
By

Nicholas Colby Currie

Bachelor of Science
Aerospace Studies
Embry-Riddle Aeronautical University
May 2007

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

Master of Science in Aviation
in
Applied Aviation Safety

Melbourne, Florida
September 2016
We the undersigned committee hereby approve the attached thesis be accepted as fulfilling in part the requirements for the degree of Master of Science in Applied Aviation Safety


By

Nicholas Colby Currie

____________________________
John E. Deaton, Ph.D.
Major Advisor
Professor, College of Aeronautics

_____________________
Stephen K. Cusick, J.D.
Committee Member
Associate Professor, College of Aeronautics

_____________________
Sohair Wastawy, Ph.D.
Committee Member
Dean of Libraries
Abstract


Author: Nicholas Colby Currie

Major Advisor: Dr. John Deaton

Trust in automation is a growing field of research that serves a vital role in understanding how humans interact with modern technology. Though automation is certainly becoming more prevalent in many professions, it has become a mainstay in the modern helicopter cockpit. One particular piece of modern aviation technology that incorporates a significant amount of automation is Enhance Flight Vision Systems (EFVS). EFVS technology provides pilots with a wealth of information that enables them to “see” under low visibility conditions, thereby increasing their situational awareness. Though the benefits of EFVS technology are easily recognizable, it is still an automated system that is susceptible to developing a poor human-automation relationship in terms of trust. When trust in automation is not properly regulated, it can result in an operator developing overreliance in system capabilities or even potentially lead to system neglect. Given the high workload demands placed upon the modern helicopter pilot, it is necessary that every automated system is designed to inculcate trust. As a result, this study sought to determine the current state of the trust-based relationship that exists between pilots
of different experience levels and EFVS technology in order to recommend strategies to improve the pilot-automation relationship. The study presented helicopter pilots of three different experience levels with two different EFVS technologies and asked them to rate their trust in the system’s capabilities under various scenarios. The results of the study indicated that pilots trust EVS displays significantly more than SVS displays. Additionally, the results suggested that a pilot’s flight experience does not impact a pilot’s trust in EFVS displays. Furthermore, the results indicated that no significant relationship exists between display type and a pilot’s flight experience in terms of trust in EFVS automation. In the end, the data collected from this study helped to develop a deeper understanding of how trust in automation impacts the modern cockpit and how EFVS technology can be designed to improve the pilot-automation relationship.
# Table of Contents

Abstract .................................................................................................................... iii  

Chapter 1 - Introduction ............................................................................................. 1  
  Problem Statement .................................................................................................. 1  
  Purpose Statement ................................................................................................... 2  
  Operational Definitions ........................................................................................... 3  
  Research Questions and Hypotheses ...................................................................... 4  
  Significance of the Study ........................................................................................ 5  
  Assumptions and Limitations ................................................................................. 7  
    Assumptions ......................................................................................................... 7  
    Limitations ........................................................................................................... 9  

Chapter 2—Literature Review ................................................................................. 11  
  Introduction ........................................................................................................... 11  
  Section One ........................................................................................................... 12  
    Enhanced Flight Vision Systems ........................................................................ 12  
    Trust in Automation ............................................................................................ 16  
    Pilot Experience ................................................................................................ 20  
    Designing Automation to Enhance Trust ........................................................... 24  
    Conclusion ........................................................................................................... 29  

Chapter 3—Methodology ........................................................................................ 30  
  Introduction ........................................................................................................... 30  
  Research Design and Approach ............................................................................ 30  
  Research Setting and Sample ................................................................................ 33  
  Power Analysis ..................................................................................................... 35  
  Research Instrumentation and Materials ............................................................... 37  
    The Study Instrument ......................................................................................... 37  
  Variables ............................................................................................................... 41  
    Independent Variable ......................................................................................... 41  
    Dependent Variable ............................................................................................ 41  
  Participants’ Eligibility Requirement ................................................................... 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Participants’ Protection</td>
<td>43</td>
</tr>
<tr>
<td>Legal and Ethical Consideration</td>
<td>44</td>
</tr>
<tr>
<td>Conclusion</td>
<td>45</td>
</tr>
<tr>
<td>Chapter 4—Analysis</td>
<td>46</td>
</tr>
<tr>
<td>Introduction</td>
<td>46</td>
</tr>
<tr>
<td>Data Analysis</td>
<td>46</td>
</tr>
<tr>
<td>Descriptive Statistics</td>
<td>46</td>
</tr>
<tr>
<td>Analysis of Variance (ANOVA)</td>
<td>51</td>
</tr>
<tr>
<td>Conclusion</td>
<td>55</td>
</tr>
<tr>
<td>Chapter 5—Conclusion</td>
<td>57</td>
</tr>
<tr>
<td>Overview</td>
<td>57</td>
</tr>
<tr>
<td>Summary of Findings</td>
<td>59</td>
</tr>
<tr>
<td>Discussion</td>
<td>60</td>
</tr>
<tr>
<td>Practical Implications</td>
<td>63</td>
</tr>
<tr>
<td>Uncontrolled Variables</td>
<td>67</td>
</tr>
<tr>
<td>Recommendation for Future Research</td>
<td>68</td>
</tr>
<tr>
<td>Conclusion</td>
<td>69</td>
</tr>
<tr>
<td>References</td>
<td>71</td>
</tr>
</tbody>
</table>
List of Figures

1 Example of the Astronic Max-Vis EVS .................................................. 14
2 Example of Rockwell Collins’ HeliSure H-SVS ................................. 15
3 Main Effects Polygon Plot .................................................................... 53
List of Tables

1 Percentages for Demographic Variables.................................................. 48
2 Descriptive Statistics by Experience Level............................................. 49
3 Descriptive Statistics by Survey Question.............................................. 50
Acknowledgement

The author of this document would like to acknowledge the support provided by Florida Institute of Technology’s College of Aeronautics faculty. Additionally, the support provided by the John H. Evan’s library staff was instrumental in conducting the background research for this particular study. Furthermore, the Federal Aviation Administration’s (FAA) Research Center of Excellence at the college known as the Partnership to Enhance General Aviation Safety, Accessibility, and Sustainability (PEGASAS) played a pivotal role in developing a well-rounded understanding of how Enhance Flight Vision Systems (EFVS) influence helicopter operations. Moreover, it is important to recognize the financial support provided by the United States Army’s Advanced Civil Schooling program that provided the opportunity for this research to be conducted in the first place. Finally, the author would like to acknowledge the love and support of his family, without which this study would not have been possible.
Chapter 1 - Introduction

Problem Statement

Over the past two decades, aircraft cockpits have gradually been redesigned to include more automated systems as well as improved instrumentation layouts in order to facilitate the pilot’s ability to manage his relatively complex work environment. Modern cockpit displays are able to present a wealth of information to the pilot in order to facilitate the pilot’s ability to maintain a high level of Situational Awareness (SA) while performing multiple simultaneous tasks. Recently, a relatively new technology called Enhanced Flight Vision Systems (EFVS) has become more prevalent in civilian aviation. EFVS provides the pilot with the ability to “see” at night and during low visibility conditions. The technology is comprised of three different systems that each have unique capabilities: Enhanced Vision Systems (EVS), Synthetic Vision Systems (SVS), and Combined Vision Systems (CVS). Though EFVS technology provides the pilot with a wealth of information and improved SA, it is still subject to the same problems that plague other automated systems found in today’s modern aircraft. Namely, technology is limited by the user’s ability to trust the automation and use it for its intended purpose. This study examined the pilot-EFVS automation relationship in order to develop an understanding of how current system designs impact the pilot’s trust in system
capabilities. In an effort to cultivate a mature understanding of the current pilot-EFVS relationship, research focused on the impact that pilot experience and individual system capabilities have on the pilot’s trust in EFVS automation.

**Purpose Statement**

In the modern cockpit, trust is a core component to the relationship between the pilot and the automated cockpit systems. The demands placed upon the modern pilot necessitate him to rely upon some form of automated aid to accomplish the litany of tasks that must be completed during every flight. This is especially true during high workload conditions such as takeoff and landing. If the pilot fails to use the automation properly due to overreliance (i.e., misuse) or neglect (i.e., disuse), then the risk of a negative outcome will undoubtably increase. Extensive research has already been conducted on the general relationship between humans and automation; however, EFVS technology has been left relatively unchecked in terms of trust-based issues. The purpose of this study was to determine the current state of the trust-based relationship that exists between pilots of different experience levels and EFVS technology in order to potentially recommend strategies to improve the pilot-automation relationship. The study focused on the current helicopter pilot population and employed the quantitative research method in the form of a nonexperimental research design. The research solicited support from several professional helicopter pilot organizations with the highest density of pilots, including the United States Army, Helicopter Association International, Curt Lewis
& Associates, and Bristow flight academy. At the conclusion of the study, a list of recommendations was developed for future system enhancements to facilitate a more cohesive working relationship between the pilot and EFVS automation.

**Operational Definitions**

For the purpose of this particular study, it is important to discuss the meaning of two important measures: trust and experience. Establishing the operational meaning of these measures early on will facilitate further discussion on how data was collected and examined by the researcher.

**Trust**

Trust is an individual’s willingness to enter a vulnerable state where he expects that an agent has the intention to help accomplish his goal (Lee & See, 2004; Hoffman et al., 2013). In terms of this particular study, trust was the measure used to describe the dynamic relationship that exists between a pilot and EFVS automation. Trust in EFVS automation was measured on a seven-point Likert scale during data collection.

**Experience**

Experience is professional knowledge developed over a period of time that provides an understanding of the relationships that exist between various cues (Schriver et al., 2008). For the purpose of this research, experience was derived from each participant’s total helicopter flight time, and it was used to categorize each participant
into one of three groups established by military regulations: low (0-1000 hours), medium (1000.1-2000), and high (2000.1+) (United States Army, 2010).

**Research Questions and Hypotheses**

**Research Questions (RQ)**

RQ1: To what extent does display type (i.e., EVS/SVS) impact an operator’s trust in automation?

RQ2: Does pilot experience have an effect on the human-automation relationship between the pilot and an EFVS display?

RQ3: What relationship exists, if any, between display type (i.e., EVS/SVS) and pilot experience level in terms of trust in EFVS automation?

**Hypotheses**

**Null Hypothesis 1**

$H_{01}$: There will be no significant differences in automation trust between the different display types used in this study.

**Alternative Hypothesis 1**

$H_{A1}$: There will be significant differences in automation trust between the different display types used in this study.
Null Hypothesis 2

$H_{02}$: There will be no significant differences in automation trust as a function of a pilot’s experience level.

Alternative Hypothesis 2

$H_{A2}$: There will be significant differences in automation trust as a function of a pilot’s experience level.

Null Hypothesis 3

$H_{03}$: There will be no significant interaction of pilot experience level and display type on automation trust.

Alternative Hypothesis 3

$H_{A3}$: There will be a significant interaction of pilot experience level and display type on automation trust.

Significance of the Study

The research performed in this study serves an important purpose in understanding the role of trust in the human-automation relationship. Though trust in automation is certainly not a new field of research, the interaction between EFVS technology and pilot trust appeared to be in need of further examination. The data collected from this study helps further the understanding of how trust in automation impacts the modern cockpit and how EFVS technology can be designed to improve the pilot-automation relationship.
As noted by the researchers Parasuraman and Riley, a core component to improving the human-automation relationship is that a current state of trust must be established (Parasuraman & Riley, 1997). As a result, it became necessary to study the current impact that trust in EFVS automation had on the modern cockpit. By studying the relationship between EFVS technology and pilot trust, researchers were able to determine if there was a significant gap in performance caused by a poor human-automation relationship. If a gap did exist, then it could be a contributing factor to many helicopter aviation accidents. Furthermore, as EFVS technology continues to evolve and FAA regulations are rewritten to allow for a reduction in helicopter approach minimums, a potential gap in performance would likely become more significant as EFVS automation slowly becomes more prevalent in helicopter cockpits.

In addition to further understanding how trust in EFVS automation impacts the modern cockpit, the research established a baseline in order to facilitate potential future redesigns of EFVS technology. With an established baseline, Parasuraman and Riley have stated that a road map can be developed that will enable system developers to improve both implicit and explicit trust through system redesigns (Parasuraman & Riley, 1997). Furthermore, research has shown that the baseline can be utilized to improve trust in automation by addressing several considerations, including training, system design, policy, and procedures (Parasuraman & Riley, 1997). As a result, the baseline established by this study provides an ideal starting point for potential redesigns of future EFVS technology to increase a pilot’s trust in
EFVS automation. As in any other scholastic research, the study did not seek to assign blame to a particular design or component. Rather, the study attempted to identify trust pitfalls that could be addressed for future system designs.

Finally, aviation safety is constantly searching for effective ways to mitigate the risk associated with flying. Mitigating risk is a rather daunting challenge for the helicopter community given the typical low altitude, obstacle rich flight environment that many helicopters operate in on a daily basis. However, it is believed that by researching the relationship between a pilot and EFVS automation, it is plausible that a positive impact will be made on helicopter safety. To be more specific, by developing a more effective and reliable EFVS system, it is believed that the helicopter accident rate associated with a breakdown in the human-automation relationship (e.g., loss of SA) can be reduced in some capacity.

Assumptions and Limitations

Assumptions

In order to commence any scientific research, some assumptions have to be made in regards to the sample audience as well as the method chosen to conduct statistical analysis. This particular study was no different, and the following section will detail the assumptions made by researcher when designing this study.

The most significant assumption for the research involved the measurement of the study’s dependent variable: a pilot’s trust in EFVS automation. Considering
data collection was conducted via a Likert scale survey, the assumption was made that participants were providing an accurate representation of their trust in EFVS automation. Though it was conceivable that participants could simply select an option on the Likert scale that was less cognitively demanding (e.g., the neutral option), the assumption was made that participants were genuinely interested in providing the study with good data. It should be noted that every scientific study that involves inferential statistics has to make a similar assumption any time it utilizes sample statistics to draw conclusions about a corresponding population parameter. This naturally occurring inconsistency, otherwise known as sampling error, is manageable through the use of unbiased random sampling.

The final assumptions made during this study were derived from the statistical analysis method that were utilized during data analysis: a 2 x 3 mixed factorial analysis of variance. This particular type of statistical analysis required the study to make the following three assumptions: (a) each case represents a random sample from the populations with test variable scores that are independent of each other, (b) the dependent variable is normally distributed for each combination of levels of the within-subjects factors, and (c) the variances of the dependent variable are the same for all populations. Homogeneity of variance and the normality assumption were tested during data analysis through the use of Statistical Package for the Social Science (SPSS) Statistics version 23 for Macintosh.
Limitations

Research in the field of trust in automation is fairly well-established and backed by several notable studies. However, the topic of trust in EFVS automation appeared to be a relatively novel field of study. As a result, this study was not able to rely solely on previous trust in automation research to guide its methodology. This minor limitation was compensated for by incorporating scientific research from several different fields, including trust in automation, pilot cognitive workload, and the effects of experience on primary job performance.

The primary limitations of this study can be found in the data collection as well as the participant recruitment tool that were utilized. All data collection was done online via a Likert scale survey. The online survey was published on the website SurveyGizmo ®. Additionally, due to limitations associated with acquiring reliable population statistics (e.g., total number of flight hours for every registered helicopter pilot) requisite for probability-based sampling techniques (e.g., stratified sampling), the study had to rely on convenience sampling to collect data remotely from several professional helicopter organizations. Seeing that some data collection and participant recruitment was done remotely, it was difficult to verify the authenticity of a participant’s experience flying helicopters. Demographic questions regarding flight experience were included in the survey, but a certain degree of control was lost due to the remote nature of data collection for this study. In order to compensate for this limitation, a question was included in the survey that requires some helicopter experience to answer correctly. The inclusion of this question enhanced the study’s
ability to isolate erroneous survey data. An additional limitation that is typical of online surveys is repeated entries by a single user. SurveyGizmo® had several tools that were utilized to mitigate this limitation including one that prevents multiple entries from one specific computer; however, it was still feasible that a tech-savvy participant could find a way to take the survey multiple times. Though both limitations could not be negated completely, the benefits associated with attaining a truly random sample through convenience sampling outweighed the potential side effects.

The final limitation of this particular study was financial and resource support. Every effort was made to ensure the study maintained an appropriate level of validity and statistical power. The sample size used by this study was supported by a power analysis calculated by G*Power 3.1.9.2, but it was close to the bare minimum required to ensure the integrity of the study (Erdfelder et al, 1992). A lack of sufficient funding and resources was the primary reason for this limitation. A larger sample size and outreach effort could have been accomplished with additional support.
Chapter 2—Literature Review

Introduction

Advancements in computer-based technology and aircraft capabilities over the past two decades seems to have continually increased the sheer volume of information available to a pilot via cockpit instrumentation. As a result of this rapid expansion in technology and the limitations of a pilot’s cognitive capabilities in terms of workload, the modern aviation cockpit has become increasingly more automated. Automation can serve exceptionally well in a litany of different roles including accomplishing tasks that humans are unable to perform (e.g., complex mathematical equations), augmenting human performance (e.g., digital transmission of required information to reduce working memory load), and compensating for human performance limitations (e.g., night vision systems) (Wickens et al., 2013). However, every automated system is susceptible to failure, which can have a significant impact on a pilot’s ability to trust the system’s capabilities (Wickens et al., 2013). It is therefore important to review the dynamic relationship that is formed between pilots and their associated cockpit automation. The following chapter will review the purpose of EFVS, the complexities of trust in automation, the role that pilot experience plays in the human-automation relationship, and how automated systems can be designed to facilitate mutual trust as well as enhance shared performance.
Section One

Enhanced Flight Vision Systems

The helicopter accident rate makes up a considerable portion (0.36 accidents per 100,000 hours) of General Aviation (GA) accidents every year (Helicopter Association International, 2015). Though there may be no one particular reason for this trend, some answers can be found when one looks at the typical flight profile of a helicopter when compared to its fixed-wing counterpart. Helicopter operators consist of organizations like the military, Helicopter Emergency Medical Service (HEMS), offshore helicopter transportation to drilling platforms, police departments, and corporate helicopter agencies. The flight environment for almost every one of the helicopter organizations rests solely in lower altitude, obstacle rich environments that typically terminate to a heliport instead of a much larger airport. Heliports are designed to provide ease of access for the company, but they also come with several challenges, such as proximity of obstacles and infrastructure support in terms of instrument approaches. This environment increases the need to provide the pilot more situational awareness during reduced visibility (e.g., nighttime or fog) conditions to improve safety for the vast majority of helicopter operators.

In the last ten years, professional aviation agencies have identified EFVS technology as a possible solution to reducing the accidents associated with the loss of situation awareness. Each of the EFVS technologies provide a graphical representation of the outside environment based on different types of radiation or
digital mapping capabilities. EFVS technologies can be classified into one of the three following categories: Enhanced Vision Systems (EVS), Synthetic Vision Systems (SVS), or Combined Vision Systems (CVS). For the purpose of this study, primary focus has been placed on EVS and SVS technology. CVS technology will not be discussed considering it essentially combines both EVS and SVS onto one display enabling the pilot to receive the benefits of both technologies and negate some of the potential disadvantages.

EVS technology works by monitoring the external environment using sensors (e.g., Forward-Looking Infrared (FLIR) and millimeter wave radar) that are able to detect an array of radiation and transmit this information to the pilot. This technology is able to display the environment more clearly in poor flying conditions (e.g., nighttime and fog), see Figure 1 for a graphical depiction. Once the information from the receiver is processed, the pilot is able to view the video on a Multi-Purpose Display (MPD)/Primary Flight Display (PFD) panel mounted display, a Heads Up Display (HUD) typically composed of a semi-transparent material placed between the pilot and the outside world, or a Helmet/Head Mounted Display (HMD) where the information is displayed on a device attached directly over the pilot’s eyes. Though EVS technology is certainly a beneficial piece of technology in the modern cockpit, it does have some notable limitations. The limitations of EVS technology include a reduction in the pilot’s situational awareness due to a limited Field of View (FOV) caused by the mechanical limitations of the visual sensor as well as a phenomenon known as thermal cross-over, a condition that occurs twice a day where
an object’s temperature is similar to the background temperature (United States Army, 2005).

Figure 1. Example of the Astronic Max-Vis EVS © (Astronics, 2016)

SVS technology uses a database of known terrain, obstacles, airport, and airway data to graphically represent the outside world to the pilot on a display device. The pilot can be presented this information with any of the same display systems described in the EVS section above; however, the image that is displayed is a graphical representation of digital database and not the real world, see Figure 2 for a graphical representation. The system does not sense the actual environment outside. Instead, the system relies on GPS coordinates, radar/pressure altitude, and the provided database to show the pilot what the surrounding area should look like. The benefit of the SVS over an EVS is that it is not limited by environmental conditions such as thermal cross-over or fog density. Additionally, some systems allow a pilot to select obstacle clearance notifications to aid in terrain avoidance. On the other hand, SVS can display a only graphical representation of an object if the data was
entered into its on-board database. If the database does not include obstacle or terrain data for a certain location, it will not be presented to the pilot.

![Figure 2. Example of Rockwell Collins’ HeliSure H-SVS © in a simulated environment (Rockwell Collins, 2016)](image)

As previously evidenced, each of the EFVS technologies significantly improves the pilot’s SA and visibility in reduced visibility environments. Due to the aforementioned capabilities, the FAA recently authorized a reduction in approach minima for having an EFVS, more specifically an EVS/CVS, installed on the fixed-wing aircraft. This operational credit is covered in 14 CFR paragraph 91.175 as well as FAA Advisory Circular 90-106 (FAA, 2010; FAA, 2015; FAA, 2015). In essence, fixed-wing aircraft are allowed to descend below Decision Height (DH) on an instrument approach to 100 feet above the runway before calling a missed approach as long as they are able to identify the runway environment (e.g., threshold markings or runway lighting) using the EVS. Once at 100 feet, the pilot must be able to identify the runway environment unaided before he can descend below 100 feet. As outlined
in FAA Advisory Circular 20-167, the FAA does not afford any operational credit for aircraft operating with only an SVS (FAA, 2010). Though the FAA has not approved the same reduction in DH for rotary-wing aircraft, interest was sparked by the Partnership to Enhance General Aviation Safety, Accessibility, and Sustainability (PEGASAS) which sought to determine how EFVS technology could safely reduce approach minimums for helicopter pilots (PEGASAS, 2015).

**Trust in Automation**

EFVS automation has the potential to significantly improve the pilot’s ability to make timely decisions in less than ideal situations. However, as previously mentioned in the purpose statement, research has proven that automation is limited by the operator’s ability to trust the technology and use it appropriately (Gells-Blair, 2013; Dzindolet et al., 2003; Hoffman et al., 2013; Lee & See, 2004; Parasuraman & Riley, 1997; Rice, 2009). An example can be found in many aviation accidents, such as the American Airlines crash in December of 1995 (Federal Aviation Administration [FAA], 1996). In this particular accident, the crew failed to select the appropriate navigational aid when flying toward Cali, Columbia. The crew entered the appropriate code for the navigational beacon and selected the first waypoint that appeared in the Flight Management System (FMS), a habit defined by overconfidence in the capabilities of the automation. What the crew did not know is that the waypoint was in a completely different direction. Shortly after the aircraft executed the appropriate maneuvers to get on course it crashed into a mountain and
ultimately took the lives of all the passengers on board. During the last half a decade of technological innovation, system designers have believed that if cockpit automation is continually increased, then human error could be eliminated and the aviation accident rate could be significantly reduced (Dzindolet et al., 2003; Thackray & Tou, 1989; Wickens et al., 2013). Though some research does support this mentality, the accident previously discussed and many others just like it highlight the fact that trust in automation must be balanced by proper implementation in order to prevent possible accidents.

In many ways, trust between automation and a human operator is similar to many other social relationships that can be found in our everyday life (Geels-Blair, 2013). When we trust another person to accomplish a task, we are confident in that person’s capabilities and fairly certain that person will complete his or her assigned duty. In many ways this is similar to our default trust, or unconditional trust that technology will accomplish its assigned task (Hoffman et al., 2013). For example, we trust that our automobile’s tire pressure monitoring system will be capable enough to accurately determine the tire pressure in each tire so the car will operate efficiently. Though this trust-based contract is similar to a human social relationship on the surface, there are several other aspects that are unique to the human-automation relationship that have a direct impact on trust: false-alarm rate, robustness, and validity (Hoffman et al., 2013).

Each of the factors previously mentioned can be a reliable means to determine the level of implicit trust in an automated system. Furthermore, if they are considered
during the design of a new system, they can have a significant impact on the explicit trust between the operator and the end product (Merritt et al., 2013; Parasuraman & Riley, 1997). Research supports this claim by showing that trust can be a reliable tool to determine the potential performance of a human-automation interface (Parasuraman et al., 2008; Parasuraman & Riley 1997). Assuming the design of a system effectively addresses each of the factors, operators will be more likely to move from a simplistic default trust to an expertise level of trust by balancing the knowledge of system limitations with proper utilization in order to achieve good results (Hoffman et al., 2013). However, if the false-alarm rate is too high or if the system fails to accomplish its assigned task, the human-automation relationship can quickly breakdown (Hoffman et al., 2013). Unlike the social relationship between two humans, an automated system cannot confess a mistake. As a result, the operator is left assuming some form of personal failure, and the operator will likely lose trust in the system’s capabilities to prevent future failures from occurring (Hoffman et al., 2013).

Another important consideration in the dynamic human-automation relationship is the operator’s growing reliance on automated systems to help accomplish the primary task of flying. The assumption is that if automation can be designed to accomplish a certain task with an acceptable level of trust, the operator is free to complete other tasks as needed (Rice, 2009; Wickens et al., 2013). To further understand this concept, it is important to review the Multiple Resource Theory (MRT) as explained by Dr. Christopher Wickens. Wickens’ Multiple
Resource Theory explains that workload should be distributed across different stages of information processing, codes of processing, and modality dimensions by regulating the usage of cognitive stores (Wickens, 2008). In the example of a pilot flying the aircraft while viewing a cockpit display, the pilot faces a difficult task of managing two concurrent tasks that utilize one common cognitive store across every dimension: visual-spatial. The stages of processing are dominated by the pilot determining the current state of the aircraft using spatial resources and responding with some form of manual spatial input. The codes of processing require the pilot to use spatial resources to interpret the environment around the aircraft as well as any SA data on the cockpit display. Finally, the modality dimension is governed by visual perception requirements to acquire the visual data throughout the environment and on the cockpit display.

As previously evidenced, visual-spatial resources are at a premium in the modern aircraft cockpit, and a pilot must trust that every automated system is accomplishing their assigned task in order to operate efficiently. Though system designers cannot directly increase a pilot’s pool of available cognitive resources, they can design automated systems that facilitate an operator’s ability to work effectively under high workload conditions. If an automated system is designed to be robust, reliable, and to minimize its false-alarm rate, system designers can effectively improve the baseline human-automation trust relationship. Ultimately, once pilots are able to establish a balanced level of trust in their automated systems, it is plausible
that a positive impact can be made on the safety of general aviation helicopter operations.

**Pilot Experience**

In an overly simplified manner, one could readily equate more experience with better performance, better decision making skills, and better knowledge of all assigned tasks. To a certain extent, the last statement is absolutely the true. Scholars and researchers have noted that experts are able to automate certain tasks that require novice operators to devote more mental resources (Damos, 1978; Wickens et al, 2013). Furthermore, research has shown that experienced operators are able to more effectively scan their environment for required information and remain focused on a critical task when faced with lower priority interruptions (Ebbatson et al., 2010; Johnson & Catano, 2013; Koh et al., 2011; Shinar, 2008; Taylor et al., 2007; Tolton et al., 2012; Wickens et al., 2013). Indeed, as experience increases, so will an operator’s knowledge of system operations. However, the relationship between experience and trust in system automation is slightly more complex. An operator’s trust in an automated system can be impacted by several different factors including the onset of initial automation failure, an operator’s previous experience with automation in general, and an operator’s experience with a particular system’s performance over time. In order to understand how pilot experience could impact trust in EFVS technology, it is necessary to review each of the aforementioned topics in more detail.
It seems logical that a perfectly reliable automated system would be the most ideal solution for establishing trust in the human-automation relationship. However, with the exception of some very simplistic automated systems, a perfectly reliable automated system is rarely achievable with modern technology (Wickens et al., 2013). Failure of an automated system is virtually inevitable due to system complexity and the possibility of software bugs (Wickens et al., 2013). Knowing and anticipating system failures is part of the battle in establishing trust in the human-automation relationship. When an operator perceives that an automated system performs without error, he is subject to an increased rate of commission errors (Bahner et al., 2008; Manzey et al., 2012). Research has shown that an individual operator’s trust in automation is directly related to when the system failures occur and how quickly the failures manifest themselves (Manzey et al., 2012). For example, if an operator experiences an automation failure in real-world conditions early on in the human-automation relationship, the operator’s trust in the system’s capabilities will decline significantly and likely never approach the same level of trust as an operator who experienced a failure later in the relationship (Manzey et al., 2012). However, if the operator is exposed to the same failure during initial training in a controlled environment, the operator develops a more complete understanding of system capabilities and a well-calibrated level of trust (Bahner et al., 2008; Parasuraman et al., 1996).

The relationship between trust, reliability, and dependence just discussed is often referred to as the calibration curve (Wickens et al., 2013). When an operator’s
subjective bias towards an automated system’s reliability matches an objective quantitative measurement of system reliability, it can be said that the team has achieved a perfectly calibrated level of trust (Carstens, 2016; Wickens et al., 2013). Ideally, every human-automation relationship should achieve a state of calibrated trust, but external influences such as operator experience can cause the two measures to become desynchronized. Research has shown that a significant amount of people have a negative automation bias due to previous experiences, stating that they are certain the automation will ultimately fail to work properly (Hoffman et al., 2013). If that is indeed the case, their subjective trust will likely be lower than the objective reliability of an automated system resulting in an under-trust relationship. Likewise, if an operator’s previous experience with automation in general has been predominately positive, the operator’s subjective trust could potentially exceed the objective reliability of an automated system leading to over-trust.

The final aspect in the human-automation relationship that is affected by experience is a qualitative shift in an operator’s interaction with a specific system over time. Research suggests that as operators gain more experience with an automated system, the monitoring rates of system performance will continue to decline and lead to a negative relationship between both parties (Bailey & Scerbo, 2007; Muir, 1987; Muir, 1994). Furthermore, an operator who has a long history of positive experiences with automation can suffer from increased reaction times to off-nominal events which could spell disaster in the helicopter aviation community (Manzey et al., 2012). During one particular study conducted at the Berlin Institute
of Technology, researchers developed an experimental scenario that required participants to monitor life support systems in a remote capsule while simultaneously recording carbon dioxide levels as well as the communications link with the capsule (Manzey et al., 2012). The experiment broke the participants into several smaller groups with various levels of automation support (e.g., low automation support to high automation support). The researchers measured the participant’s ability to accurately record the required information in a timely manner while also identifying potential faults in the system’s automation. Additionally, the participants had various levels of experience with the automated system prior to the first failure. At the conclusion of the study, the researchers found that participants who experienced an automation failure early on in the relationship were less bias and less likely to commit an error of commission by failing to adequately sample system parameters.

As mentioned previously in this section, several factors in terms of operator experience can directly impact the human-automation relationship. The Manzey et al. study from 2012 was referenced quite often throughout this section as it does an excellent job of highlighting this dynamic relationship. Though the study utilized engineering students as its only participants, the results of the study still provide broad reaching guidance on how the human-automation relationship is affected by operator experience. On one hand, if an operator experiences a system failure under less than optimal conditions (e.g., real world conditions), then the operator will likely develop an under-trust relationship, never attain a calibrated level of trust with that particular system, and possibly develop a negative bias towards automation in
general. On the other hand, if the operator is exposed to the same failure under controlled conditions early on in the human-automation relationship (e.g., initial system training), then it is entirely possible to establish a calibrated trust relationship and positively influence the operator’s understanding of system capabilities. In conclusion, the most ideal scenario is one where an operator has a wealth of experience with a *predictable* automated system so he can develop a deeper understanding of system capabilities and manage an appropriate level of automation reliance (Yuviler-Gavish & Gopher, 2011).

**Designing Automation to Enhance Trust**

Once the operator’s baseline trust has been established, how can system designers re-engineer an automated system to enhance the trust-based relationship? As mentioned in the section on trust in automation, one possible solution to enhance the human-automation trust relationship is to design the system to incorporate features that will facilitate the operator’s ability to establish an appropriate level of trust in the system from the onset. A team of researchers conducted an in-depth study with 132 airline pilots who had experience flying commercial aircraft with advanced automation systems (Tenney et al., 1998). The airline pilots were asked a series of questions that covered a litany of issues including the human-centered design philosophy, trustworthy automation, mental workload in the cockpit, levels of automation, and personal automation experience. It is important to note that only three airframes were represented during the research (i.e., Boeing 747-400, Douglas
MD-11, and Airbus A-320), potentially limiting the scope of data collection in terms of the commercial airline pilot population. Be that as it may, at the conclusion of the study, researchers found that pilots desired a system that is predictable, reliable, and simple to operate (Tenney et al., 1998). Furthermore, the pilots preferred a system that encouraged shared-performance between the pilot and cockpit automation over a fully autonomous system. Using this study as a baseline, it is logical that an automation redesign process should ideally address each of the aforementioned attributes to enhance the human-automation relationship.

In terms of predictability, researchers have found that increasing the observability of automated processes will aid in the operator’s ability to understand how the system operates, avoid mode errors, and help negate automation surprise (Salas & Maurino, 2010). An example can be found in a study conducted by Robert Sorkin that focused on the tendency of equipment operators to silence alarms that were perceived as a nuisance (Sorkin et al., 1988; Sorkin, 1989). During the study, equipment operators were observed casually acknowledging an alarm without validating the actual status of system operations. Operators dubbed these alarms as “old friends” and had become desensitized to their presence in the environment. The study found that this laissez faire attitude could be circumvented by providing the operator a likelihood display that would relay the system’s level of certainty to the operator. By providing the operator access to system certainty in an automated process, the operator can readily judge the authenticity of what is being presented to him in the cockpit.
In addition to providing the operator an ability to monitor system certainty, research has shown that the predictability of automation interruption can also impact an operator’s trust in system capabilities (Dorneich et al., 2012). Termed automation etiquette, the timeliness and importance of system interruption can be a major influence on the mental workload of a system operator (Dorneich et al., 2012; Wickens et al., 2013). Researchers developed an adaptive automated system that presented a user with high priority messages during high workload conditions, saving lower priority messages for a more suitable time (Dorneich et al., 2012). The study found that system operators performed significantly better at diagnosing system advisories under good etiquette conditions when compared to a system that interrupted operations despite the conditions. An important note regarding the high level of adaptive technology used in this study is that it requires biometric feedback from the system operator (e.g., an electroencephalogram (EEG) monitor). A lower level of automation etiquette can still be achieved by designing the system to provide cursory warning information to users and allowing them to complete a current task without further interruption (Parasuraman & Miller, 2004). By combining a likelihood display with lower level improvements in automation etiquette, system engineers will provide the operator with an ideal operational environment to make a more prudent and timely decision.

Regarding reliability, research shows that designers must be careful when establishing an automated system’s bias to avoid high false-alarm rates or misses (Rice, 2009; Parasuraman & Riley, 1997). Too many false-alarms have been directly
linked to a degradation in user compliance and disuse (Parasuraman & Riley, 1997; Rice, 2009). On the other hand, too many misses have been primarily associated with a reduction in system reliance and misuse (Parasuraman & Riley, 1997; Rice, 2009). Seeing that the possibility of building a system that has absolutely no false-alarms or misses may be farfetched, designers must make the critical decision on the false-alarm/miss bias to ensure their automated system achieves acceptable results (Rice, 2009). Furthermore, research has shown that system feedback should be provided to a user across several modalities to increase situation awareness (Wickens et al., 2013). By providing a user auditory feedback in conjunction with a visual display notification, the user is more adept to recognize changes in system status. Finally, research has shown that providing access to some raw data can potentially increase the operator’s understanding of system operations (Wickens et al, 2009; Wickens et al. 2013). In one particular study, researchers found that air traffic controllers were presented with conflict alerts that required no response almost 45 percent of the time (Wickens et al, 2009). On one hand, the study found some evidence that controllers located at air traffic control centers with the highest false alarm rate exhibited a lower response rate. On the other hand, the vast majority of the data indicated that a controller’s response to a true alert was not affected by false alarm rate. It was determined that considering these controllers had access to the system’s raw data in the form of a radar display, they were able to ascertain the reliability of system reports without degrading their trust in system operations.
The final aspect of developing a desirable automated system involves operator training. In order to ensure a system is easy to operate, the operator must be trained and calibrated appropriately to understand the capabilities of an automated system to avoid misuse/disuse (Merritt et al., 2013; Parasuraman et al., 1996; Parasuraman & Riley, 1997). One way to conduct effective training is to pre-expose the user to a failure in the automated system. Several studies have highlighted the positive impact that training system failures can have on the relationship between a user and an automated system (Bahner et al., 2008; Parasuraman et al., 1996). In one particular study, participants were asked to monitor an automated diagnostic system for the space station (Bahner et al., 2008). When a system failure was identified by automation, the fault management system presented the participant with a recommended course of action. Ideally, the participants were to verify the accuracy of the report prior to accepting the recommended course of action. The study found that those participants not pre-exposed to automation failures became more complacent to the system’s capabilities and failed to detect several inaccurate readings when compared to participants that were pre-exposed to automation failures. In the end, training system failures in a controlled environment can help establish a more ideal shared-performance relationship that effectively negates complacency and calibrates a system user to the automation’s real-world capabilities.
Conclusion

The research outlined in this chapter helps establish a solid foundation of knowledge for understanding the potential impacts of EFVS technology and the complicated relationship that is human-automation trust. Through this understanding it becomes possible for system designers, pilots, and regulatory authorities to make informed decisions on how to properly implement EFVS automation in the modern cockpit (Parasuraman & Victor, 1997). Furthermore, pilots will have the ability to develop a deeper understanding of how EFVS automation can be leveraged appropriately to make more timely decisions and avoid costly mistakes (Geels-Blair, 2013). As a result, it is plausible that by studying pilot perception of EFVS technology in an attempt to improve the human-automation relationship, a positive impact could be made on aviation operations as well as the current general aviation helicopter accident rate.
Chapter 3—Methodology

Introduction

A key component to the validity and power of any study is the methodology that is employed by the researchers. As such, the purpose of this chapter is to review the methodological process that was utilized throughout data collection and data analysis. The chapter will begin with a conversation on the research design and approach in order to provide an overview of the research’s methodology. Following this overview, each component will be explored in more detail to ensure the complete transparency of the research. In particular, a discussion of the research setting will detail the target population and the proposed sample. Additionally, a succinct discussion of the power analysis will be provided as well as a thorough description of the research instrumentation and materials. Furthermore, the study’s variables and data analysis procedures will be reviewed. Finally, the factors associated with a participant’s eligibility and protection will be explored as well as the legal considerations for the study.

Research Design and Approach

The purpose of this research was to determine the current level of trust in EFVS automation as it relates to display type (i.e., EVS and SVS) and pilot experience. In order to conduct this research on the human-automation relationship,
a quantitative study was utilized to examine each research question outlined in chapter one. The design of the study was centered around a nonexperimental repeated measures method using two quasi-independent variables. The first quasi-independent variable, pilot experience, was based on the participant’s total helicopter flight time, and it was used to assign the participants to one of three groups: low (0-1000 hours), medium (1000.1-2000 hours), and high (2000.1+ hours) (United States Army, 2010). The second quasi-independent variable was the different display types: EVS and SVS. Considering the research was comparing preexisting groups defined by pilot experience and display type, random assignment to ensure equivalent groups was not a viable option. Furthermore, due to limitations associated with acquiring population statistics (e.g., total number of low experienced helicopter pilots) requisite for probability-based sampling techniques (e.g., stratified sampling), the study had to rely on convenience sampling to collect data. Convenience sampling of participants was accomplished by soliciting participation from several different professional aviation organizations, including the United States Army, Helicopter Association International, Curt Lewis & Associates, and Bristow flight academy. Once a sufficient number of participants completed the survey, the participants were distributed equally to one of the three aforementioned pilot experience groups for further analysis. Finally, individual participant differences were negated through the use of a repeated measure design.

The statistical analysis method that was utilized during this study was a repeated measure 2 x 3 mixed factorial Analysis of Variance (ANOVA). The purpose
was to find any significance for the within subject, between subject, and within-between subject interactions that directly correlate to the proposed hypotheses. The primary focus for this study was the within-between interaction that directly correlates to both the purpose statement and research question three. As mentioned previously, two quasi-independent variables composed the overall design: pilot experience and display type. The dependent variable was the participant’s level of trust in EFVS automation measured on a 7-point Likert scale. The scale ranged from -3 (completely distrust) to +3 (completely trust) with a neutral option of zero (neither trust nor distrust). In addition to researching the impact that experience and display types had on trust in EFVS automation, the demographic information (i.e., age and sex) collected during the study was utilized to conduct a covariant analysis to further examine the dependent variable and develop a deeper understanding of trust in EFVS automation.

The study asked each participant a series of trust-based questions for both EVFS displays. The survey was conducted online. The online survey was access controlled to ensure the proper audience was participating in the research. Each participant was presented with a consent form that included an explanation of the intent of the research project, an overview of each EFVS capability, and the possible choices for survey answers to ensure all participants are operating on a common baseline. Refer to Appendix A for an example of the actual survey documents. Seeing that each participant did not necessarily have experience operating both EFVS technologies, the participant was still asked to provide feedback for every question
in order to develop an understanding of implicit/explicit trust for both systems. Following the consent form, each participant was presented with the capabilities of each system as well as an example question. Immediately following the example question, participants were assigned to one of EFVS displays (i.e., EVS or SVS) and asked a series of questions regarding their trust under several different conditions. Once the participant was complete with the first display, the participant was asked the same questions regarding the second display. In order to compensate for the possibility of order effects that is typical with a repeated measure study, the surveys were counterbalanced to distribute any outside effects over both treatments. Finally, each participant was asked to provide some demographic information (e.g., age, sex, ratings, and previous experience with EFVS technology) in order to develop a deeper understanding of the individual’s experience level and background. Once the data collection was complete, an analysis was conducted to determine the validity of each of the aforementioned hypotheses.

Research Setting and Sample

Population

According to the FAA’s 2014 Active Annual Civil Airmen Statistics, there are approximately 33,292 pilots in the United States that hold an active helicopter license (FAA, 2015). Though that information is limited in scope to only the United States, it provides a basic understanding of the potential size of the entire population of helicopter pilots on a global basis. It should be noted that the purpose of this study
was to generalize the findings from a sample of helicopter pilots to understand the
dynamics of the EFVS human-automation relationship for the population.
Considering it was not feasible to contact every possible helicopter pilot in the United
States or the world for that matter, the population of this study was defined by several
professional organizations with the highest density of helicopter pilots, including the
United States Army, Helicopter Association International, Curt Lewis & Associates,
and Bristow flight academy. All of these organizations combined together help
represent the entire helicopter community in terms of civil, government, and military
aviation. Furthermore, Helicopter Association International, Curt Lewis &
Associates, and Bristow flight academy have an international audience which helped
ensure the sample provided an accurate representation of helicopter pilots as a whole.

Sample

The sample used in this study was composed of participants found in one of
several professional organizations, including the United States Army, Helicopter
Association International, Curt Lewis & Associates, and Bristow flight academy.
The sample was achieved through the use of convenience sampling, with each
participant being categorized by pilot experience. The sampling pool was open to all
nationalities, with the only limitation being that each participant must be a qualified
helicopter pilot. A participant was not required to have previous experience or
knowledge regarding the topic of EFVS technology. The participants were asked a
series of demographic questions at the end of the survey which included several
experience related questions to develop an understanding of the individual’s experience level and background. If a participant did not have prior experience with EFVS technology, the participant’s answers still provided valuable insight into the implicit/explicit trust associated with EFVS automation.

The study utilized a repeated measures design, and it attempted to identify any significance for the within subject, between subject, and within-between subject interactions. However, the primary focus for this study was the within-between interaction. More specifically, the study was primarily interested in determining if there was a significant interaction between pilot experience level and display type on automation trust. As outlined in the power analysis section, considering it was necessary to obtain three different separate sample sizes to ensure each test maintained a sufficient level of validity, research best practices dictated that the primary sample size of the experiment be defined by the primary focus of the study (Prajapati et al., 2010). As such, the sample size associated with the within-between interaction was utilized ($n=42$), requiring the study to acquire 14 participants for each level (i.e., low experience, medium experience, and high experience).

**Power Analysis**

A priori power analysis was conducted to identify the sample size required to maintain a sufficient level of validity in the research. Considering the study included an analysis of within subject, between subject, and within-between subject interactions, it was possible to develop three different power analyses that correspond
to each interaction. Research was conducted on the topic of best practices for similar statistical analyses in order to ensure the appropriate sample size was utilized to maintain the integrity of the overall study. A peer reviewed article by Prajapati, Dunne, and Armstrong revealed that the study should select the power analysis that is associated with the primary focus of the research project (Prajapati et al., 2010). Furthermore, once data collection is complete, a post hoc power analysis can be conducted to show the resultant power associated with the remaining interactions (Prajapati et al., 2010). As a result, it was determined that the sample size for the overall study would be dictated by the within-between subject interaction power analysis, and a post hoc analysis will be conducted on the remaining two interactions (i.e., within subject and between subject).

G*Power 3.1.9.2 was utilized to conduct the all power analyses (Erdfelder et al., 1996). A priori-computation of sample size given $\alpha$, power, and effect size was applied to determine the within-between subject interaction sample size. The following values were used for each of the aforementioned parameters: level of significance ($\alpha$) .05, power ($\beta$) of .80, and effect size of .25 (medium effect). The power analysis resulted in a minimum sample size of 42 participants, with 14 participants in each stratum or level of experience.
Research Instrumentation and Materials

The Study Instrument

The primary instrument for collecting data was a survey that was conducted online via a website called SurveyGizmo ®. SurveyGizmo ® was selected due to its ability to provide ample data security, measures to prevent unauthorized access, and counterbalancing. All participants were solicited from several of the highest density professional helicopter pilot organizations, including the United States Army, Helicopter Association International, Curt Lewis & Associates, and Bristow flight academy. Each participant started the survey by reviewing and accepting an informed consent form that covered the purpose of the study as well as how the study would use the data. Following the informed consent form, participants were given an overview of both EFVS technologies (i.e., EVS and SVS) and then they were shown an example question to ensure they understood how the survey would be conducted. At the conclusion of the main study, participants were asked to provide several pieces of demographic information (e.g., age, sex, ratings, and previous experience with EFVS technology), and participants were asked a verification question to ensure they had some background helicopter flight experience. The order of the last four documents was the same for all participants. Refer to Appendix A for an example of the actual survey documents.

Immediately following the presentation of the study’s baseline documents, the participants were randomly assigned to one of two surveys. The surveys asked
the participants to rate their trust in EFVS automation under several different conditions. EVS and SVS displays were covered in both surveys. However, in order to compensate for the possibility of order effects typically found in repeated measure studies, the order of the surveys was counterbalanced. As a result, one participant’s survey began with the EVS display, while another participant’s survey began with the SVS display.

Each display’s survey had a visual depiction of the display at the top of the page. The purpose of this picture was to ensure the participant was cognizant of which display was being referenced during questioning. The participant was then asked to rank his or her trust in the system’s automation under several different conditions. For example, the participant was asked: To what extent do you trust the SVS/EVS display to provide a realistic representation of the outside world? Immediately below this question, the participant was presented with a seven-point Likert scale to rate his or her level of trust. The scale ranged from -3 (completely distrust) to +3 (completely trust) with a neutral option of zero (neither trust nor distrust). Though the validity of a Likert scale is often challenged, researchers have found that when a participant is familiar with the topic and concerned with the context of the study, the participant’s answers on a Likert test “may add to the content and construct validity of the scale” (Joshi et al., 2015). Furthermore, it has been determined by several researchers that the Likert scale is an appropriate method to measure a participant’s feelings as well as the unique characteristics of a group (Joshi et al., 2015; Murray, 2013).
Another commonly debated topic for Likert scales is the inclusion of a neutral option. It should be noted that there is research that states a neutral option could be potentially used as a “dumping ground for unsure or non-applicable responses” (Kulas et al., 2008). However, several studies show that providing a neutral option as well as more response varieties increases the probability of participants selecting an option that accurately describes their feelings (Guy & Norvell, 1977; Joshi et al., 2015; Kulas & Stachowski, 2009; Rempel et al., 1985). In fact, one study suggests that it is less cognitively demanding to simply agree/disagree with a statement than it is to select a neural option (Kulas & Stachowski, 2009). In the Guy and Norvell study (1977), researchers compared the responses of participants on a 5-point neutral option Likert scale to a 4-point non-neutral scale. At the conclusion of the study, the team found that participants who were familiar with Likert-type scales tended to become more sensitized on the 4-point scale (Guy & Norvell, 1977). As a result, the outcome of the 4-point scales tended to distort data more frequently when compared to the 5-point scale (Guy & Norvell, 1977). The team concluded that the distortion was possibly due to participants attempting to compensate for the missing neutral option by selecting middle-range values or by simply giving no response at all (Guy & Norvell, 1977).

In addition to the aforementioned example question, participants were asked to rank their trust in the ability of each display’s automation to support the following tasks: support situational awareness of the flight environment, accurately display the outside world under limited visibility conditions (e.g., fog, dust, sunrise/sunset),
accurately display the outside world under night time conditions, accurately display
the outside world under daytime unrestricted visibility conditions, provide reliable
obstacle clearance information at cruise altitude (above 200 feet), provide reliable
obstacle clearance information at low level altitude (below 200 feet), provide reliable
obstacle clearance information within urban terrain, provide reliable obstacle
clearance information over open terrain (e.g., farmland), provide reliable obstacle
clearance information on an instrument approach, provide reliable information
following a system failure (e.g., pilot is presented with a message stating “IR image
degraded” on EVS or “system confidence low” on SVS), and provide reliable
information after the system failure message is removed. Refer to Appendix B for
the actual survey questions. Following the completion of the survey, each participant
was thanked for their contributions and subsequently dismissed.

Each question in the survey was designed to help gauge the participant’s trust
in EFVS automation under both high (i.e., questions 3, 4, 7, and 8) and low (i.e.,
questions 5, 6, 9, and 10) workload conditions. Furthermore, the questions provided
an understanding of how trust was impacted by both system reliability and
observability (i.e., questions 11 and 12). The survey was exposed to several pilot
studies in order to validate survey questions and ensure participants were provided
with adequate answer selections for each question. Following the pilot studies,
several adjustments were made to the survey and it was released to the sample
audience for data collection.
Variables

Independent Variable

For the purpose of this study, two independent variables were utilized to provide insight into the EFVS human-automation relationship. The first variable, pilot experience, was composed of three different levels, which were based on the participant’s total helicopter flight experience as defined by hours flown: low (0-1000 hours), medium (1000.1-2000 hours), and high (2000.1+ hours). The criteria used to create the different levels of the pilot experience factor were derived from military guidelines that specify when helicopter pilots are awarded their basic, senior, and master flight wings (United States Army, 2010). The second independent variable was composed of two different levels defined by the EFVS display type being tested. More specifically, the two levels of the display type factor were EVS and SVS. Seeing that both of these independent variables are preexisting groups that do not support random assignment of participants to create equivalent groups, they are both considered quasi-independent variables.

Dependent Variable

The primary focus of this study was pilot trust in EFVS automation. As such, the dependent variable that was measured throughout the research was trust. The participant’s level of trust in EFVS automation was measured on a 7-point Likert scale. The scale ranged from -3 (completely distrust) to +3 (completely trust) with a neutral option of zero (neither trust nor distrust). The scale of measurement for the
dependent variable was classified as interval. Though there is an on-going debate regarding the proper scale of measurement for Likert scale data, several scholarly journals indicate that Likert responses generate empirically interval data when a composite score is generated (Carifio & Perla, 2008; Joshi et al., 2015). Considering the data collected from each participant’s survey was averaged to make trust score comparisons between pilot experience and display type, it was determined that the data could be logically measured on the interval scale.

**Data Analysis**

The statistical analysis procedure that was employed by this research was a repeated measure 2 x 3 mixed factorial Analysis of Variance (ANOVA). The study evaluated the significance for the within subject, between subject, and within-between subject interactions. Each of these sources of variance directly correlates to one of the proposed hypotheses. Though the researcher was interested in studying each of these factors, the primary focus for this study was the within-between interaction. To be more exact, the study was primarily interested in how pilot experience and EFVS display type interact to affect pilot trust in EFVS automation. The alpha level of significance was set at .05 (\( \alpha = .05 \)), a nominal significance level for this field of research. Following the completion of data collection, an ANOVA test was conducted to determine if any of the aforementioned relationships were significant. If the ANOVA test resulted in a significant outcome, a Tukey post hoc analysis was conducted to identify each significant relationship. In addition to
researching the impact that experience and display types have on trust in EFVS automation, the demographic information (i.e., age and sex) collected during the study was utilized to conduct a covariant analysis to further examine the dependent variable and develop a deeper understanding of trust in EFVS automation. All data analysis was computed using Statistical Package for the Social Science (SPSS) Statistics version 23 for Macintosh.

**Participants’ Eligibility Requirement**

The study did not impose any limitation on gender, nationality, or ethnicity as the intent of the research was to identify trends across the helicopter community. The only requirements were that the pilot must be 18 years or older, have access to the internet to take the survey, be qualified in a helicopter, and fall into one of three experience categories derived from military guidelines: low (0-1000 hours), medium (1000.1-2000 hours), and high (2000.1+ hours) (United States Army, 2010). Participants were recruited from several professional organizations that had the highest density of helicopter pilots and an international audience, including the United States Army, Helicopter Association International, Curt Lewis & Associates, and Bristow flight academy.

**Participants’ Protection**

As with any academic research, the protection of a participant’s identity and information was crucial to maintaining the integrity of the research. As a result, it
was necessary to ensure the site used for online data collection provided an appropriate level of protection to each participant. After extensive research, it was determined that all surveys would be conducted on a site called SurveyGizmo®, which provided an ample amount of protection to the participant while simultaneously meeting the requirements for maintaining the integrity of the research. All participant information was secured by a Secure Sockets Layer (SSL) connection through the utilization of a secure survey link. All of the data collected throughout the survey process is kept in a secure file located at the Florida Institute of Technology’s College of Aeronautics. Participants were not required to provide any confidential information. As a result, participants and their associated answers remained anonymous. Each participant was presented with an informed consent form that detailed this information prior to completing the survey. The participant had to agree to the form before the survey would commence.

**Legal and Ethical Consideration**

At no time did the study pose any form of mental, physical, or psychological harm to the participants. Furthermore, the study did not ask the participants to make any statements that could potentially expose them to legal actions. At no time did the study solicit the support of a person under the age of 18 years. Participants were recruited through several professional organizations via a general announcement, including the United States Army, Helicopter Association International, Curt Lewis
& Associates, and Bristow flight academy. Participation was solely voluntary. At no time was a participant forced to take the survey against their will.

Conclusion

The purpose of this chapter was to provide a succinct overview of the methodology associated with this body of research. Several important considerations were covered throughout this chapter including the research design, setting, and instrumentation. Additionally, an overview was provided on the topic of both data and power analysis, participant eligibility and protection, and the legal considerations for the study. It is important to note that every effort has been taken to ensure the information outlined in this chapter is supported by both the academic and research community. As a result of this effort, it is believed that the methodological foundation of this research is based on sound logic and reasoning.
Chapter 4—Analysis

Introduction

The following chapter provides a detailed overview of the results obtained from the trust in EFVS automation survey. The results are broken down into two main sections: descriptive statistics and analysis of variance (ANOVA). The descriptive statistics section presents the raw data in several different forms to ensure complete transparency and to facilitate discussion in chapter five. The ANOVA section is divided into two subsections: main effects analysis and covariant analysis. The main effects analysis focuses intently on the three relationships that directly correlate to the study’s research questions. The covariant analysis evaluates the impact that two demographic variables (i.e., age and sex) had on the between subject study.

Data Analysis

Descriptive Statistics

At the completion of the study, a total of 73 surveys were collected from several different professional helicopter organizations. The data obtained from the study were consolidated into a single file in order to perform data cleaning prior to conducting data analysis. While performing data cleaning, the researcher identified three surveys that incorrectly answered the helicopter pilot verification question
located at the end of the survey. As a result, the three surveys were excluded from the study.

Following data cleaning, the completed surveys were categorized by experience level based on the estimated helicopter flight experience provided by each participant. Once the surveys were distributed to an appropriate experience level, the experience level with the smallest total sample size (i.e., high experience $n = 16$) was utilized to establish the sample size for the remaining two experience levels. In order to obtain the appropriate sample size, the researcher employed a random number generator to select 16 surveys for each of the remaining two experience levels (i.e., low and medium experience). The random number generator was utilized to ensure each sample was truly random and to prevent any potential bias.

Once the final sample was identified, the data were uploaded to Statistical Package for the Social Science (SPSS) Statistics version 23 for Macintosh to conduct several forms of statistical analyses. In order to adequately present the descriptive statistics of the data, three separate tables were generated. The first table provides an overview of the demographic information provided by the surveyed pilots, see Table 1. Each demographic variable has a corresponding overall percentage value based on the entire sample ($n = 48$). Furthermore, a proportional value has been provided for each demographic variable based on individual experience levels ($n = 16$).
Table 1. Percentages for Demographic Variables of Surveyed Pilots (n=48)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Overall</th>
<th>Flight experience level</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>91.7%</td>
<td>93.7%</td>
<td>81.2%</td>
<td>100%</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>8.3%</td>
<td>6.3%</td>
<td>18.8%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-25</td>
<td>2.1%</td>
<td>6.3%</td>
<td>0%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>26-35</td>
<td>60.4%</td>
<td>81.2%</td>
<td>87.5%</td>
<td>12.5%</td>
<td></td>
</tr>
<tr>
<td>36-45</td>
<td>27.1%</td>
<td>12.5%</td>
<td>6.3%</td>
<td>62.5%</td>
<td></td>
</tr>
<tr>
<td>46-55</td>
<td>4.2%</td>
<td>0%</td>
<td>0%</td>
<td>12.5%</td>
<td></td>
</tr>
<tr>
<td>56-65</td>
<td>4.2%</td>
<td>0%</td>
<td>0%</td>
<td>12.5%</td>
<td></td>
</tr>
<tr>
<td>65+</td>
<td>2.1%</td>
<td>0%</td>
<td>6.3%</td>
<td>0%</td>
<td></td>
</tr>
<tr>
<td>Certificates/ratings</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>66.7%</td>
<td>62.5%</td>
<td>68.8%</td>
<td>68.8%</td>
<td></td>
</tr>
<tr>
<td>Instrument</td>
<td>81.3%</td>
<td>75.0%</td>
<td>81.3%</td>
<td>87.5%</td>
<td></td>
</tr>
<tr>
<td>Commercial</td>
<td>79.2%</td>
<td>68.8%</td>
<td>75%</td>
<td>93.8%</td>
<td></td>
</tr>
<tr>
<td>CFI</td>
<td>20.8%</td>
<td>0%</td>
<td>18.8%</td>
<td>43.8%</td>
<td></td>
</tr>
<tr>
<td>CFII</td>
<td>16.7%</td>
<td>0%</td>
<td>6.3%</td>
<td>43.8%</td>
<td></td>
</tr>
<tr>
<td>Military pilot</td>
<td>97.9%</td>
<td>100%</td>
<td>100%</td>
<td>93.8%</td>
<td></td>
</tr>
<tr>
<td>Military IP</td>
<td>29.2%</td>
<td>0%</td>
<td>18.8%</td>
<td>68.8%</td>
<td></td>
</tr>
<tr>
<td>Military IE</td>
<td>10.4%</td>
<td>0%</td>
<td>0%</td>
<td>31.3%</td>
<td></td>
</tr>
<tr>
<td>SVS experience(^a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>35.4%</td>
<td>50.0%</td>
<td>43.8%</td>
<td>12.5%</td>
<td></td>
</tr>
<tr>
<td>Hands on</td>
<td>22.9%</td>
<td>18.8%</td>
<td>18.8%</td>
<td>31.3%</td>
<td></td>
</tr>
<tr>
<td>Some</td>
<td>22.9%</td>
<td>12.5%</td>
<td>18.8%</td>
<td>37.5%</td>
<td></td>
</tr>
<tr>
<td>Extensive</td>
<td>18.8%</td>
<td>18.8%</td>
<td>18.8%</td>
<td>18.8%</td>
<td></td>
</tr>
<tr>
<td>EVS experience(^a)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>6.3%</td>
<td>12.5%</td>
<td>0%</td>
<td>6.3%</td>
<td></td>
</tr>
<tr>
<td>Hands on</td>
<td>12.5%</td>
<td>18.8%</td>
<td>12.5%</td>
<td>6.3%</td>
<td></td>
</tr>
<tr>
<td>Some</td>
<td>10.4%</td>
<td>12.5%</td>
<td>6.3%</td>
<td>12.5%</td>
<td></td>
</tr>
<tr>
<td>Extensive</td>
<td>70.8%</td>
<td>56.2%</td>
<td>81.3%</td>
<td>75.0%</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Display experience classified as None = no experience, Hands on = no flight experience, Some = ~100 hours or less, and Extensive = 100+ hours.

Note. CFI = Certified Flight Instructor; CFII = Certified Flight Instructor Instrument; IP = Instructor Pilot; IE = Instrument Examiner. Flight experience level is divided into three categories: low (0-1000.0 hours), medium (1000.1-2000.0 hours), and high (2000.1+ hours).
The next table provides a detailed summary of both the average flight hours and the average trust scores for the pilots surveyed in the study, see Table 2. The data are categorized by the three experience groups utilized throughout the study. The trust scores are broken down into three separate groups (i.e., SVS, EVS, and overall), and each group directly corresponds to one of the study’s research questions.

Table 2. Descriptive Statistics of Surveyed Pilots by Flight Experience Level (n=48)

<table>
<thead>
<tr>
<th>Variable</th>
<th>M</th>
<th>SD</th>
<th>X-min</th>
<th>X-max</th>
<th>Range</th>
<th>s²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours</td>
<td>576.44</td>
<td>406.74</td>
<td>18.00</td>
<td>1000.00</td>
<td>982.00</td>
<td>101.68</td>
</tr>
<tr>
<td>SVS trust</td>
<td>0.65</td>
<td>1.18</td>
<td>-1.25</td>
<td>2.25</td>
<td>3.50</td>
<td>0.30</td>
</tr>
<tr>
<td>EVS trust</td>
<td>1.96</td>
<td>0.54</td>
<td>0.58</td>
<td>2.67</td>
<td>2.09</td>
<td>0.13</td>
</tr>
<tr>
<td>Overall trust</td>
<td>1.30</td>
<td>0.78</td>
<td>-0.13</td>
<td>2.38</td>
<td>2.51</td>
<td>0.19</td>
</tr>
<tr>
<td>Medium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours</td>
<td>1423.38</td>
<td>226.68</td>
<td>1100.00</td>
<td>2000.00</td>
<td>900.00</td>
<td>56.67</td>
</tr>
<tr>
<td>SVS trust</td>
<td>0.15</td>
<td>1.54</td>
<td>-2.08</td>
<td>2.00</td>
<td>4.08</td>
<td>0.38</td>
</tr>
<tr>
<td>EVS trust</td>
<td>1.99</td>
<td>0.58</td>
<td>0.92</td>
<td>2.83</td>
<td>1.91</td>
<td>0.15</td>
</tr>
<tr>
<td>Overall trust</td>
<td>1.07</td>
<td>0.85</td>
<td>-0.13</td>
<td>2.42</td>
<td>2.55</td>
<td>0.21</td>
</tr>
<tr>
<td>High</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hours</td>
<td>3928.13</td>
<td>1560.12</td>
<td>2100.00</td>
<td>8500.00</td>
<td>6400.00</td>
<td>390.03</td>
</tr>
<tr>
<td>SVS trust</td>
<td>0.83</td>
<td>1.12</td>
<td>-2.17</td>
<td>2.25</td>
<td>4.42</td>
<td>0.28</td>
</tr>
<tr>
<td>EVS trust</td>
<td>2.02</td>
<td>0.70</td>
<td>0.67</td>
<td>2.83</td>
<td>2.16</td>
<td>0.17</td>
</tr>
<tr>
<td>Overall trust</td>
<td>1.42</td>
<td>0.74</td>
<td>-0.04</td>
<td>2.54</td>
<td>2.58</td>
<td>0.19</td>
</tr>
</tbody>
</table>

Note. \(X_{\text{min}}\) = Smallest Score; \(X_{\text{max}}\) = Largest Score; Hours = total helicopter flight time reported by participants. Flight experience level is divided into three categories: low (0-1000.0 hours), medium (1000.1-2000.0 hours), and high (2000.1+ hours). Trust is measured on a 7-point Likert scale ranging from -3 (completely distrust) to +3 (completely trust) with a neutral option of zero (neither trust nor distrust).
The final table provides an overview of the average trust score for each question on the trust in EFVS automation survey, see Table 3. Considering the range for each question was limited (i.e., completely distrust (-3) to completely trust (+3)), some regularly reported descriptive statistics have been omitted from the table (i.e., minimum value, maximum value, and range).

Table 3. Descriptive Statistics of Surveyed Pilots by Survey Question (n=48)

<table>
<thead>
<tr>
<th>Variable</th>
<th>(M)</th>
<th>(SD)</th>
<th>(s^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question 1</td>
<td>0.69</td>
<td>1.43</td>
<td>2.05</td>
</tr>
<tr>
<td>Question 2</td>
<td>1.13</td>
<td>1.55</td>
<td>2.41</td>
</tr>
<tr>
<td>Question 3</td>
<td>0.81</td>
<td>1.62</td>
<td>2.62</td>
</tr>
<tr>
<td>Question 4</td>
<td>0.92</td>
<td>1.56</td>
<td>2.42</td>
</tr>
<tr>
<td>Question 5</td>
<td>1.23</td>
<td>1.52</td>
<td>2.31</td>
</tr>
<tr>
<td>Question 6</td>
<td>1.06</td>
<td>1.44</td>
<td>2.06</td>
</tr>
<tr>
<td>Question 7</td>
<td>-0.04</td>
<td>1.60</td>
<td>2.55</td>
</tr>
<tr>
<td>Question 8</td>
<td>-0.54</td>
<td>1.50</td>
<td>2.25</td>
</tr>
<tr>
<td>Question 9</td>
<td>0.88</td>
<td>1.41</td>
<td>1.98</td>
</tr>
<tr>
<td>Question 10</td>
<td>1.13</td>
<td>1.36</td>
<td>1.86</td>
</tr>
<tr>
<td>Question 11</td>
<td>-0.88</td>
<td>1.67</td>
<td>2.79</td>
</tr>
<tr>
<td>Question 12</td>
<td>0.13</td>
<td>1.58</td>
<td>2.50</td>
</tr>
<tr>
<td>EVS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Question 1</td>
<td>2.50</td>
<td>0.58</td>
<td>0.34</td>
</tr>
<tr>
<td>Question 2</td>
<td>2.50</td>
<td>0.62</td>
<td>0.38</td>
</tr>
<tr>
<td>Question 3</td>
<td>1.63</td>
<td>1.00</td>
<td>1.01</td>
</tr>
<tr>
<td>Question 4</td>
<td>2.52</td>
<td>0.65</td>
<td>0.43</td>
</tr>
<tr>
<td>Question 5</td>
<td>2.54</td>
<td>0.65</td>
<td>0.42</td>
</tr>
<tr>
<td>Question 6</td>
<td>2.23</td>
<td>0.81</td>
<td>0.65</td>
</tr>
<tr>
<td>Question 7</td>
<td>1.63</td>
<td>1.20</td>
<td>1.43</td>
</tr>
<tr>
<td>Question 8</td>
<td>1.54</td>
<td>1.25</td>
<td>1.57</td>
</tr>
<tr>
<td>Question 9</td>
<td>2.25</td>
<td>0.79</td>
<td>0.62</td>
</tr>
<tr>
<td>Question 10</td>
<td>1.94</td>
<td>1.12</td>
<td>1.25</td>
</tr>
<tr>
<td>Question 11</td>
<td>0.88</td>
<td>1.30</td>
<td>1.69</td>
</tr>
<tr>
<td>Question 12</td>
<td>1.75</td>
<td>0.96</td>
<td>0.92</td>
</tr>
</tbody>
</table>

*Note. \(X_{\text{min}}\) = Smallest Score; \(X_{\text{max}}\) = Largest Score; Hours = total helicopter flight*
Analysis of Variance (ANOVA)

Overview

The primary purpose of this study was to determine the current level of trust in EFVS automation as it relates to display type (i.e., EVS and SVS) and pilot experience. To accomplish this goal, the study utilized a repeated measures 2 x 3 mixed factorial ANOVA to establish empirical evidence as to the true nature of trust in EFVS automation. In order to ensure the aforementioned statistical analysis conformed to the proper requirements, the following statistical assumptions were confirmed during testing: (a) each case represents a random sample from the populations with test variable scores that are independent of each other, (b) the dependent variable is normally distributed for each combination of levels of the within-subjects factors, and (c) the variances of the dependent variable are the same for all populations.

Though ensuring the sample was truly random has been addressed on several occasions throughout this document, the remaining two assumptions have yet to be discussed in terms of validation. Concerning the normality assumption, a Shapiro-Wilks test was performed and it did not return a significant value. As a result, it was determined that the dependent variable was normally distributed. Regarding homogeneity of variance, a Mauchly’s sphericity test is typically utilized to validate homogeneity for repeated measure designs; however, the design of this particular study (i.e., two levels for the within subject analysis) prevented the test from providing a significance level for the data. As a result, a Levene’s test for equality of
variance was utilized to validate homogeneity of variance. Based on a .05 significance level, the Levene’s test did not report significance, and it was determined that the variances of the dependent variable were the same.

The following section provides an overview of both the main effects and covariant analysis of the data. The main effects analysis primarily concentrates on investigating the relationships between the two independent variables (i.e., display type and pilot experience) and the dependent variable (i.e., pilot trust in EFVS automation) without controlling for any covariates. The covariant analysis, which is a between subject analysis, was used to control two covariates (i.e., age and sex) in order to determine if they influenced the dependent variable. The reason the analyses were conducted separately has to do with the fact that the within subject results can be altered when covariates are included in the statistical analysis. Logically, the within subject results should not be effected by the age or sex of the participants because those two variables remain consistent across the two measures of the dependent variable. Based on established guidance regarding repeated measure designs, a preferred solution is to conduct the covariate analysis separately when a study is primarily interested in the main effects analysis (Thomas et al., 2009). As a result, the researcher conducted each analysis separately and chose to present the results separately to ensure the data were not misleading.

**Main Effects Analysis**

The main effects analysis focused on investigating the three research questions outlined in chapter one. A polygon plot was developed to show the overall
interaction between the two independent variables as well as the dependent variable of trust, see Figure 3. Based on a cursory evaluation of the polygon plot, it is apparent that there is a noticeable difference in the average trust scores between the two display types, with EVS displays scoring higher than SVS displays. Furthermore, pilot experience appears to have a dynamic impact on a pilot’s trust in EFVS automation. Though the polygon plot provides a straightforward visual depiction of the data, additional information is required to answer the study’s research questions. In order to provide more detail, the statistical results of the main effects analysis for each research question are provided in the following section.

Figure 3. Polygon plot displaying the average trust score for two Enhanced Vision Flight Systems (EFVS) based on experience level. The two types of EFVS are Synthetic Vision Systems (SVS) and Enhanced Vision Systems (EVS). Flight experience level is divided into three categories: low (0-1000.0 hours), medium (1000.1-2000.0 hours), and high (2000.1+ hours).
Research Question 1

The results of the test were statistically significant, $F(1, 45) = 61.77, p < .001$, $\eta^2 = .34$. As a result, the null hypothesis was rejected and it was determined that there is a significant difference in automation trust between the different display types used in this study. Helicopter pilots trusted EVS displays ($M = 1.99, SD = 0.60$) more than SVS displays ($M = 0.54, SD = 1.30$). The strength of the relationship between display type and pilot trust, as determined by $\eta^2$, was strong with difference between display type accounting for 34% of the variance in pilot trust. A post-hoc power analysis determined that the power of this particular test was 0.92.

Research Question 2

The results of the test were not statistically significant, $F(2, 45) = 0.84, p = .84$, $\eta^2 = .01$. Therefore, the null hypothesis was confirmed and it was determined that there is no significant difference in automation trust based on a pilot’s experience level. Pilots who were categorized in the high experience group exhibited the most trust ($M = 1.42, SD = 0.74$) in EFVS automation when compared to both low ($M = 1.30, SD = 0.78$) and medium ($M = 1.07, SD = 0.85$) experienced aviators. However, none of the between factor relationships were significantly different and experience only accounted for 1% of variance in trust. A post-hoc power analysis determined that the power of this particular test was 0.39.
Research Question 3

The results of the test were not statistically significant, $F(2, 45) = 1.22$, $p = .31$, $\eta^2 = .01$. Consequently, the null hypothesis was confirmed and it was determined that there is no significant interaction between pilot experience level and display type on automation trust. The strength of the interaction, as determined by $\eta^2$, was weak with pilot experience level and display type accounting for only 1% of the variance in pilot trust. A post-hoc power analysis determined that the power of this particular test was 0.86.

Covariant Analysis

As outlined in the ANOVA overview, two demographic variables (i.e., age and sex) were utilized to conduct a covariant analysis of the data obtained by the trust in EFVS automation survey. The purpose of the covariant analysis was to control the effects of age as well as sex to determine if they impacted the relationship between pilot experience and a pilot’s trust in EFVS automation. The results from the test were not statistically significant, $F(1, 43) = 0.88$, $p = .42$, $\eta^2 = .02$. As a result, it was determined that there was not a significant difference in automation trust based on a pilot’s experience level after controlling for the effects of age and sex.

Conclusion

As discussed previously, the aim of this study was to investigate the current state of trust in EFVS automation in terms of pilot experience and display type. The results outlined in this chapter provide empirical evidence to suggest that display
type does impact a pilot’s trust in EFVS automation. Furthermore, the results from both the main effects analysis and the covariant analysis indicate that pilot experience does not impact a pilot’s trust in EFVS automation. Finally, the results reveal there appears to be no significant interaction between display type and pilot experience when it comes to trust in EFVS automation.
Chapter 5-Conclusion

Overview

The purpose of this study was to determine the current state of the trust-based relationship that exists between pilots of different experience levels and EFVS technology in order to potentially recommend strategies to improve the pilot-automation relationship. Though trust in automation is a fairly well-established field of research in many aspects, trust in EFVS automation appeared to be a novel subset that could benefit from further investigation. As a result, this study was conducted in order to add to the trust in automation global knowledge base and seek potential system improvements with an aim to enhance safety of flight.

The study collected a total of 73 surveys from a sample audience composed of helicopter pilots located throughout the world. The helicopter pilots were members of one of several professional organizations with the highest density of helicopter pilots, including the United States Army, Helicopter Association International, Curt Lewis & Associates, and Bristow flight academy. The participants were categorized by flight experience (i.e., low (0-1000 hours), medium (1000.1-2000 hours), and high (2000.1+ hours)) and asked a series of trust questions on two EFVS technologies. Each trust question asked the participants to score their trust in a particular system’s automation on a 7-point Likert scale ranging from -3 (completely distrust) to +3 (completely trust) with a neutral option of zero (neither trust nor distrust). Prior to
data analysis, the sample size of all three experience groups were balanced based on the smallest experience group size (i.e., high experience $n = 16$) using a random number generator. In the end, a total of 48 surveys were utilized as the primary sample for data analysis.

The study incorporated a nonexperimental research method with two quasi-independent variables (i.e., EFVS display type and pilot experience) to measure trust in EFVS automation. Data analysis was conducted through the use of a repeated measure 2 x 3 mixed factorial design with a primary focus on the within-between subject interaction. Additionally, a covariant analysis was conducted on the between subject interaction to determine if two covariates (i.e., age and sex) impacted the relationship between pilot experience and a pilot’s trust in EFVS automation. In the end, the data analysis helped establish empirical evidence to investigate the following research questions:

RQ1: To what extent does display type (i.e., EVS/SVS) impact an operator’s trust in automation?

RQ2: Does pilot experience have an effect on the human-automation relationship between the pilot and an EFVS display?

RQ3: What relationship exists, if any, between display type (i.e., EVS/SVS) and pilot experience level in terms of trust in EFVS automation?
Summary of Findings

The data analysis results were broken down into two different categories: main effects analysis and covariant analysis. The main effects analysis was primarily concerned with investigating the three research questions by evaluating the within subject, between subject, and within-between subject interactions. The covariant analysis was conducted in order to provide additional evaluation of the between subject interaction by evaluating the impact that age as well as sex had on the relationship between pilot experience level and trust in EFVS automation.

The results from the main effects analysis indicated that there is indeed a significant difference in trust in EFVS automation between the two types of EFVS technologies. More specifically, pilots tended to trust EVS displays more than SVS displays. On the other hand, the study found that the interaction between pilot experience and display type does not impact a pilot’s trust in EFVS automation. Additionally, the study found that pilot experience levels do not impact a pilot’s trust in EFVS automation. However, it is important to note that this particular analysis, the between subject interaction, failed to achieve sufficient power ($1 - \beta = 0.39$). Finally, the results from the covariant analysis provided additional confirmation that pilot experience did not have a significant impact on trust in EFVS automation by controlling two covariates: age and sex.
Discussion

Given the results of the data analysis, it is important to discuss the impact each result has on the aforementioned research questions. The following section will discuss each research question and provide insight into what influence the results will have on each topic.

Research Question 1

The first research question was concerned with investigating the relationship between EFVS display type and a pilot’s trust in EFVS automation. The hypothesis associated with this particular research question predicted that there would be a difference in trust between the two display types. The results obtained from the data analysis supported the hypothesis and it was determined that pilots trust EVS displays more than SVS displays. There are several plausible explanations why the data supported this particular finding. First, as discussed in chapter two, pilot experience with a particular piece of automation can significantly impact their trust in system capabilities (Bahner et al., 2008; Manzey et al., 2012; Parasuraman et al., 1996). Keeping that in mind, it is important to consider the prior experience the surveyed pilots had with both EFVS technologies. In terms of EVS displays, 75% of the participants indicated they had over 100 hours of experience with the technology. On the other hand, over 43% of the participants stated they had either no experience or only hands on experience with the SVS display. Considering the influence that
experience can have on implicit trust in automation, it is plausible that differences in specific display experience could have impacted the results of the study.

In addition to past experience, it is important to consider the impact that system designs could have on implicit/explicit trust. More specifically, when an operator is provided access to information that increases automation observability, trust in automation will be positively impacted (Salas & Maurino, 2010). Given the differences between how both systems create a visual representation of the flight environment, it seems plausible that a pilot’s trust could be negatively impacted by a lack of observability in the SVS display’s automated processes. While the EVS system provides an enhanced image of the flight environment, the SVS display relies on regular updates to a digital database to provide an accurate representation of the outside world. It is plausible that pilots are unsure about the system’s confidence in terrain/obstacle depiction and are therefore less trusting in its capabilities.

**Research Question 2**

The second research question was focused on determining if pilot experience had any impact on a pilot’s trust in EFVS automation. The hypothesis stated that pilot experience would influence a pilot’s trust in EFVS automation; however, the results obtained from data analysis indicated the hypothesis was not true. Though the polygon plot clearly showed a noticeable difference in average trust scores for all three experience groups in terms of the SVS display, the trust scores for the EVS display were grouped together. It is likely that the dynamic relationship formed
between the two display types and pilot experience could have accounted for a lack of significance during the data analysis.

Though the dynamic nature of pilot experience and display trust could have been a potential influence on the results, the most probable reason why this particular analysis failed to indicate significance is likely due to the fact that it did not achieve sufficient power to accurately evaluate the between subject interaction. A priori power analysis indicated that a sample size of $n = 120$ would have been necessary to achieve a power of 0.80 for the between subject interaction. Based on background research, the researcher determined that the within-between subject interaction would decide the overall sample size and a post hoc analysis would be conducted for the between subject interaction to determine the achieved power (Prajapati et al., 2010). The post hoc power analysis indicated that the between subject interaction achieved a power of only 0.39 based on a sample size of $n = 48$. Seeing that slightly over one third of the necessary sample size was utilized for the between subject study, it is likely that sample size played a significant role in the outcome of the analysis.

**Research Question 3**

The final research question was concerned with investigating the interaction between pilot experience and display type in terms of a pilot’s trust in EFVS automation. The study hypothesized that the interaction between both of the aforementioned variables would have an impact on a pilot’s trust in EFVS automation. Nevertheless, the results from the data analysis indicated that the interaction between pilot experience and display type did not significantly impact
trust in EFVS automation. Given the fact that this particular research question is essentially combining the within and between subject interactions, it is likely that many of the same factors that impacted both of those studies also influenced the results of the within-between subject analysis. However, an exception to this rule would be the power achieved by the study.

As with the other two analyses, a priori and post hoc power analysis was conducted for the within-between subject interaction. The priori power analysis indicated that a sample size of $n = 36$ was necessary to achieve a power of 0.80. Seeing as a total of 48 participants were utilized for data analysis, the achieved power of the within-between subject study was 0.86. As a result, it is likely that sample size did not contribute to a lack of significance. However, in addition to other the aforementioned factors, it is plausible that the sample did not accurately reflect the population. In other words, it is possible that sampling error was present in the study. Consequently, sampling error could have impacted not only the results of this particular interaction but also the other two analyses as well.

**Practical Implications**

In addition to enhancing the global base of knowledge for trust in automation, the study also sought to provide recommendations for potential system improvements to enhance trust in automation and flight safety. The following section provides a succinct overview of the insight obtained from the answers provided by
the participants in the survey as well as a few recommendations that could be implemented to potentially improve the implicit/explicit trust for EFVS automation.

Along with the former discussion regarding the impact that previous display experience could have had on a pilot’s trust, there are two more factors that could have impacted the pilot’s trust in EFVS automation. First, as discussed in chapter two, workload plays a significant role in an operator’s inclination to trust an automated system (Wickens, 2008). As a result, the survey incorporated four questions to replicate low workload (i.e., questions 5, 6, 9, and 10) and high workload (i.e., questions 3, 4, 7, and 8) conditions. The average trust score for the EVS and SVS displays were calculated for both workload conditions. In terms of low workload, pilots tended to trust EVS displays \( (M = 1.83, SD = 0.46) \) more than SVS displays \( (M = 0.29, SD = 0.70) \). Regarding high workload conditions, pilots once again placed more trust in EVS displays \( (M = 2.24, SD = 0.25) \) when compared to SVS displays \( (M = 1.08, SD = 0.15) \). Though the fact that EVS was trusted more than SVS on both occasions is not surprising considering the results of the within subject analysis, it is interesting pilots tended to have a higher trust rating for both systems under high workload conditions. This finding tends to agree with previous research in the field of workload and automation. Researchers noted that operators are more likely to rely on automation during high workload conditions (Rice, 2009; Wickens et al., 2013). It is likely that Wickens’ resource theory can provide context to this finding as well (Wickens, 2008). Considering high workload conditions necessitate more cognitive resources, it is likely the pilots are more willing to trust a system to
accomplish its task in order to conserve cognitive resources for other tasks. Likewise, when workload is low, pilots could be more likely to scrutinize system accuracy and therefore less likely to trust EFVS automation.

In addition to the impact that workload has on trust, a pilot’s exposure to a failure condition in an automated system can have a serious impact on trust. The survey incorporated two questions (i.e., questions 11 and 12) to study this particular condition. A comparison was made between the average trust score of both displays on each question. When the pilots were exposed to the initial failure in question 11, pilots rated their trust in both EVS ($M = 0.88, SD = 1.30$) and SVS ($M = -0.88, SD = 1.67$) at the lowest point throughout the survey. Though this finding is not necessarily surprising considering the circumstances, the average trust ratings when the failure condition was removed in question 12 were intriguing. Even though the pilot was told both systems were operating normally, the trust scores were still substantially lower for EVS ($M = 1.75, SD = 0.96$) and SVS ($M = 0.13, SD = 1.58$) when compared to every other question. The only exception to this statement would be when a comparison is made between question 12 and some of the high workload questions for EVS displays (i.e., question 3 ($M = 1.63, SD = 1.00$), 7 ($M = 1.63, SD = 1.20$), and 8 ($M = 1.54, SD = 1.25$)). The implications of this particular analysis provide additional empirical evidence to support the findings of other research regarding the impact automation failures have on pilot trust. More specifically, when pilots are exposed to a failure in an automated system under less than optimal conditions, it is
likely that they will never place the same amount of trust in the system (Manzey et al., 2012).

Given the previous discussion and the results of the study, it is necessary to outline a few enhancements that can be made to potentially improve trust in automation as well as safety of flight for helicopter pilots. First, in terms of observability and reliability, SVS displays should consider integrating some form of system confidence. Integrating system confidence into an SVS display would provide the pilot a reference point to quickly ascertain the system’s ability to adequately represent the outside world. System confidence could be based on the currency of terrain/obstacle data, the level of detail in terrain data, and Global Positioning System (GPS) accuracy.

In addition to improving system reliability and observability, adjustments should be made to pilot training programs to cultivate a calibrated trust between the pilot and EFVS automation. To be more specific, EFVS training should be conducted in a simulator to pre-expose pilots to system failures. The simulator flight should be accompanied by a subsequent flight where the pilot is exposed to actual system performance in a real aircraft. The flight should incorporate system failure training to ensure the pilots understand how the failure would impact their workload under actual flight conditions. Finally, the training should focus on highlighting both the benefits and limitations of the system to encourage the pilot to develop a calibrated trust in EFVS automation.
Uncontrolled Variables

As with any research topic, there were a few notable uncontrolled variables that could have impacted the results of the study. Though a few have been named in various sections, the following discussion will provide a consolidated list in the essence of maintaining complete transparency.

To begin with, as mentioned in chapter one’s discussion on limitations, the study utilized inferential statistics to make observations about the entire population of helicopter pilots. Assuming a study is able to achieve a sufficient level of power (0.80), the possibility of Type II error can be reduced and accurate conclusions can be made by the researcher. In this particular study, the between subject analysis failed to achieve sufficient power (0.39) and, as a result, could have subsequently suffered from Type II error. In addition to the possibility of Type II error, sampling error is a distinct possibility for any inferential statistics study, and this study is no exception to the rule.

Another uncontrolled variable would be the sampling technique utilized by the study. As discussed in chapter one, a distinct lack of population statistics for helicopter pilots prevented the researcher from incorporating probabilistic sampling techniques. Ideally, probabilistic sampling techniques would have been employed by the researcher to increase the power of the study and provide additional measures to avoid potential bias. However, it is believed that, given the circumstances, convenience sampling provided the most ideal solution to attain a sample that was truly random and void of researcher bias.
Next, the study instrument utilized for data collection (i.e., an online survey) accounted for a few uncontrolled variables that could have impacted the results of the study. For instance, though every effort was made to ensure participants could not take the survey multiple times, it is still feasible that a savvy user could have bypassed the measures in place and skewed the data. Additionally, it is also possible that people not qualified in helicopter flight operations could have taken the survey. Though a verification question was incorporated into the survey to identify such participants, it is feasible that an unqualified individual accurately answered the question and thereby skewed the data.

Finally, as discussed in chapter five’s discussion section, individual pilot experience with both EFVS technologies was not evenly balanced. In other words, the overwhelming majority of the surveyed pilots reported extensive experience with EVS displays and little to no experience with SVS displays. As a result, it is feasible that variations in pilot experience could have impacted the results of the within as well as the within-between study.

**Recommendation for Future Research**

The research outlined in this document was exploratory in nature due to the fact that trust in EFVS automation is a relatively novel field of study. When combined with several previously mentioned resource limitations (e.g., time and finances), the study was able to scratch only the surface of understanding how pilots interact with EFVS automation in the cockpit. Consequently, it is necessary that
future research be conducted in order to validate the findings outlined in this document and provide more detail on how several uncontrolled variables could impact a pilot’s trust in EFVS automation. Future research should focus on validating the analysis of the relationship between a pilot’s experience level and a pilot’s trust in automation by attaining a sufficient sample size \((n = 120)\). Furthermore, research should also be conducted to verify the legitimacy of the results from the within-subject and within-between subject study by balancing pilot display experience between both display types. Finally, future research should attempt to utilize stricter verification methods to truly prevent multiple responses from one participant, prevent unqualified participants from participating in the survey, and to validate participant responses in terms of demographic questions (e.g., flight hour experience).

**Conclusion**

One of the main reasons this research was conducted was to develop a current understanding of the trust based relationship between helicopter pilots and EFVS automation. Furthermore, the study also intended to investigate potential ways to improve trust in EFVS automation and thereby improve flight safety. By analyzing the survey responses of 48 helicopter pilots, the researcher was able to establish empirical evidence to answer three research questions regarding the current state of trust in EFVS automation. Additionally, based on background research and the empirical evidence obtained through data analysis, the researcher was able to develop
a succinct list of recommendations to potentially improve trust as well as flight safety. Though the research does have its limitations, it provides a solid foundation to support future trust in EFVS automation research and adds to the global knowledge of trust in automation.
References


Appendix A – Survey Documents

Informed Consent Page

Enhanced Flight Vision Systems

Consent

INFORMED CONSENT FORM

I agree to participate in this session which is part of an effort to develop an understanding of how current Enhanced Flight Vision System (EFVS) designs impact a pilot’s trust in automated system capabilities. The session will last approximately 10 minutes. The session will consist of the following:

1. Reading and signing of the informed consent
2. Overview of EFVS technology
3. Answering of demographic information (e.g., age, sex, pilot ratings, & previous experience with EFVS technology) to provide deeper understanding of the human-automation relationship
4. Review an example question
5. Answering survey questions
6. Debriefing

Should there be any problems, please contact Nicholas Currie at 386-463-7750 or ncurrie2015@my.fit.edu.

The risk level of the interview is no more than the risk involved with personal conversation via internet protocol. The benefit gained by the study is to understand how current Enhanced Flight Vision System (EFVS) designs impact a pilot’s trust in automated system capabilities.

I understand that in the EVENT OF PHYSICAL INJURY resulting from the research procedures in which I am to participate, no form of compensation is available. Medical treatment may be provided at my own expense or at the expense of my health care insurer (i.e., Medicare, Medicaid, private payor) which may or may not provide coverage. If I have questions, I understand that it is my responsibility to contact my insurer.

I understand that PARTICIPATION IS VOLUNTARY. Refusal to participate will involve no penalty or loss of benefits to which I am otherwise entitled. I understand that I may discontinue participation at any time without penalty or loss of benefits to which I am otherwise entitled.

1. By selecting agree below, I verify that I have read the above statement, I agree to the conditions outlined in said statement, and that I am at least 18 years of age or older.
   ○ Agree
Enhanced Flight Vision Systems

System Overview

Enhanced Vision System (EVS)

- Commonly known as Forward-Looking Infrared (FLIR) or millimeter wave radar
- Able to "see" through light obscurations (e.g., fog) and at night without a need for a light source
- Pilot is able to view the video on a Multi-Purpose Display (MPD)/Primary Flight Display (PFD) panel mounted display, a Heads Up Display (HUD), or a Helmet/Head Mounted Display (HMDS)
- Typically has a fixed Field of View (FOV) caused by the visual sensor and mechanical capabilities
- Subject to a phenomenon known as thermal cross-over, a condition that occurs twice a day where an object's temperature is similar to the background temperature

Example of the Astronic Max-Vis EVS ©

Synthetic Vision Systems (SVS)

- Computer generated flight video using a database of known terrain, obstacles, airport, and airway data
- System relies on GPS coordinates, radar/pressure altitude, and the provided database to replicate the surrounding environment for the pilot
- Pilot is able to view the video on a Multi-Purpose Display (MPD)/Primary Flight Display (PFD) panel mounted display, a Heads Up Display (HUD), or a Helmet/Head Mounted Display (HMDS)
- System visibility is not limited by environmental conditions such as thermal cross-over or fog density
- Some systems allow a pilot to select obstacle clearance notifications to aid in terrain avoidance
- System relies on pilot to update obstacle database on a regular basis to ensure system accuracy
Survey Instructions Page

Enhanced Flight Vision Systems

Survey Instructions

The following section provides a quick overview of how the survey will be conducted. You will be randomly assigned to one of the two EFVS technologies (i.e., EVS or SVS). You will then be asked to answer twelve (12) questions regarding your trust in the system’s automation. A sample question is provided below. Please answer each question honestly based on your experience and understanding of system capabilities.

Once you complete the first twelve (12) questions, you will then be redirected to the other EFVS technology and asked to answer the same twelve (12) questions. At the completion of that section, the survey will be complete.

1. To what extent do you trust that your personal automobile will start every morning?

<table>
<thead>
<tr>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

Example of Rockwell Collins’ HeliSure © H-SVS
Enhanced Flight Vision Systems

Demographic Information

The following questions will provide general demographic information necessary for the research. All information will be kept close-held and at no time will the study collect identifiable information (e.g., name). Please answer the following questions to the best of your knowledge.

1. Please select an age range that best describes you:
   - 18-25
   - 26-35
   - 36-45
   - 46-55
   - 56-65
   - 65 or older

2. Please select your gender:
   - Male
   - Female

3. Please select the helicopter pilot ratings that you currently hold:
   - Private
   - Instrument
   - Commercial
   - Certified Flight Instructor (CFI)
   - Certified Flight Instructor Instrument (CFII)

4. Please enter an approximate total number of flight hours you have flying helicopters:

5. Please select an experience level that best describes your familiarity with each EPVS system:

<table>
<thead>
<tr>
<th>Experience Level</th>
<th>Enhanced Vision System</th>
<th>Synthetic Vision System</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Experience</td>
<td>[ ]</td>
<td>[ ]</td>
</tr>
<tr>
<td>Hands On Flight Experience</td>
<td>[ ] [ ]</td>
<td>[ ] [ ]</td>
</tr>
<tr>
<td>Some Flight Experience (~100 hours or less)</td>
<td>[ ] [ ] [ ]</td>
<td>[ ] [ ]</td>
</tr>
<tr>
<td>Extensive Flight Experience (100+ hours)</td>
<td>[ ] [ ] [ ] [ ]</td>
<td>[ ] [ ]</td>
</tr>
</tbody>
</table>

6. Which helicopter control is used to make directional inputs (e.g., right, left, and forward) to the main rotor system?
   - Throttle
   - Collective
   - Cyclic
Appendix B – Questions Measuring Trust

Enhanced Vision System (EVS) Condition

**Enhanced Vision Systems (EVS)**

The following section will ask you to rank your trust in the automation of *Enhanced Vision System (EVS) technology* (e.g., *Forward Looking Infrared*). You will be asked a total of twelve questions regarding the application of this technology under several conditions. Please read each question carefully and provide your honest answer regarding your trust in the system's automation.

<table>
<thead>
<tr>
<th>Question</th>
<th>Options</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. To what extent do you trust an Enhanced Vision System (EVS) technology (e.g., Forward Looking Infrared) display to provide a realistic representation of the outside world?</td>
<td>Completely Distrust</td>
</tr>
<tr>
<td>2. To what extent do you trust an Enhanced Vision System (EVS) technology (e.g., Forward Looking Infrared) display to support your situational awareness of the flight environment?</td>
<td>Completely Distrust</td>
</tr>
<tr>
<td>3. To what extent do you trust an Enhanced Vision System (EVS) technology (e.g., Forward Looking Infrared) display to accurately display the outside world under limited visibility conditions (e.g., fog, dust, sunrise/sunset)?</td>
<td>Completely Distrust</td>
</tr>
<tr>
<td>4. To what extent do you trust an EVS (e.g., Forward Looking Infrared) display to accurately display the outside world under night time conditions?</td>
<td>Completely Distrust</td>
</tr>
<tr>
<td>Question</td>
<td>Completely Distrust</td>
</tr>
<tr>
<td>----------</td>
<td>--------------------</td>
</tr>
<tr>
<td>5. To what extent do you trust an EVS (e.g., <strong>Forward Looking Infrared</strong>) display to accurately display the outside world under daytime, unrestricted visibility conditions?</td>
<td>☐</td>
</tr>
<tr>
<td>6. To what extent do you trust an EVS (e.g., <strong>Forward Looking Infrared</strong>) display to provide reliable obstacle clearance information at cruise altitudes (above 200 feet)?</td>
<td>☐</td>
</tr>
<tr>
<td>7. To what extent do you trust an EVS (e.g., <strong>Forward Looking Infrared</strong>) display to provide reliable obstacle clearance information at low level altitudes (below 200 feet)?</td>
<td>☐</td>
</tr>
<tr>
<td>8. To what extent do you trust an EVS (e.g., <strong>Forward Looking Infrared</strong>) display to provide reliable obstacle clearance information within urban terrain?</td>
<td>☐</td>
</tr>
<tr>
<td>9. To what extent do you trust an EVS (e.g., <strong>Forward Looking Infrared</strong>) display to provide reliable obstacle clearance information over open terrain (e.g., farmland)?</td>
<td>☐</td>
</tr>
<tr>
<td>10. To what extent do you trust an EVS (e.g., <strong>Forward Looking Infrared</strong>) display to provide reliable obstacle clearance information during an instrument approach?</td>
<td>☐</td>
</tr>
<tr>
<td>11. To what extent do you trust an EVS (e.g., <strong>Forward Looking Infrared</strong>) display to provide reliable information following a system failure (e.g., pilot is presented with an &quot;Image Degraded&quot; message and FLIR image is still being presented)?</td>
<td>☐</td>
</tr>
<tr>
<td>12. To what extent do you trust an EVS (e.g., <strong>Forward Looking Infrared</strong>) display to provide reliable information once the system failure message has been removed indicating the system is operating nominally?</td>
<td>☐</td>
</tr>
</tbody>
</table>

**Synthetic Vision System (SVS) Condition**

84
Enhanced Flight Vision Systems

Synthetic Vision Systems (SVS)

The following section will ask you to rank your trust in the automation of Synthetic Vision System (SVS) technology (e.g. computer generated flight video). You will be asked a total of twelve questions regarding the application of this technology under several conditions. Please read each question carefully and provide your honest answer regarding your trust in the system's automation.

1. To what extent do you trust an Synthetic Vision System (SVS) technology (e.g. computer generated flight video) display to provide a realistic representation of the outside world?

<table>
<thead>
<tr>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

2. To what extent do you trust an Synthetic Vision System (SVS) technology (e.g. computer generated flight video) display to support your situational awareness of the flight environment?

<table>
<thead>
<tr>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

3. To what extent do you trust an Synthetic Vision System (SVS) technology (e.g. computer generated flight video) display to accurately display the outside world under limited visibility conditions (e.g., fog, dust, sunrise/sunset)?

<table>
<thead>
<tr>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

4. To what extent do you trust an Synthetic Vision System (SVS) technology (e.g. computer generated flight video) display to accurately display the outside world under night time conditions?

<table>
<thead>
<tr>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>
5. To what extent do you trust an *Synthetic Vision System (SVS)* technology (e.g., computer generated flight video) display to accurately display the outside world under daytime, unrestricted visibility conditions?

<table>
<thead>
<tr>
<th></th>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

6. To what extent do you trust an *Synthetic Vision System (SVS)* technology (e.g., computer generated flight video) display to provide reliable obstacle clearance information at cruise altitudes (above 200 feet)?

<table>
<thead>
<tr>
<th></th>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

7. To what extent do you trust an *Synthetic Vision System (SVS)* technology (e.g., computer generated flight video) display to provide reliable obstacle clearance information at low level altitudes (below 200 feet)?

<table>
<thead>
<tr>
<th></th>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

8. To what extent do you trust an *Synthetic Vision System (SVS)* technology (e.g., computer generated flight video) display to provide reliable obstacle clearance information within urban terrain?

<table>
<thead>
<tr>
<th></th>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

9. To what extent do you trust an *Synthetic Vision System (SVS)* technology (e.g., computer generated flight video) display to provide reliable obstacle clearance information over open terrain (e.g., farmland)?

<table>
<thead>
<tr>
<th></th>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

10. To what extent do you trust an *Synthetic Vision System (SVS)* technology (e.g., computer generated flight video) display to provide reliable obstacle clearance information during an instrument approach?

<table>
<thead>
<tr>
<th></th>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

11. To what extent do you trust an *Synthetic Vision System (SVS)* technology (e.g., computer generated flight video) display to provide reliable information following a system failure (e.g., pilot is presented with an "Image Degraded" message and FLIR image is still being presented)?

<table>
<thead>
<tr>
<th></th>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>

12. To what extent do you trust an *Synthetic Vision System (SVS)* technology (e.g., computer generated flight video) display to provide reliable information once the system failure message has been removed indicating the system is operating nominally?

<table>
<thead>
<tr>
<th></th>
<th>Completely Distrust</th>
<th>Moderately Distrust</th>
<th>Somewhat Distrust</th>
<th>Neutral</th>
<th>Somewhat Trust</th>
<th>Moderately Trust</th>
<th>Completely Trust</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
</tbody>
</table>