmen experience could lead to better interaction with the HMI thereby, improving operational performance. Unfortunately, these differences could not be directly correlated to prior levels of training. The level of skill transfer from prior flight training to UAS flight may precisely depend on the similarity that exists between the manned and unmanned vehicle and the level of training and expertise the operator has previously received (Williams, 2012).

Knowledge, Skills, and Abilities

Unmanned aircraft flight offers unique challenges and requires specific equipment knowledge and skill competencies when compared to traditional manned flight. In order to enhance the principles for UAS, Pavlas et al. (2009) recommended practitioners identify constraints inherent to UAS systems and their operation as these deficiencies could identify a set of knowledge and skills not captured in current training paradigms. By identifying the inherent challenges with UAS flight and understanding the associated key competencies required for UAS operations, industry officials can define standards for commercial UAS applications (Pavlas et al., 2009). At present and based on the high level of variability in UAS systems design, it appears that each UAS system may require a unique set knowledge and skills training guided by the mode of operation, level of automation, GCS design, and mission capability (Pavlas, et al., 2009).

Knowledge

Knowledge as it pertains to UAS operator can be considered both declarative and procedural in nature. Declarative knowledge refers to factual knowledge (i.e., knowing what) while procedural knowledge relates to cognitive skills associated with accomplishing a task (i.e., knowing how; Driscoll, 2005). Knowledge in the framework for UAS include the ability to recognize and recall previously learned information to accomplish a set of operational tasks based on information presented in a specific setting (Pavlas, et al., 2009).

In an attempt to circumvent the irregularities associated with UAS operator training, the CJCS (2011) provided a common category list of general aviation knowledge and academic content that should at a minimum be included in any UAS training paradigm. The following categories can serve as initial criteria to develop commercial UAS curriculum and training regimens. These general categories included: (1) airspace design and operating requirements, (2) air traffic control (ATC) procedures, rules, and regulations, (3) aerodynamics, including effects of controls, (4) aircraft systems and emergency procedures, (5) performance, (6) navigation (7) meteorology, (8) communication procedures, and (9) mission preparation. As regards UAS, equipment-specific knowledge includes elements such as monitoring and understanding the spatial and temporal state of the system, level of automation, and command and control feedback during normal and emergency situations (Pavlas et al., 2009).

Skills

The term skill is associated with a specific level of performance and is related with the accuracy and speed of performing a task or set of tasks (Cunningham, 2008; Winterton, Le Deist, & Stringfellow, 2005). The acquisition of advanced skill corresponds with comprehension of new knowledge facilitated by the recall of cognitive prerequisites to create new schema (Schunk, 2011). Ultimately, higher levels of skill competency are associated with an increased level of declarative and procedural knowledge (Winterton et al., 2005). For UAS operators, the most basic skill is the ability to input and verify system parameters to command and control the UAS (Stulberg, 2007). Other pertinent skills for UAS operators include but are not limited to problem solving, decision making, collaboration, coordination, risk assessment, flight instrument and system monitoring, and crew resource management (Tvaryanas, et al, 2005; Pavlas et al., 2009).

Abilities

The term ability is often used interchangeably with competence and describes the motivation to learn and perform. Ability is comprised of personal qualities and emotional values exhibited on an individual level and may be directly associated with the affective states and differences of individuals (Driscoll, 2005; Winterton et al., 2005; Pavlas et al., 2009). Consequently, based on term interchangeability, it may be difficult to arrive at a coherent theory to define the term; however, Winterton, Le Deist, and Stringfellow (2005) suggested the term may be useful to bridge the gap between knowledge, skills, and the requirements of a specific job task. Essentially, ability relies on the underlying motivational characteristics of an individual and their willingness to learn. This can be viewed from a competence-performance approach. According to Winterton, Le Deist, and Stringfellow (2005) the competence-performance approach relates to domain specific knowledge and is categorized into three distinct components: (1) conceptual competence, (2) procedural competence, and (3) performance competence (Winterton, Le Deist, & Stringfellow, 2005). They defined conceptual competence as rule based knowledge about an entire domain. Procedural competence denotes the ability to apply procedures and skills in standard situations to perform domain specific tasks. Performance competence represents the ability to apply knowledge in the presence of a problem and derive an appropriate strategy to formulate a solution (Winterton, Le Deist, & Stringfellow, 2005).

Overall, the development of competence or ability depends on the learning and training opportunities presented to an individual and how the individual accepts and learns from these opportunities (Driscoll, 2005). From a UAS perspective, ability includes the cognitive and psychomotor prerequisites presented to a trainee in an education and training paradigm and how the trainee uses these conditions to incrementally reach an expert level criterion. Unfortunately, system usability is often poor in UAS designs and attaining the knowledge and skills required to reach a level of proficiency to perform near error free interaction may be substantive. The researcher postulated that experience in the manned domain might not be sufficient to improve UAS system learnability as UAS necessitate system specific training.

Chapter Summary

Chapter 2 introduced credible literature to support the human factors concerns associated with the command and control for unmanned aircraft systems and also provided a background of information pertaining to the usability of these systems and the complexities associated with the teleoperation of UAS. The researcher described system design inadequacies and their impact on system usability as they related to the command and control of a UAS from a ground control station. Insufficient systems design coupled with disjointed training methods have resulted in poor systems integration. Poor system usability has been identified as a leading cause for sub-optimal human performance in accidents associated with UAS applications.

Chapter 3

Methodology

Introduction

This chapter presented the methodology to support the experimental design and procedures applied to investigate the system usability of a UAS HMI representation. The addressable problem of the proposed study evaluated the system learnability of an industry-standard UAS HMI as system usability is often reported as poor and attaining the cognitive knowledge and psychomotor skills required to reach an expert level of proficiency to perform near error free task interaction with UAS HMI is considered substantive from a training perspective.

The following research questions were investigated as part of this study:

1. How accurately did task completion rates such as task completion time, time until failure, total time on task, and errors (Sauro & Lewis, 2012) serve to measure the learnability of the UAS HMI representation?
2. Were participants satisfied with the level of interaction to perform the specific set of operational UAS tasks as regards the System Usability Scale (Brooks, 1996)?
3. Based on the System Usability Scale as scored by Sauro and Lewis (2012), did participants find the UAS HMI usable and learnable?
4. Was incremental learning exhibited as participants become more familiar with the HMI (i.e., reduction in terms of task completion rates and errors)?

5. To what degree did the level of conventional flight experience (i.e., subsequent learning) impact system learnability as regards the dependent variables and perceived satisfaction when compared to those without any conventional flight experience?

Experimental Design

A causal-comparative or Ex Post Facto research design was established for this experiment. Gay, Mills, and Airasian (2012) suggested that causal-comparative research studies attempts to determine the cause for existing differences in the behavior or actions of individuals or groups. The grouping variable for this research, experience, had three levels or factors: (1) no conventional flight or UAS operational experience (2) low conventional flight experience and no UAS operational experience and (3) high conventional flight experience and no UAS operational experience.

According to Gay, et al. (2012) causal-comparative studies do not necessitate the randomization of participants but instead researchers select participants with specific characteristics or attributes that define the differences among the groups on some variable. In this study, the level of conventional flight experience served as the major factor to differentiate among the participant groups. Unlike traditional experimental studies, Gay, et al. (2012) determined the independent variable is not manipulated because of the fact that it has already occurred. For this study, the researcher implemented the ISO measures for effectiveness, efficiency and satisfaction using three primary methods for data collection across the three participant groups (i.e., no pilot experience, *n*=15, low pilot experience, *n*=15, high pilot experience, *n*=15). Data were collected for five dependent variables which included: (1) task completion time, (2) time until failure, (3) total time on task, (4) number of errors, and (5) satisfaction. These data served to rationalize the learnability of the UAS HMI based on user interactions with the system.

The first method implemented to capture each participant’s session and the interaction with the representative HMI under investigation was audio and video recording. Each participant’s trial iteration was recorded and maintained for post-hoc evaluation. Three independent subject matter experts examined the video recordings in a side-by-side manner and scored the videos for the dependent variables of task completion time and number of errors committed. Any scoring discrepancy amongst the three SMEs was rectified on a case-by-case basis by re-examining the specific video(s) and discussing the findings until a unified consensus was achieved. Second, the simulator used for this study provided a time-stamp data log for each participant’s simulator activity. Data collected from each participant’s simulator activity afforded the researcher the ability to extract data for the dependent variable total time on task.

The data associated with task completion or task success were coded as binary. The inverse of the completion rate afforded the researcher to rapidly generate group failure rates. The number of errors committed by a participant was scored as a count. The raw data set for total time on task as extracted from the data logs indicated a positive skew with a heavy tail. Data transformations were executed for total time on task to normalize the data in preparation for statistical analysis. Last, the System Usability Scale was implemented to gather measures of perceived satisfaction. The instrument was administered in the original format as defined by Brooks (1996).

Sauro and Lewis (2012) described the SUS as a popular questionnaire for end-of-test subjective assessment to offer a valid and reliable tool for evaluating user perception in different types of interfaces. The SUS is discussed in a latter section titled measurement and instrumentation and format for presenting the results.

Causal-Comparative Research Design		
No Pilot Experience *n=15*	Low Pilot Experience *n=15*	High Pilot Experience *n=15*
Dependent Variables	**Task Success/Completion Rate**	
	Number of Errors	
	Failure Rate	
	Total Time on Task	
	Satisfaction	

Figure 1. Experimental Design Matrix

Participants

Forty-five individuals were solicited to participate in this study. For this investigation, the researcher had access to a large sample population of undergraduate students at ERAU to include students rated and experienced as private pilots, commercial pilots, and certified flight instructors as prescribed by the Federal Aviation Administration (FAA).

The demographics obtained from the Institutional Research Department (2016) at ERAU describe the characteristics and diversity within the university. For instance, 16% of the undergraduate population is from foreign countries and female students are represented at 17%. The total undergraduate enrollment is 5,278. The average age of full-time undergraduate students is 21 years of age. Of the 5,278, 1,178 students are enrolled in the Aeronautical Science degree program and possess the desired level of certification and experience (i.e., low versus high) for this study based on FAA criteria. The target population for this study was extracted from the sample pool of 1,178 students. These students were queried via electronic solicitation through the academic advisement department within Aeronautical Science. The participant recruitment briefing served to fulfill the context of the electronic solicitation. The demographic survey was designed to capture the eligibility requirements for the sample population in an effort to aggregate results and define the 45 participants for this study. The demographic survey was administered as part of the participant recruitment process. The researcher implemented a judgment or purposive sample to select the appropriate candidates with the desired level of experience

Clear criteria defined by Title 14 of the Code of Federal Regulations (CFR) Part 61 as it pertained to the training and certification of pilots, flight instructors, and ground instructors was used to validate the independent variable, experience, for this study. The regulations afforded guidance for eligibility requirements, aeronautical knowledge requirements, and practical test standards (PTS) for each type of pilot certificate issued by the FAA (United States Government Printing Office, 2013b). According to the Department of Transportation (2008) the type of intended flying will typically determine the type of pilot certificate granted by the FAA as the eligibility, training, experience, privileges, and testing requirements differ based on the level of FAA certification.

The researcher’s goal was to minimize any bias that could hinder the internal validity of this experimental design. Further, participant qualification criterion was determined based on responses to question four, six, and seven of the demographic survey. Questions four and six determined the eligibility requirements for the conventional flight groups. Of the sample population, the first fifteen that qualified for the no, low and high experience were included in this study. As regards question seven, those who answered yes were excluded from this study as they had a level of familiarity with UAS applications associated with medium-altitude long endurance platforms.

IRB Considerations

The researcher submitted an application to the Institutional Review Board (IRB) at both Nova Southeastern University and Embry-Riddle Aeronautical University. Applications at both institutions have been reviewed, expedited and approved (see Appendix C). Further, the researcher was proficient and up to date with certification as regards the Collaborative Institutional Training Initiative (CITI) on human subject's research. He understands the importance of consent, confidentiality, anonymity, and the overall importance of participant safety. The researcher ensured participants and participants' data remained anonymous and safe.

Evaluation Procedures

The first procedure in the study gathered literature as it pertained to human-computer interaction, human factors, usability, and learnability in UAS. The review of the literature afforded ample evidence to support the constructs presented in this study and to validate the methods in which data was collected. The second procedure incorporated the development of the demographic questionnaire (see Appendix B) to gather participant information for determining the purposive sample.

The third procedure described the experimental session. In a simulation experiment, 45 participants interacted with a high fidelity device that modeled the ground control station for a medium-altitude long endurance UAS. The initial segment of the experimental session consisted of the researcher summarizing the specific purpose and procedures applied. The consent form was administered during the initial session. After signing the consent form, participants underwent a ten-minute instructional session to provide them with some insight on the functionality of the ground control station. After completing the training session, the participants were allocated a ten-minute independent free flight session to become familiar with the system. Upon completion of the ten-minute free flight session, the researcher allotted a five-minute break. Session one lasted approximately thirty minutes.

Session two served as the experimental session. Participants were pre-assigned to one of the three experimental groups commensurate to their level of experience. Session two lasted approximately one hour. Participants performed a specific cognitive and psychomotor task three consecutive times during this session. The first attempt at the task was utilized as a measure of initial learnability while the third attempt was used to evaluate extended learnability and whether any incremental performance gains were exhibited by the independent participant groups between trials. Last, the SUS questionnaire was administered upon completion of session two and before the debrief.

Measurement and Instrumentation

This section discussed the measurement and instrumentation used in this study. The data collected in this study was defined as quantitative. According to Lazar et al. (2010) the three most common quantitative measurements for usability testing are task performance, time performance, and user satisfaction. Task performance pertains to how many tasks were correctly completed. Time performance often serves as an indication of how long each task took to complete while user satisfaction is often measured by a standardized and validated survey tool (Lazar et al., 2010). The System Usability Scale

(SUS; see Appendix A) served as the standardized and validated survey tool to measure user satisfaction in a post experiment fashion.

System Usability Scale

The International Standard Organization (ISO) provides a class of usability measures as documented in ISO 9241-11. According to ISO 9241-11, the three classes of usability measures are: (1) effectiveness, (2) efficiency, and (3) satisfaction and correspond to the quantitative description as described by Lazar et al. (2010). According to Brooke (1996) these classes provide a general idea for usability measures; however, the precise measures used within these classes can vary widely based on the specific system characteristics. Similarly, in a meta-analysis Grossmann, Fitzmaurice, and Attair (2009) found that learnability metrics are diverse across various usability studies and consensus for a set of well-accepted metrics for measuring learnability is lacking. The lack of a well-defined and standardized metric may not exist as Brooke (1996) suggested usability should be viewed from the lens of how appropriate the tool or system is to a specific purpose. Therefore, the diversity of metrics across HCI literature could simply imply and correspond to the diversity of tools or systems evaluated and their intended use.

Brooke (1996) suggested that it was impossible to specify the usability of any system (i.e., its purpose) without first defining the intended users of the system, the tasks those users will perform with the system, and the characteristics of the physical, organizational and social/environmental context in which the system will be used. Brooke (1996) claimed the SUS provided practitioners and researchers the ability to evaluate usability of any tool or system in terms of the context in which it is used, with its intended users, and its appropriateness to that context. Similarly, Sauro and Lewis (2012) described the SUS as a unidimensional measure of usability and yields a single composite measure for the overall learnability of the specific system being examined.

Validity and Reliability

Maximizing internal validity and reliability was essential to this research in examining the system learnability for a medium altitude long endurance UAS HMI representation. Gay, et al. (2012) described internal validity as the degree to which observed differences on the dependent variable result solely from the manipulation of the independent variable and not from any uncontrolled extraneous variables. Of primary importance was to collect data in a manner that was credible and reliable to minimize any threat to the internal validity of this experiment (Gay, et al., 2012).

Lazar et al. (2010) suggested that establishing internal validity corresponded to the development of a multi-faceted argument that supported the interpretation of collected data. Instituting standardized procedures, a fixed location, and consistent data collection methods enhanced the internal validity of this study. Lazar et al. (2010) described the first step in this process is to construct a database that includes all the materials that a researcher collects during the study. These materials include any notes, documents, tables, procedures and products of analysis.

The researcher collected each participant's performance data by extracting system state files from the simulator in a .csv format. Extracting the raw data for each participant from the simulator in a .csv format after each task offered consistency in data collection and will enhance the internal validity of this study. Data extracted in this format afforded the researcher to present the tabular presentation of the collected data. The raw data were stored electronically and all physical documents were kept in a secured filing cabinet. The researcher maintained authenticity of data in the raw form in an effort to trace any and all analytic results back to the original raw data, if desired. This method increased the reliability to repeat the study based on the original protocol and analytic steps (Lazar et al. 2010). Last, by enhancing the internal validity, external validity was also enhanced allowing the results to serve as a baseline representation of usability data in the UAS industry.

As regards the validity and reliability of the SUS, in a review of 2,324 SUS surveys, Bangor, Kortum, and Miller (2008) suggested the SUS is a highly reliable, valid, and sensitive psychometric useful for a variety of tools and interfaces. They found the reliability coefficient to be over .90. Similarly, Sauro and Lewis (2012) examined the sensitivity of the psychometric under a two-item sub-scale (i.e., Usable and Learnable) and found the coefficient alpha for Usable at 0.91 and 0.70 for Learnable. This is consistent with Bangor, et al. (2008). Therefore, the SUS served as a reliable metric to assess usable and learnable aspects of a system (Sauro & Lewis, 2012). The SUS was implemented as defined by Brook (1996) and without any modification.

Environment, Setting, and Apparatus

The research had access to a large controlled UAS laboratory environment used to train students enrolled in the Unmanned Aircraft Systems Bachelor of Science degree program at Embry-Riddle Aeronautical University (ERAU; see Figure 2) with different categories and representations of simulated UAS and their associated subsystems used for command and control.

Figure 2. UAS Simulation Center at ERAU depicting the X-Gen Simulator.

High fidelity simulators (see Figure 2) representative of the HMI for a medium altitude long endurance UAS served as the apparatus. The simulator depicted in figure two coined X-Gen incorporates the latest technology for UAS training from basic programming of flight procedures through full operational mission simulation and crew training. The X-Gen offered scalable high fidelity image generation software for real-time scene rendering. The open architecture design delivered a programmable six degree of freedom (DOF) aerodynamics model to closely mimic the air vehicle and controls. Figure 3 depicts the menu structures, moving map and system health display. The primary heads-up display HUD depicted in Figure 4 was the target display.

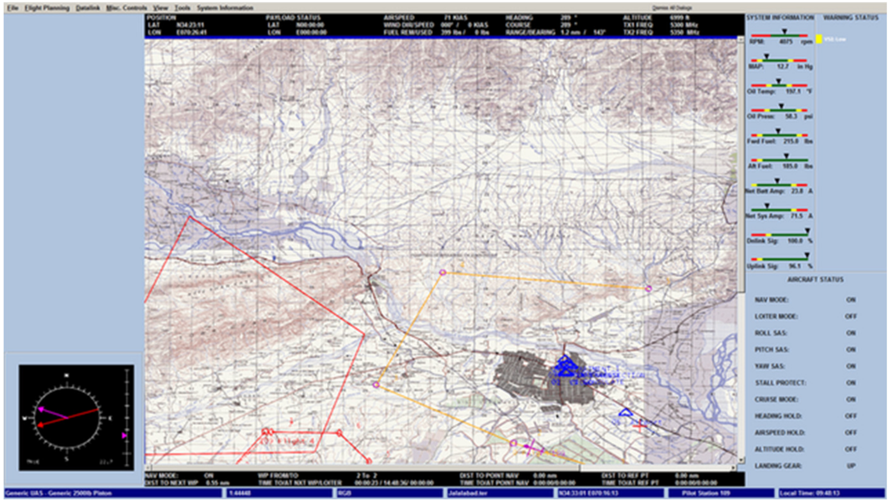


Figure 3. Map Tracker and System Health Display.

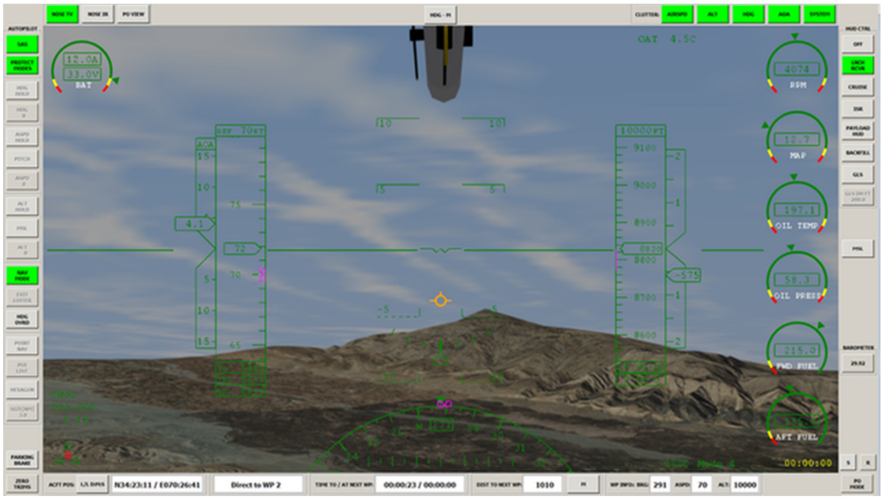


Figure 4. Primary Heads-Up Display.

The laboratory environment depicted in figure two is the UAS segment of the Advanced Flight Simulation Center (AFSC) at ERAU. The AFSC is home to a collection of high fidelity flight training devices (FTD) as well as a FAA Level D Full-Flight Simulator (FFS) with a six degree of freedom motion platform used for airline transport pilot training. In addition to the simulators, there are many training aids that provide students with the opportunity to engage in a variety of aircraft related cognitive and psychomotor exercises. Similarly, the UAS segment of the AFSC provides the prospect for students to practice an array of UAS related tasks which include crew resource management (CRM), autopilot programming, operations under various simulated aerodynamic effects, flight planning, and much more. Students' activities further range from the assembly, programming, and testing of small UAS (sUAS) to determining aircraft performance models for a specific operational mission using data from a medium altitude long endurance UAS and other platforms. Overall, the AFSC is a welcoming environment that provided the opportunity to familiarize and train for flight UAS specific operational tasks in a manner which is rather effective, efficient, and safe.

Format for Presenting Results

Data were gathered in numeric form and the researcher used Microsoft Excel to organize, sort, and analyze the data. The analysis of the dependent variables collected was evaluated to determine whether significant differences exist between the three participant groups. As for the SUS, Brooks (1996) suggested that it is neither cost-effective nor practical to perform full-scale context analysis as the precise measures within the basic metrics of usability vary significantly. He suggested selecting metrics based on effectiveness, efficiency and satisfaction that do not expend significant effort or cost to collect and analyze. Therefore, the SUS served as a reliable and valid metric across the HCI community to attain usability data quickly and in a very simple and low cost manner.

As regards data collection and scoring of the SUS, the SUS yielded a single number that served as a composite measure of the overall usability of the system. SUS scores have a range from 0-100 with participant response ranges based on a 5-point Likert scale. The scale ranges from 1 (i.e., strongly disagree) to 5 (i.e., strongly agree).

Once the SUS was completed, the researcher determined each item's score contribution, which ranged from 0 to 4. It is important to note that scores for individual items are not meaningful on their own (Brooke, 1996). Therefore, the researcher pooled the SUS scores based on the participants grouping variable for comparison analysis and also generated a composite SUS score consisting of all 45 participant surveys.

Sauro and Lewis (2012) suggested that for positively-worded items (i.e., 1, 3, 5, 7 and 9), the score contribution is the scale position minus 1. For negatively-worded items (i.e., 2, 4, 6, 8 and 10), the score contribution is the scale position minus 5. The researcher next calculated the overall SUS score and multiplied the sum of the item score contributions by 2.5. After completion, the researcher achieved SUS scores that ranged from 0 to 100 in 2.5 point increments as described by Sauro and Lewis (2012) Data were presented using tables, graphs, and charts in addition to a thorough annotated explanation of the results.

Chapter Summary

In Chapter 3 the methodology was described. The experimental research design included the identification of the dependent and independent variables. The sample population, IRB considerations, evaluation procedures, and the format for presenting the results was also stated. The System Usability Scale (SUS) served as a post experiment survey to better understand the learnability and overall usability of the HMI representation. Subsequently, the researcher addressed validity and reliability, the apparatus and laboratory environment, and the methods in which data were collected.

The researcher evaluated the system learnability of an industry-standard UAS HMI representation for the command and control of a medium altitude long endurance UAS as system usability has been poor for decades. Many suggested current HMI for UAS command and control are antiquated and require significant attention. The goal was to better understand usability testing for UAS man-machine systems and to provide the UAS industry with baseline usability data as regards the learnability of these HMI representations currently in place. From the researcher’s perspective, the study furnished important information to the UAS industry regarding the system usability of current HMI representation and furnished a foundation to expand further research in usability testing to advance the state-of-the art for UAS HMI designs.

Chapter 4

Results

Overview

This chapter presented the analysis and results of an ex-post facto experiment applied to investigate system learnability as it pertained to a representative human-machine interface (HMI) used in the command and control loop for medium altitude long endurance unmanned aircraft system (UAS). Usability as defined by the International Standard Organization (ISO) is the extent to which a product can be used by specific users to achieve task specific goals effectively, efficiently and with a high level of satisfaction. Correspondingly and as prescribed by ISO 9241-11, the three classes of objective usability measures include: (1) effectiveness, (2) efficiency, and (3) satisfaction (ISO, 1998).

The ISO measures for effectiveness, efficiency and satisfaction were attained using three primary methods for data collection across the three participant groups (i.e., no pilot experience, *n* = 15, low pilot experience, *n* = 15, high pilot experience, *n* = 15). The first method implemented to capture each participant's session and the interaction with the representative HMI under investigation was audio and video recording. Each participant's trial iteration was recorded and evaluated post-hoc. Second, the simulator provided a time-stamp data log for each participant's trial iteration. Data collected from each participant's trial iteration afforded the ability to extract data for the dependent variable total time on task. The data associated with task completion were coded as binary. The number of errors committed by a participant was scored as a count. The raw data set for total time on task as extracted from the data logs indicated a positive skew with a heavy tail. Data transformations were executed for total time on task to normalize these data in preparation for statistical analysis. Last, the System Usability Scale was implemented to gather participant perceived satisfaction measures. The instrument was administered in the original format as defined by Brooks (1996).

Analysis and Findings

This section describes the characteristics of the raw data and the modifications to the data to correct for the assumption of normality. Data were formatted to consist of three segments that presented the results for the dependent variables related to the ISO measures of effectiveness, efficiency and satisfaction. Measures for total time on task, task completion rate or success rate, errors, and post experiment satisfaction were examined using tables, graphs, descriptive and parametric inferential statistics.

In retrospect, the data for effectiveness were handled in a binary manner as this dependent variable corresponds to a participant's success or the ability to complete the task. Therefore, task completion or success rate was coded as either: (1) success or (0) failure. An error was defined as any unintended actions, slips, mistakes, or omissions a user made while attempting the task. Errors served as a measure of effectiveness and were recorded as a count of the total number of errors committed by a participant per trial iteration. As regards the numeric values associated with errors, some participant's resulted in no errors at all. Values of zero were handled by applying a 0.5 constant to the entire data set. The researcher then generated the descriptive statistics and subtracted the 0.5 constant from the mean score. The method of applying a constant is common to handle fields that contain missing data, negative data or zeros that interfere with the results of statistical analysis.

Efficiency typically refers to the way a product supports users in carrying out their tasks and typically relates to the task speed once users become familiar with the design. Sauro and Lewis (2012) described three methods for measuring task time: (1) task completion time, (2) time until failure, and (3) total time on task. Efficiency was determined by calculating the total time on task for each participant's iteration taking into consideration simulator start-up time and observable simulator lag or latency. The researcher then calculated the descriptive statistics for this dependent variable and calculated one-way analysis of variance (ANOVA). Bonferroni corrected t-test was calculated when ANOVA results indicated significant findings.

As regards the final dependent variable satisfaction, satisfaction referred to the user's perception on how enjoyable the computer-application served for the specific job task. Satisfaction was measured by implementing the System Usability Scale (SUS). The SUS instrument afforded a unidimensional measure of usability in the form of a single composite score. The SUS data were handled and analyzed based on the recommendations provided by Sauro and Lewis (2012) and Brooke's (1996). The final section of this chapter addressed the research questions postulated in Chapters 1 and 3.

Effectiveness

The task results and in particular, the task completion rates for the three participant grouping variables across the three trial iterations are presented in Table 1. Completion rates as described in Table 1 are graphically depicted in figure 5. The descriptive statistics are presented thereafter in Table 2. A one-way analysis of variance

(ANOVA) was also computed to determine whether any statistically significant differences existed with completion rates between and within subjects.

Table 1

Task Completion Results for Participant Grouping Variables

	No Pilot/No UAS			Low Pilot/No UAS			High Pilot/No UAS		
Participants	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3
1	1	1	1	1	1	1	1	1	1
2	0	0	0	0	0	0	1	1	1
3	1	1	1	0	1	1	0	0	0
4	1	0	0	1	1	1	1	1	1
5	1	1	1	0	0	0	1	1	1
6	1	1	1	1	1	1	0	1	1
7	1	1	1	0	0	0	1	1	1
8	0	0	0	1	1	0	1	1	1
9	0	1	1	0	0	0	0	0	1
10	1	1	0	1	1	1	1	1	1
11	1	1	1	0	0	0	1	1	1
12	0	0	0	0	0	0	1	1	1
13	1	1	1	0	0	0	0	0	0
14	1	1	1	1	1	1	0	1	1
15	1	1	1	0	0	0	1	1	1
Completion Rate	73%	73%	66%	40%	46%	40%	66%	80%	86%

Table 2

Descriptive Statistics for Task Completion Rates

	No Pilot/No UAS			Low Pilot/No UAS			High Pilot/No UAS		
	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3
Mean	0.73	0.73	0.67	0.40	0.47	0.40	0.67	0.80	0.87
Standard Error	0.12	0.12	0.13	0.13	0.13	0.13	0.13	0.11	0.09
Median	1.00	1.00	1.00	0.00	0.00	0.00	1.00	1.00	1.00
Mode	1.00	1.00	1.00	0.00	0.00	0.00	1.00	1.00	1.00
Standard Deviation	0.46	0.46	0.49	0.51	0.52	0.51	0.49	0.41	0.35
Sample Variance	0.21	0.21	0.24	0.26	0.27	0.26	0.24	0.17	0.12

Figure 5 illustrates the completion rates for all trial iterations across the three participant grouping variables. As depicted in Figure 5, the high pilot experience group exhibited higher levels of incremental performance gains when compared to the no pilot experience group and the low pilot experience participant groups. Alternately, completion rates for the no pilot experience group plateaued by their second iteration and declined slightly by their third iteration. When visually interpreting the data for the low participant group, it appears the completion rate for this group was observably lower than both the no pilot experience and the high pilot experience group for trial one and trial three.

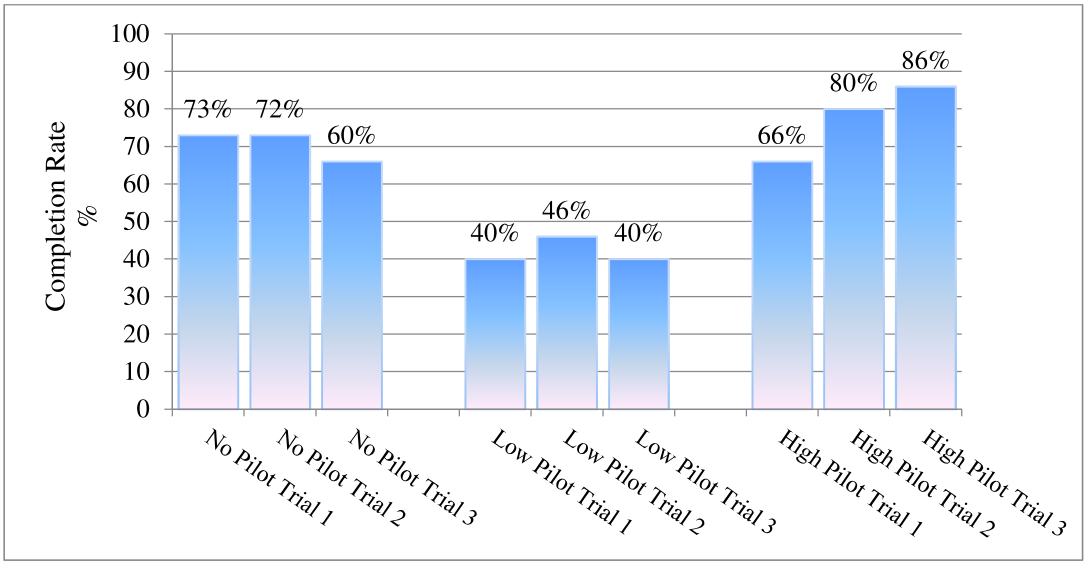


Figure 5. Participant Group Completion/Success Rates for Trial One, Two, and Three

The results of the analyzed data suggested the high pilot experience group performed the task most effectively than both the no pilot experience and the low pilot experience group. A one-way analysis of variance was calculated between subjects for trial one and trial three completion rates to discern whether any statistically significant differences were present between levels of participant experience and success rates for their first and third iteration. The one-way ANOVA computations for completion rates depicted below in Table 3 revealed statistical significance for trial three between subjects at [F (2, 42) = 3.98, p = 0.02]. Post hoc t-test using a Bonferroni correction revealed statistical significance in completion rates t (28) = -2.92, p<0.016 between the low experience group (M = 40%, SD =. 50) and high experience group (M = 86%, SD = .39). There were no other statistically significant findings for trial one or trial three between or within subjects.

Table 3

ANOVA: Between Subjects for Trial 3

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	1.644444	2	0.822222	3.984615	0.02603	3.219942
Within Groups	8.666667	42	0.206349			
Total	10.31111	44				

An assessment of the error rates in parallel with the completion rates for trial three also shows that the high pilot experience group committed less errors (M= 2.44, SD=3.9) during their third iteration when compared to the low pilot experience group (M= 9.53, SD= 12.63). Bonferroni corrected t-test was performed on participant error rates; however, the t-test statistic at t (27) = 1.95 with alpha set to p<. 05 and at p<.016 for the Bonferroni adjusted *p* value revealed no statistically significant differences between the high pilot experience group and the low experience group with regard to the number of errors committed during trial three.

Overall, the high pilot experience group (M = 86%, SD = .39) performed better than both the no pilot experience group (M = 66%, SD = .48) and low pilot experience group (M = 40%, SD =.50) with regard to task success and the number of errors committed. An assessment of completion rates for the low pilot experience group suggested no observable increase in performance between their first and third iteration. In essence, this group performed least effectively when compared to the two aforementioned groups. The no pilot experience group had a decrease in performance over time. The decrease in performance over time was computed statistically. The findings revealed no statistical significance.

As regards the high pilot experience group, there was an improvement in performance between their first and third iteration but the within subjects one-way ANOVA calculations indicated no statistical significance between trial one and trial three mean scores. However, when visually analyzing the graphed data, it is certainly apparent the high pilot experience group performed better overtime when compared to both the no pilot experience and low pilot experience group. The increase in performance for the high pilot experience group may be related to the fact that all participants were trained as commercial pilots in accordance to the training regimen prescribed by the Federal Aviation Administration (FAA). Aspects of conventional pilot training and at the commercial level (i.e., high pilot experience) could have influenced the performance increase as denoted between the trial iterations. This level of training may inherently provide individuals with the pre-requisite knowledge that translate skills more effectively from the conventional flight environment to the operation of medium altitude long endurance UAS. In this instance, the completion rates for the high pilot experience group suggested the system is quite learnable using task specific training; especially, for those who have a high level of conventional flight training and level of certification.

This is an important finding for the UAS training community as the representative medium altitude high endurance simulator revealed desirable task specific training outcomes for participants certified as commercial pilots. Further research is certainly warranted to determine which aspects of conventional pilot training may translate effectively to the operation of medium altitude long endurance UAS and the extent of simulator training versus training in the real world for operational effectiveness is highly desired.

Errors

Errors were defined as unintended actions, slips, mistakes, or omissions a user makes while attempting a task. Errors committed when interacting with the HMI reveal certain user interface problems and served as excellent diagnostic information to investigate the usability of this particular system. The descriptive statistics for errors committed by participants across the three trial iterations as a measure of effectiveness is presented in Table 4. All data fields that contained zero values were counterbalanced by implementing a 0 .5 constant across the data set followed by subtracting the 0.5 constant from the calculated mean.

Table 4

Descriptive Statistics for Participant Errors per Trial

	No Pilot/No UAS			Low Pilot/No UAS			High Pilot/No UAS		
	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3
Mean	9.80	5.13	3.20	15.40	9.87	9.53	12.20	5.13	2.44
Standard Error	2.37	1.26	0.90	4.58	3.00	3.26	4.19	2.18	1.01
Median	7.50	4.50	2.50	12.50	6.50	2.50	5.50	2.50	1.50
Mode	3.50	0.50	2.50	0.50	0.50	0.50	2.50	1.50	0.50
SD	9.18	4.88	3.49	17.73	11.64	12.63	16.23	8.43	3.89
Variance	84.31	23.84	12.17	314.40	135.41	159.55	263.38	70.98	15.16
Kurtosis	1.09	0.66	-0.80	1.57	5.67	3.45	3.03	9.50	5.35
Skewness	1.27	1.09	0.86	1.40	2.12	1.71	1.94	2.94	2.26

The task completion rates and the mean number of errors committed by each participant group across the three trial iterations are illustrated in Figure 6. In this case, error data graphed alone across trial iterations was not informative but when observed in parallel with completion rates, the data from the two dependent variables complemented one another to generate a more meaningful representation of these data.

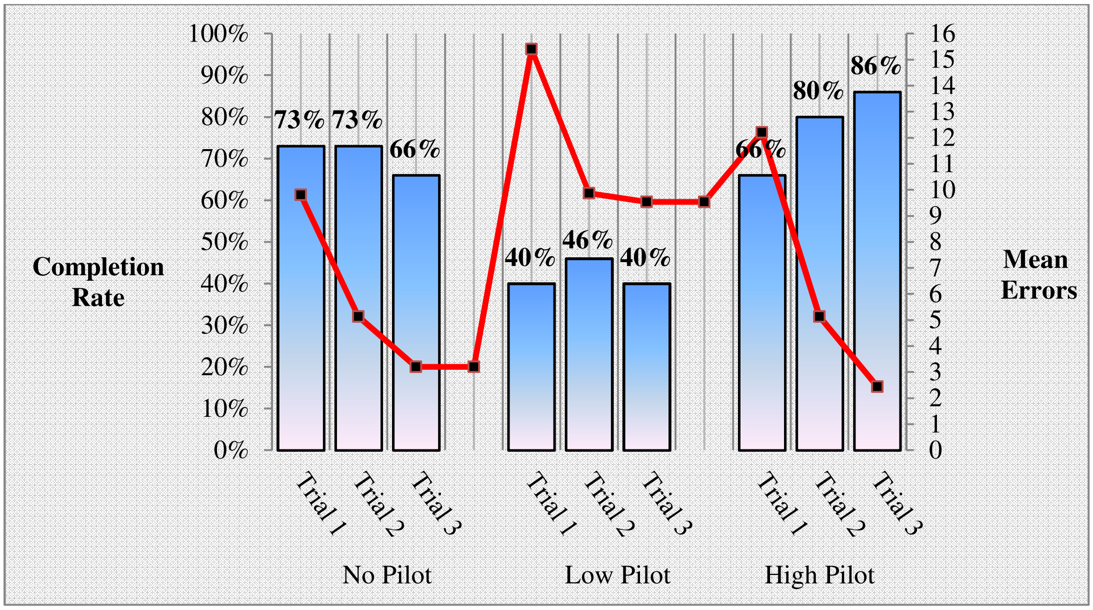


Figure 6. Task Completion Rate versus Error Rate

When observing the mean completion rates against the mean number of errors committed across the trial iterations, it is evident there was a reduction in committed errors across the three trials for all participant groups albeit, this finding is not significant as both the no pilot experience and low pilot experience group had a reduction in errors due to an increased failure rate rather than enhanced performance. In fact, performance declined between trial iteration one and trial iteration three for the no pilot experience group whereas the low pilot experience group revealed no performance gains. Alternately, when the trial iterations for the high pilot group were visually examined, there was an observable and desirable trend between task success and the number of errors committed. The inverse relationship of completion rate versus error rate signified improved performance overtime as completion rates increased and error rates decreased. This suggested that the high pilot experience group advocated a higher level of system learnability when compared to the no pilot experience and low pilot experience participant groups for this particular task on the representative device.

The ANOVA computations for a between subjects comparison of trial three error rates for the low experience group versus the high experience group is presented in Table 5. The purpose was to determine whether any statistically significant differences existed between the two groups related to the number of error committed for trial iteration three. The results suggested a significant effect on the number of errors committed on task between the two groups at the $p<.05$ level for the two grouping variables [$F(1, 28) = 4.31$, $p = 0.046$]. The high pilot experience group committed significantly less errors during trial iteration number three when compared to the low participant grouping variable for the same iteration. In fact, the low experience group committed the most errors during trial three (M = 9.53, SD = 12.63) when compared to the no pilot experience group (M = 3.20, SD = 3.49) and the high pilot experience group (M = 2.44, SD = 3.89).

It is anticipated that all users slip and make mistakes along the way. Unfortunately, mistakes in the UAS environment have significant ramifications and therefore, more intuitive display designs may aid to circumvent and reduce the number of errors or the criticality of the errors committed when performing operational tasks with this category of UAS. Understanding the variables associated with human performance degradation in the operational UAS environment may aid to design UAS interface with a representative model that promotes effective performance.

Table 5

ANOVA Output Source Table:

Number of Errors: Trial Three Low Experience vs. High Experience

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	377.3653	1	377.3653	4.319818	0.046952	4.195972
Within Groups	2445.989	28	87.35676			
Total	2823.355	29				

Efficiency

The descriptive statistics for the raw data set total time on task Table 6. The graphical representation of total time on task (raw) as depicted in Figure 7 suggested a positively skewed data set with a heavy tail to the right. Sauro and Lewis (2010) suggested the distribution of task time data is almost always positively skewed in usability studies.

Positively skewed data values tend to generate a mean score larger than the median as the mean is strongly influenced by participant task times with extended duration. Long task times inflate and attract the arithmetic mean thereby skewing the results of an analysis. Therefore, a data modification was required to normalize the raw data in preparation for statistical analysis.

Table 6

Descriptive Statistics for Total Time on Task (Raw)

	No Pilot/No UAS			Low Pilot/No UAS			High Pilot/No UAS		
	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 2
Mean	118.6	84.5	70.4	135.1	103.5	94.1	111.7	73.9	55.7
Standard Error	14.1	13.6	10.4	21.9	17.0	16.7	15.2	11.3	6.7
Median	103.0	64.0	68.0	94.0	90.0	77.0	91.0	76.0	42
Mode	66.0	#N/A	81.0	#NA	96.0	36.0	61.0	#NA	26
SD	54.4	52.8	40.2	84.6	65.9	64.5	58.8	43.9	674.5
Variance	2961.4	2791.0	1612.8	7162.9	4348.6	4160.7	3461.4	1929.3	-1.9
Kurtosis	-1.10	1.57	5.51	0.49	3.26	0.07	-0.92	5.70	-1.87
Skewness	0.56	1.43	1.99	1.23	1.91	1.13	0.40	1.91	0.25
Range	151.00	177.00	158.00	283.00	238.00	195.00	187.00	183.00	68.00

The skewness statistic in this particular case also indicated values greater than zero and in one particular case (i.e., no pilot group, trial 3) the skewness statistic reported a value of 1.99. Terrell (2012) suggested that skewness values greater than 2 interfere with the ability to use parametric inferential statistics. In this case, the researcher had the option to calculate nonparametric alternatives (Terrell, 2012) or to perform a logarithmic transformation to compensate for the positively skewed distribution (Sauro, 2011), or both. Within the scope of this research, the data were stabilized using the base-10 logarithmic transformation. Correspondingly, Sauro (2011) and Rummell (2017) ascertained the logarithmic transformation affords a convenient yet reliable method to normalize data for classical inferential statistical analysis.

The descriptive statistics for the transformed data set total time on task are presented in Table 7. As evident by the skewness statistic, the skewness was compensated for, as the values are evidently much closer to zero when compared to the skewness statistics for the raw data set. Also, a closer look at the sample variance also indicates an adjustment and a more normalized distribution when compared to the relatively high variance scores in the raw data set. In most cases, the greater the variance, the more the data is dispersed (Terrell, 2012).

Table 7

Descriptive Statistics for Total Time on Task (Log)

	No Pilot/No UAS			Low Pilot/No UAS			High Pilot/No UAS		
	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3	Trial 1	Trial 2	Trial 3
Mean	2.0	1.9	1.8	2.1	1.9	1.9	2.0	1.8	1.7
Standard Error	0.052	0.063	0.056	0.065	0.063	0.074	0.066	0.063	0.055
Median	2.0	1.8	1.8	2.0	2.0	1.9	2.0	1.9	1.6
Mode	1.8	#N/A	1.9	#N/A	2.0	1.6	1.8	#N/A	1.4
SD	0.2	0.2	0.2	0.3	0.2	0.3	0.3	0.2	0.2
Variance	0.0	0.1	0.0	0.1	0.1	0.1	0.1	0.1	0.0
Kurtosis	1.58	0.66	0.08	0.42	1.47	0.80	-0.73	0.17	-1.81
Skewness	0.16	0.50	0.48	0.32	0.17	0.21	-0.41	-0.06	0.00
Range	0.53	0.77	0.76	0.89	0.95	0.91	0.847	0.952	0.558

The graphical representation for total time on task (raw) is depicted in Figure 7. A histogram for total time on task (log) is presented in Figure 8. A polynomial trend line was added to both the raw and logarithmic transformed histograms as a method to visually enhance the distribution of the representative data. As depicted in Figure 7 the data set for total time on task is positively skewed. The distribution of these data violates the rules of parametric inferential statistics. The data set post logarithmic transformation is depicted in Figure 8.

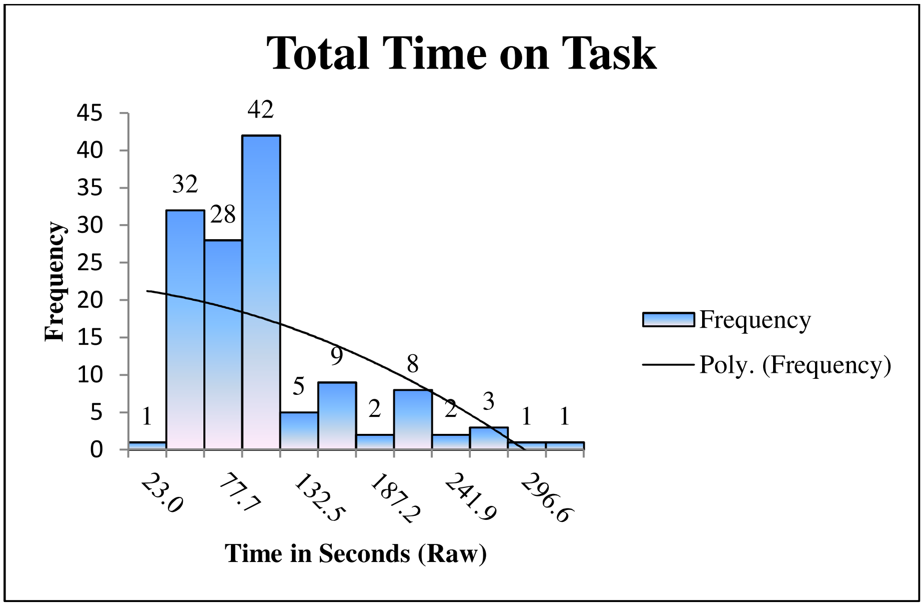


Figure 7. Total Time on Task in Seconds (Raw)

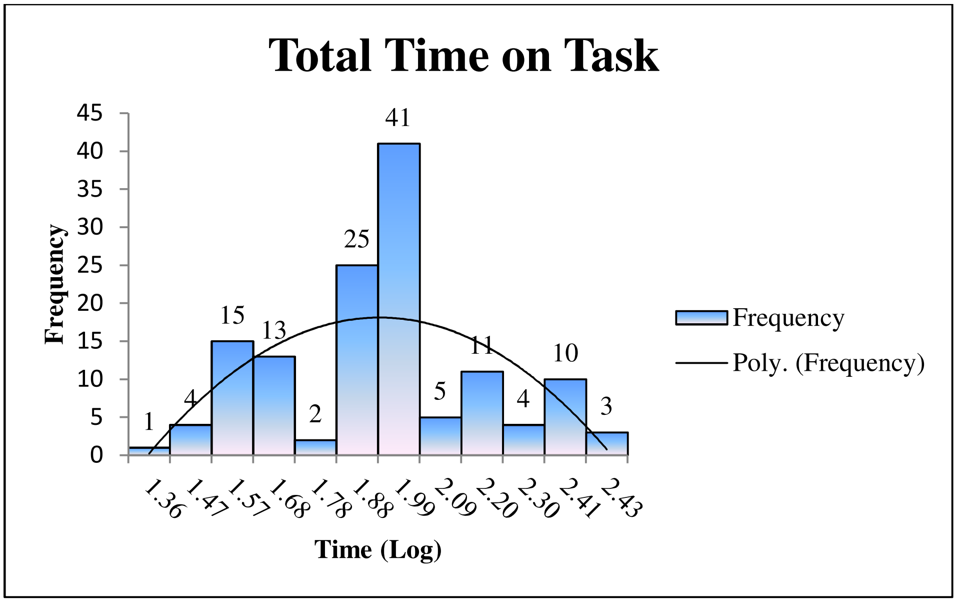


Figure 8. Total Time on Task (Log)

A bar graph of the raw mean scores for total time on task across the grouping variables (i.e., no, low, high) for trial one and trial three is illustrated in Figure 9. As a comparison, a bar graph of the base-10 logarithmically transformed mean scores for total time on task is presented in Figure 10.

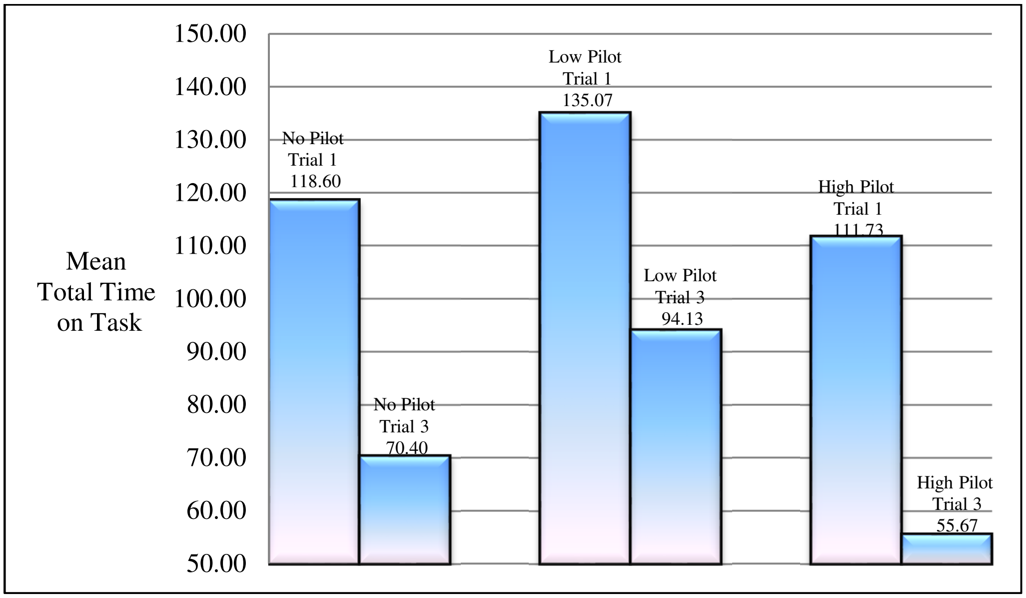


Figure 9. Mean Total Time on Task (Raw) Comparison for Trial One and Trial Three

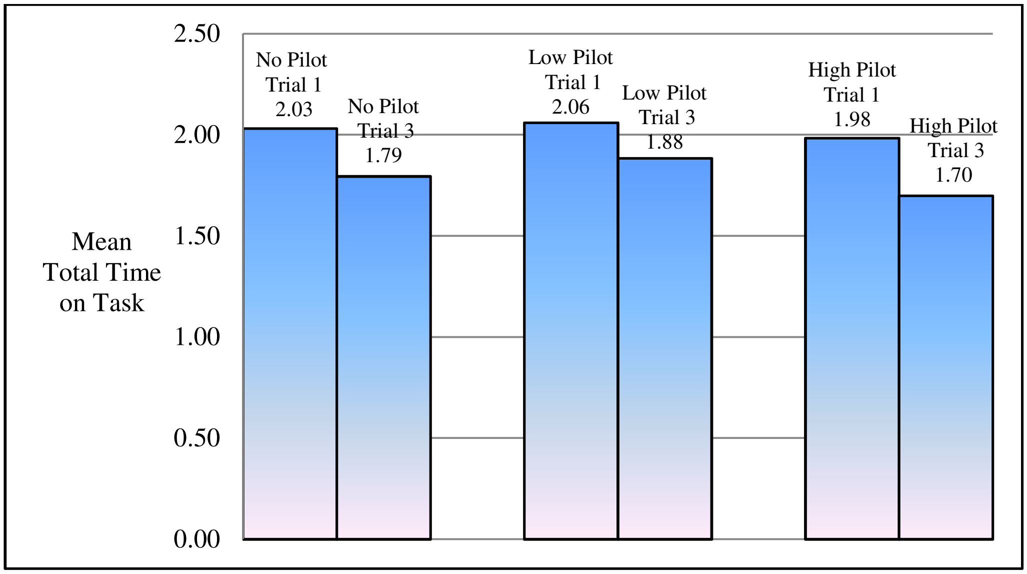


Figure 10. Mean Total Time on Task (Log) Comparison for Trial one and Trial three.

A one-way ANOVA was used on the logarithmic transformed data to determine whether any statistically significant differences existed between the mean scores for trial one and trial three of the three independent groups for total time on task. The results of the one way ANOVA indicated no significant difference between subjects for trial one or for trial three but did reveal statistical significance between trial one and trial in a within subjects comparison. The analysis of variance output table for a within subjects comparison of trial one and trial three mean scores for the low pilot no UAS experience group is presented in Table 8.

The goal was to determine whether participants exhibited a significant reduction in total time on task between their first and third iteration. There was a significant effect identified on the total time on task between trial one and trial three for the no pilot experience group at the p<.05 level for the two conditions [F (1, 28) = 9.71, p = 0.004]. However, this finding is related to an increased failure rate between trial one and trial three for the no pilot experience group rather than enhanced performance indicated by higher completion rates.

Table 8

Analysis of Variance Output Table:

Within Subjects Trial One vs. Trial 3: No Pilot Experience

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.421267	1	0.421267	9.712818	0.004202	4.195972
Within Groups	1.214424	28	0.043373			
Total	1.635691	29				

The reduction in completion rate although not statistically significant from (M= 73%, SD = .45) for trial one to (M = 66%, SD = .48) for trial three may be indicative of low system learnability. Similarly, it is important to note that participants classified as low pilot experience exhibited the lowest completion rate at (M = 40%, SD = .50) for trial one with no improved performance over time (M = 40%, SD = .50) for trial three. The low pilot group also committed a higher degree of errors (see Table 4 and Figure 5) than the no pilot and high pilot experience groups.

After examining success rate versus the number of errors committed by the no pilot experience group, it appears the total time on task reduction is prominent as fewer participant's completed the task during the third iteration (i.e., 66% success) versus their first iteration (i.e., 73%). In essence, a higher failure rate was exhibited during the third iteration (i.e., 34%) versus the first (i.e., 27%) for this participant group. This lends the researcher to conclude a low level of system learnability with regard to the interactions exhibited by the no pilot experience group.

The ANOVA computations for the within subjects comparison of trial one and trial three mean scores pertaining to the low pilot no UAS experience group exhibited no significant effect on the total time on task between trial one and trial three**.** However, and as depicted in Table 9, the analysis of variance output for a within subjects comparison of trial one and trial three mean scores for the high pilot experience group indicated a significant effect at $p<.05$ level between the two iterations [F(1, 28) = 10.9, p = 0.002].

Table 9

Analysis of Variance Output Table:

Between Subjects Design: High Pilot vs. No Pilot Experience

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.609979	1	0.609778	10.95447	0.002575	4.195972
Within Groups	1.559125	28	0.055683			
Total	2.169103	29				

Overall, the high pilot experience group revealed a significant reduction in total time on task between their first and third iteration. A significant effect was also determined within subjects for the high pilot experience group. The increase in success rate for the high pilot group for trial three (i.e., 86% success) versus trial one (i.e., 66% success) may indicate that extended learnability was achieved between the first and third iteration as participant's completed the task in a shorter span of time while committing less errors during trial three (M = 2.94, SD = 3.89) versus the number of errors committed during trial one (M = 12.7, SD = 16.23). Post hoc t-test using a Bonferroni correction revealed a decrease in the number of errors committed between trial one and trial three t (28) = 2.14, p<0.04 for the high pilot experience group.

The findings lend the researcher to conclude the high pilot experience group exhibited a higher level of system learnability, as improved interaction with the system was evident. This participant group's previous knowledge as it relates to flying conventional aircraft may have influenced their ability to quickly learn to become proficient with this system for this specific task.

Last, the researcher also compared the mean scores for level of experience on the dependent variable total time on task for trial one and trial three against the mean score of an expert who is proficient with executing the task. The average total time on task for an expert user was 35 seconds. To ensure consistency in comparisons, the researcher calculated the logarithmic value for the mean score of 35 seconds. The value of 1.53 served as a benchmark for total time on task. The results of the one-way ANOVA indicated a significant effect on the dependent variable total time on when compared with an expert benchmark for total time on task. A significant effect for this condition was revealed at p<.05 is [F (3, 56) = 21.0, p = 0.000000003] for trial one (see Table 10) and p<.05 is [F (3, 56) = 7.44, p = 0.000277912] and for trial three (see Table 11).

Table 10

Analysis of Variance Output Table:

Expert versus Levels of Experience for Trial One

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	2.687081	3	0.895684	21.08333	0.000003	2.769431
Within Groups	2.384729	56	0.042584			
Total	5.07181	59				

Bonferroni corrected t-test revealed significant effects for all trials when compared to the benchmark total time on task. All participant groups spent significantly longer on task than an expert on the same task. The comparison is important from a training perspective, as a significant amount of time and monetary resources are typically required to train individuals to operate these types of systems effectively and efficiently. Individuals who already possess a conventional flight rating at the commercial level may be the ideal candidates to select as UAS operators for medium altitude long endurance platforms as their level of aeronautical knowledge and performance is advanced and could translate effectively to the UAS operational environment. The selection of commercial rated pilots to operate medium altitude long endurance platforms may be beneficial and reveal itself in terms of cost savings associated with initial and recurrent training for agencies who implement medium-altitude long endurance UAS for operational use.

Table 11

Analysis of Variance Output Table:

Expert versus Levels of Experience for Trial Three

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.97598	3	0.32533	7.44857	0.00028	2.76943
Within Groups	2.44587	56	0.04368			
Total	3.42185	59				

Most often effective interaction with these types of systems requires a significant amount of system specific training to perform effectively and without errors. In some instances, error-free interaction may be compromised by inadvertent mistakes or slips often referred to as errors along the way rather than task completion errors, which ultimately could affect the success of a desired outcome.

Errors associated with the specific task at hand were not isolated or classified; instead, the researcher collected a total count of the number of errors a participant committed during each trial iteration and aggregated an average. Future research should isolate errors and attempt to classify these errors related to the specific HMI used for the operational environment. Standardized designs and manufacturing practices may aid to incorporate more intuitive displays; systems similar to the intuitive systems and displays embedded in today's sophisticated transport category and private business jet aircraft.

Satisfaction

Satisfaction was measured using the System Usability Scale (SUS) in the original format and without modification as defined by Brooks (1996). To analyze the SUS data, the researcher aggregated the data in two ways. First, the researcher separated the SUS data by participant grouping variable (i.e., no, low, and high) to evaluate the overall perceived satisfaction in a between subjects manner. Second, the researcher consolidated the SUS data for all 45 participants and compiled an overall composite SUS score.

The SUS instrument provided measures for a composite SUS score and two sub-scale scores: (1) usability and (2) learnability. To determine the learnability subscale, the researcher calculated the total of the SUS scores for items 4 and 10 and multiplied the total by 12.5. The data modification scaled the SUS scores from 0 to 100. To determine the usability subscale, the researcher calculated the total of the SUS scores for the remaining 8 items and multiplied the total by 3.125. Again, the data modification scaled the SUS data from 0 to 100 as prescribed by Sauro (2012).

Sauro (2012) suggested that there are 41 possible SUS score combinations that range from 0 to 100. He suggested the SUS is best understood when compared to an industry benchmark. The industry benchmark for an average SUS score is a 68. Sauro (2012) described anything above a 68 is considered above average and a 76 is considered a good SUS score.

The descriptive statistics for the SUS scores across the independent grouping variables are presented in Table 12. A one-way analysis of variance was executed to determine main effects on satisfaction across level of experience. The ANOVA output statistic for a between subjects comparison indicated no significant effects at $p<.05$ with [$F(2, 44) = .733$, $p = 0.486$].

Table 12

Descriptive Statistics:

SUS Scores across the Three Participant Grouping Variables

	No Pilot	Low Pilot	High Pilot
Mean	72.0	62.7	67.3
Standard Error	4.757	5.409	6.097
Median	77.5	55	75
Mode	82.5	75	75
Standard Deviation	18.42	20.95	23.61

The mean SUS scores between subjects are illustrated in Figure 11. Two trend lines were added to the representative bar graph. One trend line represents the SUS benchmark score of 68 and the second trend line denotes the score of 76 for what is considered a good SUS score (Sauro, 2012). When examining the SUS data between subjects, SUS data for two (i.e., low pilot and high pilot) out of the three groups fell below the industry benchmark of 68 (see Figure 11). The no pilot experience group rated the HMI at a 72, which was closest to the “good SUS score” of a 76.

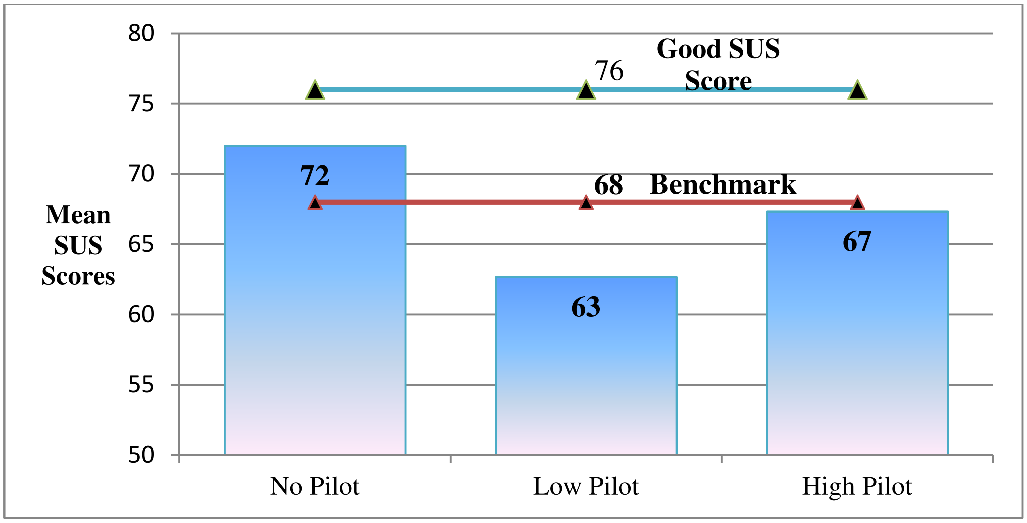


Figure 11. Mean SUS Score Comparison by Participant Grouping Variable

The mean scores for the primary scale SUS and the usability and learnability subscale are graphically depicted in Table 13. On a curved grading scale letter interpretation based on Sauro and Lewis (2012) the no pilot experience group measured at 72 which correlated to a C+. The low pilot experience group measured at 62.7 which correlated to a C- and the high pilot experience measured at 67.3 which correlated to a C. Overall, the letter grade interpretation revealed the system to be average with regards to user perception and satisfaction according to Sauro and Lewis (2012) whereas the rating scaled defined by Bangor, et al. (2008) issued a letter grade of a D for this system.

Table 13

Mean Scores for SUS, Usability and Learnability

	No Experience	Low Experience	High Experience
SUS	72.0	62.7	67.3
Usability	74.8	63.8	68.3
Learnability	60.8	58.3	63.3

Figure 12 illustrates the overall composite SUS score of a 67.3 and the two subscale scores for usability and learnability, 69.0 and 60.8 respectively on a scale from 0 to 100. This graph depicts the representative HMI based on independent participant responses and the HMI measures approximately at the 50th percentile when compared to the benchmark SUS score.

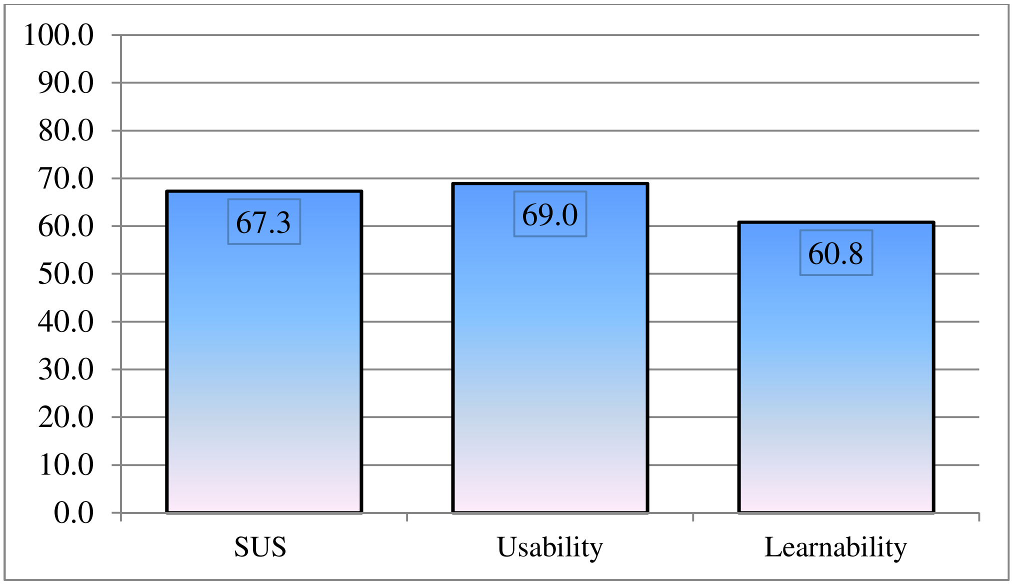


Figure 12. Mean System Usability Scale and Subscale Scores for all Participants

Next, the confidence intervals for the independent group SUS scores at a 95% confidence level are illustrated in Figure 13. The margin of error for each group and the overall SUS mean score are presented. The confidence interval provides a range of values that describe the uncertainty as it relates to a sample population. For this research, there is a 95% probability that the mean SUS scores could range between (61% -10.2 and 82% + 10.2) for the no experience group, (51% -11.06 and 74% +11.06) for the low experience group and (54% -13 and 84% +13) for the high experience group.

According to Sauro (2012) the confidence interval width and sample size have an inverse square root relationship which suggested that a significant increase in sample size relates to a reduction of the margin of error by its half. The confidence interval around the mean SUS score for the independent groups is presented in Table 14. The confidence intervals for the independent group SUS are graphically depicted in Figure 13.

Table 14

Confidence Intervals around the Mean SUS Score

CI around a SUS Score	
SUS Mean	67.3
SUS Standard Deviation	21
Sample Size	45
Low	60.9
High	73.6
Margin of Error	9%
Confidence Interval	95%

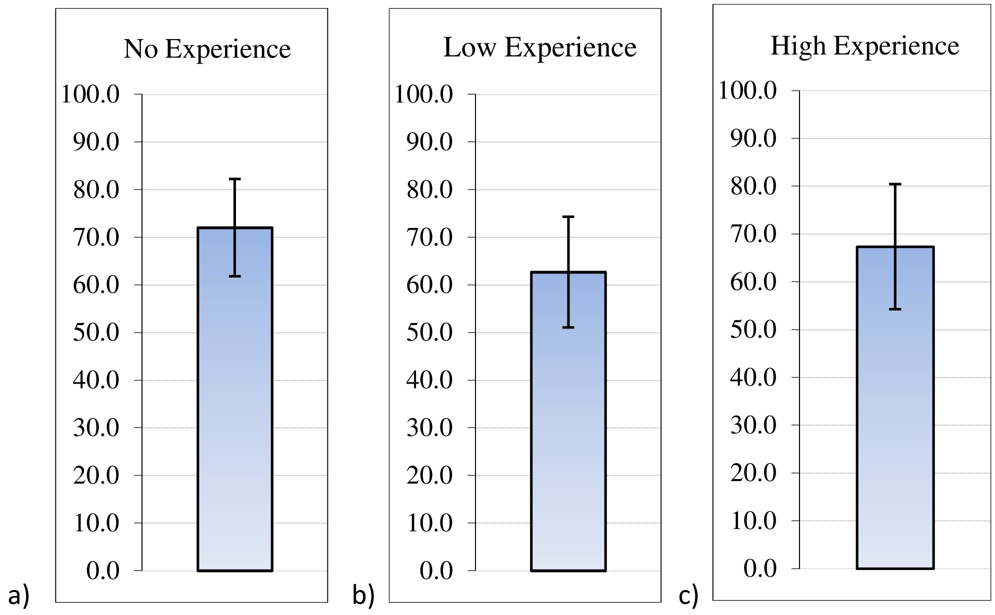


Figure 13. Confidence Intervals for the Independent Group SUS Scores

Overall, perceived satisfaction based on the results of the SUS survey suggested participants were somewhat satisfied using the HMI to reprogram a flight task. The overall composite SUS score was a 67.3 and the two subscale scores for usability and learnability were 69.0 and 60.8 respectively. These results indicate the representative HMI based on independent participant responses measures approximately at the 50th percentile and fell short from earning what is considered a good SUS score. Further research should investigate procedural tasks on expert users in an effort to collect SUS data specific to the operation of medium altitude long endurance UAS as expert perceived satisfaction is desired as an initial construct to build a mental representation of user needs for future HMI design.

Chapter Summary

Chapter 4 presented the statistical analysis and results of a causal comparative experimental design and procedures applied to investigate the system usability of a representative UAS HMI associated with medium altitude long endurance systems. The results revealed significant effects for task completion rates and the number of errors committed. Graphs and tables were used to visually present the data and parametric inferential statistics were used to objectively analyze outcomes as regards objective measures related to ISO 9241-11.

Chapter 5

Conclusions, Implications, Recommendations and Summary

Introduction

This causal comparative study examined system learnability of a representative medium altitude long endurance unmanned aircraft system (UAS) human-machine interface (HMI) on participants with varying levels of conventional flight experience and no previous experience with UAS. Data were collected as they related to the three classes of objective usability measures prescribed by the International Organization for Standardization (ISO) 9241-11. The three ISO classes included: (1) effectiveness, (2) efficiency, and (3) satisfaction.

Data collected for the dependent variables incorporated methods of video and audio recordings, a time stamp data log for participant time on task, and the SUS survey instrument. The results of this study provide baseline usability data as regards the system learnability of current HMI representations for medium altitude long endurance UAS from both a human-cognitive processing perspective and from a machine design perspective. The researcher implemented a purposeful sample to formulate the participant grouping variables. Usability data for this class of UAS are obsolete within the UAS industry and across the HCI community, which signified the importance of this study.

The remainder of this chapter presented the analysis of the findings. The analysis has been constructed to incorporate four primary sections. Section one presents the conclusions and addresses key elements regarding UAS HMI designs. Section two presents an analysis of the findings in the context of the research questions.

Subsequently, section three describes the implications and section four includes recommendations for future research. The chapter concludes with a summary of the research study.

Conclusions

The physical separation of human and machine regarding unmanned aircraft systems (UAS) imposes many elements of concern related to human factors and the full-scale integration of sophisticated UAS to operate in the National Airspace System (NAS). In fact, Vincenzi, et al. (2015) ascertained HMI for UAS command and control has been identified as a primary human factors concern as current HMI representations are a direct reflection of early development and rapid prototyping concentrating solely on the air vehicle and sensor payload with little emphasis placed on the HMI for command and control. Typically, UAS designs related to this category of UAS have been exclusively developed for the Department of Defense (DoD). The design of these systems has been service-branch specific and numerous systems have been developed in the last two decades for the various military branches. Unfortunately, coordination among UAS manufacturers has not been evident (Vincenzi, et al., 2015). These UAS vary considerably in systems design and therefore, practitioners often see a high variability in systems design when examining this class of UAS.

The element of inconsistency across systems also suggested inconsistencies across training regimes, as no two UAS are alike. The lack of consistency across systems design necessitates system specific training, which often results in significantly long training regimens to become proficient with this type of system. Long duration training regimens also translate into high monetary costs to develop operators' proficiency in terms of knowledge and skills to command and control a UAS of this class effectively and efficiently.

Of primary concern were the inconsistencies presented across UAS HMI designs as current models impose significant inadequacies that affect human performance and lead to task performance degradation. Current modalities for command and control are suboptimal as designs include: hierarchical menu structures to change system state parameters, QWERTY keyboards, joysticks, and trackballs to input commands for vehicle state, control and navigation. Holden, et al. (2013) suggested these modalities for vehicle command and control generate a significant level of overhead tasks for the operator which allocates the user's attention resources towards many secondary and tertiary tasks and not the primary task of operating the air vehicle.

Therefore, it is imperative to better understand key task elements, the level of automation required (LOA) for UAS command and control and user requirements to design interfaces that capture the mental representations of the user in an effort to optimize future UAS HMI designs. Congruently, Terwilliger et al. (2014) and Damliano et al. (2012) suggested that established design principles focused on UAS HMI coupled with a sophisticated level of automation could significantly improve usability of these systems and counterbalance the effect of degraded human performance. At present, the complexities of these sophisticated systems necessitate system specific training, as a high level of variability exists across UAS HMI. Further usability testing is required to investigate key task elements associated with this category of system, expert user requirements, and pre-requisite training models, as they presently exist for conventional pilot training to serve as a foundation for establishing guidelines in the form of policies and procedures for the UAS industry.

Damliano et al. (2012) suggested that even though the UAS ground control station (GCS) varies significantly from the conventional aircraft flight deck, practitioners should not disregard the established design practices and models used in conventional flight deck designs. These established guidelines incorporate human factors epistemology to enhance operator performance and have been validated as models over time. The same design guidelines could translate effectively as an underpinning to enhance design practices for the UAS industry. The notion is sound and consistent among UAS literature related to UAS HMI designs. Damliano et al. (2012) ascertained that when compared to conventional aircraft, UAS in today's arena have very low reliability as aspects of LOA, automation interaction, display design and operating training have yet to be considered at a level required to advance the state-of-the art for UAS HMI designs and ultimately, the full-scale integration of UAS into the NAS.

This study examined system learnability as it related to medium altitude long endurance UAS HMI in an effort to provide the industry with a form of baseline usability data regarding current HMI representations for this category of UAS. This study found the system learnability associated with the representative HMI to be sub-optimal as the representative HMI representations in the industry. The following section presents the analysis of the results obtained from this study constructed in the framework of the research questions as presented in Chapter 1 and 3.

Analysis in the Framework of Research Questions

Question 1: How accurately did task completion rates such as task completion time, time until failure, total time on task, and errors (Sauro & Lewis, 2012) serve to measure the learnability of the UAS HMI representation?

The dependent measures served well to examine system learnability related to participants with varying levels of conventional flight experience. Data collected from each participant’s simulator activity afforded the researcher the ability to extract data for the dependent variable total time on task. Total time on task served as a measure of efficiency. In this particular case, the data was solely used to verify how long each collective group spent on the task. The data collected from the audio and video recordings afforded the ability to accurately extract the task completion time and the number of errors committed by each participant for each trial iteration. The data associated with task completion and the number of errors committed appeared to be the most meaningful in this examination of system learnability.

Analyzing the task completion rates and errors combined, these data suggested that higher levels of incremental performance gains were exhibited for the high pilot experience group when compared to the no pilot experience group and the low pilot experience participant groups. The high pilot experience group performed significantly better than the low pilot experience group indicated by a significant increase in task performance by their third trial iteration. A steep decline in the number of errors committed was also evident when examining this group’s error rate. The no pilot experience group outperformed the low pilot experience group but the findings were not statistically significant. With regard to the low pilot experience group, this group performed the least effective and committed the most errors on the task. Further research is deemed necessary to investigate the knowledge, skills and abilities associated with the private pilot certificate on UAS flight tasks as this certificate is required for some UAS employment opportunities to command and control UAS. Determining whether this certification corresponds to these UAS activities would be beneficial for the UAS industry.

As an educator of aeronautical science, it is often observed that commercial rated pilots exhibit an advanced level of aeronautical knowledge and psychomotor skill set that affords the ability to quickly perform and improve their performance in an air vehicle that holds characteristics of a conventional flight deck. The medium altitude long endurance UAS; especially the General Atomics Predator/Reaper UAS has often been identified as the most airplane like application. Although different, there are basic elements that correspond to general conventional flight and navigation. Therefore, it is imperative to determine which characteristics are shared between the conventional flight deck and the UAS HMI for medium-altitude long endurance platforms. This will aid developers to refine current UAS HMI towards more effective and intuitive command and control applications.

Future research should also examine the knowledge, skills, and abilities of conventional commercial rated pilots on UAS command and control to determine which components of conventional pilot training transfer effectively to the UAS flight environment. The investigation of this construct may lead the UAS industry to establish the criterion to establish a training paradigm for UAS operators. This recommendation is also consistent with Williams (2012) as he found those rated as civilian pilots exhibited better performance on a UAS task than those not rated as civilian pilots. Last, the findings associated with task completion rate is deemed most valuable as an observable and desirable trend between task success and the number of errors committed was observed for the high pilot experience group. The researcher concludes that the high pilot experience group advocated a higher level of system learnability when compared to the no pilot experience and low pilot experience groups for this particular task on the representative device.

Question 2 Were participants satisfied with the level of interaction to perform the specific set of operational UAS tasks as regards the System Usability Scale (Brooks, 1996)?

The descriptive statistics for the composite SUS score was (M = 67.3, SD = 21.0). This overall composite score indicated that the representative HMI measured approximately at the 50th percentile when compared to an industry benchmark score of 68 on the SUS. Overall, system learnability fell below average. According to the Sauro and Lewis (2012) rating scale, the score of a 69 earns the HMI a letter grade of a C- whereas the rating scale defined by Bangor, et al. (2008) issued a letter grade of a D- for this system. Future work should examine UAS usability on expert users to devise a baseline SUS score specific to this category of UAS and its users. This will afford researchers to specifically evaluate and compare a multitude of UAS designs using the SUS instrument based on expert operator perceptions. Baseline SUS scores would serve invaluable to the industry.

Question 3 Based on the System Usability Scale as scored by Sauro and Lewis (2012), did participants find the UAS HMI usable and learnable?

Data collected as it related to the SUS sub scores of usability and learnability suggested a below average usability sub score of 69% and a low learnability sub score of 60.8% on a scale from 0 to 100. As regards the learnability subscore, this system exhibited low levels of learnability based on the overall composite score. Participants with no previous flight experience appeared to be the most satisfied with the system whereas the low pilot experience group was most dissatisfied. Low SUS scores typically translate into ineffective and inefficient HMI designs. Further research is warranted to determine design inadequacies in current UAS HMI to aid in the development of more usable devices that offer intuitive characteristics that may help to improve the safety of operation by circumventing degraded performance. Overall, system learnability based on the composite SUS score ascertained an HMI representation that exhibited low poor system learnability.

Question 4: Was incremental learning exhibited as participants become more familiar with the HMI (i.e., reduction in terms of task completion rates and errors)?

After close examination of the completion rates for the high pilot group trial three (i.e., 86% success) versus trial one (i.e., 66% success), it appeared extended learnability was achieved between the first and third iteration for the high pilot experience group as this group completed the task in a shorter span of time while committing less errors during trial three (M = 2.94, SD = 3.89) when compared to the number of errors committed during trial one (M = 12.7, SD = 16.23).

The results of the one-way ANOVA suggested a significant effect on the number of errors committed on task at the $p<.05$ level [$F(1, 28) = 4.31$, $p = 0.046$]. Post hoc tests suggested the high pilot experience group committed significantly fewer errors during their third trial iteration when compared to the low pilot experience participant-grouping variable for the same trial iteration. In fact, the low pilot experience group committed the most errors during trial three ($M = 9.53$, $SD = 12.63$) when compared to the no pilot experience group ($M = 3.20$, $SD = 3.49$) and the high pilot experience group ($M = 2.44$, $SD = 3.89$).

The findings lend the researcher to conclude that the high pilot experience participant group exhibited a higher level of learnability as improved interaction with the system was evident. Last, upon examining the trial iterations for the high pilot experience group, there was an observable and desirable trend between task success and the number of committed errors for this group as completion rates increased and the number of errors committed decreased between their first and third iteration. Incremental performance gains were exhibited by the high pilot experience group only.

Question 5 To what degree did the level of conventional flight experience (i.e., subsequent learning) impact system learnability as regards the dependent variables and perceived satisfaction when compared to those without any conventional flight experience?

Overall, the high pilot experience group performed more effectively and efficiently than both the low pilot experience group and the no pilot experience group with regard to task completion rates and the number of errors committed. A level of significance was noted between the low pilot experience group and the high pilot experience group but no statistical significance was exhibited between any other group comparisons. With regards to level of satisfaction based on SUS measures, this research found no statistically significant differences between perceived satisfaction based on independent SUS and level of conventional flight experience. The results of the ANOVA did not reveal any statistical significance between the three groups on the level of satisfaction.

Last, the final conclusion points were presented. The final points suggested a dire need for a method to generate comprehensive guidelines and policies to streamline industry efforts towards the full-scale integration of UAS into the NAS.

- Study results suggested that the level of a participant's previous flight experience might have had an impact on system learnability from a human performance perspective when evaluating objective data based on ISO 9241-11.
- Based on the experimental design of this study, task completion rates and the number of errors committed across trial iterations were most valuable for evaluating previous levels of conventional pilot training on UAS system learnability.
- Study results indicated that the low pilot participant group (i.e., private pilots) performed the least effective when interacting with this system while the high pilot experience group (i.e., commercial pilots) performed most effective.
- Surprisingly, the no pilot experience group outperformed the low pilot experience group which provides more reason to determine the knowledge, skills, and abilities for UAS operators as the traditional manned aviation paradigm for pilot certification is often used as a method to quantify the level of experience for operators across the UAS industry as no established guidelines to steer the

industry are in place. Congruently, this study also supported the notion on the importance of established design principles for HMI command and control interface. This notion is consistent across literature examining the inadequacies of UAS HMI and is reiterated as a final comment in this study.

Implications

This study ascertained a critical need for future research in the domain of unmanned aircraft systems user interface designs and operator requirements as the UAS industry is experiencing revolutionary change at a very rapid rate. Small UAS platforms have been incorporated into uncontrolled segments of the NAS but many safety-related issues remain unresolved threatening the expansion of UAS in the NAS. Safety-related issues stem from the absence of common policies and regulations and disseminate into to design, manufacturing, and operating inadequacies, training inefficiencies, inconsistencies in physical and logical control orientation across vehicles (Maybury, 2012) and irregularities in display design technology, terminology, and symbology (Terwilliger, Ison, Vincenzi & Liu, 2014).

The anticipated milestone of 2025 for full-scale integration of UAS into the NAS as noted by the FAA and reported by the GAO (2012) is on the horizon. Limited research has been conducted to investigate the aforementioned inadequacies. Policies and regulations to guide UAS design manufacturing and the training requirement to operate more sophisticated systems other than small UAS is obsolete in the civilian environment. The information provided in this study aids the necessary attention for research in the UAS domain. The development of comprehensive policies and regulations to guide systems design, and the development of comprehensive guidelines to steer UAS operator training to meet the demands of consumers of UAS is imperative and necessitated.

Recommendations

The safe, effective, and efficient integration of UAS into the NAS highly depends on the development of policy and legislation to establish airworthiness certification programs, a classification framework for air vehicles and associated subsystems, to identify the knowledge, skills and abilities/attitudes required of UAS operators based on vehicle classification, to define operator training and certification requirements, and to institute detect and avoid (DAA) collision avoidance solutions among many more. Of particular interest was the design of usable HMI that promote good learnability characteristics as a mechanism to enhance UAS operator performance. Usability testing is a common method to evaluate the extent in which a user can operate a product but unfortunately this method of evaluation appears to be obsolete in the UAS industry. Maybury (2012) ascertained the need for usability testing but suggested the high variability in systems designs complicates usability testing for UAS applications placing them on the "avant-garde" for usability evaluations.

At present, comprehensive policy and legislation to define guidelines that translate into effective and usable HMI criteria have yet to be established. This closely implicates significant safety concerns for the full-scale integration of UAS in the NAS (Jimenez, et al., 2016). Current HMI representations reflect early design characteristics of rapid prototyping in which manufacturers concentrated solely on the air vehicle and sensor payload to meet military demands. These design characteristics have seeped into today's commercial UAS market as little emphasis is placed on the design of the HMI for

UAS command and control (Vincenzi, et al., 2015). In fact, the majority of current UAS HMI representations are non-intuitive, non-effective and non-efficient (Vincenzi, et al., 2015; Terwilliger, et al., 2014; Maybury, 2012; Williams, 2004). This is of particular interest as the success of any UAS mission depends highly on the operator's ability to effectively use the information attained from the HMI.

Jimenez, et al. (2016) suggested a well-designed HMI is one that enhances the flying experience by increasing an operator's situational awareness with regard to system state parameters and the surrounding flight environment. They described a well-designed HMI, as one that improves a user's response time if consistency in terms of shared characteristics exists between the system design and the user's mental representation of the system and system state throughout a flight regime. The notion ascertains the fact that information displayed on the HMI should move consistently with the operator's mental representation and match the dynamic operational environment to reduce the stimulus-response time associated with operational task performance.

Jimenez, et al. (2016) described this model as the dimensional overlap model and a mechanism to design UAS HMI where both the stimulus set and the response set share identical elements or characteristics to improve the usability of the UAS HMI. The model affords the ability to trigger automatic responses in users through design characteristics that incorporate the same stimulus and response sets. In turn, this may extenuate system learnability revealed by an increased probability for users to perform tasks successfully and with a reduction in the number of errors committed.

Holden, et al. (2013) suggested that inadequate design of displays and controls significantly impacts a user's problem solving capability by creating discord between the user's mental representations versus the external representation of the system. User interfaces that do not hold similar characteristics between the systems design and the user's mental representation often promote poor usability. This was certainly evident in this study as pilots with a high level of flight experience performed the operational task most effectively during their third iteration when compared to the no pilot and low pilot experience groups. This performance gap might suggested that the commercial pilots had the greatest mental representation of the system as incremental performance gains were observed over time. Further research is deemed necessary to investigate whether the mental models associated with aviation flight training at the commercial level corresponds to UAS flight.

Based on the results attained from this study, the following recommendations are proposed relevant to the state-of-the industry as regards unmanned aircraft systems:

1. *Establish comprehensive guidelines and standards for airworthiness certification in the design and development of UAS and UAS HMI for command and control.* Damliano, Guglieri, Quagliotti, and Sale (2011) and Terwilliger et al. (2014) suggested that considerable research investment has been vested in the development of new HMIs for modern conventional aircraft displays in an effort to improve operator performance through user-centered designs; however, little research attention has been expended to investigate the potential problems associated with current UAS HMI for command and control. Similarly, Jimenez, et al. (2016) ascertained the need to establish guidelines that aid manufacturers of

UAS in developing HMI for UAS command and control as industry guidelines are yet to be set in place.

A regulatory framework to guide industry standards for UAS systems design is highly desired as poor human factors and ergonomics (HFE) present in today's HMI designs directly stem from the absence of common policy and regulations (Maybury, 2012). Williams (2004) ascertained that improved design practices established on aviation display concepts and development focused around the tasks of the user may aid to reduce human error by improving overall system usability. At present, minimal research has been conducted. The same inconsistencies as presented in literature more than a decade ago still loom in the industry.

2. *Establish comprehensive guidelines to classify the complexity associated with UAS systems design.* At present, there are two industry classification systems in place (i.e., European and Department of Defense) to describe and define the complexity associated with the variety of UAS systems on the marketplace. Neither the European nor the Department of Defense classification system has become an industry standard. In fact, when practitioners discuss and describe UAS system capabilities, both modalities for classification are discussed in a compare and contrast fashion. An accurate classification model is required to establish design guidelines that correspond directly to the complexity of vehicle traits, as this is imperative to steer the direction and characteristics of the design. Chimbo et al., (2011) suggested the aim of design guidelines, standards and principles is to help designers improve the usability of their products by designing

in accordance to rules that aid developers to make successful design decisions. Design guidelines are set in place to restrict the range of aircraft design decisions that may negatively affect a product's overall usability (Chimbo, et al., 2011).

3. *Investigate mechanisms to develop comprehensive guidelines and regulations to guide UAS operator training.* According to Pavlas et al. (2009) future training paradigms should incorporate both human-focused knowledge and equipment-focused knowledge to ensure that cognitive and psychomotor skills could be shared across the dynamic elements dictated by specific mission criteria using a variety of UAS platforms. Similarly, Jimenez, et al. (2016) suggested a standard for UAS display would aid the UAS community in many ways. They suggested a standardized display for UAS applications will simplify training initiatives and allow operators to transfer their training across multiple platforms without the need to learn a new system reducing the potential for mishaps caused by human error.
4. *Develop methods and processes to optimize UAS interface design through automation integration and adaptive display technologies.* Optimization characteristics to generate user-friendly applications are necessitated to expand the state-of-the-art for UAS HMI display designs. Design features that enable efficient operator inputs (i.e., short task duration) for command and control and navigation, a level of automation scalable to assist with primary, secondary and tertiary tasks, sensory feedback in the form of visual, vestibular and tactile to enhance levels of situational awareness of the operating environment and displays that provide effective and efficient information exchange between human and

machine is highly desired. Last, optimization can be achieved by harnessing the disparity with regard to the level of automation presented by current UAS designs. Damliano et al., (2012) ascertained that the level of automation corresponds directly to the HMI design as the type of feedback, the display layout, and control interface may vary in accordance to the level of implemented automation. Therefore, a framework to define characteristics associated with LOA and function allocation between human and automation must also be described as a vehicle towards display optimization.

5. *Adopt methods and metrics to evaluate human-computer interface for UAS applications with regard to system usability and system learnability.* When a process development model is not implemented iteratively throughout the design of an HMI, the design entails characteristics often noted as inconsistent, ineffective, inefficient, and dissatisfying to users (Holden, et al., 2013). The design of displays and controls is of significant importance especially since these displays and controls are the primary interfaces with which operators command and control the air vehicle. Holden, et al. (2013) ascertained the design of these components is of critical importance. Incorporating standard methods and metrics to evaluate HMI representations for system usability is highly necessitated.

Final Summary

A causal-comparative design was applied to investigate system learnability as it pertained to a representative human-machine interface (HMI) used in the command and control of medium altitude long endurance unmanned aircraft system (UAS). The International Standard Organization (ISO) defines usability by the extent to which a product can be used by specific users to achieve task specific goals effectively, efficiently and with a high level of satisfaction. Correspondingly, data were collected to examine the results based on the three classes of objective usability measures as prescribed by the ISO 9241-11. The ISO measures for effectiveness, efficiency and satisfaction were attained using three primary methods for data collection across the three participant groups (i.e., no pilot experience, *n* = 15, low pilot experience, *n* = 15, high pilot experience, *n* = 15)**.** Metrics for total time on task, task completion rate or success rate, errors, and post experiment satisfaction were presented using tables, graphs, descriptive and parametric inferential statistics. The System Usability Scale questionnaire was distributed post-hoc as a method to collect participant's satisfaction ratings.

The results pertaining to efficiency indicated a significant reduction in total time on task for the high pilot experience group when comparing their first and third iteration at a $p<.05$ level between the two iterations [F (1, 28) = 10.9, p = 0.002]. Additionally, the reduction in total time on task was accompanied by an increase in success for the high pilot group and complemented by a reduction in the number of errors committed between trial three (M = 2.94, SD = 3.89) and trial one (M = 12.7, SD = 16.23). The findings lend the researcher to conclude that this participant group exhibited a higher level of system learnability, as improved interaction with the system was evident. This participant group’s previous knowledge relating to flying conventional aircraft might have influenced their ability to quickly learn to become proficient with this system for this specific task.

Interestingly, the low pilot experience group performed in a least effective manner when compared to the high pilot and no pilot group. Overall, the results related to task success suggested the high pilot experience group exhibited higher levels of incremental performance gains when compared to the no pilot experience group and the low pilot experience participant groups. Data associated with completion rates and errors suggested the no pilot experience grouped performed better than the low pilot experience group. In fact, the low pilot experience group indicated very low learnability of the system, as completion rates appeared to be the lowest of the three groups. This group also committed the highest rate of errors on the prescribed task.

Data associated with the SUS were analyzed using a commercially available SUS calculator. The SUS instrument provided measures for a composite SUS score and two sub-scale scores: (1) usability and (2) learnability. Sauro and Lewis (2012) suggested anything above a 68 is considered above average and a 76 is often considered a good SUS score. At a score of 76, the notion is the interface scored higher than 75% of all products on the market tested using the SUS. On a curved grading scale letter interpretation based on Sauro (2012), the no experience group measured at 72 which correlated to a C+, the low experience group measured at 62.7 which correlated to a C-, and the high experience group measured at 67.3 which correlated to a C. The overall composite SUS score was a 67.3 and the two subscale scores for usability and learnability were 69.0 and 60.8 respectively. The industry benchmark for an average satisfaction on the SUS is a score of 68. These results indicated the representative HMI based on independent participant responses measures approximately at a 50th percentile short from earning a good SUS score. Further research should investigate procedural tasks on expert users in an effort to collect SUS data specific to the operation of medium altitude long endurance UAS as expert perceived satisfaction is desired as an initial construct to build a mental representation of user needs for future HMI design.

This study ascertained a critical need for future research in the domain of unmanned aircraft systems designs and operator requirements as this industry is experiencing revolutionary change at a very rapid rate. The lack of legislation in the form of policy to guide the scientific paradigm of unmanned aircraft systems has generated significant discord within the UAS industry leaving many facets associated with the teleportation of UAS in dire need of research attention. As regards the current state for user interface, practical HCI usability testing is obsolete from the industry (Maybury, 2012). Last, the researcher believes this study furnished important information on the criticality for sound HCI principles in UAS applications and introduced the HCI community to a facet of usability testing related to complex UAS user interface as poor system usability has been identified as a leading cause for sub-optimal human performance in UAS operations.

Recommendations for future work included a need to (1) establish comprehensive guidelines and standards for airworthiness certification in for the design and development of UAS and UAS HMI for command and control, (2) establish comprehensive guidelines to classify the complexity associated with UAS systems design, (3) investigate mechanisms to develop comprehensive guidelines and regulations to guide UAS operator training, (4) to develop mechanisms that lead to UAS HMI design optimization, and (5) to adopt methods and metrics to evaluate human-machine interface related to UAS applications for system usability and system learnability.

The proliferation of UAS is on the horizon as this technology is well suited to accommodate both commercial and public industries by providing robust and flexible capabilities for a range of applications often considered dull, dirty or dangerous (Austin, 2010). The final recommendation to the safe integration of this technology is for UAS stakeholders to provide initiatives that encourage research in areas requiring answers to fundamental questions as we move forward towards UAS integration into the NAS.

Appendix A

System Usability Scale

System Usability Scale (Brooke, 1996)

	Strongly disagree				Strongly agree
1. I think that I would like to use this system frequently	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
2. I found the system unnecessarily complex	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
3. I thought the system was easy to use	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
4. I think that I would need the support of a technical person to be able to use this system	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
5. I found the various functions in this system were well integrated	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
6. I thought there was too much inconsistency in this system	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
7. I would imagine that most people would learn to use this system very quickly	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
8. I found the system very cumbersome to use	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
9. I felt very confident using the system	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
10. I needed to learn a lot of things before I could get going with this system	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

Appendix B

Demographic Survey

Introduction

The purpose of this survey is to establish a participant sample population for a study investigating the usability and in particular, the learnability associated with the Human-Machine Interface (HMI) for the command and control of a simulated medium altitude long endurance unmanned aircraft system (UAS).

As a potential participant in this study, you will have the opportunity to interact with a simulated UAS representative of a system often used by public agencies. To ensure that the most recent information is documented as regards flight hours and experience, please have your pilot log book available when completing this survey. This survey should take no longer than 5 minutes to complete.

If you meet the desired criteria for the sample population, the principal investigator for this project will contact you. Please provide a current email and phone number at the end of the survey.

Demographic Information

1. Age
   ____

2. Current year in college
   ___ Freshman
   ___ Sophomore
   ___ Junior
   ___ Senior
   ___ Graduate Student

3. Gender
   ___ Male
   ___ Female

4. Are you certificated as a rated pilot by the Federal Aviation Administration (FAA)?
   ___ Yes
   ___ No

   4a. If yes, what certificate do possess?
   ___ Student Pilot for Single Engine Land Airplane
   ___ Private Pilot for Single Engine Land Airplane
   ___ Commercial Pilot for Single Engine Land Airplane
   ___ Other (Please List):
   ____
   ____
   ____
   ____

   4b. Do you have an instrument rating?
   ___ Yes
   ___ No

5. Are you certified as a flight instructor?
   ___ Yes
   ___ No

6. What is your total flight time?
   ____

7. Have you ever flown a medium altitude long-endurance unmanned aircraft system in the real world or in a simulated environment?
   ___ Yes
   ___ No

8. How many hours a day to you spend on a computer?
   ___ Less than 1 hour
   ___ 2-4 hours
   ___ 5- 7 hours
   ___ 8-10 hours

9. How many hours a day do you spend playing video games?
   ___ Less than 1 hour
   ___ 2-4 hours
   ___ 5- 7 hours
   ___ 8-10 hours

10. Have you ever experienced motion sickness when focusing on a fixed point such as when looking at a computer monitor or playing video games?
    ___ Yes
    ___ No

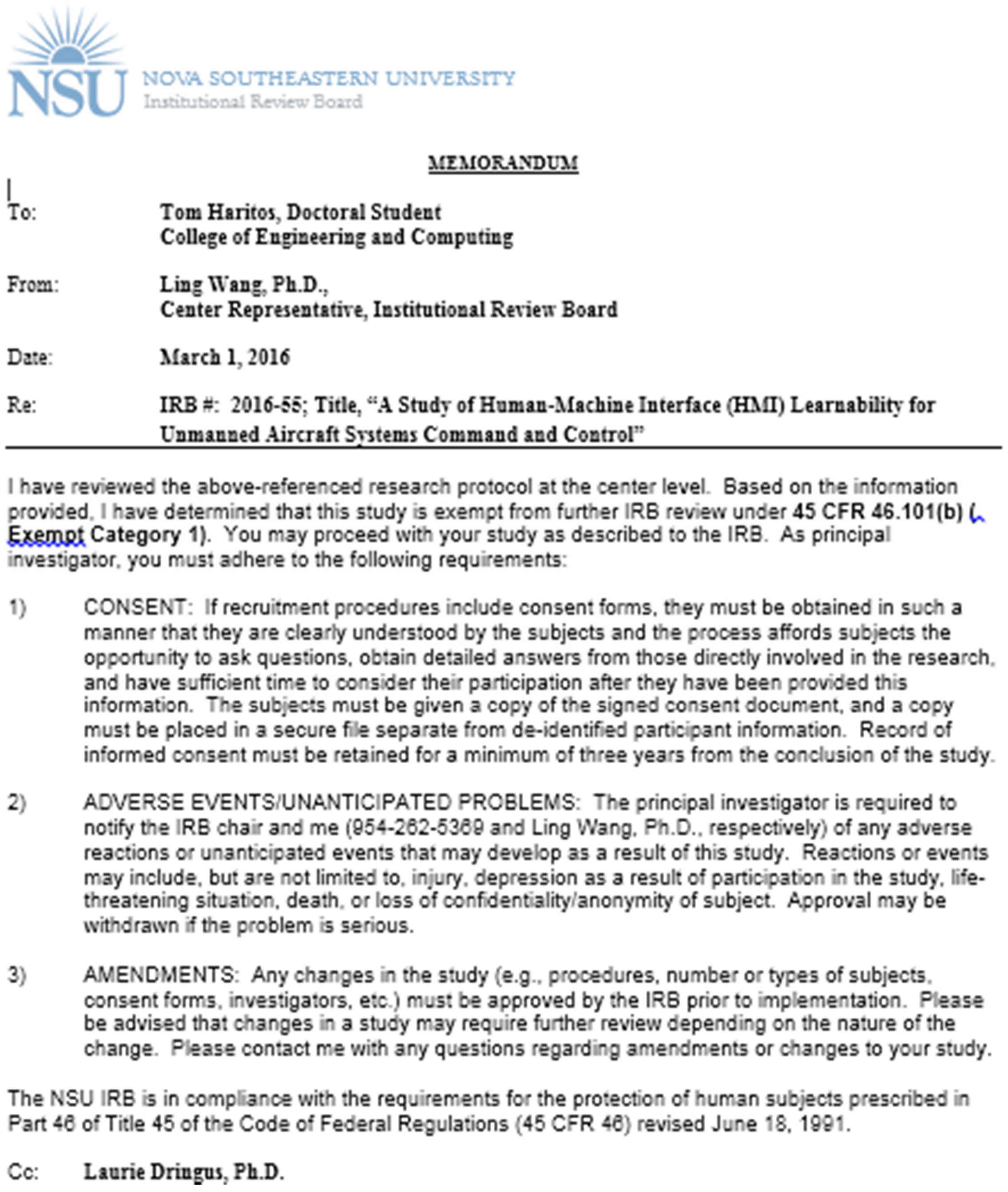
    10a. If yes, please explain____
    ____
    ____
    ____
    ____

Contact Information

1. Name: ___________________________________
2. Email :___________________________________
3. Phone: ___________________________________
4. What is the best time to contact you?
   ____

Appendix C

Institutional Review Board Documentation



NOVA SOUTHEASTERN UNIVERSITY
Institutional Review Board

MEMORANDUM

To: **Tom Haritos, Doctoral Student**
College of Engineering and Computing

From: **Ling Wang, Ph.D.,**
Center Representative, Institutional Review Board

Date: **March 1, 2016**

Re: **IRB #: 2016-55; Title, "A Study of Human-Machine Interface (HMI) Learnability for Unmanned Aircraft Systems Command and Control"**

I have reviewed the above-referenced research protocol at the center level. Based on the information provided, I have determined that this study is exempt from further IRB review under **45 CFR 46.101(b) (Exempt Category 1)**. You may proceed with your study as described to the IRB. As principal investigator, you must adhere to the following requirements:

1) CONSENT: If recruitment procedures include consent forms, they must be obtained in such a manner that they are clearly understood by the subjects and the process affords subjects the opportunity to ask questions, obtain detailed answers from those directly involved in the research, and have sufficient time to consider their participation after they have been provided this information. The subjects must be given a copy of the signed consent document, and a copy must be placed in a secure file separate from de-identified participant information. Record of informed consent must be retained for a minimum of three years from the conclusion of the study.

2) ADVERSE EVENTS/UNANTICIPATED PROBLEMS: The principal investigator is required to notify the IRB chair and me (954-262-5369 and Ling Wang, Ph.D., respectively) of any adverse reactions or unanticipated events that may develop as a result of this study. Reactions or events may include, but are not limited to, injury, depression as a result of participation in the study, life-threatening situation, death, or loss of confidentiality/anonymity of subject. Approval may be withdrawn if the problem is serious.

3) AMENDMENTS: Any changes in the study (e.g., procedures, number or types of subjects, consent forms, investigators, etc.) must be approved by the IRB prior to implementation. Please be advised that changes in a study may require further review depending on the nature of the change. Please contact me with any questions regarding amendments or changes to your study.

The NSU IRB is in compliance with the requirements for the protection of human subjects prescribed in Part 46 of Title 45 of the Code of Federal Regulations (45 CFR 46) revised June 18, 1991.

Cc: **Laurie Dringus, Ph.D.**

Embry-Riddle Aeronautical University
Application for IRB Approval
Expedited Determination

Principle Investigator: Tom Haritos **Other Investigators:** Laurie Dringus **Role:** Faculty
Campus: Daytona Beach **College:** COA

Project Title: ***A Study of Human-Machine Interface (HMI) Learnability for Unmanned Aircraft Systems Command and Control***

Submission Date: 11/18/2015 **Determination Date:** 2/17/2016

Review Board Use Only

Exempt: No

Approved: M.B. McLatchey Digitally signed by M.B. McLatchey DN: cn=M.B. McLatchey, o=Embry-Riddle Aeronautical University, ou=Institutional Review Board for the use of Human Subjects, email=mai.gabriel@erau.edu, c=US Date: 2016.02.17 13:28:07 -05'00'

Chair of the IRB Signature

Brief Description: The purpose of this study falls within the realm of human-computer interaction (HCI) and aims to examine design inadequacies related to the usability of a human-machine interface (HMI) for the command and control of a medium altitude long endurance (MALE) UAS by evaluating participant interactions with a HMI in terms of effectiveness, efficiency, and user satisfaction. Of particular interest is the level of learnability exhibited with the design of industry standard UAS HMI for MALE systems as these interfaces have been highly scrutinized for being non-intuitive, difficult to learn, and difficult to use. Modes for information presentation and information exchange have been cited as a leading cause in human performance errors (Cooke, 2008). High fidelity simulators representative of the ground control station for a medium altitude long endurance vehicle will serve as the apparatus for this study. The simulator coined X-Gen incorporates the latest technology for UAS training of basic programming and flight procedures through full operational mission simulation.

This research falls under the **expedited** category as per 45 CFR 46.110 under:

> 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)

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

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

Appendix D

Officials Granting Permission to Use Facilities

Officials Granting Permission to Use Facilities

This is a formal request by Assistant Professor **Tom Haritos** of the Department of Aeronautical Science/Unmanned Aircraft Systems Science to use the **Unmanned Aircraft System (UAS) Laboratory** and specifically, the **X-Gen UAS simulators to include the flight data recording capabilities** located at the **Advance Flight Simulation Center** on the **Daytona Beach Campus** for the purpose of preparing and conducting the experimental research segments of his dissertation during the 2016-2017 academic calendar year. All events will be coordinated as to not interfere with any scheduled class activities.

Guided by the principles of usability research, this study aims to investigate learnability in UAS interface design as the current state of human-computer interaction (HCI) for UAS applications is at its infancy. The goal of this research is to establish baseline usability data of the learnability of current human-machine interface (HMI) representations for the command and control of medium-altitude long endurance (MALE) UAS systems. The study will furnish relevant information to the aviation community of the significance for sound HCI principles in future UAS HMI designs and introduce the HCI community to usability testing in complex UAS applications.

All data collected for this research is subject to confidentiality and anonymity as per the Institutional Review Board (IRB).

Authorized Signatures:

Michael Wiggins, Ed.D.
Chair
Department of Aeronautical Science
Daytona Beach Campus

Date: 11/30/15

Nikolas "Dan" Macchiarella, Ph.D.
Interim Dean
College of Aviation
Daytona Beach Campus

Date: 12-8-2015

Appendix E

Participant Recruitment Briefing

Participant Recruitment Briefing

Principal investigator
Tom Haritos, M.S., Ed.S.
600 S. Clyde Morris Blvd.
Daytona Beach, FL 32114-3900

(386) 226-6447 (Office)
(321) 960-8551 (Cellular)

Co-investigator
Laurie Dringus, Ph.D.
3301 College Avenue
Fort Lauderdale, FL 33314

(954) 262-2073 (Office)

What is the study about?

You are invited to participate in a research study designed to investigate the learnability of a human-machine interface (HMI) used for the command and control of a medium-altitude long endurance (MALE) unmanned aircraft system (UAS). The simulator used in this study is a representative model for this category of UAS. As a participant you will interact with this high fidelity system to aid in addressing elements of system effectiveness, efficiency and user satisfaction.

Learnability is one of the five quality attributes that formally define the term “usability”. The learnability of a system refers to ease of use for first time users to perform basic functions or tasks while engaged with a computing application or interface.

The goal of this study is to better understand the design characteristics of these HMI through participant interaction in the form of usability testing as industry systems have yet to be investigated thoroughly from a HMI perspective. The goal of this research is to provide the UAS industry with usability data of the learnability for current HMI representations used in the command and control of MALE UAS. The study aims to furnish important information to the fields of aviation of the significance for sound HCI principles in future UAS HMI designs and to introduce the HCI community to usability testing in complex UAS applications. This study is not evaluating your performance but instead evaluating the ease of use of the UAS HMI representation for command and control.

Why are you asking me?

We are inviting you to participate in this study for one of three reasons. There will be 45 participants for this research. Of the 45 participants, 15 will have no conventional flight experience and no previous UAS experience, 15 will have low levels of conventional flight experience with no previous UAS experience, and 15 participants will have a high level of conventional flight experience with no previous UAS experience. Based on one these criteria you have been selected to participate in this study as you fit the participant profile demographic for one of the three groups.

Appendix F

Consent Form for Participation

NOVA SOUTHEASTERN UNIVERSITY
College of Engineering and Computing
NSU

Consent Form for Participation in the Research Study Entitled
A Study of Human-Machine Interface (HMI) Learnability for Unmanned Aircraft Systems Command and Control

Funding Source: None.

IRB protocol #

Principal investigator	Co-investigator
Tom Haritos, M.S., Ed.S.	Laurie Dringus, Ph.D.
600 S. Clyde Morris Blvd.	3301 College Avenue
Daytona Beach, FL 32114-3900	Fort Lauderdale, FL 33314
(386) 226-6447 (Office)	(954) 262-2073 (Office)
(321) 960-8551 (Cellular)	

For questions/concerns about your research rights, contact:
Human Research Oversight Board (Institutional Review Board or IRB)
Nova Southeastern University
(954) 262-5369/Toll Free: 866-499-0790
IRB@nsu.nova.edu

Site Information
Embry-Riddle Aeronautical University
Department of Aeronautical Science
Advanced Flight Simulation Center (AFSC)
Unmanned Aircraft Systems Laboratory
600 S. Clyde Morris Blvd.
Daytona Beach, FL 32114-3900

What is the study about?

You are invited to participate in a research study designed to investigate the learnability of a human-machine interface (HMI) used for the command and control of a medium-altitude long endurance (MALE) unmanned aircraft system (UAS). The simulator used in this study is a representative model for this category of UAS. As a participant you will interact with this high fidelity system to aid in addressing elements of system effectiveness, efficiency and user satisfaction.

Learnability is one of the five quality attributes that formally define the term “usability”. The learnability of a system refers to ease of use for first time users to perform basic functions or tasks while engaged with a computing application or interface.

The goal of this study is to better understand the design characteristics of these HMI through participant interaction in the form of usability testing as industry systems have yet to be investigated thoroughly from a HMI perspective. The goal of this research is to provide the UAS industry with usability data of the learnability for current HMI representations used in the command and control of MALE UAS. The study aims to furnish important information to the fields of aviation of the significance for sound HCI principles in future UAS HMI designs and to introduce the HCI community to usability testing in complex UAS applications. This study is not evaluating your performance but instead evaluating the ease of use of the UAS HMI representation for command and control.

Why are you asking me?

We are inviting you to participate in this study for one of three reasons. There will be 45 participants for this research. Of the 45 participants, 15 will have no conventional flight experience and no previous UAS experience, 15 will have low levels of conventional flight experience with no previous UAS experience, and 15 participants will have a high level of conventional flight experience with no previous UAS experience. Based on one these criteria you have been selected to participate in this study as you fit the participant profile demographic for one of the three groups.

What will I be doing if I agree to be in the study?

If you agree to participate in the study, you will initially undergo a 15 minute training session to provide you with some initial information and instruction on the functionality of the ground control station (i.e., the HMI) commensurate for engaging the experimental task(s) using the UAS HMI. After completing the initial training session, you will be allocated a fifteen minute independent free flight session with written documentation (i.e., operation’s manual) for the HMI investigated in this study. Upon completion of the 15 minute independent flight session, the researcher will answer any questions you may have and will provide you with a five minute break. Session one is estimated to last approximately 30 minutes.

Session two will serve as the experimental session. Participants will have been pre-assigned to one of the three experimental groups commensurate to their level of experience. Session two is estimated to last approximately .5-1.0 hour. Participants will perform a specific cognitive and psychomotor task three consecutive times during this session. The first attempt at the task will be used to measure initial learnability while the third attempt will be used to evaluate extended learnability between first and last interaction.

Upon completion of the experimental session two, the researcher, Mr. Tom Haritos will ask you to complete a 10 question survey coined the System Usability Scale (SUS). The SUS is designed to rate your experience and satisfaction in terms of usability as regards your interaction with the HMI. The survey should take you no more than 10 minutes to complete followed by a 5-10 minute debrief scheduled to last no more than 10 minutes.

Is there any audio or video recording?

This research project may include audio and video recording of the entire experimental session. This recording will be available to be viewed and heard only by the researcher, Mr. Tom Haritos, personnel from the IRB if required, the dissertation chair, Dr. Laurie Dringus, and the members of this dissertation committee, if deemed necessary. The audio/visual recordings will be reviewed and transcribed, if necessary by Mr. Tom Haritos. He will use earphones while reviewing the experimental session in a locked and private office to guard your privacy. The recording and all data will be kept securely in Mr. Haritos's locked private office and in a locked filing cabinet. The recording will be kept for 36 months from the end of the study. The recording will be destroyed after that time by properly reformatting the hard drive of the device in which the digital data is contained and shredding any paper-based documents. Because your voice will be potentially identifiable by anyone who hears the recording, your confidentiality for things you say on the recording cannot be guaranteed although the researcher will limit access to the recording using the parameters as described in this paragraph.

What are the dangers to me?

Risks to you are minimal, meaning they are not thought to be greater than other risks you experience every day in home, school, or work. The experiment for this study will be conducted in the advanced flight simulation center (AFSC) at Embry-Riddle Aeronautical University (ERAU). The simulators used in this study are found in a large classroom located in the AFSC. The classroom is separated as part simulation laboratory and part traditional classroom for face-to-face instruction. The risks associated with the interaction of the simulators used for this study and in this classroom is no greater than the risk of entering a building and sitting down at an office desk with multiple computer interface presented before you. The devices are fixed based and offer no relative physical motion. The participants will essentially interact with four computer interface and peripherals to include: joystick, mouse, keyboard, and throttle quadrant).

Some may experience slight simulator sickness caused by vection but the vection provided in these simulated scenes is no different or no greater than the vection found in today's home gaming domain. However, if you feel ill at any time, please notify the researcher immediately. Under these circumstances the session will end.

Note: As regards being recorded means that confidentiality cannot be promised.

If you have questions about the research, your research rights, or if you experience an injury because of the research please contact Mr. Tom Haritos at (386) 226-6447. You may also contact the IRB at the numbers indicated above with questions about your research rights.

Are there any benefits to me for taking part in this research study?

There are no benefits to you for participating in this research study.

Will I get paid for being in the study? Will it cost me anything?

There are no costs to you or payments made for participating in this study.

How will you keep my information private?

All information obtained in this study will remain strictly confidential and will be disclosed only with your permission or as required by law. The data collected for this study will be coded using pseudonyms, numeric, and/or alphanumeric techniques. Identifiable participant information will be maintained in a locked filing cabinet within the Aeronautical Science Department at Embry-Riddle Aeronautical University. This information will be kept separately from the rest of your experimental data.

The questionnaire will not ask you for any information that could be linked to you. The transcripts of the visual and digital recordings will not have any information that could be linked to you. As mentioned, all data will be destroyed 36 months after the study ends. All information obtained in this study is strictly confidential unless disclosure is required by law. The IRB, regulatory agencies, or Dr. Laurie Dringus may review research records.

What if I do not want to participate or I want to leave the study

You have the right to leave this study at any time or refuse to participate. If you do decide to leave or you decide not to participate, you will not experience any penalty or loss of services you have a right to receive. If you choose to withdraw, any information collected about you **before** the date you leave the study will be kept in the research records for 36 months from the conclusion of the study and may be used as a part of the research.

Other Considerations:

If the researcher learns anything which might change your mind about being involved, this information will be provided to you.

Voluntary Consent by Participant:

By signing below, you indicate that

- this study has been explained to you
- you have read this document or it has been read to you
- your questions about this research study have been answered
- you have been told that you may ask the researchers any study related questions in the future or contact them in the event of a research-related injury
- you have been told that you may ask Institutional Review Board (IRB) personnel questions about your study rights
- you are entitled to a copy of this form after you have read and signed it
  you voluntarily agree to participate in the study entitled *A Study of Human-Machine Interface (HMI) Learnability for Unmanned Aircraft Systems Command and Control*

Participant's Signature: ____________________________ Date: ________________

Participant’s Name: _______________________________ Date: ________________

Signature of Person Obtaining Consent: ______________________________

Date: ____________________________

References

Austin, R. (2010). *Unmanned aircraft systems: design, development, and deployment*. Reston, VA: American Institute of Aeronautics and Astronautics.

Bangor, A. Kortum, P., & Miller, J. A. (2008). The System Usability Scale (SUS): An empirical evaluation. *International Journal of Human Computer Interaction, 24*(6), 574-595.

Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In P. W. Jordan, B. Thomas, B. A. Weerdmeester, & A. L. McClelland (eds.), *Usability Evaluation in Industry*. London: Taylor and Francis.

Chairman of the Joint Chiefs of Staff (2011). *Joint unmanned aircraft systems minimum training standards* (Report No. CJCSI 3255.01). Washington D.C. Retrieved from http://www.dtic.mil/ cjcs_ directives/cdata/unlimit/3255_01.pdf.

Chimbo, B., Gelderblom, J. H., & DeVilliers, M. R. (2011). Engageability: A new sub-principle of the learnability principle in human-computer interaction. *Journal for Transdisciplinary Research in South Africa*, *7*(December), 383–406.

Cooke, N.J. (2008). Unmanned aircraft systems: Human factors issues. *NTSB Forum on the Safety of Unmanned Aircraft Systems, April, 2008, Washington DC*.

Cooke, N. J., & Pedersen, H. K. (2010). Unmanned aerial vehicles. In J. A. Wise, V. D. Hopkin, & D. J. Garland (Eds.), *Handbook of Aviation Human Factors* (pp. 18–1–18–7). Boca Raton.

Cunningham, I. (2008). Are “skills” all there is to learning in organizations? The case for a broader framework. *Development and Learning in Organizations*, *22*(3), 5–8. doi:10.1108/14777280810861749.

Dalamagkidis, K., Valavanis, K. P., & Piegl, L. A. (2008). On unmanned aircraft systems issues, challenges and operational restrictions preventing integration into the National Airspace System. *Progress in Aerospace Sciences*, *44*(7-8), 503–519. doi:10.1016/j.paerosci.2008.08.001.

Dalamagkidis, K., Valavanis, K. P., & Piegl, L. A. (2012). On integrating unmanned aircraft systems into the national airspace system: Issues, challenges, operational restrictions, certification, and recommendations. In S. G. Tzafestas (Ed.), *Intelligent Systems, Control and Automation: Science and Engineering* (2nd ed., pp. 1–301). London: Springer. doi:10.1007/978-94-007-2479-2.

Damilano, L., Guglieri, G., Quagliotti, F., & Sale, I. (2011). FMS for unmanned aerial systems: HMI issues and new interface solutions. *Journal of Intelligent & Robotic Systems*, *65*(1-4), 27–42. doi:10.1007/s10846-011-9567-3.

Department of Transportation (2008). *Pilot's handbook of aeronautical knowledge*. (FAA-H-8083-25A). Washington, DC: Federal Aviation Administration. Retrieved from http://www.faa.gov/library/manuals/aviation/pilot_handbook/.

Driscoll, M. P. (2005). *Psychology of Learning for Instruction* (3rd ed.). Boston, MA: Pearson.

FAA (2013). NextGen Implementation Plan. http://www.faa.gov/nextgen/library/media/nextgen_implementation_plan_2014.pdf (Retrieved).

Gay, L.R., Mills, E. G., & Airasian, P. (2012). *Educational research: Competencies for analysis and application.* Upper Saddle River, NJ: Pearson Education Inc.

Goldberg, B. (2010). Unmanned aerial systems: The role of the operator and human factors implications. *Proceedings of the 2010 Simulation Multiconference: SpringSim'10:* (pp. 1–4). Orlando, FL.

Grossman, T., Fitzmaurice, G., & Attar, R. (2009). A Survey of software learnability: metrics, methodologies and guidelines. In *CHI '09: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 649–658). Boston, MA: ACM digital Library.

Gundlach, J. (2012). *Designing unmanned aircraft systems: A comprehensive approach*. Reston, Virginia: American Institute of Aeronautics and Astronautics.

Haritos, T., & Robbins, J. M. (2012). The use of high fidelity simulators to train pilot and sensor operator skills for unmanned aerial systems. *New Learning Technologies*. (pp. 1–6). Orlando, FL: Society for Applied Learning Technologies.

Hammer, J. M. (2010). Intelligent Interfaces. In J. A. Wise, V. D. Hopkin, & D. J. Garland (Eds.), *Handbook of Aviation Human Factors* (pp. 24–1 – 24–17). Boca Raton. Florida: CRC Press.

Holden, K., Vos, G., & Martin, L. (2013). *Evidence Report : Risk of Inadequate Human-Computer Interaction Human Research Program Space Human Factors and Habitability* (pp. 1–46). Houston, TX.

ISO/IEC 25000 (2014). *Software and system engineering: software product quality and requirements and evaluation (SQuaRE.).* Geneva, Switzerland: International Organization for Standards.

ISO/IEC 25010 (2011). *Systems and software engineering: systems and software product quality requirements and evaluation*. Geneva, Switzerland: International Organization for Standards.

ISO 9241-11. (1998). *Ergonomic requirements for office work with visual display terminals (VDTs), Part 11: Guidance on usability specification and measures.* Geneva, Switzerland: International Organization for Standards.

Jimenez, C., Faerevaag, C. L., & Jentsch, F. (2016). User Interface Design Recommendations for Small Unmanned Aircraft Systems (sUAS) User Interface Design Recommendations for Small Unmanned Aircraft. *International Journal of Aviation, Aeronautics, and Aerospace*, *3*(2), 1–13.

Karwowski, W. (2012). A review of human factors challenges of complex adaptive systems: Discovering and Understanding chaos in human performance. *Human Factors*, *54*(6), 983–995. doi:10.1177/0018720812467459.

Lazar, J., Feng, J. H., & Hochheiser, H. (2010). *Research methods in human-computer interaction* (pp. 1–426). United Kingdom: John Wiley & Sons Inc.

Macdonald, C. M., & Atwood, M. (2013). Changing perspectives on evaluation in HCI : past, present, and future. In *CHI 2013: changing Perspectives. Paris, France.* (pp. 1969–1978). doi:10.1145/2468356.2468714.

Maybury, M. T. (2012). Usable advanced visual interfaces in aviation. In *Advanced Visual Interfaces 2012* (Vol. 1, pp. 2–3). Capri Island, Italy. doi:10.1145/2254556.2254558.

McCarley, J.S. & Wickens C.D. (2005) *Human factors implications of UASs in the national airspace.* University of Illinois Institute of Aviation Technical Report (AHFD-05-5/FAA-05-01)Savoy, IL:Aviation Human Factors Division.

Nielsen, J. (2012). *Usability 101: Introduction to usability*. Retrieved from http://www.nngroup.com/articles/usability-101-introduction-to-usability/.

Nullmeyer, R., Herz, R., Montijo, G., & Leonik, R. (2007, November). *Birds of prey: Training solutions to human factors issues*. Paper presented at the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Orlando, FL.

Pavlas, D., Burke, C. S., Fiore, S. M., Salas, E., Jensen, R., Fu, D., Florida, C. (2009). Enhancing unmanned aerial system training: A taxonomy of knowledge, skills, attitudes, and methods. *Proceedings of the Human Factors and Ergonomics Society 53rd Annual Meeting* (pp. 1903–1907). San Antonio, TX.

Pederson, H.K., Cooke, N.J., Pringle, H.L. and Conner, O. (2006). UAS human factors: Operator perspectives, In P.H. Cook (Ed.) *Human factors of remotely operated vehicles,* pp. 21-33. New York, NY: Elsevier.

Peschel, J. M., & Murphy, R. R. (2013). On the human-machine interaction of unmanned aerial system mission specialists. *IEEE Transactions on Human-Machine Systems*, *43*(1), 53–62. doi:10.1109/TSMCC.2012.2220133.

Preece, J., Rogers, Y., & Sharp, H. (2015). *Interaction design. Beyond human-computer interaction* (4th ed.). Somerset, NJ: John Wiley & Sons.

Rafique, I., Weng, J., Wang, Y., Abbasi, M. Q., & Lew, P. (2012a). Software Learnability Evaluation: an overview of definitions and evaluation methodologies for GIS applications. In *The Seventh International Multi-Conference on Computing in the Global Information Technology* (pp. 212–217).

Rafique, I., Weng, J., Wang, Y., Abbasi, M. Q., Lew, P., & Wang, X. (2012b). Evaluating software learnability: a learnability attributes model. *2012 International Conference on Systems and Informatics, ICSAI 2012*, (Icsai), 2443–2447. doi:10.1109/ICSAI.2012.6223548.

Reynolds, C., Liu, D., & Doherty, S. (2011). Effect of operator experience and automation strategies on autonomous aerial vehicle task performance. *Human Factors and Ergonomics in Manufacturing & Service Industries*, *12*(15), 481–490. doi: 10.1002/hfm.20330

Rummel, B. (2017). Beyond Average: Weibull Analysis of Task Completion Times. *Journal of Usability Studies*, *12*(2), 56–72.

Sauro, J., & Lewis, J. R. (2009). Correlations among prototypical usability metrics: evidence for the construct of usability. In *Proceedings of the 27th International Conference on Human factors in Computing Systems - CHI 09* (p. 1609). https://doi.org/10.1145/1518701.1518947

Sauro, J., & Lewis, J. R. (2012). *Quantifying the user experience: practical statistics for user research*. Waltham, MA: Elsevier Inc.

Shamsuddin, N. A., Sulaiman, S., Syed-Mohamad, S. M., & Zamli, K. Z. (2011). Improving learnability and understandability of a web application using an action-based technique. *2011 5th Malaysian Conference in Software Engineering, MySEC 2011*, 245–250. doi:10.1109/MySEC.2011.6140678.

Simmons, Liu, D., & Vincenzi, D. A. (2008). Effect of pilot experience and automation strategy on unmanned aerial vehicle mission performance. In Fowler J. and Mason, S. (Ed.), *Proceedings of the 2008 Industrial Engineering Research Conference* (pp. 98–104).

Stulbeerg, A.N. (2007). Manning the unmanned revolution in the U.S. Air Force. *Foreign Research Policy Institute, 51*(2), 251-265.

Su, K.-W., & Liu, C.-L. (2012). A mobile nursing information system based on human-computer interaction design for improving quality of nursing. *Journal of Medical Systems*. doi:10.1007/s10916-010-9576-y.

Tvaryanas, A.P., Thompson, W.B., & Constable, S.H. (2005). *U.S. military unmanned aerial vehicle mishaps: Assessment of the role of human factors using HFACS* (Report No. 20050711064). San Antonio, TX: 311th Performance Enhancement Directorate, Performance Enhancement Research Division, Brooks City-Base.

Tvaryanas, A. P. (2006). Human systems integration in remotely piloted aircraft operations. *Aviation, Space, and Environmental Medicine*, *77*(12), 1278–1282.

Terwilliger, B. A., Ison, D. C., Vincenzi, D. A., & Liu, D. (2014). Advancement and application of unmanned aerial system human-machine-interface (HMI) technology. *Human Interface and the Management of Information in Applications and Services: Lecture Note Sin Computer Science*, *8522*, 273–283.

United States Government Accountability Office (2012). *Unmanned aircraft systems: Measuring progress and addressing potential privacy concerns would facilitate integration into the national airspace system* (GAO Report No. GAO-12-981). Retrieved from http://www.gao.gov/assets/650/648348.pdf.

United States Government Accountability Office (2015). *Unmanned aerial systems: Actions needed to improve DOD pilot training* (GAO Report No. GAO-15-461). Retrieved from https://www.gao.gov/products/GAO-15-461.pdf.

United States Government Publishing Office (2013a). *Electronic Code of Federal Regulations Title 14: Aeronautics and Space Part 91: General Operating and Flight Rules*. Washington, DC: Federal Aviation Administration. Retrieved from http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&sid=3efaad1b0a259d4e48f11 50a34d1aa 77&rgn=div5&view=text&node=14:2.0.1.3.10&idno=14.

United States Government Publishing Office (2013b). *Electronic Code of Federal Regulations Title 14: Aeronautics and Space Part 61: Certification: Pilots, Flight Instructors, and Ground Instructors*. Washington, DC: Federal Aviation Administration. Retrieved from http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&SID =66016d6a8e77cf73c0413727e 0405621&tpl=/e cfrbrowse/ Title14/14 cfr6 1ma in02.tpl.

United States Government Printing Office (2013c). *Electronic Code of Federal Regulations Title 14: Aeronautics and Space Part 141: Pilot Schools*. Washington, DC: Federal Aviation Administration. Retrieved from http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&SID =66016d6a8e77cf73c0413727e 0405621&tpl=/e cfrbrowse/ Title14/14 cfr6 1ma in02.tpl.

United States Government Publishing Office (2013d). *Electronic Code of Federal Regulations Title 14: Aeronautics and Space Part 142: Training Centers*. Washington, DC: Federal Aviation Administration. Retrieved from http://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&SID =66016d6a8e77cf73c0413727e 0405621&tpl=/e cfrbrowse/ Title14/14 cfr6 1ma in02.tpl.

United States Government Publishing Office (2017). *Electronic Code of Federal Regulations Title 14: Aeronautics and Space Part 107: Small Unmanned Aircraft Systems*. Washington, DC: Federal Aviation Administration. Retrieved from https://www.ecfr.gov/cgi-bin/text-idx?SID=6861307c3d2900d0ee3d2e376 cdf2219&mc=true&tpl=/ecfrbrowse/Title14/14cfr107_main_02.tpl.

Vincenzi, D. A., Terwilliger, B. A., & Ison, D. C. (2015). Unmanned Aerial System (UAS) Human-machine Interfaces: New Paradigms in Command and Control. In *6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences* (Vol. 3, pp. 920–927). Elsevier B.V. doi:10.1016/j.promfg.2015.07.139

Williams, K. W. (2005). Unmanned aircraft pilot medical and certification requirements, In: Krebs, W.K. (ed.), *Unmanned aerial vehicles human factors,* FAA-AAR-100, FY-05. Washington, D.C.

Williams, K. W. (2004). *A summary of unmanned aircraft accident/incident data: Human factors implications*. (Technical report DOT/FAA/AM-04/24). Washington, DC: Office of Aerospace Medicine, FAA.

Williams, K. W. (2012). *An investigation of sensory information, levels of automation, and piloting experience on unmanned aircraft pilot performance*. (Technical report DOT/FAA/AM-12/4). Washington, DC: Office of Aerospace Medicine, FAA.

Willits, P., Abbott, M., & Kailey, L. (Eds.). (2004). *Guided flight discovery: Private pilot*. Englewood, CO: Jeppesen Sanderson, Inc.

Yin, S., Wickens, C. D., Helander, M., & Laberge, J. C. (2014). Predictive displays for a process-control schematic interface. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 1–15. doi:10.1177/0018720814542104.

Winterton, J., Le Deist, F. D., & Stringfellow, E. (2005). *Typology of knowledge, skills and competences: Clarification of the concept and prototype* (pp. 1–103). Toulouse. Retrieved from http://www.cedefop.europa.eu/en/Files/3048_EN.

Zhang, J., & Walji, M. (2011). TURF: Toward a unified framework of EHR usability. *Journal of Biomedical Informatics*, *44*(6), 1056–1.