AN EVALUATION OF CONFLICT RESOLUTION DISPLAY FORMATS IN DETECT-AND-AVOID (DAA) TASKS IN UNMANNED AERIAL SYSTEMS

A Dissertation by

Michael Anthony Castaneda

Master of Arts, Psychology, University of West Florida, 2007

Bachelor of Arts, Psychology, Texas A&M University, 2004

Submitted to the Department of Psychology
and the faculty of the Graduate School of
Wichita State University
in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

December 2016

AN EVALUATION OF CONFLICT RESOLUTION DISPLAY FORMATS IN DETECT-AND-AVOID (DAA) TASKS IN UNMANNED AERIAL SYSTEMS

The following faculty members have examined the final copy of this dissertation for form and content, and recommend that it be accepted in partial fulfillment of the requirement for the degree of Doctor of Philosophy with a major in Psychology.

____
Alex Chaparro, Committee Chair

____
Barbara Chaparro, Committee Member

____
Jibo He, Committee Member

____
Paul Ackerman, Committee Member

____
Jeremy Patterson, Committee Member

Accepted for the College of Liberal Arts and Sciences

____
Ron Matson, Dean

Accepted for the Graduate School

____
Dennis Livesay, Dean

ACKNOWLEDGMENTS

I would like to sincerely thank my advisor, Dr. Alex Chaparro. He was able to stick with me when many probably believed I would never get to this point. His guidance throughout my years in graduate school has had a profound impact on how I approach problems in my everyday life. Over the years I have been able to work with him outside of the classroom developing real-world products where he has taught me how to effectively bridge the gap that often exists between academia and industry. I don't think I would be where I am in my career today without this invaluable skill which is often lost in other graduate programs.

I would also like to thank my defense committee members, Dr. Barbara Chaparro, Dr. Jibo He, Dr. Paul Ackerman and Dr. Jeremy Patterson. Without their willingness to accommodate my short timeline for proposing and defending I would not have been able to complete my degree.

Finally, I would like to thank Brandon Suarez, Dr. Erik Theunissen and all of the individuals at the NASA Ames Research Center. Without their dedication and long hours of work I would not have been able to collect the data necessary to complete my dissertation. I truly believe these individuals will one day revolutionize how UAS operate within the National Airspace.

ABSTRACT

An investigation was undertaken to evaluate three different display presentation types of the Conflict Prediction Display System (CPDS). The CPDS application is designed to provide operators with readily actionable Level 3 SA (Endsley, 1995a). The different display configurations were designed to evaluate the effects of information level and display format on pilot performance in order to inform future design decisions. A secondary objective was to utilize archival data to identify differences in workload and performance between integrated and standalone displays. Subject perception and objective performance data were collected from 10 participants who performed two detect-and-avoid tasks using a simulated ground control station (GCS). Each participant completed a 38-minute trial with each of the display configurations followed by a questionnaire. Several significant differences were identified in both objective performance and subjective perception between CPDS display configuration and between standalone and integrated displays. Overall, operators were significantly faster at responding to alerts on 3 of 5 performance metrics when conflict probes were used to display self-separation and collision avoidance information. In addition, operators preferred the integration of conflict probes when displaying surrounding traffic; however there was no clear preference regarding whether or not the probes display self-separation information. Standalone versus integrated comparisons indicated operators were significantly faster on all performance metrics when using an integrated display format.

Given the applied nature of the research it is important to go beyond just the statistical comparisons and look at the real-world implications of the data. The practical implications of the differences in display configurations are discussed in the framework of flight management in the national airspace (NAS).

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES (continued)

LIST OF FIGURES (continued)

Figure Page

CHAPTER 1

INTRODUCTION

An Unmanned Aerial Vehicle (UAV) is an aircraft controlled remotely by a pilot that is not physically on board the aircraft and is one component of an Unmanned Aerial System (UAS) which includes the UAV, Ground Control Station (GCS) and the datalink required to communicate between the GCS and the UAV. Although there is debate as to the first deployment of UAVs in military applications many will agree the 1994 deployment of the MQ-1 Predator to support military operations in the Balkans marked the beginning of the modern era of UAVs for military operations. Since their initial role as intelligence, surveillance and reconnaissance (ISR) aircraft the role of UAVs has grown considerably to include close air support (CAS), dynamic targeting, strike coordination and reconnaissance (SCAR), and combat search and rescue (CSAR). In November of 2010 UAVs had accumulated one million combat hours and by October of 2013 they had already surpassed two million combat hours according to the *United States Air Force RPA Vector: Vision and Enabling Concepts 2013-2038* (2014). As conflicts across the globe continue it is inevitable that the roles and responsibilities of UAVs will increase as technologies advanced and lessons are learned.

Although the foundations of UAV technology were born out of military applications civil and commercial applications for these aircraft have become increasingly apparent. In 2013 the Federal Aviation Administration (FAA) released the *Integration of Civil Unmanned Aircraft Systems (UAS) in the National Airspace System (NAS) Roadmap* in which they identified several civil and commercial applications including security awareness, disaster response, communications and broadcast, cargo transport, spectral and thermal analysis, commercial photography and aerial mapping. Unfortunately, prior to any of these utilities coming to fruition there are significant hurdles which must be overcome. In 2013 the Association for Unmanned Vehicle Systems International (AUVSI) attempted to quantify the potential economic impact of UAVs in a report titled *The Economic Impact of Unmanned Aircraft Systems Integration in the United States*. AUVSI's findings indicated that within the first three years of integration more than 70,000 jobs would be created in the United States with an economic impact of more than $13.6 billion. They also predicted that the benefit would continue to grow through 2025 at which time they predicted more than 100,000 jobs being created and an economic impact of $82 billion. Today UAVs are not allowed to operate within the National Airspace System (NAS) without first obtaining special authorization. Current operations within the NAS require a Certificate of Authorization (COA) from the FAA; however, the process of obtaining a COA can be laborious taking several months and even when granted they typically have restrictions on where, when and how flight operations can take place. As long as this requirement is in place the FAA will quickly be overcome with a backlog of COA requests.

In order to meet the growing demand of military, civilian and commercial UAV applications it will be necessary for them to be able to operate in all classes of airspace with the same access as aircraft with a pilot onboard. Over the past several years National Aeronautics and Space Administration (NASA) in conjunction with industry partners has started a concerted effort to allow seamless integration of UAVs into the NAS. As part of this effort they have begun to conduct a series of experiments using extensively developed real-world scenarios in real-world environments using instrument-rated pilots in order to provide valid data from which inferences can be extracted to make assumptions regarding how UAV integration into the NAS would or would not change the current airspace. The present study investigated archival data which was collected as part of NASA's ongoing effort.

The primary limiting factor preventing UAVs from unrestricted access to the NAS is the inability to meet the intent of the FAA's "see-and-avoid" requirement. According to Title 14 Code of Federal Regulations Part 91, section 91.113, "when weather conditions permit, regardless of whether an operation is conducted under instrument flight rules or visual flight rules, vigilance shall be maintained by each person operating an aircraft so as to see and avoid other aircraft." Since UAVs do not maintain a pilot on board to see-and-avoid the FAA has mandated UAVs be capable of performing the see-and-avoid function and thereby achieving an "equivalent level of safety" (ELOS), comparable to an aircraft with a pilot onboard; however, there has been no formal definition of what defines an ELOS. Although these aircraft have onboard video systems that enable the pilot to view their surrounding environment they are rendered useless in the event that the pilot losses link with the aircraft. Since UAVs are flown without a pilot onboard, applying these requirements to unmanned operations will require a "detect-and-avoid" (DAA) system with a self- separation (SS) function that allows pilots to comply with minimum separation standards remotely while in-the-loop under normal operating conditions (RTCA, 2013; Santiago & Mueller, 2015). At a minimum, the information displayed by a DAA system must provide pilots with sufficient awareness to detect potential threats, determine appropriate responses to detected threats, and execute maneuvers using the GCS's command and control interface (Fern, Rorie, Pack, Shively, & Draper, 2015). However, what has yet to be determined is what types and to what level information should be provided to a pilot through a DAA system. For example, a display system which provides predictive information may provide pilots advanced warning of potential conflicts; however, predicting too far in the future may cause pilots to perform preventative maneuvers which may have not been necessary, thereby creating unplanned deviations from their filed flight plans.

In addition to an inability to meet see-and-avoid requirements there is currently no formal method for UAVs to obtain airworthiness certification which will be required for full integration in the NAS. Although there is a well delineated process for manned aircraft to obtain airworthiness certification many of the regulations which are outlined in the process are not applicable to UAVs. For example, current regulations address security requirements for cockpit doors; however, these regulations do not provide a legal definition of what a cockpit is or where it is located. This presents a challenge for UAVs considering that the cockpit, or GCS, may be located in a building, in a vehicle, or even outside with no physical boundaries.

Full integration of UAVs in the NAS will require a consolidated effort amongst military, governmental, civil and commercial agencies. Technologies such as Traffic Collision Avoidance System (TCAS), Automatic Dependent Surveillance-Broadcast (ADS-B) and air-to-air radars will need to advance to the point where they can be integrated onboard UAVs in order to enable UAVs to meet an equivalent level of safety as the see-and-avoid requirement of manned platforms. Regulations such as airworthiness certification, datalink latency and lost link requirements will need to be established in order to provide UAV manufacturers a pathway for developing aircraft capable of meeting these requirements. Finally, all of this information will need to concisely and efficiently be provided to the operator. Today most UAV operations take place in special use airspace or combat environments where all flight operations are coordinated thought flight operations centers. The amount of information provided to operators operating within the NAS will inherently increase workload and the amount of situational awareness required to safely and efficiently navigate their aircraft thought increasingly congested airspace..

1.1 Enablers

1.1.1 Technology

Traffic Collision Avoidance System II (TCAS-II) and Automatic Dependent Surveillance System – Broadcast (ADS-B) are examples of current technologies which provide situational awareness to pilots about traffic within their operating area. TCAS is a cooperative surveillance system that requires an aircraft to have onboard TCAS equipment. TCAS works by interrogating the transponders of nearby aircraft to identify imminent risks of collision. If the TCAS computer considers another aircraft to be a potential threat a traffic advisory (TA) is initiated in the aircraft equipped with TCAS. Aircraft with a transponder only and no TCAS equipment would not receive any notification. Traffic advisories issued by the TCAS system are provided as situational awareness information only and do not require the pilot to maneuver. Traffic advisories are issued when a predicted collision is within 20 to 48 seconds depending on aircraft altitude. Should the projected encounter remain and the predicted collision comes within 15 to 35 seconds a resolution advisory (RA) will be issued which will direct the pilot to initiate a vertical maneuver. When a RA is issued a pilot is required to initiate the directed maneuver. Encounters between two TCAS equipped aircraft will always initiate coordinated RAs between aircraft. Although TCAS has proven to be an effective flight coordination system it is most effective when both aircraft are TCAS equipped and have similar performance parameters.

ADS-B is a two part system consisting of ADS-B In and ADS-B Out. ADS-B Out broadcasts precise information regarding the aircraft's location, ground speed and other flight data to satellites at a rate of 1 Hz. This data is then distributed to a network of ADS-B ground stations which allow air traffic controllers and other aircraft equipped with ADS-B In to receive this information. In contrast to typical radar surveillance technology which sweeps for position information every 5 to 12 seconds, ADS-B Out will provide the ability to track aircraft at a much higher data rate. In addition, the radio waves used by radar surveillance are limited to line-of-site (LOS) meaning the signals cannot travel long distances or penetrate solid objects. ADS-B ground stations are smaller and more adaptable than radar towers and have the ability to be placed in locations where radar towers cannot. In 2014 the FAA completed the deployment of 660 ADS-B ground stations across the United States which now allows for complete coverage across U.S. airspace. With ground stations in place throughout the country ADS-B will provide better visibility regardless of terrain or other obstacles.

ADS-B In provides operators of with ability to receive Flight Information Service-Broadcast (FIS-B) and Traffic Information Service-Broadcast (TIS-B). FIS-B provides a wide range of weather products and flight information including Aviation Routine Weather Reports (METARs), NEXRAD precipitation maps, status of Special Use Airspace (SUA), Temporary Flight Restrictions (TFRs) and winds and temperatures aloft. TIS-B is an advisory-only service designed to provide traffic information on all transponder-based aircraft within the vicinity of the ADS-B In equipped aircraft.

Although TCAS II and ADS-B provide insight into surrounding air traffic, they are not currently mandated for all aircraft operating in the NAS. In order to fully meet the FAA's see-and-avoid requirement UAVs must be able to autonomously detect aircraft operating without transponders or surveillance equipment, and avoid them in worst case scenarios such as when the GCS loses link with the aircraft. In order to meet this need aircraft must be outfitted with advanced air-to-air radar as well as complex processors which can take inputs from multiple sources such as TCAS, ADS-B and onboard radars and correlate aircraft which are being tracked by multiple sources. Although air-to-air radars do not represent a new technology integrating them into UAVs can represent a unique challenge. In general, UAVs are typically smaller, lighter and less powerful than their manned counterparts. Given the limited space available within these smaller aircraft future air-to-air radars will need to be lighter and smaller than previous generations. Several companies are already in the process of developing, integrating and flight testing air-to-air radar solutions into their respective platforms including Northrup Grumman and General Atomics Aeronautical Systems Incorporated.

1.1.2 Regulatory

As part of the FAA's ongoing transformation of the NAS, referred to as NextGEN, the FAA has mandated that by January 1, 2020, all aircraft be equipped with ADS-B Out. Although the ground infrastructure is already in place to support ADS-B Out across the United States there is currently no mandate for aircraft to be equipped with ADS- B In. Should the FAA decide to issue such a mandate for UAVs in order to mitigate the aircraft's inability to see-and-avoid, this could potentially minimize the need for onboard air-to-air radars in UAVs, but not until 2020.

The Radio Technical Commission for Aeronautics (RTCA) is a private, not-for-profit association made up of public, private and government members which has been chartered by the FAA to operate federal advisory committees. Although the RTCA does not set policy or issue mandates their recommendations are often used as the basis for government and private sector decisions as well as forming the foundation for many FAA Technical Standard Orders (TSOs). In 2013, the RTCA set up Special Committee (SC) 228 in order to develop the Minimum Operational Performance Standards (MOPS) for Unmanned Aircraft Systems (UAS). Since its' inception the committee has come to the realization that developing MOPS for a UAS must be addressed at the component level rather than at the system level. In September 2015 SC-228 issued its' first draft MOPS for air-to-air radars for detect-and-avoid (DAA) systems. Once officially released, these MOPS will provide guidance to manufactures regarding the specifications that may one day be required as part of the airworthiness certification process that would enable full access in the NAS. In May of 2016 the committee released a draft of its' second MOPS designed to address Command and Control (C2) data link standards. This document is designed to layout the link capabilities necessary to support UAV and ground station information exchange that allows the pilot to safely control, monitor and manage the UAV. Although there is currently no timeline set for developing the necessary MOPS required to define airworthiness certification criteria for UAS it is likely that the compendium of MOPS set forth by the RCTA will play a vital role in assisting the FAA's ongoing efforts to define airworthiness requirements specific to UAS.

1.1.3 Research and development

In 2011 NASA began a project to address barriers related to unrestricted access to the NAS by UAS. The project identified three key technical subproject areas to focus on: (a) Separation Assurance Sense and Avoid Interoperability (SSI); (b) Human Systems Integration (HSI); and (c) Communication. Regarding HSI, two overarching and interrelated objectives were identified. The first was to develop GCS guidelines for UAS access to the NAS and the second was to develop a prototype display setup (Fern, Shively, Johnson, Trujillo, Pestana & Hobbs, 2011). The first objective was identified to address the lack of GCS design standards or requirements for UAS operations in the NAS. The development of a prototype was identified to serve three purposes: (a) serve as a test-bed for UAS procedures and displays; (b) provide data for informing guidelines; and (c) provide a proof of concept of those guidelines. The primary concern of the HSI subproject was to identify how to present information in the GCS in a unified and intuitive manner while ensuring manageable pilot workload, and increasing situational awareness.

1.2 Situational Awareness (SA)

Within the last 25 years there has been a significant amount of research conducted on the psychological construct of Situational Awareness (SA). Although this construct has been formalized in both the research and operational communities over the past two decades, pilots as far back as World War II have been able to cite the importance of SA in their operational environment (Press, 1986, as cited in Endsley, 2015a, p. 163). Over the years some (Dekker, 2015) have argued against the notion of SA as a construct citing concerns about circularity in the SA construct. Dekker's (2015) arguments appear to ignore or diminish the last 25 years of research which have provided a clear definition of SA, theoretical models to explain its formation and development through cognitive processes and mechanisms, as well as validated metrics that provide independent and objective measurements of SA and its processes. Dekker and Hollnagel (2004) dismiss SA as a ‘‘folk model’’, however, Parasuraman, Sheridan and Wickens (2008) soundly refuted that characterization by citing a large body of empirical research on SA offering evidence of the concepts diagnosticity, prescriptive usefulness, and both theoretical and empirical distinctions from performance and other mental constructs. In 2008, Wickens provided an overview of significant research on SA and progress in the areas of measurement, training, error analysis, team work, automation, and workload, with advances in both theory and applications providing a strong argument for its feasibility as a construct in the human factors domain.

Mica Endsley's (1995a,b) seminal papers on situational awareness theory and measurement have formed the basis for the most accepted theory of situational awareness. Although her theory has not been without its detractors (Chiappe, Rorie, Moran & Vu, 2012; Chiappe, Strybel & Vu, 2011; Chiappe, Vu & Strybel, 2012, Dekker, Hummerdal & Smith, 2010; Klein, Moon & Hoffman, 2006a,b; Salmon, Stanton & Young, 2012; Sorensen, Stanton & Banks, 2010); Uhlarik & Comerford, 2012). Endsley (2015b) has vigorously refuted the misconceptions and misunderstandings posed to the theory as well as identify how alternative theories of SA are in fact encompassed and accounted for in her 1995 theory of situation awareness, thereby reiterating the robustness of the theory.

After years of working with pilots and the design of advanced aircraft Endsley endeavored to provide a comprehensive account for the construct of Situational Awareness. Her account, as it would turn out, rests on many other cognitive mechanisms and behaviors and is heavily embedded in the challenges and messiness of complex environments (Endsley, 2015c). Although other researchers (Sarter & Woods, 1995) had argued that developing a definition of SA was futile and not constructive, Endsley set forth to clearly define the construct of her model. Accordingly, she defined situation awareness as “the perception of the elements in the environment within a volume of space, the comprehension of their meaning, and the projection of their status in the near future” (p. 36). In its simplest form, Endsley’s model rests on the premise that there are 3 levels of Situation Awareness (Figure 1) which map to her definition.

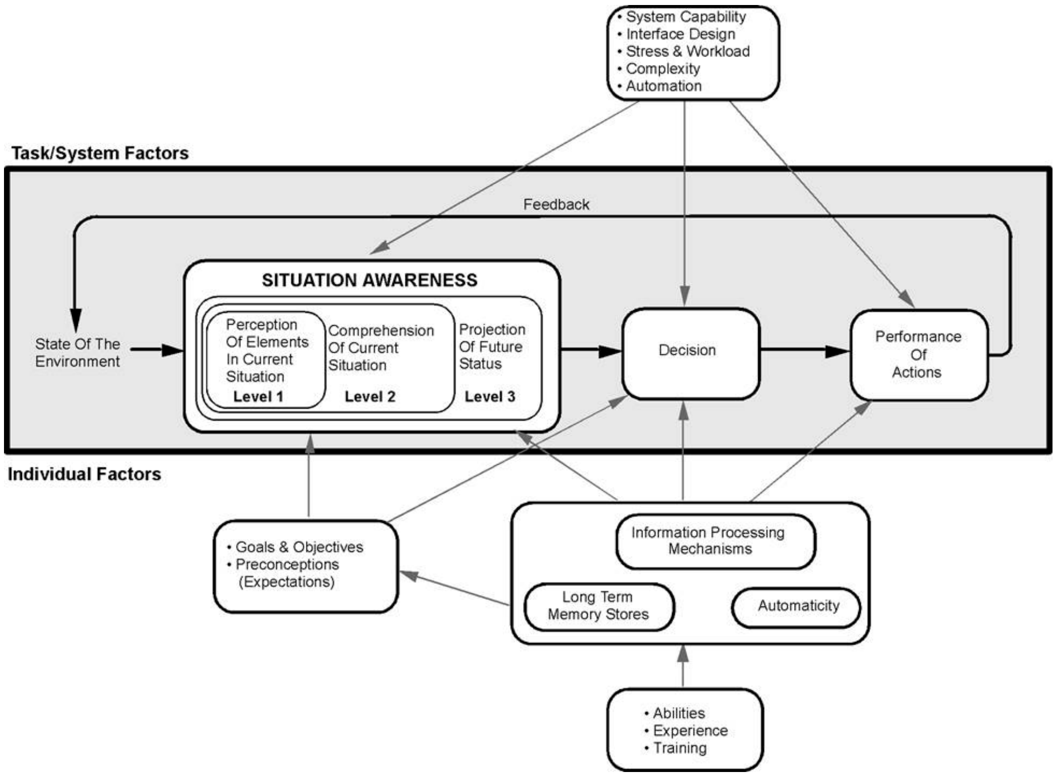


Figure 1. Endsley's (1995a) Model of Situation Awareness

Level 1 SA is essentially the perception of cues in the environment. The lack of or incorrect perception of information dramatically increases the likelihood of forming an incorrect picture of a given situation. In one particular study (Jones & Endsley, 1996) 76% of SA errors in pilots were linked to difficulties in perception.

Level 2 SA is the integration of elements perceived in Level 1 SA to form a comprehensive understanding of the environment in reference to a particular goal. According to Endsley (2000) when Level 2 SA is achieved a person has been able to develop operationally relevant meaning and significance from the Level 1 data. Jones and Endsley (1996) traced 20% of SA errors to problems involving the formation of Level 2 SA.

The highest level of SA (Level 3) represents one’s ability to predict future situation events and dynamics based on the information perceived and integrated in Levels 1 and 2. Endsley (2000) notes that in the multitude of fields she has studied experienced operators rely heavily on future projections and it is often a mark of an expert.

In order to fully understand Endsley's model there are several aspects which must be made explicit in Figure 1. First, the levels of SA were never meant to be conveyed as strictly linear stages, such that one must obtain Level 1 and Level 2 to get to Level 3. Rather, the levels are presented in the model as embedded within each other in order to convey that there is an ongoing iterative process between the levels in order to update all three levels in real-time. Second, there is also a temporal aspect to SA. SA can be thought of as ones understanding of a given situation for a given state in time, while the model actually represents the process for maintaining and updating the given state for any one point in time. The dynamic nature of the environment dictates that a situation is always changing and to prevent SA from becoming out-dated or inaccurate one must constantly adapt their cognitive strategies to maintain an up-to-date and accurate state of SA. Finally, SA as construct should be viewed separately from decision making and performance as indicated by the model. Although the model dictates that SA is the main precursor to decision making and performance it is possible to have perfect SA yet make incorrect decisions resulting in poor performance. For theoretical purposes Endsley presents decision making and performance as distinct stages from SA; however, they can all affect each other in an ongoing circular cycle, yet then can be decoupled through various other factors.

1.3 Cognitive Mechanisms of SA

1.3.1 Attention and working memory

One's ability to achieve and maintain good SA is constrained by limited attention and working memory. As a limited resource, how attention is directed will have a profound impact on which elements of an environment are incorporated in the formation of SA. The perceived salience of cues and the meaningful direction of attention both play important roles in the formation of SA. Attention should be prioritized based on how important a particular cue is perceived to be. The ability to appropriately prioritize attentional demands has been found to be a marked trait in experienced users in a variety of domains (Endsley & Smith, 1996; Endsley & Rogers, 1998; Gugerty (1998).

The limits of working memory also pose a constraint on the formation of SA as novice decision makers or operators in novel situations are required to combine information, interpret it and make future projections all in limited working memory (Endsley, 2000). Gugerty and Tirre (1997) provided strong evidence on the importance of working memory in distinguishing between people with higher and lower SA.

1.3.2 Long-term memory and working memory

In order to circumvent the limitations of working memory on developing SA, Endsley's (1995a) model relies on Cowan's (1988) model of cognition which proposes working memory to be an activated subset of long-term memory. According to Cowan, information can proceed directly from sensory memory to long-term memory which is necessary for pattern recognition and coding. Salient portions of the environment remain in working memory as a highlighted subset of long-term memory through localized attention or automatic activation.

1.3.3 Mental models, schemas and scripts

A mental model is an understanding of how something works based on semantic and system knowledge. While semantic knowledge includes knowing what as opposed to how, system knowledge includes an understanding of how systems function (Endsley & Jones, 2004). Mental models provide direction on what information is important and form expectations which enable higher levels of SA without increasing strain on working memory. Limiting strain on working memory makes mental models enablers of Levels 2 and 3 SA (Endsley & Jones, 2004).

Functioning like shortcuts, schemas allow for very rapid classification and understanding of perceived information without necessarily accessing a mental model. A schema can provide both comprehension and projection based on well-known or recognized classes of situations. Mental models and schemas are advantageous in developing SA in that they do not require the current situation do be exactly like a previously experienced situation to recognize it which allows individuals to generalize from past experiences to new ones.

Finally, individuals may develop scripts associated with a particular schema. Scripts are sequences of actions on what do to in each case a schema represents. Although scripts can be acquired through experience or directed by a particular domain, they cannot exist on their own. An individual must still use a mental model or schema to understand a situation well enough to know when to apply a particular script. The presence of scripts can further reduce the strain on working memory when deciding the best course of action for a given situation (Endsley & Jones, 2004).

1.3.4 Goals

Essential to the formation of SA is the concept that human information processing in complex systems is an alternating process between data driven (bottom-up) and goal driven (top-down) processing (Endsley, 1988, 1995a). According to Endsley's model, goals serve three critical functions:

1. Direct the selection of a mental model.
2. Direct attention in selecting information from the environment.
3. Interpret and integrate the information to achieve comprehension.

As goal driven processing is occurring data driven processing is also occurring as information is perceived in the formation of SA. At any given time the information obtained through data driven processing can serve to change or modify the goals which will in turn modify the goal driven (top-down) processing. This alternating process is one of the most important mechanisms underlying Endsley's model.

1.3.5 Expectations

Expectations also influence how attention is distributed and the perception of information processed. While expectations may be advantageous when they support the current goal, they can have negative consequences in novel situations or when unanticipated information is present. Taylor, Endsley and Henderson (1996) demonstrated both the positive and negative effects of expectations when assessing SA in teams. Teams given a well-developed set of expectations tended to perform poorly when events and situations did not meet expectations. They also did not develop the same type of group skills for active problem solving as teams that were not provided well-developed of expectations.

1.3.6 Automaticity

Repeated exposure to highly similar environments can often lead operators to develop a level of automaticity for certain routine tasks. Over time operators can develop a highly routinized pattern-recognition/action-selection sequence (Logan, 1988) which can provide an adaptive quality in well-understood environments by allowing for good performance with a relatively low level of attentional demand. However, when the environment deviates from the normal routine automaticity can lead to detrimental effects. As automaticity increases it leads operators to be less receptive to novel events and less likely to effectively integrated novel information resulting in an increase of errors when a change from the learned pattern is required.

1.4 Designing for SA

Endsley and Jones (2004) identified eight general factors which can undermine situational awareness (SA) in a variety of systems and environments, referring to them as SA demons. The list of SA demons included:

1. Attentional tunneling
2. Requisite memory trap
3. Workload, anxiety, fatigue and other stressors (WAFOS):
4. Data overload
5. Misplaced salience
6. Complexity creep
7. Errant mental models
8. Out-of-the-loop syndrome

In order to avoid these SA demons, Endsley and Jones point to six design principles which can help designers develop interfaces to that will allow operators to avoid these traps. These principles include:

1. Organize information around goals
2. Present Level 2 information directly – Support comprehension
3. Provide assistance for Level 3 SA projections
4. Support global SA
5. Support trade-offs between goal-driven and data-driven processing
6. Make critical cues for schema activation salient

Since situational awareness is a construct distinct from decision making and performance is it ever possible to answer the question, “how much SA is enough?” Although Endsley (2000) points out that SA and performance are probabilistically linked, high levels of one does not guarantee high levels of the other. While a high level of SA will increase the probability of a high level of performance, it is completely possible to have perfect SA, yet perform poorly based on decision making which lies between SA and performance in Endsley's model. Endsley (2000) argues that there is no such thing as too much SA, more SA is always better. However, in this argument she refers back to the operational definition of SA which defines SA as "that which one really needs to know". Given this argument it could be argued that too much SA is not really SA in a strict sense, but rather one of the SA demons (e.g., data overload) acting to increase SA in a certain regard at the cost of a more important aspect. It should be noted however that in designing operator interfaces a distinction must be made between data provided and data perceived or extracted. As advancements in technology allow for greater amounts of data fusion in to singular displays designers must be cautious not to provide too much data which can make it harder for data extraction and the development of situational awareness (Duggan, Banbury, Howes, Patrick, & Waldron, 2004).

CHAPTER 2

INTEGRATED VERSUS STANDALONE DISPLAYS

2.1 Research in Related Domains

Although unable to find of any systematic research investigating display formats in the UAS domain, two studies were identified which compared integrated and standalone displays in a laboratory-based context (Gillie & Berry 1994; Hansen, 1995) and two additional studies have compared integrated and standalone displays in realistic simulation environments (Sauer et al., 2002; Friedman-Berg & Ahlstrom, 2005). Gillie and Berry found participants showed better performance with an integrated display in a control task but no difference on a fault detection task. A year later, in 1995, Hansen found participants had a higher hit rate on a fault detection task using an integrated display, but also had a higher false alarm rate concluding there was no definitive advantage of an integrated display compared to a standalone display.

Sauer et al. (2002) investigated integrated and standalone displays in the maritime environment from the perspective of a ship's bridge displays. Similarly to the aviation domain ships have access to displays which provide information regarding surrounding traffic (Advanced Radar Plotting Aids (ARPA)) and displays which provide navigational information (Electronic Chart Display Information Systems (ECDIS)). Although these display systems are not integrated, Lee and Sanquist (1996) identified two primary motivations for the need to integrate the information provided by the separate displays: 'perceptual augmentation' and 'control integration'. Perceptual augmentation aims to improve an operator's understanding of the ship's environment (i.e. situation awareness), while control integration aims to enhance the operator's ability to navigate the ship. Similar to traffic displays and navigation displays in aviation, both ARPA and ECDIS are displays of the same geographical region and could therefore be superimposed on a single display. Sauer et al. argued the integration of information should be technologically feasible as well as a natural and appealing advancement. In particular, they argued the integration into a single display would agree with the proximity compatibility principle, which states that if a task requires high processing proximity (i.e. different information sources need to be sampled), display integration should be high (Wickens & Carswell, 1995). In addition, Wickens (2000) also points out the heavy demand on cognitive resources related to visually scanning separate displays and the mental integration of information. However, at the time of publication Sauer et al. noted that the maritime community was divided over the issue, with some accepting the case for integration of the two functions, and others fearing it would lead to information overload, resulting in difficulty identifying targets on the integrated, and potentially cluttered, display (Smeaton, Dineley, & Tucker, 1995).

Similar to the NASA's HSI subprojects, Sauer et al. (2002) were looking to conduct an ecologically valid experiment that could easily translate to real-world maritime operations. Seeking conclusions which would be transferable to real-world operations, Sauer et al. developed their simulation environment guided by general workload theory (Wickens, 1992) and a more specific framework provided by Hockey's model of cognitive-energetical control (Hockey, 1997). The primary argument of Hockey's model is that task performance is goal-regulated, and operators will typically maintain adequate levels of performance on high priority (primary) tasks (e.g., collision avoidance), even during periods of peak work demands and stress. This is achieved by a compensatory process in which additional resources are recruited to the primary task, maintaining performance on top level goals (e.g., safety of the aircraft) at the expense of lower priority activities (e.g., responding to radio calls). Although this compensatory process may succeed in maintaining primary performance, there is a psychological cost to be paid in terms of increased effort and strain, particularly an increase in fatigue (Hockey, 1997).

In all, the study looked at 3 display configurations, fully integrated, functionally-separate, and spatially-separate displays. The fully integrated display contained both ARPA and ECDIS information superimposed on a single display, while the functionally-separate display contained both ARPA and ECDIS on a single display but as two separate views which required the operator to switch between which meant the operators could never view both pieces of information simultaneously. Finally, the spatially-separate display contained the ARPA and ECDIS information on two separate vertically stacked displays allowing operators to view both pieces of information simultaneously. Using three measures of primary task performance, Sauer et al. (2002) reported superior performance of the integrated display on only one of three primary task dependent measures. In the integrated display conditions operators deviated significantly less from their planned course than in the separate display conditions. There was no difference between display presentation types in safe distance error, as measured in terms of the percentage of time that the closest-point-of-approach fell below the safe distance of 0.25 nautical miles, or expert ratings, as measured by subject matter experts subjectively rating the operator’s performance. Three measures of secondary task performance were collected, including detection of oil temperature drifts, display resets and cargo temperature recordings. None of the secondary task measures varied significantly as a function of display configuration. The results did not support Hockey’s model which predicted that superior performance on the integrated display should have come at a cost reflected in worse secondary task performance on the integrated display. However, the subjective data collected did support Hockey’s model (1997) indicating that anxiety, fatigue, and mental workload were all significantly higher in the integrated display condition while situational awareness was also significantly lower in the integrated display condition.

In 2005, Friedman-Berg and Ahlstrom investigated integrated and standalone advanced weather displays using Terminal Radar Approach Control (TRACON) controllers in a realistic simulated air traffic control environment. Six advanced weather display tools were provided to controllers on either a standalone display or integrated into the standard Terminal Control Workstation (TCW). The study investigated the utilization of the advanced tools as well as behavioral performance data in terms of completed flights. Regarding tool utilization only one of the six advanced tools (gust fronts) used more frequently in the integrated display condition, while a different advanced tool (storm motion) was used for a longer duration in the standalone display condition. In terms of performance, operators performed better in both display configurations than a control condition, but there was not significant difference between performance on the integrated or standalone displays indicating that the additional weather resources did improve performance, but it did not necessarily matter whether the tools were presented on an integrated or a standalone display.

Although intuitively it may seem an integrated approach is a better solution, as it does not require the pilot to mentally integrate information from standalone displays to form a complete picture of the airspace and potential conflicts, a major concern for integrating displays is the potential increase in clutter as well as the potential of other performance-degrading effects (Ververs & Wickens, 1998). Some may argue this concern can be mitigated by simply providing declutter capabilities, however, it has been demonstrated that simply hiding irrelevant information is not necessarily better than alternative measures, such as dimming redundant or unnecessary information (Yeh & Wickens, 2000). In addition, providing operators with declutter capabilities comes at the additional decision-making costs of what kind of information to hide. In addition to these concerns, the current research does not necessarily support this anecdotal theory. On the other hand, standalone displays provide the capability to rapidly integrate the application into new and developing UAV platforms. Unfortunately, investigating software applications for integration in a GCS is not as simple as looking at performance based metrics. Other metrics such as integration time, level of rigor for testing and stability must be considered when choosing the optimal application for inclusion in the GCS. As the FAA moves towards establishing airworthiness certification criteria for UAVs there is little doubt that the software used to control these aircraft will come under increased scrutiny and more rigorous testing. In order to integrate additional software code embedded within the same software used to control a UAV it is undoubtable that the software application would be required to undergo the same level of testing as if only the software code used to control the aircraft had been modified. On the other hand, a standalone software application can simply receive the data necessary to plot the UAVs flight path data along with surrounding aircraft which can be received directly from radar feeds or ADS-B ground stations if necessary. This type of integration would not require any changes to the software code used to control the UAV and thereby alleviate the need to undergo additional testing. Similarly, it may seem more information is always better when attempting to make a well informed decision. However, additional information displays require display space within a GCS and information which maybe supplemental but not necessary can often cause clutter and make the necessary information harder to extract.

2.2 Research in UAS Detect-and-Avoid (DAA) Tasks

2.2.1 Background

Several recent studies examined the minimum visual information requirements necessary to perform UAS pilot-in-the-loop DAA tasks. For example, predictive displays which include color-coded alerting structures, airspace warning zones, and closest-point-of-approach (CPA) indicators along with intruder state information and directionality have reduced near mid-air collisions (NMACs), well clear violations, and time spent within the well clear boundary compared to displays with less information (Bell, Drury, Estes, & Reynolds, 2012; Friedman-Berg, Rein & Racine, 2014). In 2014, Draper, Pack, Darrah, Moulton and Calhoun conducted a survey which revealed that the majority of pilots prefer intruder state information and visual alerts be displayed at all times; however respondents also designated flight restrictions, weather, and DAA maneuver recommendations as critical information elements.

These studies were used by the HSI subproject to provide a framework from which to build toward specifying the minimum DAA display elements required in a GCS for safe integration into the NAS. With this framework in mind the HSI subproject set out to conduct a series of experiments known as Part Task Simulations, to investigate the effects of information level, display location, and maneuver guidance using a prototype display suite.

Investigating the effects of these variables on pilot performance required the HSI subproject team to identify a suitable GCS software application that would provide the core functionality of the simulated GCS and display suite. The Air Force Research Laboratory's (AFRL) Vigilant Spirit Control Station (VSCS) (Feitshans, Rowe, Davis, Holland & Berger, 2008) was selected as a mature software application which had been previously used to control multiple UAS in both simulation and flight tests. The VSCS provided a robust yet flexible interface in addition to providing critical technology which could support potential flight-testing activities. A specific version of VSCS was developed by HSI in conjunction with AFRL to support its research, prototyping, and guidelines development activities. Figure 2 shows the primary VSCS display which displays the UAS ownship as well as the planned flight path over a moving map.

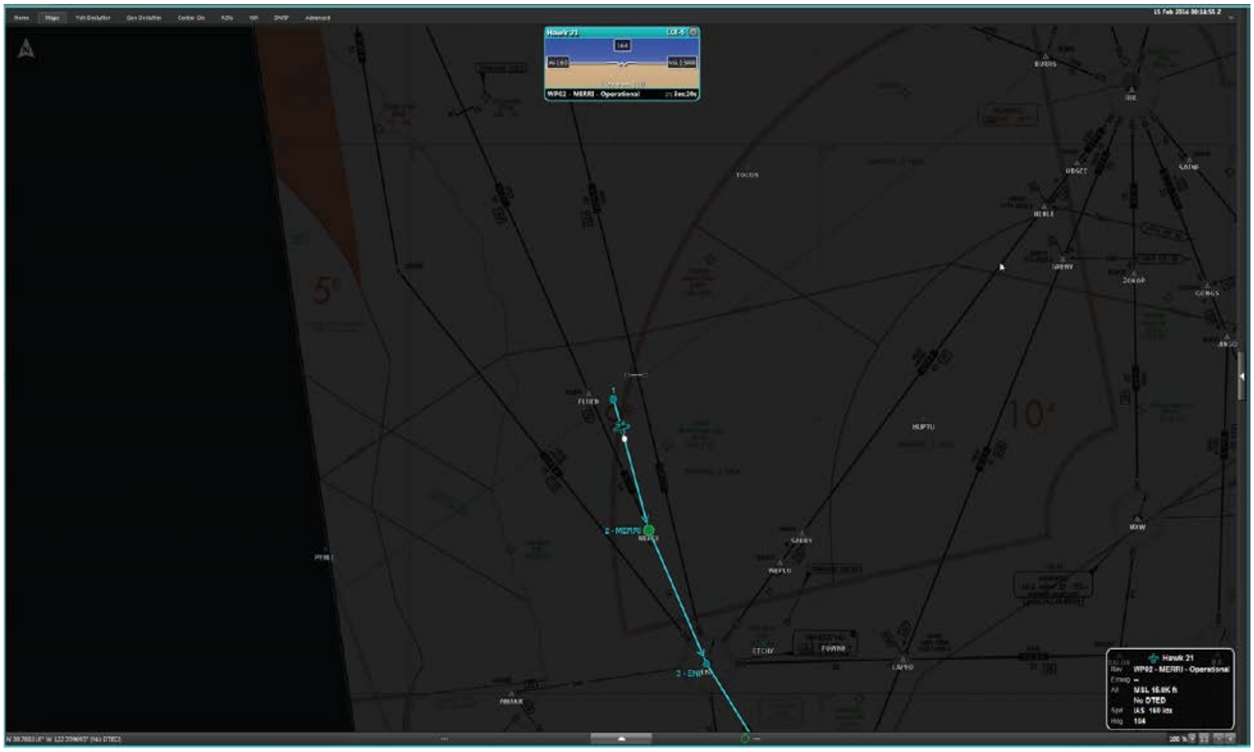


Figure 2. Vigilant Spirit Control Station (VSCS).Reprinted from “An Evaluation of Detect-and-Avoid (DAA) Displays for Unmanned Aircraft Systems: The Effect of Information Level and Display Location on Pilot Performance,” by L. Fern, R.C. Rorie, J. Pack, R.J. Shively, M. Draper, 2015, Proceedings of the 15th AIAA Aviation Technology, Integration, and Operations Conference. Copyright 2015 by the American Institute of Aeronautics and Astronautics.

In order to investigate differences between integrated and standalone displays the team also selected the Cockpit Situation Display (CSD) (Figure 3) as a separate software application capable of displaying traffic information. The CSD is a standalone 3D application developed by the Flight Deck Display Research Laboratory at NASA Ames Research Center capable of displaying the locations and trajectories of ownship and surrounding traffic (Granada, Dao, Wong, Johnson, & Battiste, 2005). During the investigation the display was limited to a 2D orientation. The CSD also contained logic for displaying conflict alerts and resolutions.

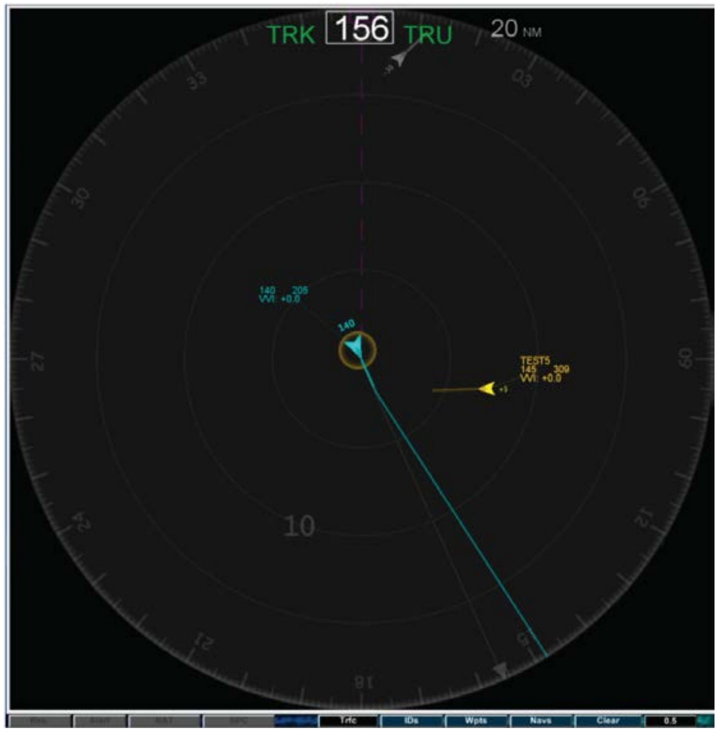


Figure 3. Cockpit Situation Display (CSD).Reprinted from “An Evaluation of Detect-and-Avoid (DAA) Displays for Unmanned Aircraft Systems: The Effect of Information Level and Display Location on Pilot Performance,” by L. Fern, R.C. Rorie, J. Pack, R.J. Shively, M. Draper, 2015, *Proceedings of the 15th AIAA Aviation Technology, Integration, and Operations Conference*. Copyright 2015 by the American Institute of Aeronautics and Astronautics.

2.2.2 Findings

Display location and information level were investigated as NASA’s Part Task Simulation 4 and reported in Fern et al. (2015) and Monk et al. (2015). Using the prototype GCS Fern et al. investigated the effects of information level and display location on pilot DAA performance and perception in an experiment referred to as Part Task Simulation 4. Information level was distinguished as basic and advanced and display location was distinguished as integrated or standalone. The basic display configuration contained a minimum set of intruder traffic information including location, range, bearing, heading, relative altitude, vertical trend, heading predictor, and threat level (Friedman-Berg et al., 2014; Draper et al., 2014).

The advanced configuration contained the basic information along with additional information and features designed to aid pilots during potential conflicts. Advanced features included the predicted CPA location and time-to-CPA, an additional alert level for traffic predicted to violate the well clear volume, a 0.8 nautical mile ring around ownship, recommended maneuvers, a vertical profile display, and tools which predicted the threat level at a given heading or altitude. When dragging the planning tools to a desired heading or altitude, the tool's arrow would change color according to the predicted alert level if the maneuver were to be executed (as shown in Figure 4).

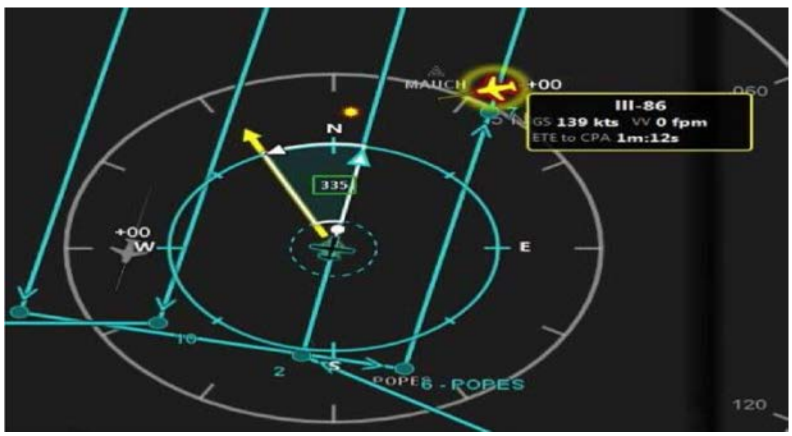


Figure 4. Lateral Vertical Planner in VSCS.Reprinted from "Effects of Display Location and Information Level on UAS Pilot Assessments of a Detect-and-Avoid System," by K. Monk, R.J. Shively, L. Fern & R.C. Rorie, 2015, *Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting* p. 51. Copyright 2015 by the Human Factors and Ergonomics Society.

A total of 12 pilots participated in a within-subjects design, experiencing both advanced and basic levels of information display on both a standalone and an integrated display. The standalone display conditions (basic and advanced) presented pilots with traffic information and conflict resolution tools (when applicable) on the CSD, while manipulation of ownship flightpath was executed using the VSCS on a separate monitor to the right. The integrated conditions presented all display features on the VSCS. The trial planning tools in the integrated condition were coupled to the auto-pilot interface, so that proposed headings and altitudes were automatically modified in the steering window. This automatic modification of the steering window was not available in the standalone condition.

Five measures of performance were assessed which related to response time: (a) notification time, (b) initial response time, (c) initial edit time, (d) total edit time, and (e) total response time. Of the five response time measures assessed none of them resulted in significant main effect of display location (i.e., standalone vs. integrated).

In addition to the objective measures collected operators were also administered questionnaires after each trial as well as after the entire study (Monk, Shively, Fern, and Rorie, 2015). The post-trial questionnaires aimed to evaluate training sufficiency, initial alert response, ease of use, clutter and performance degradation. Regarding training sufficiency, operators perceived the basic integrated display training to be significantly more sufficient for safe operations compared to the advanced standalone display. Operators also indicated they were significantly better at responding immediately to collision avoidance (CA) threats while using the advanced integrated display than when using the basic standalone and basic integrated displays. Display configuration was found to have only marginal effects on the ease ratings of both trial planning tools. Operators indicated the lateral planning tool was significantly easier to use on the advanced integrated display compared to the advanced standalone display. Similarly, the vertical planning tool was also identified as significantly easier to use on the advanced integrated display compared to the advanced standalone display. Operator perceived clutter or performance degradation did not significantly vary amongst display conditions.

The post-experiment questionnaires aimed to evaluate alerting logic, display location, information sufficiency, and advanced display features. With regard to alerting logic, 83% of operators indicated they had sufficient time to assess the situation and maintain well clear across all displays overall. All operators reported a preference for traffic information being located on an integrated display as well as perceiving operations to be safer with an integrated display. Overall, the majority of operators indicated that all display configurations provided sufficient information to remain well clear; however, the strongest consensus (92%) was reported for the advanced integrated display agreement among pilots, while only 64% of operators agreed the basic standalone display information was sufficient. Regarding tool preference, operators ranked the predictive outlining which indicated potential collision avoidance conflicts as the most preferable advanced display feature followed by the vertical situation display, the lateral and vertical trial planners, and time-to-CPA.

Given neither paper reported conclusive evidence of the advantage of an integrated display over a standalone display a follow-on study reported here was undertaken to investigate a unique standalone application referred to as the Conflict Prediction Display System (CPDS). CPDS is an alternative software application which presents operators information regarding the current and future predicted states of potential conflicts. Due to extenuating circumstances CPDS was not included in the original Part Task Simulation 4 investigation; however, a follow-on study referred to as Part Task Simulation 4B was conducted in order to investigate the potential benefits of the CPDS as an alternative to support operator decision making in DAA tasks. Using the prototype GCS developed by the HSI subproject, the original Part Task Simulation 4 was partially replicated in Part Task Simulation 4B using the CPDS as a standalone display in order to continue growing the knowledge base of information display elements which may help operators perform the DAA functions required to safely integrate UAVs in the NAS. No additional data was collected regarding the integrated display conditions during Part Task Simulation 4B; instead the archival data collected using integrated displays during Part Task

Simulation 4 was used for comparison purposes when investigating the CPDS as a standalone display relative to an integrated display.

CHAPTER 3

THE CONFLICT PREDICTION DISPLAY SYSTEM (CPDS) IN DAA TASKS

3.1 Background

Since 2001, a collaborative research project between Delft University of Technology (TU Delft), the Royal Netherlands Air Force (RNLAF) and the Netherlands Defence Academy (NLDA) has been on-going to address several aspects of UAV mission management. The UAV Laboratory of the NLDA provides a test bed with the capability to evaluate a range of both vehicle-specific and mission-level UAV functions. The project's main focus has been developing prototype software applications which can be used beyond the simulation phase and implemented in real-world applications. The fact that software interfaces of a UAV can be identical for both simulations and real world operation provides advantages regarding the transfer of training skills to real-world operation. Additionally, compared to the validation of software interfaces for manned aircraft which can require million dollar simulators which are outside the financial capabilities of many research institutions, the differences between well-designed simulated evaluations and actual flight tests for unmanned systems are relatively small given unmanned systems are operated remotely from the actual aircraft. As part of the project's effort to address the ever growing concern of integrating unmanned aircraft into the same airspace as manned platforms the Conflict Prediction Display System (CPDS) was developed in order to provide a deployable software application for UAVs that would support pilot DAA decision making.

In 2008, Tadema and Theunissen began investigating the idea of integrating traffic probes into standard cockpit displays of traffic information (CDTIs) and primary flight displays

(PFDs). Standard CDTIs typically display ownship as well as surrounding air traffic at an operator selectable distance and will often include traffic velocity and altitude. Although these display types do well in supporting Level 2 situational awareness (SA) they lack the information necessary to readily support Level 3 SA. In originally developing the CPDS, Tadema and Theunissen believed by integrating traffic probes into a typical CDTI they could provide pilots Level 3 SA which would in turn support better decision making during DAA tasks. Traffic probes present operators with a spatially integrated depiction of airspace where a loss of separation, or conflict, is predicted to occur. These probes provide readily actionable conflict information, allieviating operators from the conflict estimation task and contributing to Level 3 SA. The resulting user interface concept was intended to support operators in the assessment of potential traffic conflicts, while providing an understanding of the consequences of maneuvering on the future separation with traffic. Additionally, traffic probes support operators in monitoring the adherence to traffic ‘right-of-way’ priority rules (Tadema, 2011).

Prior research has addressed the potential of conflict probing to support UAV pilots with DAA tasks (Tadema & Theunissen, 2008) as well as the integration of vehicle maneuvering constraints to define more accurate traffic probes (Lambregts, Tadema, Rademaker, & Theunissen, 2009; Tadema & Theunissen, 2009). Additionally, pilot-in-the-loop simulations, in which traffic probes were integrated into a standard CDTI to support level 3 SA were able to demonstrate advantages in terms of safer and more efficient maneuvering decisions compared to a standard CDTI alone (Tadema & Theunissen, 2008). However, prior to Part Task 4B the CPDS had never been evaluated against alternative software interfaces which were also designed to support pilot decision making during DAA tasks. Part Task 4B was undertaken in order to assess the potential benefits of the CPDS relative to alternative user interfaces designed to support the same decision making. Although the displays in Part Task 4 and 4B were all designed to support the same task, each provides a unique and distinct interface. Conducting Part Task 4B allowed the HSI subproject to continue to compile a growing data set using the same methodology in order to begin assessing the unique contributions to pilot performance of each interfaces unique method of supporting level 3 SA during DAA tasks. In addition, all trials conducted as part of Part Task 4 displayed both self-separation and collision avoidance information to the pilots although self-separation events do not require a pilot to maneuver or report. As part of Part Task 4B the investigators looked to assess any performance differences when only providing pilots with collision avoidance and near mid-air collision (NMAC) information, which both require incident reports. Finally, the investigators hoped to be able to identify the unique contribution of the conflict probes as opposed to simply providing conflict bands. Conflict bands only provide pilots with maneuver guidance information and declutter the actual probes which provide the entire predicted collision space. The proposed research will evaluate the archival data collected during Part Task 4B as well as data collected during Part Task 4 which has not been previously reported.

3.2 CPDS Display

The CPDS software application displays traffic information in an intuitive way with the results of a conflict prediction algorithm, known as conflict probes, integrated into the same air traffic display as other air traffic, allowing the pilot to quickly assimilate the current situation and anticipate future situations. The CPDS is made up of a 2D plan view display (5) as well as a vertical profile display (VPD) (Figure 6).

The 2-D plan view display area provides the operator with a two-dimensional bird's-eye view of ownship and surrounding traffic. This display area also provides situational awareness of potential conflicts that will result in a loss of separation of a near midair collision. Together with the 2-D plan view display the VPD provides the operator with a three-dimensional picture of ownship and the surrounding airspace.

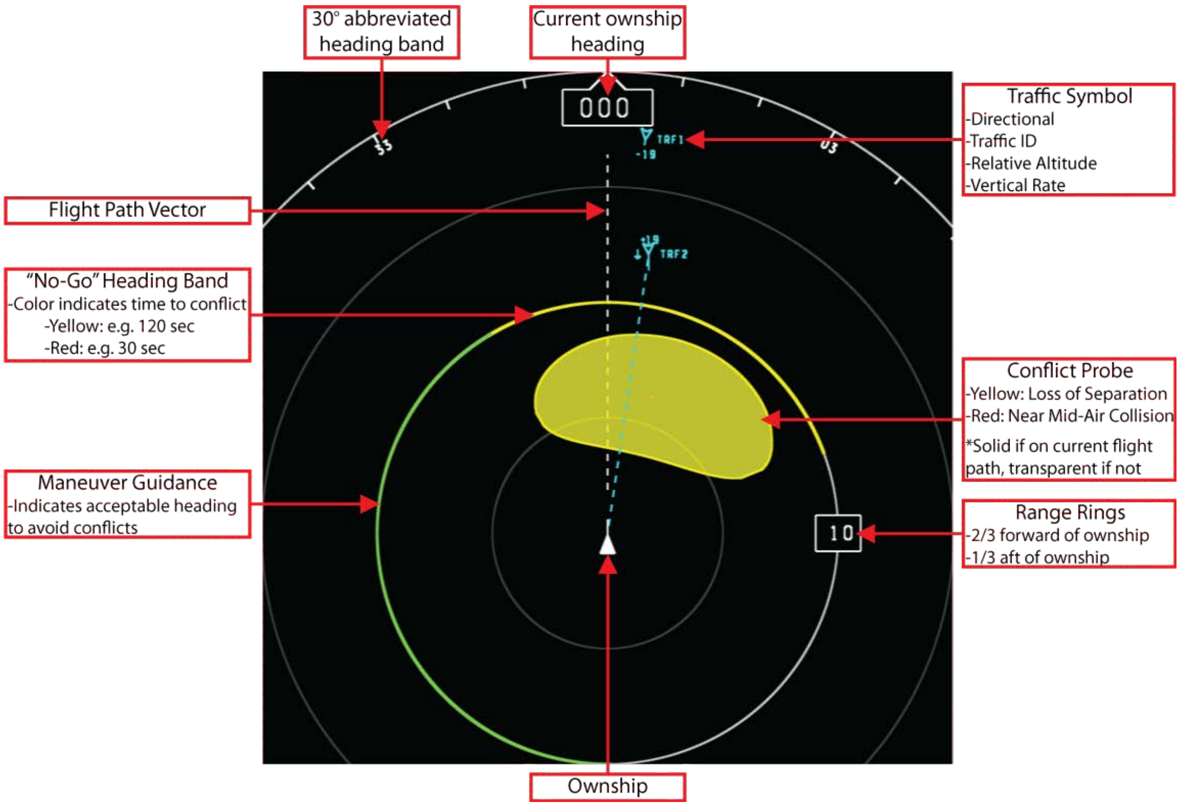


Figure 5. CPDS 2D Plan View

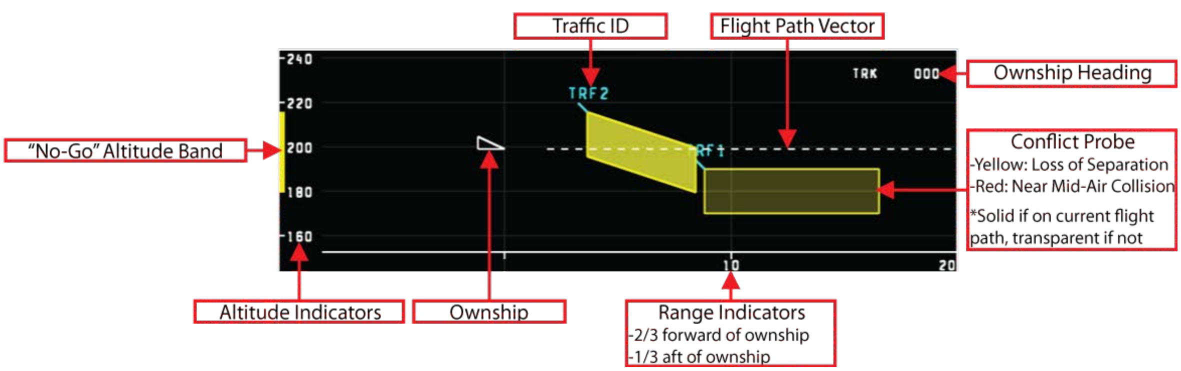


Figure 6. CPDS Vertical Profile Display

3.3 Conflict Probing

Conflict probing algorithms are designed to predict the current and future separation between ownship and other traffic for a set of ownship velocity vectors (i.e., potential combinations of Track, Flight Path Angle (FPA) and Speed) up to a specified look-ahead time. Using predefined separation criteria the probing algorithms indicate which ownship velocity vectors will lead to a future conflict and what the corresponding time to loss of separation will be.

To illustrate the concept of probing, Figure 7 shows a top-down view of a conflict resulting from converging traffic. Vehicle A (center) represents ownship, vehicle B is the converging traffic and the dashed lines represent the current tracks. Initial bearing of the traffic is 290, range is 5 nautical miles; track is 050, flying level at a Speed of 250 knot and both aircraft are at the altitude. The conflict band and probes result from probing a range of variations of the ownship track angle, for the current FPA and Speed. The separation criteria (yellow) is 1 nautical mile laterally and 1,000 feet vertically and the collision hazard criteria (red) is 0.25 nautical miles laterally and 500 feet vertically. The depicted probe represents the conflict space while anything outside of this space represents potential conflict resolution maneuvers. Provided this information operators can make decisions regarding the trade-offs and interdependencies of changes of the components. Since probing is performed in real-time, the conflict space is continuously updated as the situation develops. Should conflicting traffic maneuver, this will be reflected by corresponding changes of the conflict space.

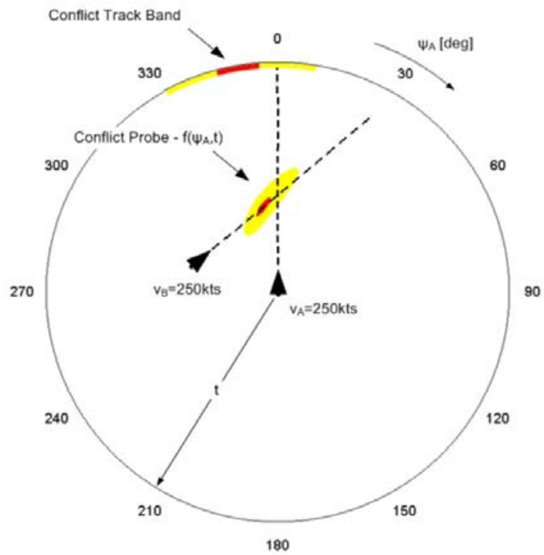


Figure 7. Conflict Probing Example.Reprinted from, "Beyond Traffic Depiction: Conformally Integrating the Conflict Space to Support Level 3 Situation Awareness," by J. Tadema, E. Theunissen, & K.M. Kirk, 2010, *Proceedings of SPIE: Enhanced and Synthetic Vision 2010,* p. 768902-2. Copyright 2010 by the International Society for Optical Engineering.

CHAPTER 4

EXPERIMENTAL QUESTIONS

In order to further refine the knowledgebase available to support HSI subproject's primary goal of identifying how to present new information in the GCS in a unified and intuitive manner while ensuring manageable pilot workload and increasing situational awareness the present study investigated the data collected during Part Task Simulation 4B on the utility of three standalone CPDS display configurations. In addition to analyzing the three CPDS configurations in relation to each other the archival data available on the VSCS integrated display configurations collected during Part Task Simulation 4 was used to compare the CPDS standalone display configurations in relation to the VSCS integrated display configurations. The same methodology was used during this study (Part Task Simulation 4B) as was used in the original Part Task Simulation 4 study in order to provide comparative data that can also be analyzed in conjunction with the previously collected data.

The experimental questions are ordered to first address the utility of the three CPDS display configurations in relation to each other. Questions 1 through 5 will only address comparisons across the three CPDS display configurations. The experimental questions will then address the utility of the CPDS as a standalone display in relation to integrated display configurations. Questions 6 through 8 will only address comparisons between the three CPDS standalone display configurations and the two VSCS integrated display configurations.

4.1 CPDS Comparisons

4.1.1 Question 1

Is there a performance difference between display presentation types?

Hypothesis 1: It is hypothesized that Display 1 will result in the best overall performance. Display 1 provides only the information necessary to allow operators to rapidly identify the collision avoidance conflict space which requires maneuvering without providing additional precautionary self-separation information. Display 2 which requires operators to mentally extrapolate the conflict space should result in additional information processing demand and burden on working memory capacity may slow response time to potential conflicts. Display 3 which provides the additional layer of precautionary self-separation information will require operators to expend additional mental processing resources on the decision of whether or not to maneuver to avoid self-separation alerts which are only precautionary in nature. The additional resources required in this decision making process should result in slower response time.

4.1.2 Question 2

Is there a performance difference on secondary tasks between display presentation types?

Hypothesis 2: It is hypothesized Display 2 will result in the largest performance decrement on the secondary tasks. All three display types will be presented on the same display which is located the furthest from the display presenting the secondary tasks. However, unlike Display 1 and Display 3, Display 2 requires the operator to mentally extrapolate the potential conflict space, which is anticipated to reduce secondary task performance. Based on Multiple Resource Theory (Wickens, 2002) tasks which demand more resources are more likely to interfere with concurrent tasks. In order to appropriately identify and manage the potential conflict space in the Display 2 condition operators will be required to mentally extrapolate this information placing an additional burden on working memory and the attentional resources available to perform concurrent secondary tasks.

4.1.3 Question 3

Which display presentation type improves perceived accuracy in determining, negotiating and executing traffic avoidance maneuvers?

Hypothesis 3: It is believed that the display presentation type which provides the most information will result in the greatest perceived accuracy. Although few studies have studied the relationship and subjective accuracy Lichacz (2002) was able to demonstrate a positive correlation between situational awareness and confidence ratings. As situational awareness increased so did participant judgments on the confidence of their responses. Display presentation type 3 will provide pilots with both conflict probes and bands which is similar to display presentation type 1; however, display 3 will provide information regarding both self-separation alerts and collision avoidance alerts. Provided the pilots will have more information available as they determine if and when to perform avoidance maneuvers thereby increasing the amount of readily actionable Level 3 SA it is hypothesized that Display 3 will result in the greatest perceived accuracy.

4.1.4 Question 4

Which display presentation type improves perceived efficiency in determining, negotiating and executing traffic avoidance maneuvers?

Hypothesis 4: Generally display presentations which reduce visual clutter on the display are anticipated to result in the greater perceived efficiency. Display 2 is designed to only provide operators with conflict bands and declutters the conflict probes. In this presentation type pilots will have the least information on the display to process in order to determine if and when to perform an avoidance maneuver. It is hypothesized that Display 2 will result in the greatest perceived efficiency.

4.1.5 Question 5

Is there a perceived workload difference between display presentation types?

Hypothesis 5: Although all three display presentation types provide varying levels of saliency for Level 3 SA which should in-turn reduce workload on the operator it is hypothesized that Display 1 will result in the lowest self-report ratings of workload. Display 2 provides acceptable heading and altitude bands in which to avoid conflicts; however, since it does not provide conflict probes it will require operators to mentally extrapolate where the conflict space actually exists. This extrapolation will but an increased working memory burden on the operator in order to achieve the same degree of Level 3 SA achieved inherently in Display 1. Although Display 3 does not require the operator mentally extrapolate the conflict space it does provide information regarding self-separation alerts which are only precautionary and do not require pilots to maneuver. This potentially undue information could lead the three of Endsley's SA demons (i.e., Data overload, misplaced salience, and complexity creep) which put an additional strain on working memory to filter our unnecessary information.

4.2 Integrated Versus Standalone Comparisons

4.2.1 Question 6

Is there a performance difference between standalone and integrated displays?

An exploratory analysis is proposed in order to investigate operator performance differences between an integrated display configuration and the CPDS as a standalone display in of detect-and-avoid (DAA) tasks. Earlier studies (Sauer et al., 2002; Friedman-Berg & Ahlstrom, 2005; Fern et al., 2015) reported no significant differences between standalone and integrated display configurations. It is possible that lack of effects could be due to the inability of the particular standalone software to provide the information necessary to aid pilots in DAA tasks.

Providing readily actionable Level 3 SA the CPDS may be able to provide operators the information necessary to obtain statistically significant performance differences compared to an integrated display.

4.2.2 Question 7

Is there a performance difference on secondary tasks between standalone and integrated displays?

An exploratory analysis is proposed in order to compare secondary task performance between an integrated and standalone display when the CPDS is used in detect-and-avoid (DAA) tasks. Secondary task performance data in Part Task Simulation 4 has not been reported so it is not known if there was a performance difference in secondary tasks when NASA's Cockpit Situation Display (CSD) was used a standalone display for DAA tasks. However, based on the findings of Sauer et al. (2002) it is expected that secondary task performance will not be effected by the use of an integrated or standalone display. Although Hockey's (1997) model of cognitive-energetical control would predict a decrement in secondary task performance when primary task performance remains constant across experimental manipulations the evidence from Sauer et al. did not support this even though primary task performance did not differ across conditions in 2 or 3 measures.

4.2.3 Question 8

Is there a perceived workload difference between standalone and integrated displays?

Hypothesis 8: Workload data in Part Task Simulation 4 has not been reported; however, based on Hockey's (1997) model of cognitive-energetical control as well as the findings of Sauer et al. (2002) it is hypothesized that the integrated display condition will result in higher self-report levels of workload.

CHAPTER 5

METHODS

Part Task Simulation 4B was designed to be a follow-on study to Part Task Simulation 4 which was originally reported in Fern et al. (2015) and Monk et al. (2015). The methods employed were designed to align with the methods used in Part Task 4 in order to continue building to the database of information available regarding display information and location as it relates to pilot abilities to safely navigate congested airspace. The raw data from Part Task 4 was available and was be used for exploratory analyses with the data collected from Part Task 4B. Part Task 4B was not originally planned as part of NASA's ongoing HSI Subproject and was limited to two weeks of data collection due to external circumstances which resulted in the ability to collect data from 10 operators (1 per day).

5.1 Participants

Ten Instrument Rated pilots (*M* = 45 years of age) participated in the study. Six operators had previous experience flying UAVs and five participated in the previous Fern et al. (2015) study. All operators were male. A retired air traffic controller participated as a confederate to ensure the minimum number of encounters occurred during each trial.

5.2 Experimental Setup

Operators were seated at a simulated GCS (Figure 8) running the Vigilant Spirit Control Station (VSCS) software and the Conflict Prediction Display System (CPDS) software. During the experiment VSCS generated three displays in the ground station, a Tactical Situation Display (TSD) on the lower center display, a health and status display on the outer right display, and a simulated nose camera on the upper center display. The TSD served as the operator's primary display, providing ownship and planned flight path information, a moving map, and navigation controls. Aircraft navigation was managed via one of two possible control interfaces. A steering command interface allowed operators to modify the assigned altitude or heading of the aircraft and also provided a means to return the aircraft to the preplanned route. Alternatively, operators could directly interact with the compass rose surrounding the aircraft to manipulate aircraft heading. Using the compass rose operators were able to drag a heading bug to a desired heading instead of entering a specific heading manually in the steering command interface. The compass rose could not be used to modify altitude. Once the heading bug was released the steering command window would automatically open. Pilots uploaded changes to the aircraft by pressing a “Send” button within the steering command window.

Figure 8. Simulated GCS. Reprinted from “An Evaluation of Detect-and-Avoid (DAA) Displays for Unmanned Aircraft Systems: The Effect of Information Level and Display Location on Pilot Performance,” by L. Fern, R.C. Rorie, J. Pack, R.J. Shively, M. Draper, 2015, Proceedings of the 15th AIAA Aviation Technology, Integration, and Operations Conference. Copyright 2015 by the American Institute of Aeronautics and Astronautics.

A health and status display was located to the right of the TSD and contained a chat client and subsystem information used for secondary tasks, as well as telemetry data. The third

VSCS display was a simulated nose camera video located above the TSD. Operators interacted with the TSD and health and status display using a standard mouse and keyboard.

A fourth monitor provided the CPDS (Figure 9). The CPDS was located to the left of the TSD on a separate monitor, and was configured to display ownship information, surrounding traffic and alerting information. Operator interaction with the CPDS was limited to adjusting the display range of the CPDS using the same standard mouse used for interaction with the VSCS. The CPDS had a minimum display range of 10 nautical miles and a maximum display range of 160 nautical miles. Along with ownship and intruder information, the CPDS provided pilots with range rings and heading markers along the edge of the display. The upper section of the CPDS was restricted to a 2-D plan view with ownship located two-thirds down on the display and a track-up orientation. The lower section of the CPDS provided a vertical profile display (VPD) with ownship located two-thirds to the left on the display.

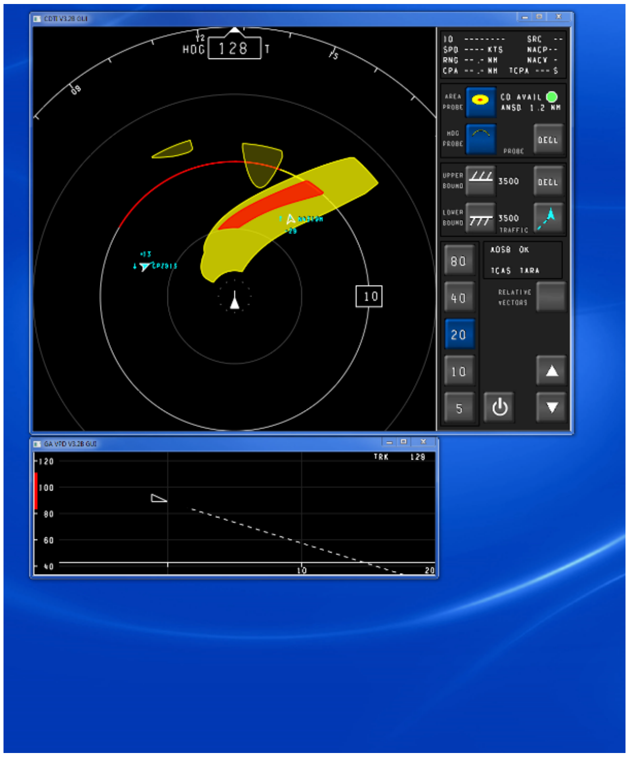


Figure 9. CPDS

5.3 Display Configurations

In order to assess the usability of the CPDS in presenting readily actionable Level 3 situational awareness (SA) three display configurations were evaluated which varied the presence of conflict probes as well as the level of information available to the pilots (Figure 10). More specifically, the display configurations were chosen in an attempt to evaluate the separate effects of the conflict probes and information level on pilot performance. Although all three configurations provided actionable Level 3 SA the methodology for presenting this information was manipulated in order to identify an optimal display configuration. In order to identify the effects of information level on operator performance the Display 1 configuration was designed to be compared against the Display 3 configuration. Both the Display 1 and Display 3 configurations provide operators with conflict probes in the plan view and the vertical profile view as well as "No-Go" heading and altitude bands. However, the level of information available to the operator was manipulated. Display 1 only provided operators with alerts regarding collision avoidance (CA) threats depicted in yellow and near mid-air collision (NMAC) threats depicted in red. In contrast, Display 3 provided operators with alerts regarding self-separation (SS) threats in yellow and collision avoidance (CA) threats in red. As depicted in Figure 10 and Table 1 the Display 3 configuration provides operators with an additional level of alert resulting in a larger conflict space being depicted on the display. Results of this comparison should help identify the advantage (or disadvantage) of displaying this additional level of self-separation information to the operator.

In order to separate out the effects of the conflict probes on operator performance the Display 2 configuration was compared against the Display 3 configuration. Information level was held consistent across the Display 2 and Display 3 configurations while the visual representation of this information was manipulated. Both Display 2 and Display 3 provided operators with alerts regarding self-separation (SS) threats in yellow and collision avoidance (CA) threats in red; however, Display 2 only provided operators with “No-Go” heading and altitude bands while Display 3 provided operators with conflict probes in the plan view and vertical profile view in addition to the “No-Go” heading and altitude bands. Results of this comparison should help identify the advantage (or disadvantage) of displaying conflict probes in the plan view and vertical profile view to the operator.

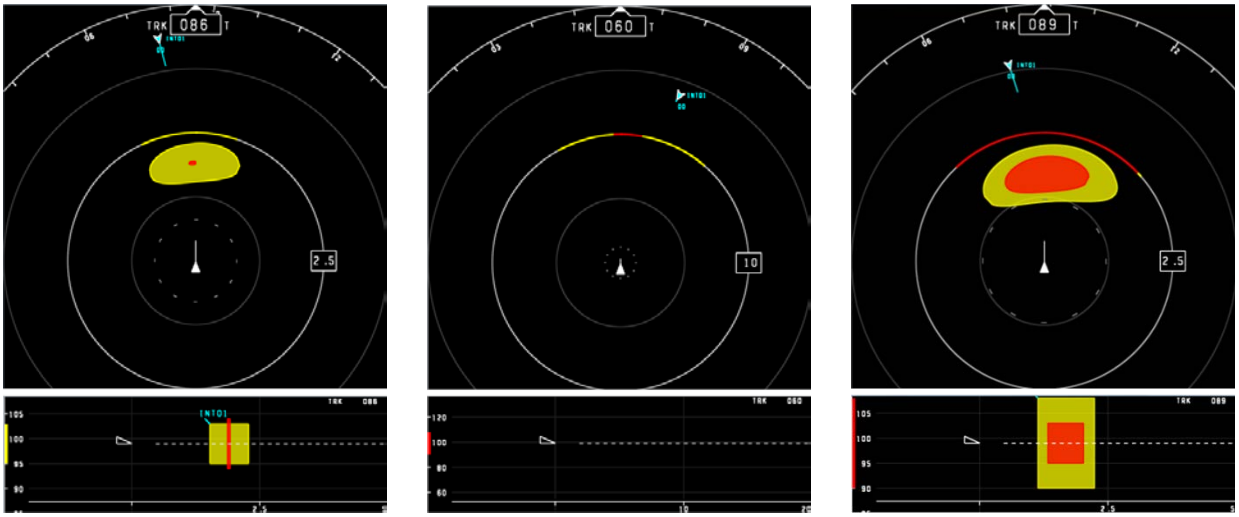


Figure 10. Display Configuration 1(L); 2(C); 3(R)

Table 1. Display Configuration Alerting Values

Alert	Distance	Altitude	Time
Self-Separation	< 1.2 NM	< 900 ft	< 110 seconds
Collision Avoidance	< 0.8 NM	< 400 ft	< 40 seconds
NMAC	< 500 ft	< 500 ft	N/A

5.4 Encounters

Air Traffic was developed to provide high fidelity traffic flow and encounters without excessive workload for the air traffic controller. Traffic load in the sector was equal to moderate traffic for the combined Oakland Air Route Traffic Control Sector 41/40. Three types of traffic were utilized in the scenarios (Bridges, 2014).

5.4.1 Background traffic

Visual Flight Rules (VFR) background traffic included aircraft squawking a beacon code of 1200 and simulated traffic at sector 40 during Visual Meteorological Conditions (VMC). The traffic flew random tracks over and through the Sonoma Valley between the Santa Rosa and

Napa airports. The majority of these aircraft flew at altitudes up to 9,500 feet.

Additional background traffic consisted of commercial traffic arriving and departing from the San Francisco Bay area. Arrival traffic flew the Golden Gate 6 Standard Arrival Route (STAR) to the San Francisco airport and the RAIDR3 STAR to the Oakland airport. San Jose airport arrivals flew the Point Reyes 1 STAR. Departure traffic flew into sector 41/40 via Skaggs Island (SGD) VHF Omni Directional Radio Range (VOR) enroute to the Red Bluff (RBL) VORTAC (navigational aid consisting of a co-located VOR beacon and a tactical air navigation system (TACAN) beacon). A few VFR aircraft were also added to the background traffic on random routes to increase the fidelity level (Bridges, 2014).

5.4.2 Encounter traffic

Previous and ongoing simulations have identified four critical encounter characteristics: the encounter angle (relative headings of the two aircraft: head-on, crossing, in-trail); the relative velocity (approximately equal or substantially different); the relative altitude and altitude rate (both level, one level and the other climbing/descending, both climbing/descending); and the predicted closest point of approach (a potential well clear violation only, or a collision avoidance encounter). These characteristics and the associated values are listed in Table 2. For a graphical depiction of the encounter angle definitions see Figure 11.

Table 2. Encounter Characteristics and Values

Encounter Characteristic	Values
Encounter Angle	Overtaking (0±45°), Crossing (90±45°), Reciprocal (180±45°)
Relative Velocity	Equal (difference ≤ 20 knots), Different (difference > 20 knots)
Relative Altitude Rate	Level vs. level, level vs. climb/descent, climb/descent vs. climb/descent
Closest Point of Approach	Potential well clear violation, collision avoidance

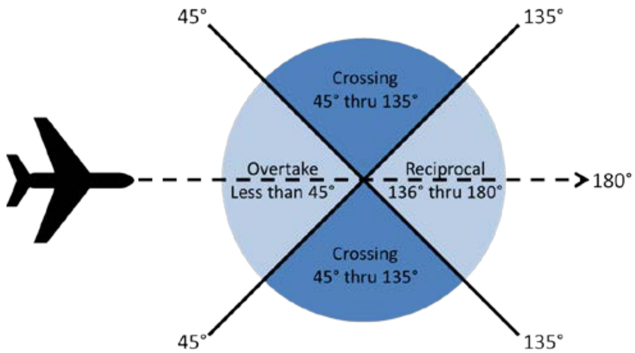


Figure 11. Encounter Characteristics

The encounters used in the scenarios are shown in Table 3 and represented an expert judgement of the most common encounter characteristics and those most important to evaluate for Detect-and-Avoid (DAA) tasks and pilot acceptability as determined by the Separation assurance/Sense-and-Avoid Interoperability (SSI) Traffic Advisory and Safety Alerting Threshold Simulation (TASATS) research team. Due to the dynamic nature of the scenarios the encounters may not have be identical in all scenarios but were only allowed to vary within the bounds identified for each variable (Bridges, 2014).

Table 3. Scenario Encounters

	Encounter Angle	Relative Velocity	Altitude Rate	Initial Alert Level
Encounter 1	Crossing	Different	D/L	Well Clear Violation
Encounter 2	Reciprocal	Different	C/L	Well Clear Violation
Encounter 3	Overtaking	Different	L/L	Well Clear Violation
Encounter 4	Crossing	Different	L/L	Well Clear Violation
Encounter 5	Crossing	Equal	L/L	Well Clear Violation
Encounter 6	Reciprocal	Equal	L/L	Well Clear Violation
Encounter 7	Crossing	Equal	C/L	Well Clear Violation
Encounter 8	Reciprocal	Equal	D/L	Well Clear Violation

5.5 Procedures

5.5.1 Pilot task

The operator's primary task was to control a simulated MQ-9 Reaper, call sign "HAWK21," along one of two predefined flight paths. Operators flew Instrument Flight Rules (IFR) and were responsible for maneuvering the aircraft in addition to responding to health and status tasks. Secondary tasks included responding to information requests (e.g., range and bearing to a specified waypoint) in a chat window and acknowledging system alerts. Operators were directed to monitor their traffic for potential safety conflicts. If a potential conflict was identified, they were to request permission to maneuver from air traffic control (ATC). Operators were instructed to minimize deviations from the predefined path and to coordinate with ATC to return to the predefined flight path as soon as practical.

5.5.2 Scenarios

Operators flew one of two predefined routes in each trial, a "Fire Line" mission and a "Coastal Watch" mission. In both scenarios the aircraft started at mission altitude, transiting towards the second waypoint of the predefined route. The Fire Line route was to be flown at 12,000 feet and the Coastal Watch route at 14,000 feet. In each scenario, eight encounters were scripted to progress to a self-separation then collision avoidance alert, without any pilot action. Although the encounters were designed to occur between ownship and a single intruder, dynamic changes to the surrounding traffic in addition to pilot maneuvers made it possible for multiple encounters to occur simultaneously.

5.5.3 Training

Operators completed an informed consent form provided by the NASA Ames Research Center (Appendix A) and a demographic questionnaire (Appendix B), regarding their experience in manned and unmanned aviation. Operators then underwent extensive training on the basic functionality of VSCS including practice on how to use the TSD's vehicle control interfaces and perform the secondary tasks. Operators then completed a 20-minute practice scenario. Operators received additional training on each display condition and completed a 20- minute practice scenario prior to each display condition.

5.5.4 Experimental trials

Operators completed three, 38-minute experimental trials. All operators received all three display configurations described above. The presentation of the display conditions was counterbalanced across operators to account for order and learning effects. Following each experimental trial, operators completed the NASA Task Load Index (TLX) and a post-trial subjective questionnaire, which focused on the unique display elements of the preceding condition. A post-simulation questionnaire and debrief followed the final experimental trial.

5.6 Objective Measures

5.6.1 Response times

Response times collected were designed to match the response time collected in Fern et al. (2015) in order to be able to directly compare the data collected. Table 4 presents the eight stages of the detect-and-avoid (DAA) interaction timeline. The time stamps for each stage were collected from a variety of data sources including raw VSCS output files, voice recordings and logs, and screen recordings of the pilot display.

Table 4. Stages of DAA Interaction Timeline

Stage	Description
T_0	Self-Separation or Collision Avoidance alert appears on display
T_1	Pilot notifies ATC of request to maneuver
T_2	ATC provides maneuver clearance
T_3	Pilot initiates an edit in VSCS to maneuver
T_{4a}	Pilot uploads first maneuver to aircraft
T_{4b}	Pilot uploads final maneuver to aircraft
T_5	Traffic alert is resolved
T_6	UAV completes maneuver

Five response time metrics were calculated based on the stages presented in Table 4. The metrics were selected based on the applicability to understanding the effect of display configuration on performance. The metrics were selected to focus on operator interaction with the DAA system, including the traffic displays and the control interfaces. Figure 12 depicts the relationship between metrics.

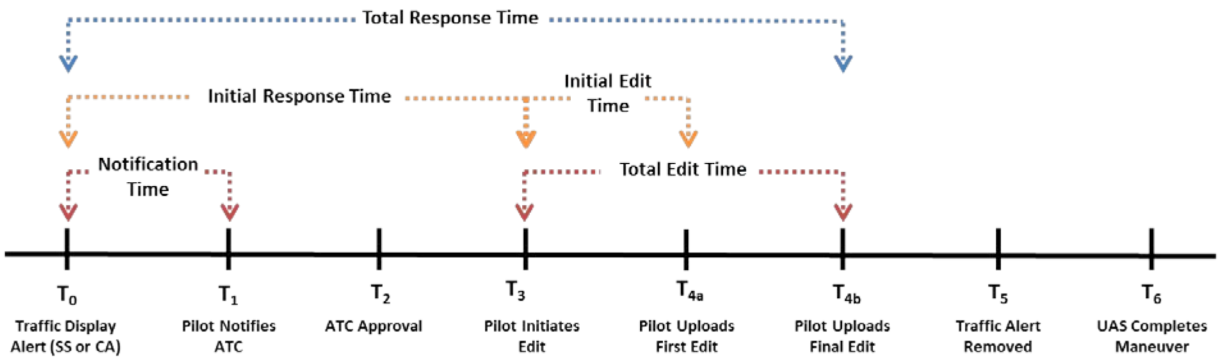


Figure 12. DAA Interaction Timeline

5.6.1.1 Notification time ($T_1 - T_0$)

Notification time was a measure of the time it took operators to notify ATC of an alert and the desire to execute a maneuver. This value represented the difference between the first appearance of the alert and the beginning of the operator's radio transmission to notify the controller.

5.6.1.2 Initial response time ($T_3 - T_0$)

Initial response time was a measure of the time it took operators to initiate a maneuver in the VSCS control interface in response to an alert. This value represented the difference between the first appearance of an alert and the start of an edit.

5.6.1.3 Initial Edit Time ($T_{4a} - T_3$)

Initial edit time was a measure of the time it took operators to input and upload an initial edit in VSCS. This value represented the time between initiating an edit and the first upload to the aircraft.

5.6.1.4 Total Edit Time ($T_{4b} - T_3$)

Total edit time was a measure of the time it took operators to complete all edits in VSCS in relation to a particular alert. This value represented the time between initiating an edit and the final upload to the aircraft. This metric was only relevant if an operator uploaded multiple edits to the aircraft. When a pilot made only one edit, total edit time was equivalent to initial edit time.

5.6.1.5 Total Response Time ($T_{4b} - T_0$)

Total response time was a measure of the time it took operators to upload a final maneuver in relation to a particular alert. This value represented the time between the initial appearance of the traffic alert and the final edit uploaded to the aircraft.

5.6.2 Secondary tasks

In addition to maintaining safe separation from other aircraft operators were asked to complete a simple and a complex secondary task. Both tasks were presented on the far right display. The simple secondary task required operators to use the mouse to acknowledge a number of alerts. Four types of alerts were presented, two warnings and two cautions indicating varying levels of importance. Each trial presented the operators with six warnings and six cautions. Each presentation was generated at a random time throughout the 38 minute trial in order to prevent operators from anticipating the alerts. Alerts were allowed to timeout after 25 seconds. Number of alerts responded to and response time were assessed.

The complex secondary task required operators to interact with the mouse and keyboard as well as the VSCS display for certain tasks. The complex task presented operators with a simulated chat window with several operators on the far right display. Throughout each 38 minute trial operators received a total of 12 chat messages at random intervals directing them to take a particular action. Required responses included changing radio frequencies, identifying remaining fuel, and providing current range and bearing to a unique point of interest on the VSCS display. In addition to the 12 messages requiring participant action chat messages directed at other chat room operators were interjected periodically through the trial. In addition to these distractor messages for other operators several distractor tasks were also directed at the participant but did not require any action. For example, “Hawk 21, good job staying on route, we're getting some good intel.” Number of chats responded to, response time and accuracy of response were assessed.

5.7 Subjective Measures

5.7.1 Post-trial

After each trial operators were provided a post-trial questionnaire (Appendices C-E) as well as the NASA-TLX (Appendix F). All post-trial questionnaires provided overlapping questions, such as "Rate your ability to maintain safe flight operations without conflicting with intruders" as well as questions unique to the particular trial, such as rating the degree to which unique features contributed most to their ability to appropriately assess and avoid conflicts. All post-trial questionnaires were designed to align with the subjective data collected in Monk et al. (2015).

Following each experimental trial, operators completed a post-trial questionnaire with subjective ratings pertaining to the preceding display configuration.

5.7.1.1 Training Sufficiency

Operators rated their agreement to the statement, "I felt I had enough training with this display to be able to operate it safely during this trial," on a 5-point Likert scale (1 - Strongly Disagree to 5 – Strongly Agree).

5.7.1.2 Initial Alert Response

Operators rated their agreement to the statements, "This display allowed me to respond immediately to self-separation threats," and "This display allowed me to respond immediately to collision avoidance threats," on a 5- point Likert scale (1 - Strongly Disagree to 5 – Strongly Agree).

5.7.1.3 Conflict Assessment and Avoidance

Operators also rated "the degree to which each of the following features contributed to your ability to most appropriately assess and avoid conflicts" on a 4-point scale (1 - Not at All to

4 - Quite a Bit): Plan view conflict probes, VPD conflict probes, plan view “No-Go” heading band, VPD “No-Go” altitude band, VPD, Intruder closest point of approach (CPA), and time to CPA.

5.7.1.4 Ease of Comprehension

Operators also rated their agreement to a number of statements stating that each of the aforementioned features were “easy to understand” on a 5-point Likert scale (1 - Strongly Disagree to 5 – Strongly Agree).

5.7.1.5 Clutter

Additionally, operators rated their agreement to the statement, “I did <u>not</u> find this display overly cluttered,” on a 5- point Likert scale (1 - Strongly Disagree to 5 – Strongly Agree).

5.7.1.6 Performance Degradation

Operators were also asked to rate, on a 5-point scale (1 - Unacceptable to 5 - Excellent), their ability to “maintain safe flight operations”, “minimize deviations from the planned path”, “handle all pilot tasks (including chat, radio changes, alerts, range/bearing)”, and acceptability of the alert timing.

5.7.2 Post-experiment

The post-study questionnaire was designed to allow operators to directly compare each display configuration against the others. In addition operators were asked to indicate their level of agreement with statements pertaining to display elements which were consistent across display configurations such as, “The intruder alerting (i.e. Probes & Bands) was easy to understand”. The post-block/study questionnaire was designed to align with the subjective data collected in Monk et al. (2015).

CHAPTER 6

RESULTS

6.1 Data Cleaning

There is little consensus among researchers regarding how to best handle outliers (Leys, Ley, Klein, Bernard & Licata, 2013). In a review of the *Journal of Personality and Social Psychology* and *Psychological Science* between 2010 and 2012 Leys et al. (2013) found that only about half use some rule of standard deviation (SD) while almost 11 percent use some rule of interquartile range (IQR). The authors argue there is a more robust method for identifying outliers which uses the median absolute deviation (MAD), however, this technique was not used in any of the articles they reviewed.

Prior to any analyses several methods of outlier identification were analyzed in order to try and identify the most appropriate method for the current data set. Outliers were identified using 1.5 times the IQR (Tukey, 1977), 2.2 times the IQR (Hoaglin & Iglewicz, 1987), 3 times the SD, and 2.5 times the MAD (Leys et al. 2013). Results of the analysis are included in Appendix G.

All four methods resulted in different classifications of outliers although there was quite a bit of overlap. According to Ratcliff (1993) the goal for models and empirical research should be to account for the middle 85 to 95 percent of observations in reaction time distributions since these are the data most likely to come from real processes under consideration and most likely to be critical in testing hypotheses and models. Due to the fact that MAD removed approximately 25 percent of the data in a particular condition it was not selected as a method for outlier identification in order to maintain a consistent method in all conditions. Reviewing the distributions of the data for the remaining three options 1.5 times the IQR was chosen due to the fact that is resulted in the closest approximation to a normal distribution in the greatest amount of experimental conditions (17 of 24).

After removing outliers the data for each condition was reviewed for each dependent measure. Histograms, probability plots, measures of variance as well as skewness and kurtosis were all reviewed as generated by SPSS. Results of the Kolmogorov-Smirnov and Shapiro-Wilk tests of normality were also reviewed. If these tests indicated a significant deviation from normality skewness and kurtosis were converted to *z*-scores for further analysis. Fields (2005) suggests in small samples *z*-score values with of skewness and kurtosis less than 2.58 should be considered for inclusion. Bootstrapping which provides a robust method for handling non-normal data was also employed during each analysis in order to calculate. While some suggest using non-parametric tests when violations of normality or homogeneity are encountered Fields (2010) suggests using bootstrapped confidence intervals instead which are robust to sources of biases. Bootstrapping was used to calculate bias corrected and accelerated confidence intervals (Efron & Tibshirani, 1993) which are reported with each test.

Initial review of response time data revealed there was insufficient data available to analyze average participant response times. Since the number of alerting events was not controlled, some operators provided more data points than others and not all operators contributed response times to all five metrics. For example, in the Display 1 condition participant 1 contributed four response times while participant 3 only contributed three response times; however in the Display 3 condition participant 1 contributed three response times while participant 3 contributed five response times. Since the primary purpose of the study was to investigate display differences all data points for response time were reviewed for inclusion in the final data set rather than including average response time per participant. For example, 37 data points were available for total response time in the Display 3 condition so all data points were reviewed for inclusion rather than only reviewing the average response time per participant which would set the maximum number of data points available to 10. Finally, all response times were compiled by display condition rather than by participant. For each comparison the sample sizes were equated by removing the furthest values from the median. Although statistical analyses would have only utilized what is common across conditions, resulting in the same amount of data points being included in the analyses, this would have resulted in the last values (most likely participant 9 or 10) always being dropped, so removing the furthest values from the median was deemed as a more optimal solution.

Response times to secondary tasks were highly skewed so several transformations were investigated. It is believed response time to secondary tasks was significantly more skewed than other response time metrics due to the nature of the task. Operators were directed to responds to the tasks as soon as possible. This direction created a condition similar to traditional *reaction time* tasks as opposed to the judgement decisions that were required during the alert conditions for which other response time metrics were collected. Kirk (2013) suggests several types of transformations for reaction time data in order. In order to select the most appropriate type of transformation for the data Kirk recommends applying each transformation type to the largest and smallest scores in each treatment level. Then determine the range for each treatment level and compute the ratio of the largest to the smallest range. The transformation that produces the smallest ration should be selected as the most appropriate. Transformation comparisons are provided in Appendices H.

Table 5 provides a summary of all the experimental questions as well as the notable findings identified during the analyses. A detailed list of the analyses conducted and the specific findings for each follow in the remainder of this section.

Table 5. Notable Findings

Experimental Questions	Notable Findings
1. Is there a performance difference between display presentation types?	• When using Display 2 operators took significantly longer to notify ATC of a request to maneuver than when using Display 3. • When using Display 1 operators took significantly longer to upload an initial edit the aircraft than when using Display 3. • When using Display 1 operators took significantly longer to upload all edits for a particular alert than when using Display 3. • When using Display 2 operators took significantly longer to upload all edits for a particular alert than when using Display 3.
2. Is there a performance difference on secondary tasks between display presentation types?	• No significant differences found.
3. Which display presentation type improves perceived accuracy in determining, negotiating and executing traffic avoidance maneuvers?	• Operators believed they were better able to appropriately respond to self-separation alerts when using Display 3 compared to Display 2. • Operators believed they were better able to appropriately respond to collision avoidance alerts when using Display 3 compared to Display 2. • Operators believed they were better able to handle all pilot tasks when using Display 3 compared to Display 2. • The Plan View Conflict Probes contributed significantly more to an operator's ability to appropriately respond to alerts than the Vertical Profile Display (VPD), the VPD "No-Go" Altitude Bands, and the Time to CPA.
4. Which display presentation type improves perceived efficiency in determining, negotiating and executing traffic avoidance maneuvers?	• Operators believed they were better able to immediately respond to alerts when using Display 3 compared to Display 2. • Operators believed Display 1 to be less cluttered than Display 3. • Operators believed Display 2 to be less cluttered than Display 3.
5. Is there a workload difference between display presentation types?	• No significant differences found.

Table 5 (continued)

Experimental Questions	Notable Findings
6. Is there a performance difference between standalone and integrated displays?	• Operators took significantly less time to initiate an edit in the VSCS control interface in the integrated Basic condition than in the standalone CPDS Display 1 and 2 conditions. • Operators took significantly less time to upload an initial edit to the aircraft in the integrated Advanced condition than in all three standalone CPDS display conditions. • Operators took significantly less time to upload all edits for a particular alert in the integrated Advanced condition compared to all three standalone CPDS display conditions. • Operators took significantly less time from the initial appearance of an alert to upload all edits for a particular alert in the integrated Basic condition than in the standalone CPDS Display 1 condition. • Operators took significantly less time from the initial appearance of an alert to upload all edits for a particular alert in the integrated Advanced condition than in all three standalone CPDS display conditions.
7. Is there a performance difference on secondary tasks between standalone and integrated displays?	• Operators were significantly faster at responding to complex secondary tasks in the integrated advanced condition than in all three standalone CPDS display conditions.
8. Is there a workload difference between standalone and integrated displays?	• Operators reported significantly less performance degradation in both integrated display conditions than in the standalone CPDS Display 2 condition.

6.2 CPDS Comparisons

6.2.1 Question 1

<u>Is there a performance difference between display presentation types?</u>

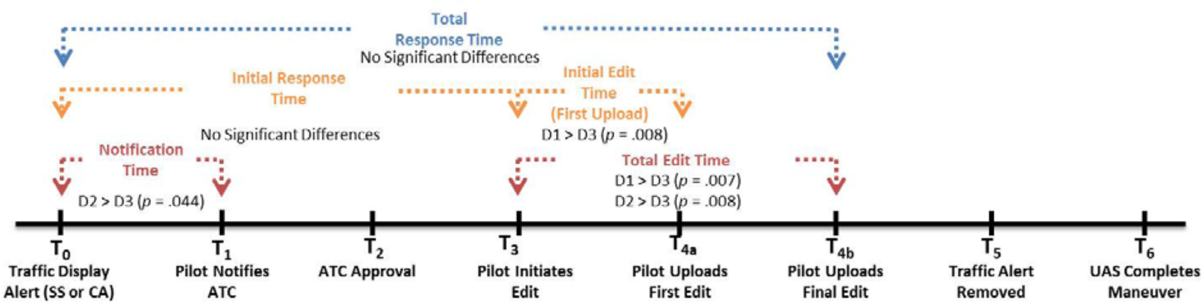


Figure 13. CPDS Display Configuration Response Time Comparisons

Notification time (T_1 – T_0). Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against Display 3. On average, when using Display 1 operators took longer to notify ATC of a request to maneuver than when using Display 3. The difference, 11.19 seconds, BCa 95% CI[-1.15, 23.94], was not significant $t(30) = 1.62$, $p = .116$, $d = .39$. On average, when using Display 2 operators took longer to notify ATC of a request to maneuver than when using Display 3. The difference, 9.55 seconds, BCa 95% CI[.57, 18.46], was significant $t(30) = 2.10$, $p = .044$, $d = .39$.

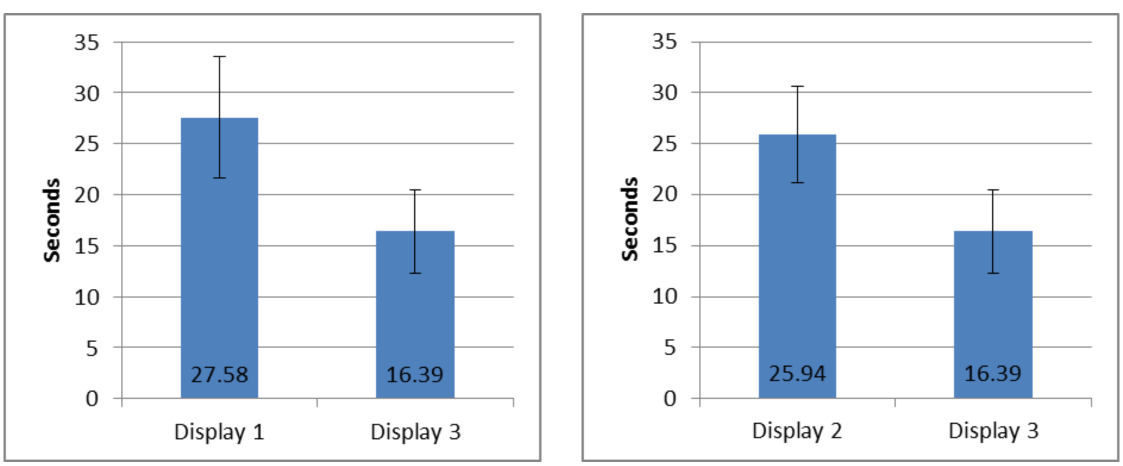


Figure 14. Notification Time Comparisons: D1 vs. D3 (L); D2 vs. D3 (R)

Initial response time (T_3 – T_0). Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against Display 3. On average, when using Display 1 operators took longer to initiate an edit in the VSCS control interface than when using Display 3. The difference, 7.91 seconds, BCa 95% CI[-4.61, 19.06], was not significant $t(32) = 1.22$, $p = .233$, $d = .30$. On average, when using Display 2 operators took longer to initiate an edit in the VSCS control interface than when using Display 3. The difference, 9.33 seconds, BCa 95% CI[-2.29, 21.37], was not significant $t(32) = 1.54$, $p = .134$, $d = .38$.

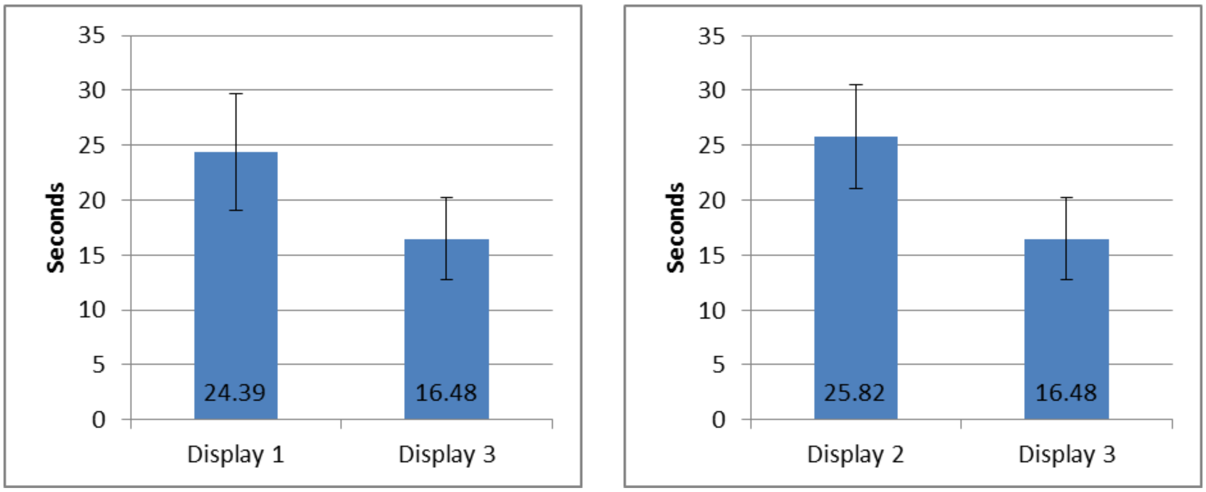


Figure 15. Initial Response Time Comparisons: D1 vs. D3 (L); D2 vs. D3 (R)

Initial Edit Time (T_{4a} – T_3). Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against Display 3. On average, when using Display 1 operators took longer to upload an initial edit to the aircraft than when using Display 3. The difference, 6.67 seconds, BCa 95% CI[2.12, 11.34], was significant $t(32) = 2.83$, $p = .008$, $d = .75$. On average, when using Display 2 operators took longer to upload an initial edit to the aircraft than when using Display 3. The difference, 1.09 seconds, BCa 95% CI[-1.21, 3.39], was not significant $t(32) = .858$, $p = .397$, $d = .20$.

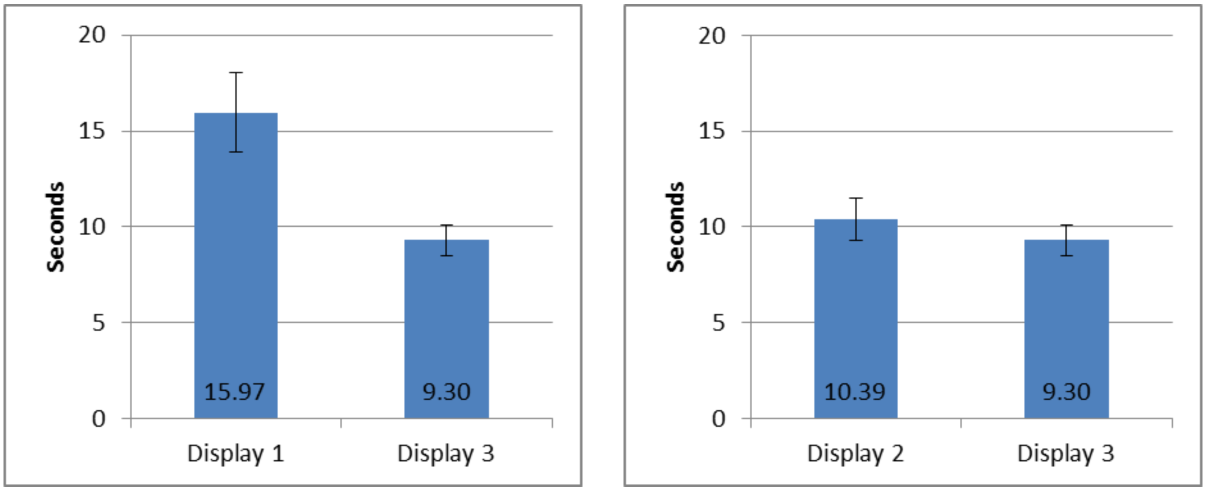


Figure 16. Initial Edit Time Comparisons: D1 vs. D3 (L); D2 vs. D3 (R)

Total Edit Time (T_{4b} – T_3). Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against Display 3. On average, when using Display 1 operators took longer to upload all edits for a particular alert than when using Display 3. The difference, 10.76 seconds, BCa 95% CI[3.32, 18.52], was significant $t(33) = 2.85$, $p = .007$, $d = .68$. On average, when using Display 2 operators took longer to upload all edits for a particular alert than when using Display 3. The difference, 17.32 seconds, BCa 95% CI[6.85, 27.40], was significant $t(33) = 2.82$, $p = .008$, $d = .72$.

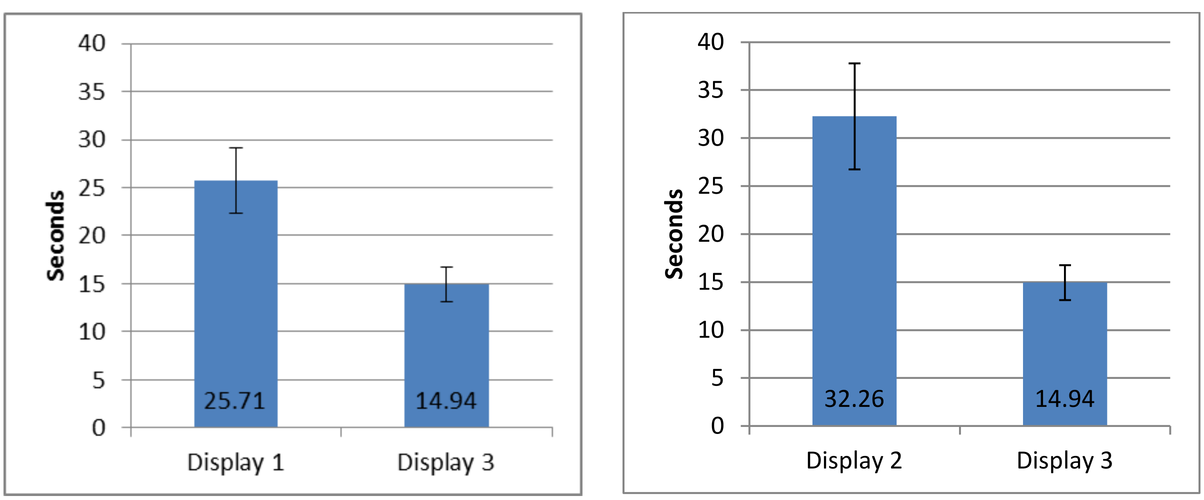


Figure 17. Total Edit Time Comparisons: D1 vs. D3 (L); D2 vs. D3 (R)

Total Response Time (T_{4b} – T_0). Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against Display 3. On average, when using Display 1 operators took longer from the initial appearance of an alert to upload all edits for a particular alert than when using Display 3. The difference, 12.45 seconds, BCa 95% CI[-1.41, 26.48], was not significant $t(32) = 1.77$, $p = .086$, $d = .46$. On average, when using Display 2 operators took longer from the initial appearance of an alert to upload all edits for a particular alert than when using Display 3. The difference, 6.00 seconds, BCa 95% CI[-7.80, 19.75], was not significant $t(32) = .827$, $p = .415$, $d = .21$.

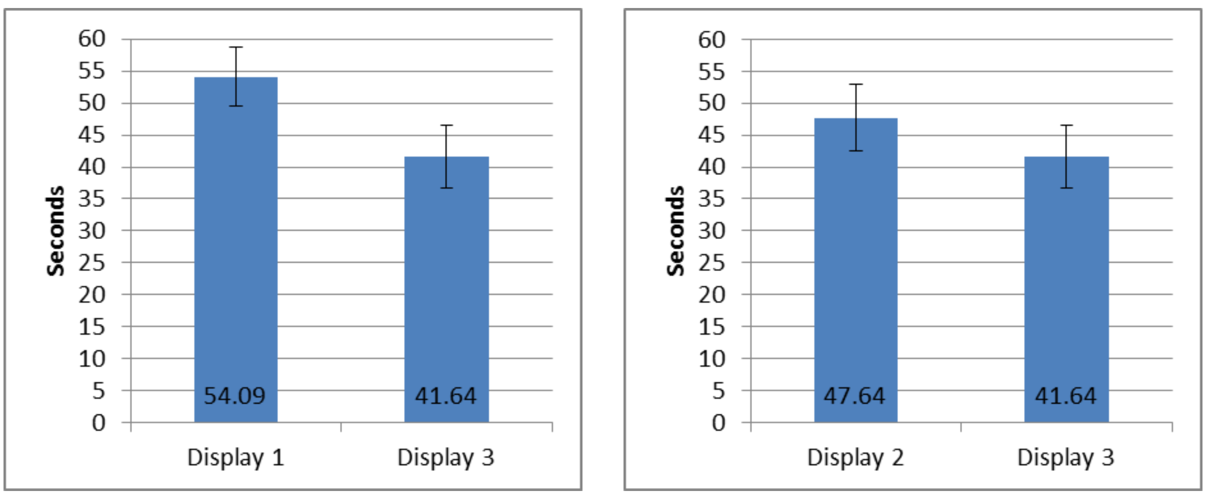


Figure 18. Total Response Time Comparisons: D1 vs. D3 (L); D2 vs. D3 (R)

Results from the analyses of primary task performance across the three display configurations did not support the proposed hypothesis that operators would demonstrate the best performance while using Display 1.

6.2.2 Question 2

<u>Is there a performance difference on secondary tasks between display presentation types?</u>

Simple Secondary Tasks. A total to 12 simple secondary tasks were presented per trial. Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against

Display 3. On average, when using Display 1 operators responded to as many simple secondary tasks as when using Display 3. The difference, .20 hits, BCa 95% CI[-2.20, 1.70], was not significant $t(9) = -.192$, $p = .852$, $d = .05$. On average, when using Display 2 operators responded to about one less simple secondary tasks than when using Display 3. The difference, .90 hits, BCa 95% CI[-2.5, .80], was not significant $t(9) = -1.01$, $p = .337$, $d = .23$.

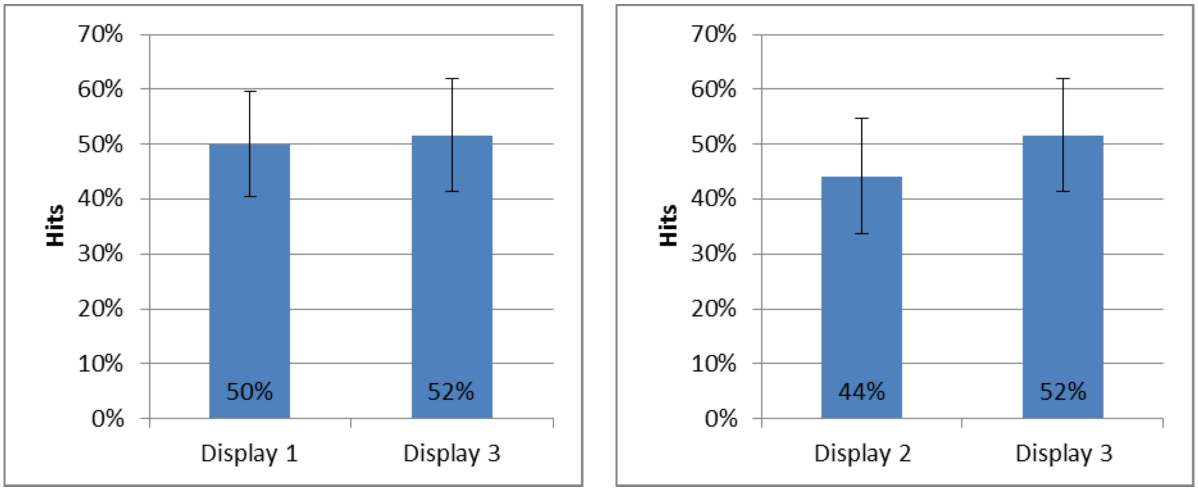


Figure 19. Simple Secondary Task % Hit Comparisons: D1 vs. D3 (L); D2 vs. D3 (R)

Based on Kirk (2013) a square root transformation was chosen for analyzing response time to simple secondary tasks between display conditions (See Appendix H for transformation comparisons). Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against Display 3. On average, when using Display 1 operators responded faster to simple secondary tasks than when using Display 3. The difference, 1.62 seconds was not significant once the data was transformed, BCa 95% CI[-.68, .17], $t(52) = -1.12$, $p = .270$, $d = .22$. On average, when using Display 2 operators responded slightly slower to simple secondary tasks than when using Display 3. The difference, .67 seconds was not significant once the data was transformed, BCa 95% CI[-.39, .51], was not significant $t(52) = .321$, $p = .749$, $d = .06$.

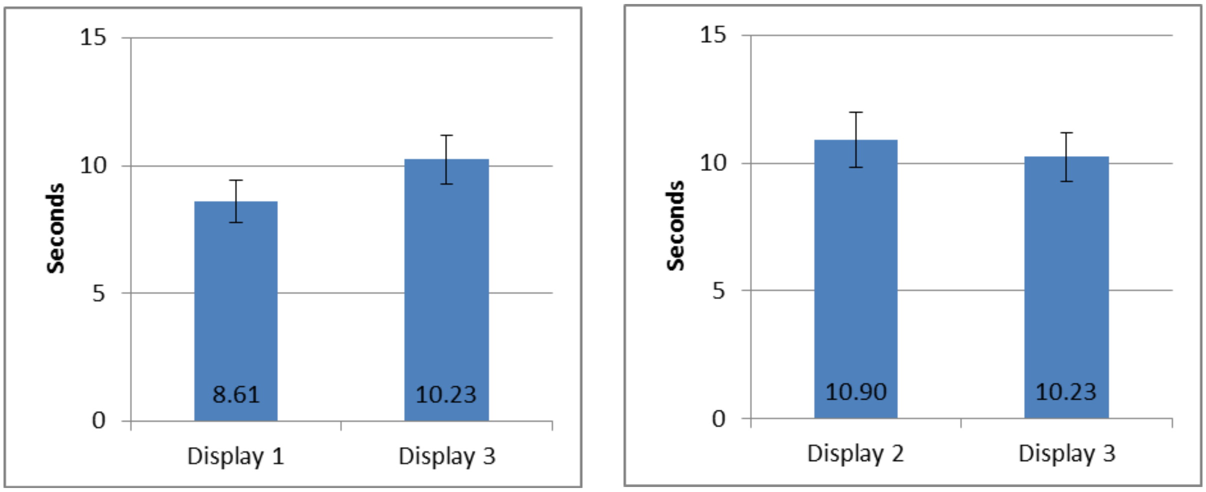
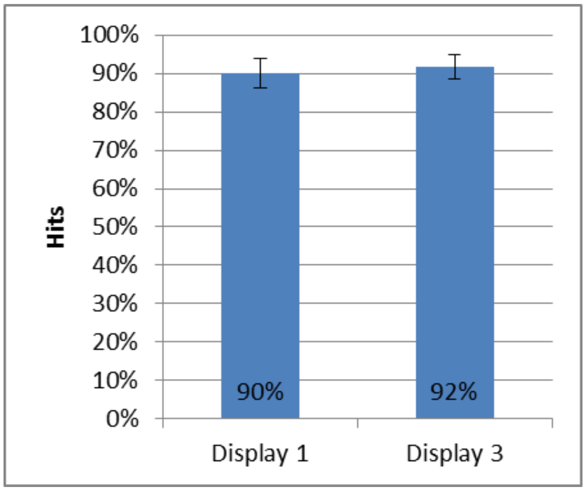


Figure 20. Simple Secondary Task Response Time Comparisons: D1 vs. D3 (L); D2 vs. D3 (R)

Complex Secondary Tasks. A total to 12 complex secondary tasks were presented per trial. Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against Display 3. On average, when using Display 1 operators responded to as many simple secondary tasks as when using Display 3. The difference, .20 hits, BCa 95% CI[-.90, .60], was not significant $t(9) = -.514$, $p = .619$, $d = .15$. On average, when using Display 2 operators responded to as many simple secondary tasks as when using Display 3. The difference, .40 hits, BCa 95% CI[-.10, .90], was not significant $t(9) = 1.50$, $p = .168$, $d = .40$. Accuracy of responses was not analyzed as all operators responded accurately to all tasks. This is not surprising given the tasks were very directed questions such as, fuel remaining or changing a radio to specified channel.



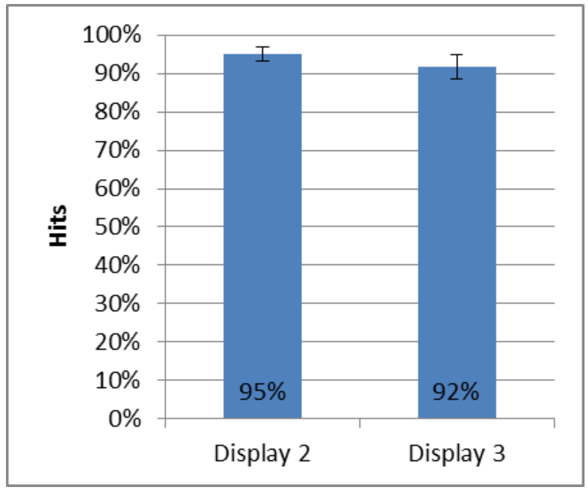


Figure 21. Complex Secondary Task % Hit Comparisons: D1 vs. D3 (L); D2 vs. D3 (R)

Based on Kirk (2013) a Log(Y+1) transformation was chosen for analyzing response time to complex secondary tasks between display conditions (See Appendix H for transformation comparisons). Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against Display 3. On average, when using Display 1 operators responded faster to complex secondary tasks than when using Display 3. The difference, 3.26 seconds was not significant once the data was transformed, BCa 95% CI[-.13, .07], $t(99) = -.519$, $p = .605$, $d = .10$. On average, when using Display 2 operators responded slightly slower to complex secondary tasks than when using Display 3. The difference, 1.02 seconds was not significant once the data was transformed, BCa 95% CI[-.09, .09], was not significant $t(99) = -.106$, $p = .916$, $d = .03$.

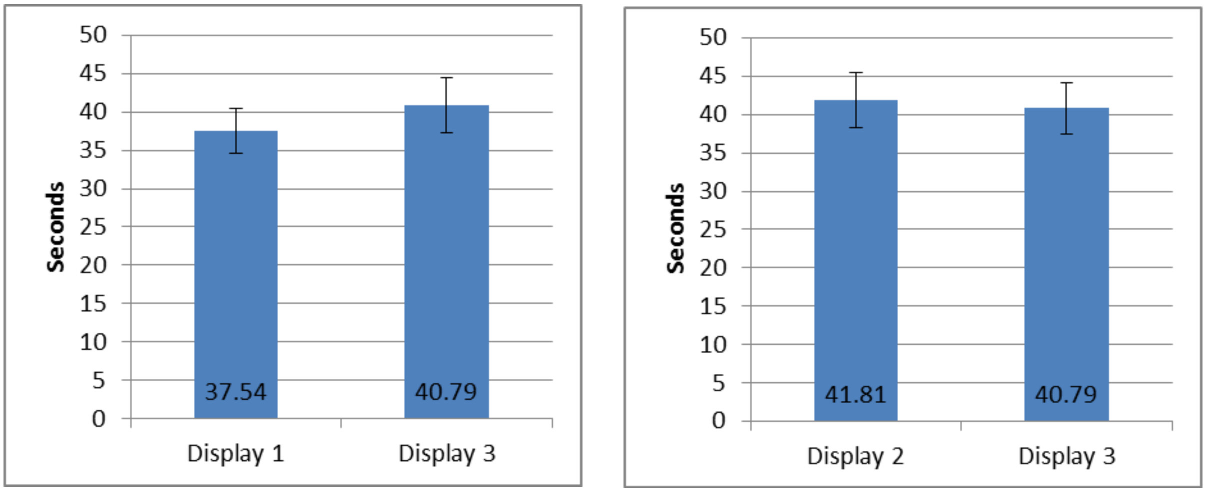


Figure 22. Complex Secondary Task Response Time Comparisons: D1 vs. D3 (L); D2 vs. D3 (R)

Results from the analyses of secondary task performance across the three display configurations did not support the proposed hypothesis that operators would suffer the largest performance decrement in secondary task performance while using Display 2.

6.2.3 Question 3

<u>Which display presentation type improves perceived accuracy in determining, negotiating and executing traffic avoidance maneuvers?</u>

Post-trial questionnaire data was analyzed in order to assess perceived differences in accuracy between the three display presentation types. After each trial operators were asked to respond on a 5-point Likert scale (1 - Strongly Disagree to 5 – Strongly Agree) to the statements:

- This display allowed me to respond <u>appropriately</u> to <u>self-separation</u> threats.
- This display allowed me to respond <u>appropriately</u> to <u>collision avoidance</u> threats.

Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against Display 3. On average, when using Display 1 operators agreed to the same extent with the displays ability to allow them to respond to self-separation (SS) alerts appropriately as when using Display 3. The was actually no difference between the mean ratings so significant difference was found, BCa 95% CI[-2.00, 2.00], $t(9) < .000$, $p = 1.00$, $d = .00$. On average, when using Display 2 operators agreed slightly less with the displays ability to allow them to respond to SS alerts appropriately than when using Display 3. The difference, .80, BCa 95% CI[-1.40, -.23], was significant $t(9) = -2.45$, $p = .037$, $d = .95$.

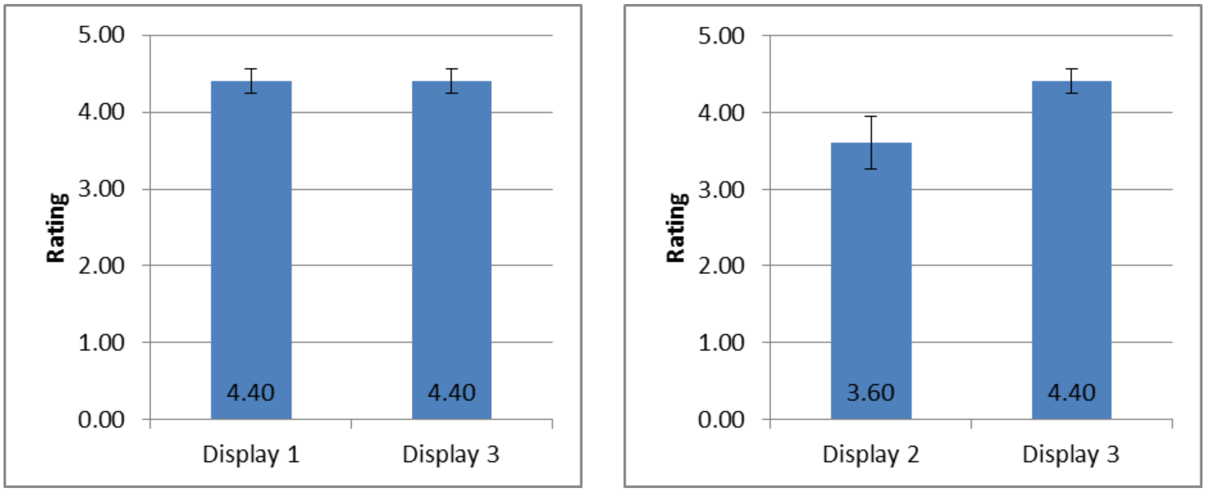


Figure 23. Appropriateness of Self-Separation Responses: D1 vs. D3 (L); D2 vs. D3 (R)

On average, when using Display 1 operators agreed to the same extent with the displays ability to allow them to respond to collision avoidance (CA) alerts appropriately as when using Display 3. There was only a .1 difference in ratings between the display conditions. No significant difference was found, BCa 95% CI[-.60, .40], $t(9) = -.318$, $p = .758$, $d = .12$. On average, when using Display 2 operators agreed less with the displays ability to allow them to respond to CA alerts appropriately than when using Display 3. The difference, 1.0, BCa 95% CI[-1.80, -.20], was significant $t(9) = -2.54$, $p = .032$, $d = .99$.

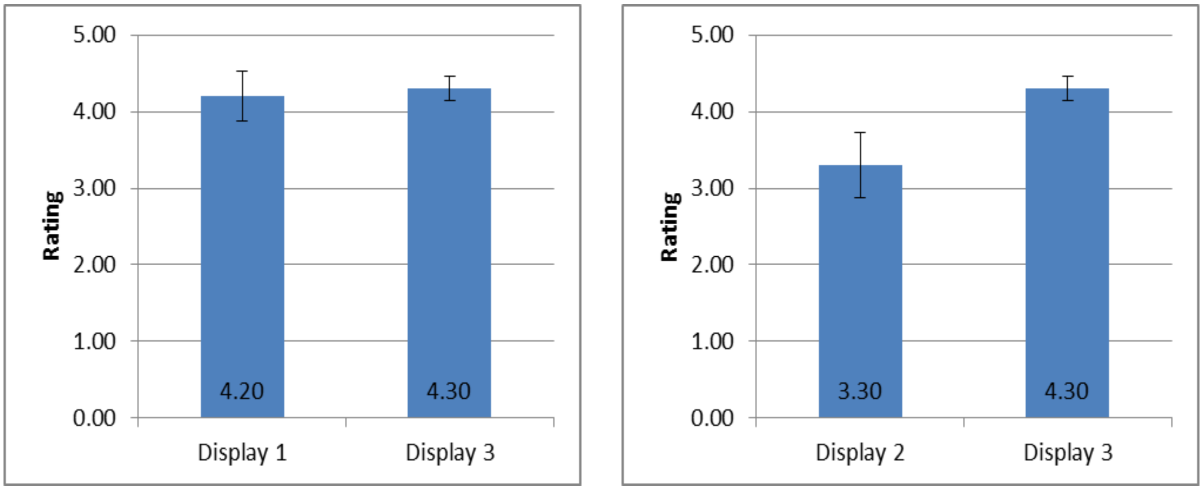


Figure 24. Appropriateness of Collision Avoidance Responses: D1 vs. D3 (L); D2 vs. D3 (R)

In order to further assess differences between display presentation types on operator's perceived accuracy in determining, negotiating and executing traffic avoidance maneuvers, operators were asked after each trial to respond on a 5-point Likert scale (1 - Unacceptable to 5 – Excellent) to the following statements:

- Rate your ability to maintain safe flight operations **without conflicting with intruders.**
- Rate your ability to **minimize deviations from the planned path.**
- Rate your ability to **handle all pilot tasks** (including chat, radio changes, alerts, range/bearing).

Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against Display 3.A significant effect was identified on only one of the three questions. On average, when using Display 2 operators agreed slightly less with the displays ability to allow them to handle all pilot tasks than when using Display 3. The difference, .7, BCa 95% CI[-1.10, -.40], was significant $t(9) = -2.69$, $p = .025$, $d = .69$.

Results from the analyses of perceived accuracy across the three display configurations did not support the proposed hypothesis that operators would report the greatest degree of accuracy in the Display 3 condition.

Finally, in order to try to assess which display feature contributed most to the operator's ability to appropriately respond to alerts, regardless of display condition, operators were asked to rate the degree to which each of the following display features contributed to their ability to most appropriately assess and avoid conflicts (1 - Not at All to 4 - Quite a Bit).

- Plan View Conflict Probes
- Vertical Profile Display Probes
- Plan View "No-Go" Heading Bands
- Vertical Profile Display "No-Go" Altitude Bands
- Vertical Profile Display
- Distance to Closest-Point-of-Approach (CPA)
- Time to Closest-Point-of-Approach (CPA)

A Friedman test was used to analyze the data. The initial analysis identified a significant effect between display features, $\chi^2(2) = 31.022$, $p < .001$. Follow-up pairwise comparisons were conducted using a Dunn-Bonferroni test. After adjusting for multiple comparisons the Plan View Conflict Probes ($M = 3.60$; $SE = .112$) were identified as contributing significantly more than the Vertical Profile Display ($M = 2.90$; $SE = .194$) ($p = .031$), the Time to CPA ($M = 2.47$; $SE = .178$) ($p = .009$) and the Vertical Profile Display "No-Go" Altitude Bands ($M = 2.70$; $SE = .160$) ($p = .003$). No other significant effects were identified.

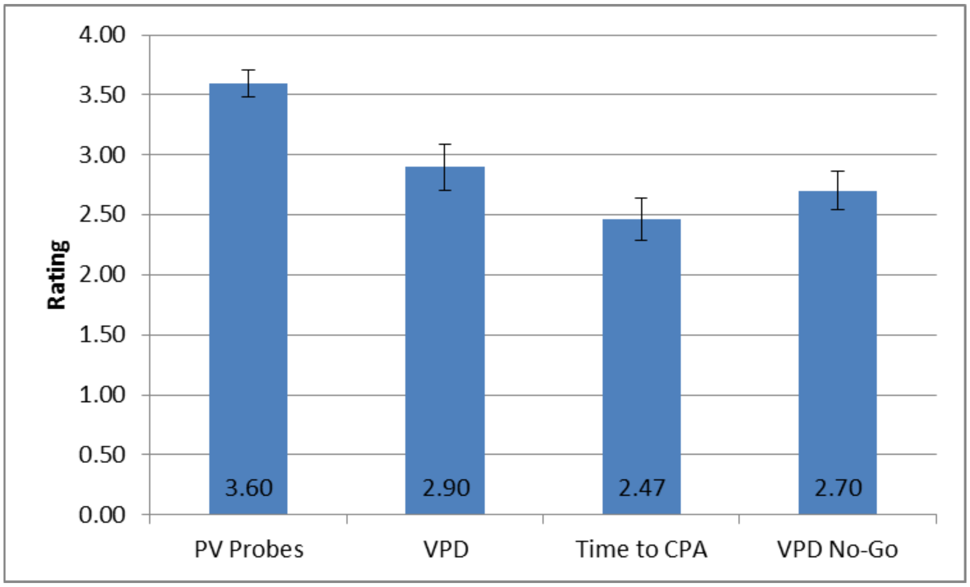


Figure 25. Utility of Display Features

6.2.4 Question 4

Which display presentation type improves perceived efficiency in determining, negotiating and executing traffic avoidance maneuvers?

Post-trial questionnaire data was analyzed in order to assess perceived differences in efficiency between the three display presentation types. After each trial operators were asked to respond on a 5-point Likert scale (1 - Strongly Disagree to 5 – Strongly Agree) to the statements:

- This display allowed me to respond immediately to self-separation threats.
- This display allowed me to respond immediately to collision avoidance threats.

Paired-samples *t* tests were used to compare Display 1 against Display 3 and Display 2 against Display 3. On average, when using Display 1 operators agreed to about the same extent with the displays ability to allow them to respond to self-separation (SS) alerts immediately as when using Display 3. The difference, .20, BCa 95% CI[-.20, .60], was not significant, $t(9) = 1.00$, $p = .343$, $d = .40$. On average, when using Display 2 operators agreed slightly less with the displays ability to allow them to respond to SS alerts immediately than when using Display 3. The difference, .50, BCa 95% CI[-.90, -.20], was not significant *t*(9) = -1.86, *p* = .096, *d* = .68.

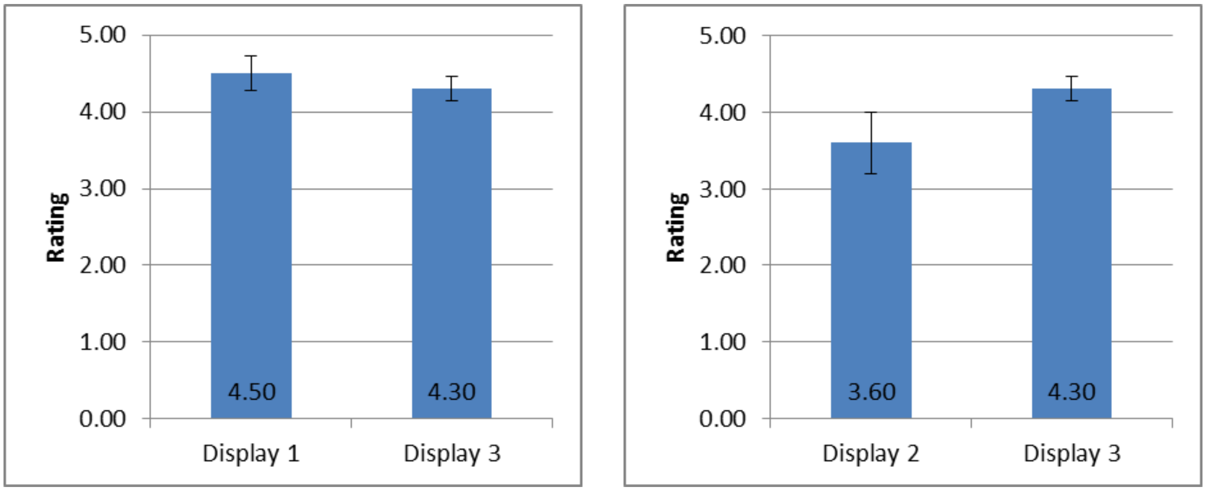


Figure 26. Immediacy of Self-Separation Responses: D1 vs. D3 (L); D2 vs. D3 (R)

On average, when using Display 1 operators agreed to about the same extent with the displays ability to allow them to respond to collision avoidance (CA) alerts immediately as when using Display 3. The difference, .20, BCa 95% CI[-.20, .60], was not significant, *t*(9) = 1.00, *p* = .343, *d* = .33. On average, when using Display 2 operators agreed slightly less with the displays ability to allow them to respond to CA alerts immediately than when using Display 3. The difference, .70, BCa 95% CI[-1.30, -.10], was not significant *t*(9) = -1.77, *p* = .111, *d* = .73.

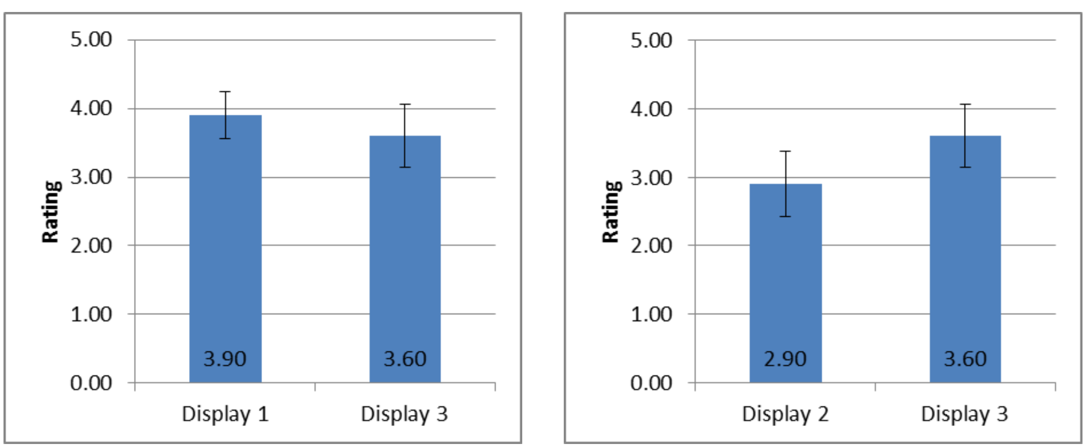


Figure 27. Immediacy of Collision Avoidance Responses: D1 vs. D3 (L); D2 vs. D3 (R)

A review of the data indicated a trend between the two questions. In both questions, operators reported stronger agreement to the statements after the Display 1 (SS: *M* = 4.50; *SE* = .167; CA: *M* = 4.50; *SE* = .224) trial followed by Display 3 (SS: *M* = 4.30; *SE* = .153; CA: *M* = 4.30; *SE* = .153) and Display 2 (SS: *M* = 3.80; *SE* = .291; CA: *M* = 3.60; *SE* = .400). Since both questions addressed the immediacy of responses to alerts the response data was consolidated and follow-up comparisons were conducted. On average, when using Display 1 operators agreed to about the same extent with the displays ability to allow them to respond to immediately to alerts as when using Display 3. The difference, .20, BCa 95% CI[0.00, .40], was not significant, $t(9) = 1.45$, $p = .163$, $d = .37$. On average, when using Display 2 operators agreed slightly less with the displays ability to allow them to respond to immediately to alerts than when using Display 3. The difference, .60, BCa 95% CI[-1.10, -.15], was significant, $t(9) = -2.57$, $p = .019$, $d = .72$.

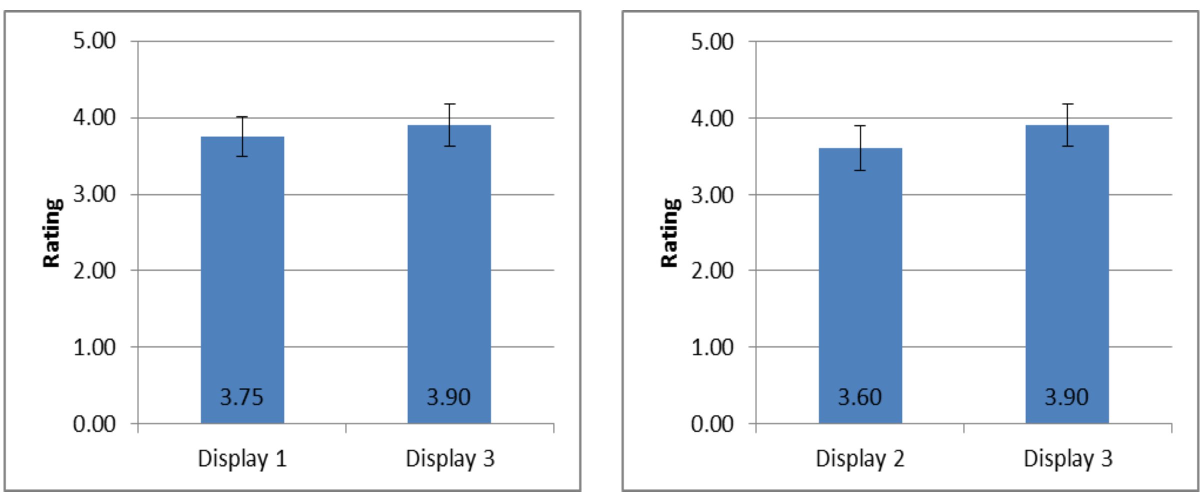


Figure 28. Immediacy of Responses: D1 vs. D3 (L); D2 vs. D3 (R)

In order to further assess perceived efficiency differences between display presentation types operators were also asked to respond on a 5-point Likert scale (1 - Strongly Disagree to 5 – Strongly Agree) to the following statement:

- I did not find this display overly cluttered.

On average, when using Display 1 operators more strongly agreed with the statement, “I did not find this display overly cluttered” than when using Display 3. The difference, .90, BCa 95% CI[.30, 1.5], was significant, $t(9) = 2.86$, $p = .019$, $d = .89$. On average, when using Display 2 operators also more strongly agreed with the statement than when using Display 3. The difference, 1.0, BCa 95% CI[.50, 1.5], was significant, $t(9) = -2.74$, $p = .023$, $d = .98$.

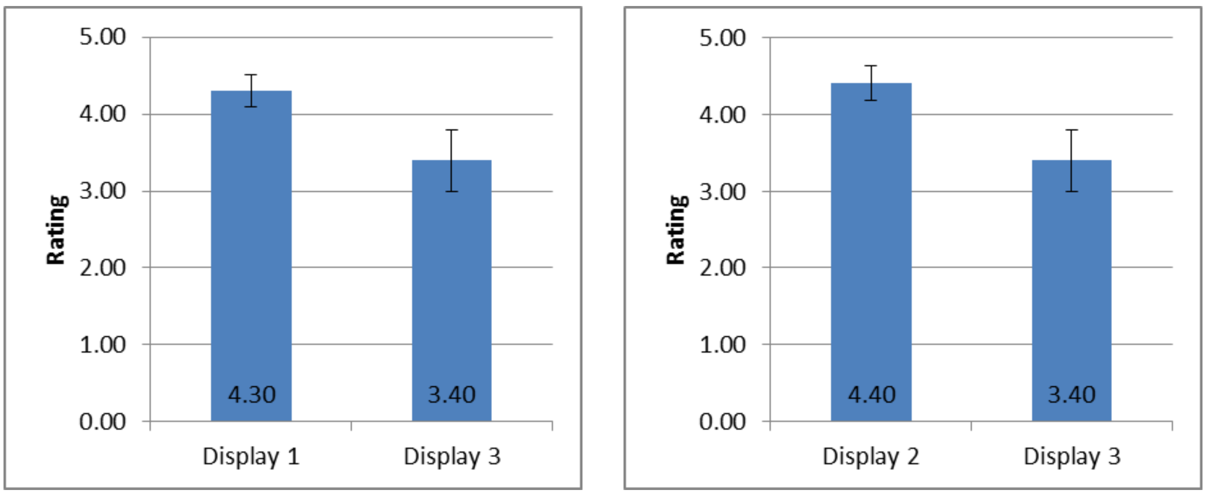


Figure 29. Clutter Ratings: D1 vs. D3 (L); D2 vs. D3 (R)

Results from the analyses of perceived efficiency across the three display configurations did not support the proposed hypothesis that operators would report the greatest degree of efficiency in the Display 2 condition.

6.2.5 Question 5

<u>Is there a workload difference between display presentation types?</u>

Each of the six NASA-TLX scales were analyzed using the raw data on a 1 to 7 scale for differences between Display 1 and Display 3 and between Display 2 and Display 3.

Mental. When using Display 1 operators rated the task slightly less mentally demanding than when using Display 3. The difference, .30, BCa 95% CI[-1.50, .80], was not significant, $t(9)$

= -.461, p = .656, d = .22. On average, when using Display 2 operators rated the task slightly more mentally demanding than when using Display 3. The difference, .20, BCa 95% CI[-.40, .80], was not significant, $t(9)$ = .557, p = .591, d = .14.

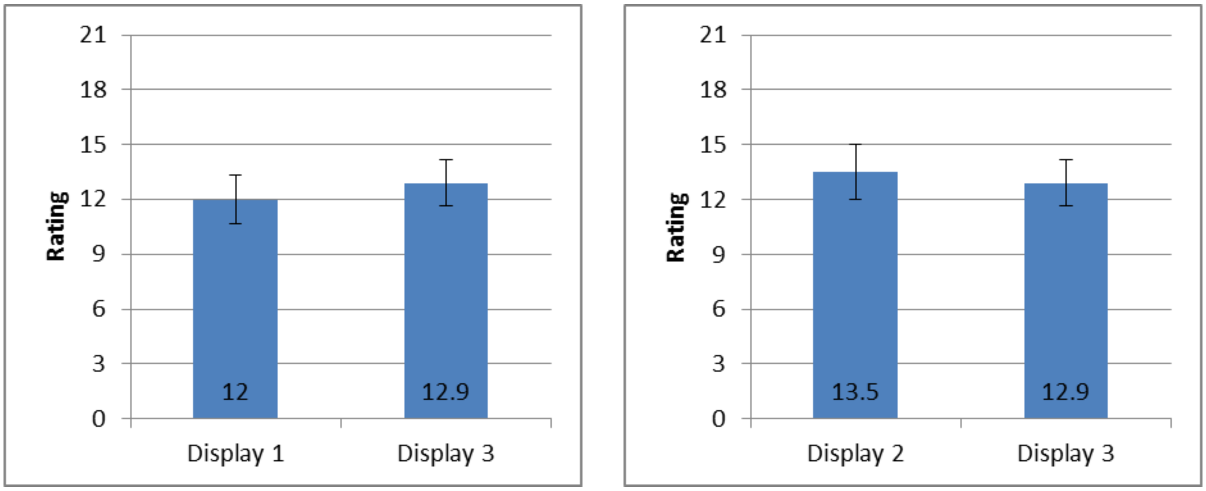


Figure 30. NASA-TLX – Mental Workload Ratings: D1 vs. D3 (L); D2 vs. D3 (R)

Physical. When using Display 1 operators rated the task slightly less physically demanding than when using Display 3. The difference, .50, BCa 95% CI[-1.00, .20], was not significant, $t(9)$ = -1.46, p = .177, d = .32. On average, when using Display 2 operators rated the task slightly more physically demanding than when using Display 3. The difference, .10, BCa 95% CI[-.20, .0], was not significant, $t(9)$ = .429, p = .678, d = .05.

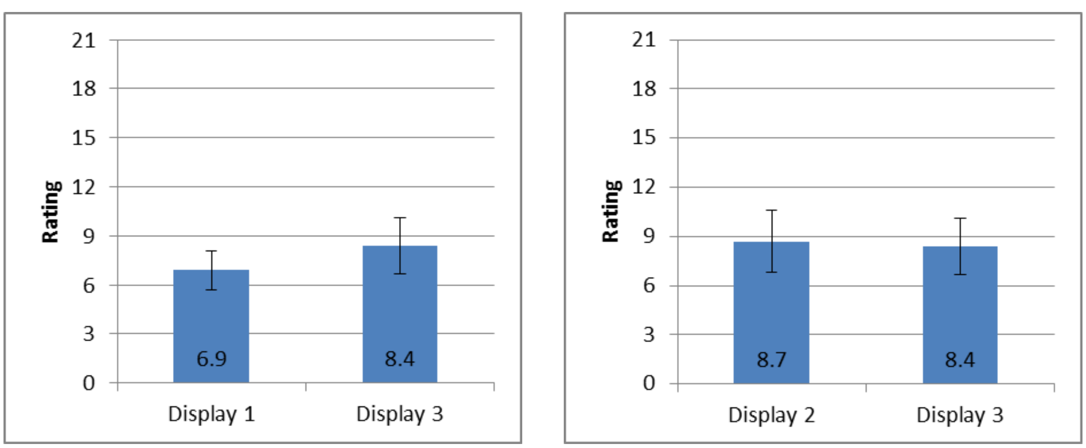


Figure 31. NASA-TLX – Physical Workload Ratings: D1 vs. D3 (L); D2 vs. D3 (R)

Temporal. When using Display 1 operators rated the task slightly more temporally demanding than when using Display 3. The difference, .10, BCa 95% CI[-.70, 1.00], was not significant, $t(9) = .183$, $p = .859$, $d = .08$. On average, when using Display 2 operators rated the task more temporally demanding than when using Display 3. The difference, .80, BCa 95% CI[.00, 1.60], was not significant, $t(9) = 1.809$, $p = .104$, $d = .48$.

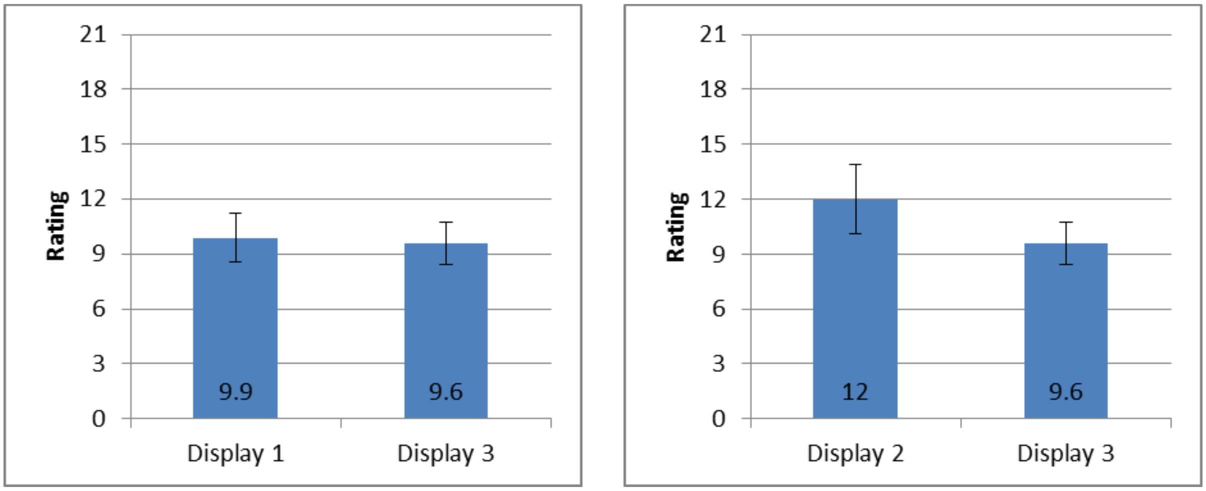


Figure 32. NASA-TLX – Temporal Workload Ratings: D1 vs. D3 (L); D2 vs. D3 (R)

Performance. When using Display 1 operators reported lower levels of performance degradation than when using Display 3. The difference, .60, BCa 95% CI[-1.00, .20], was not significant, $t(9) = -2.25$, $p = .051$, $d = .52$. On average, when using Display 2 operators reported higher levels of performance degradation than when using Display 3. The difference, .60, BCa 95% CI[.10, 1.10], was not significant, $t(9) = 1.765$, $p = .111$, $d = .39$.

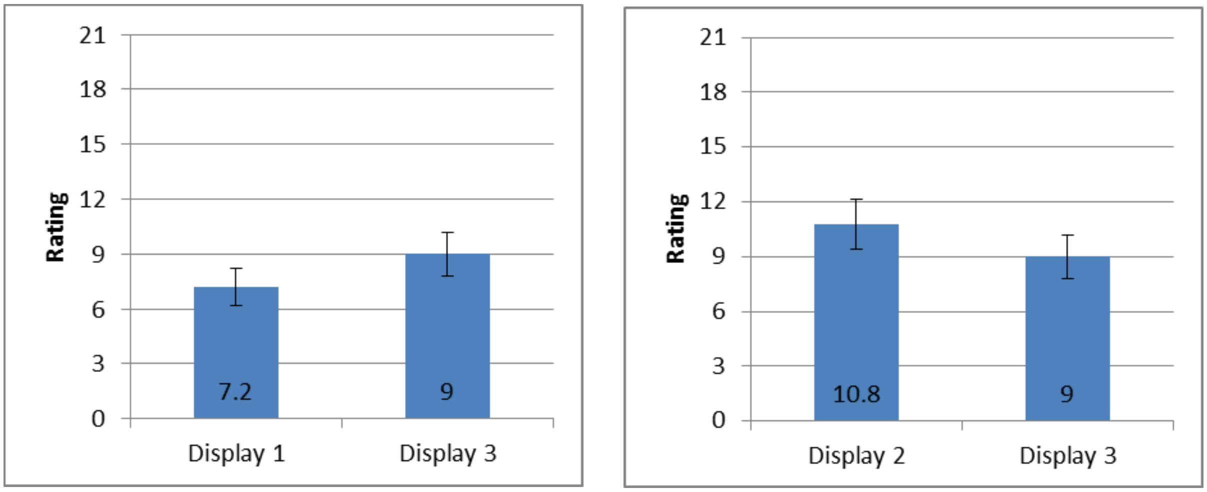


Figure 33. NASA-TLX – Performance Degradation Ratings: D1 vs. D3 (L); D2 vs. D3 (R)

Effort. When using Display 1 operators rated the task slightly less effortful than when using Display 3. The difference, .30, BCa 95% CI[-1.00, .20], was not significant, $t(9) = -.818$, $p = .434$, $d = .28$. When using Display 2 operators rated the task more effortful than when using Display 3. The difference, .50, BCa 95% CI[-.10, 1.20], was not significant, $t(9) = 1.246$, $p = .244$, $d = .45$.

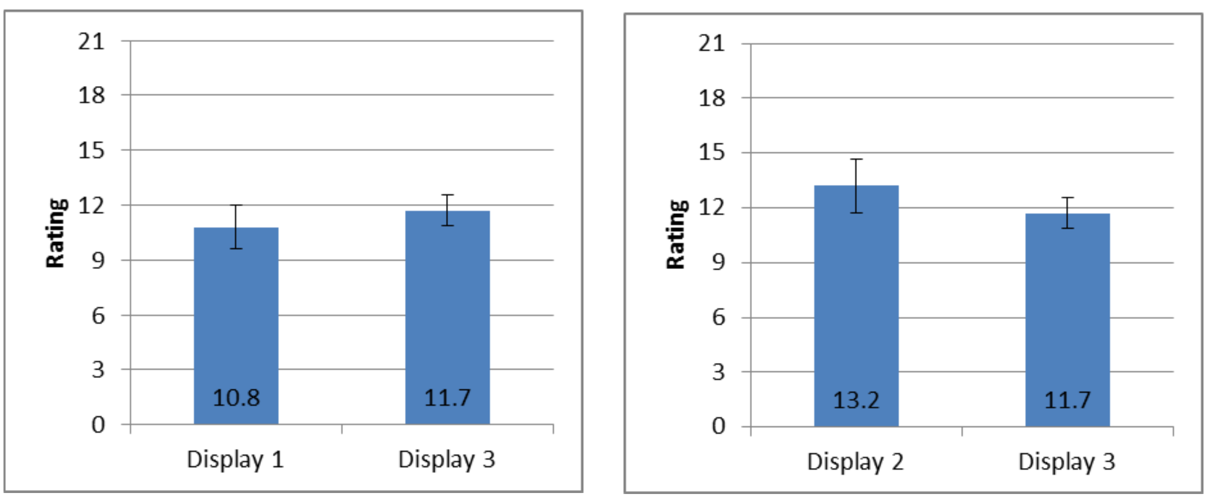


Figure 34. NASA-TLX – Effort Workload Ratings: D1 vs. D3 (L); D2 vs. D3 (R)

Frustration. When using Display 1 operators reported lower levels of frustration than when using Display 3. The difference, .50, BCa 95% CI[-2.00, 1.10], was not significant, $t(9) = -$

.808, *p* = .440, *d* = .38. When using Display 2 operators higher levels of frustration than when using Display 3. The difference, .60, BCa 95% CI[-.50, 1.40], was not significant, *t*(9) = 1.260, *p* = .239, *d* = .36.

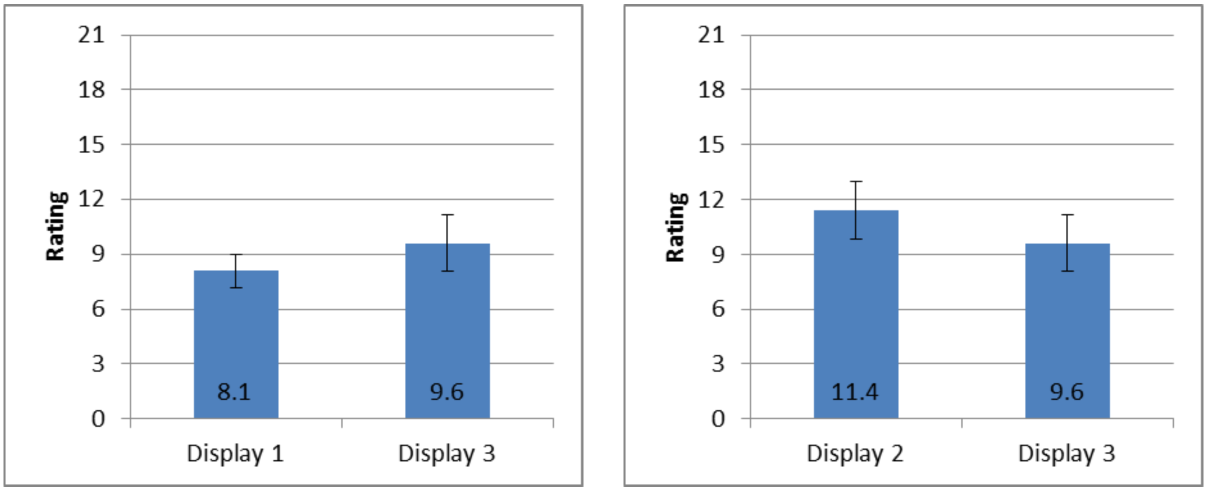


Figure 35. NASA-TLX – Frustration Ratings: D1 vs. D3 (L); D2 vs. D3 (R)

Results from the analyses of the NASA-TLX across the three display configurations did not support the proposed hypothesis that operators would report the lowest levels of workload in the Display 1 condition.

6.3 Integrated Versus Standalone Comparisons

6.3.1 Question 6

Is there a performance difference between standalone and integrated displays?

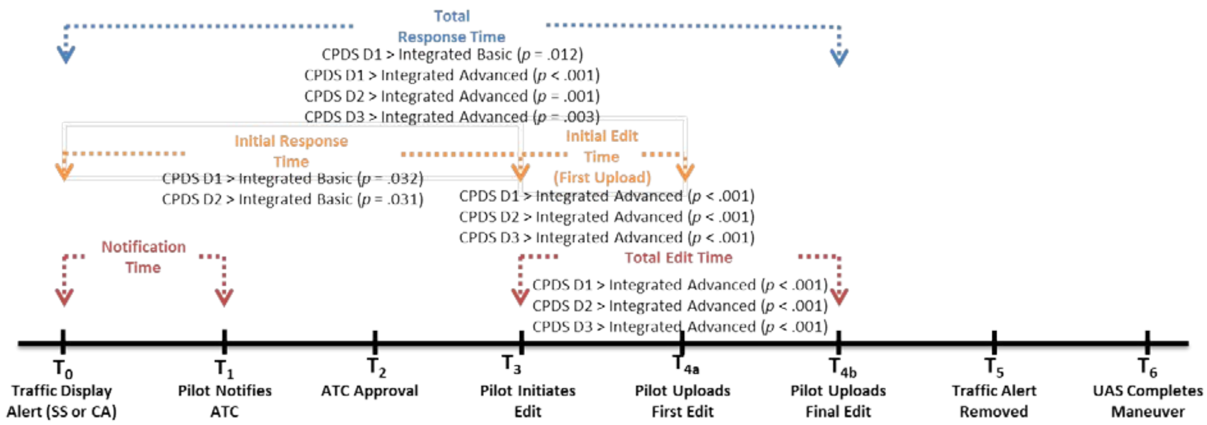


Figure 36. Integrated vs. Standalone Response Time Comparisons

All five response time metrics for the VS Basic and VS Advanced conditions significantly deviated from normality so all analyses were conducted using Mann-Whitney tests. In the case of marginal findings Kruskal-Wallis and Friedman tests were conducted to investigate equivalency of effect.

Notification time ($T_1 – T_0$). Mann-Whitney tests were conducted between each of the Part Task 4b CPDS display conditions (standalone) displays and each of the Part Task 4 integrated displays. No significant differences between standalone and integrated displays were identified. Table 6 displays the results of all of the comparisons. Since there was an overlap of operators in both experiments the data was also analyzed using a Kruskal-Wallis test (i.e., treating the data as between-subjects and a Friedman test (i.e., treating the data as within-subjects). No significant results were identified.

Table 6. Integrated V. Standalone Displays (Notification Time)

Comparison	*U*	*z*	Sig. (2-tailed)
CPDS D1 / VS Basic	798.50	-.213	.831
CPDS D1: N = 31; *M* = 27.58; *SE* = 5.93		VS Basic: N = 53; *M* = 26.43; *SE* = 2.64	
CPDS D2 / VS Basic	777.50	-.408	.683
CPDS D2: N = 31; *M* = 25.94; *SE* = 4.75		VS Basic: N = 53; *M* = 26.43; *SE* = 2.64	
CPDS D3 / VS Basic	1,026.00	1.346	.178
CPDS D3: N = 33; *M* = 20.67; *SE* = 4.89		VS Basic: N = 53; *M* = 26.43; *SE* = 2.64	
CPDS D1 / VS Advanced	811.00	-.879	.379
CPDS D1: N = 31; *M* = 27.58; *SE* = 5.93		VS Advanced: N = 59; *M* = 22.69; *SE* = 1.67	
CPDS D2 / VS Advanced	778.00	-1.160	.246
CPDS D2: N = 31; *M* = 25.94; *SE* = 4.75		VS Advanced: N = 59; *M* = 22.69; *SE* = 1.67	
CPDS D3 / VS Advanced	1,077.00	.843	.399
CPDS D3: N = 33; *M* = 20.67; *SE* = 4.89		VS Advanced: N = 59; *M* = 22.69; *SE* = 1.67	

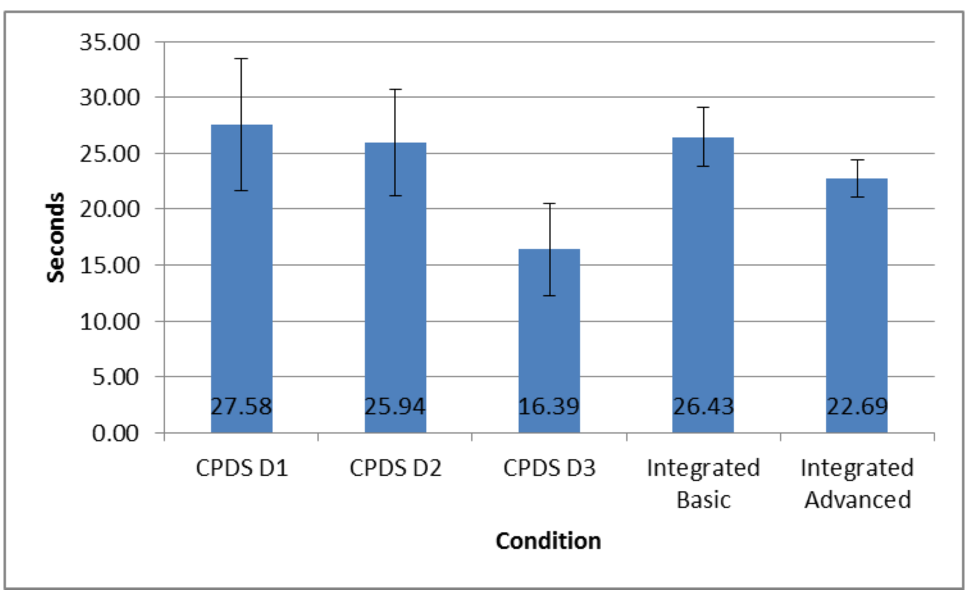


Figure 37. Notification Time

Initial response time ($T_3 - T_0$). Mann-Whitney tests were conducted between each of the Part Task 4b CPDS display conditions (standalone) displays and each of the Part Task 4 integrated displays. Results indicated that the integrated VS Basic condition was significantly different than the standalone CPDS Display 1 condition (U = 608.00, z = -2.14, p = .032) and Display 2 condition (U = 606.50, z = -2.153, p = .031). Operators in the integrated VS Basic display condition (*Mdn* = 12.00) took less time to notify ATC of a request to maneuver than operators in the standalone CPDS Display 1 (*Mdn* = 26.00) Display 2 (*Mdn* = 31.00) conditions. Table 7 displays the results of all of the comparisons. Results of the Kruskal-Wallis test did not confirm this finding, $H(4) = 8.170$, p = .086 nor did the Friedman test, $\chi^2(4) = 6.188$, p = .186.

Table 7. Integrated V. Standalone Displays (Initial Response Time)

<table>
<tr><th>Comparison</th><th>U</th><th>z</th><th>Sig. (2-tailed)</th></tr>
<tr><td>CPDS D1 /
VS Basic</td><td>608.00</td><td>-2.14</td><td>.032</td></tr>
<tr><td colspan="2">CPDS D1: N = 33; M = 24.39; SE = 5.30</td><td colspan="2">VS Basic: N = 51; M = 15.63; SE = 1.95</td></tr>
<tr><td>CPDS D2 /
VS Basic</td><td>606.50</td><td>-2.153</td><td>.031</td></tr>
<tr><td colspan="2">CPDS D2: N = 33; M = 25.81; SE = 4.70</td><td colspan="2">VS Basic: N = 51; M = 15.63; SE = 1.95</td></tr>
<tr><td>CPDS D3 /
VS Basic</td><td>794.50</td><td>-1.065</td><td>.287</td></tr>
<tr><td colspan="2">CPDS D3: N = 36; M = 20.92; SE = 4.23</td><td colspan="2">VS Basic: N = 51; M = 15.63; SE = 1.95</td></tr>
<tr><td>CPDS D1 /
VS Advanced</td><td>841.50</td><td>-1.419</td><td>.156</td></tr>
<tr><td colspan="2">CPDS D1: N = 33; M = 24.39; SE = 5.30</td><td colspan="2">VS Advanced: N = 62; M = 19.55; SE = 1.46</td></tr>
<tr><td>CPDS D2 /
VS Advanced</td><td>803.00</td><td>-1.720</td><td>.085</td></tr>
<tr><td colspan="2">CPDS D2: N = 33; M = 25.81; SE = 4.70</td><td colspan="2">VS Advanced: N = 62; M = 19.55; SE = 1.46</td></tr>
<tr><td>CPDS D3 /
VS Advanced</td><td>1,086.00</td><td>-.221</td><td>.825</td></tr>
<tr><td colspan="2">CPDS D3: N = 36; M = 20.92; SE = 4.23</td><td colspan="2">VS Advanced: N = 62; M = 19.55; SE = 1.46</td></tr>
</table>

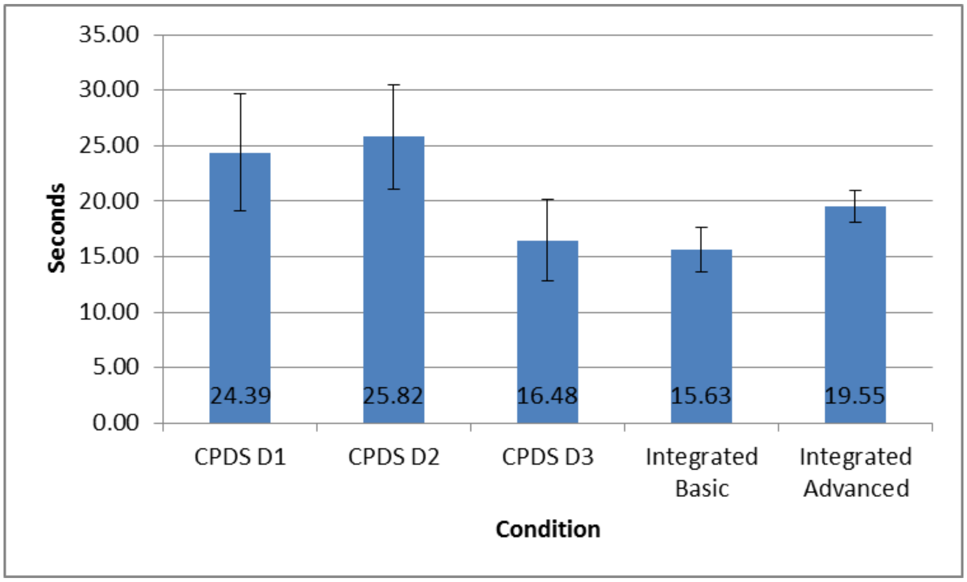


Figure 38. Initial Response Time

Initial Edit Time ($T_{4a} – T_3$). Mann-Whitney tests were conducted between each of the Part Task 4b CPDS display conditions (standalone) displays and each of the Part Task 4 integrated displays. Results indicated that the integrated VS Advanced condition was significantly different than all three standalone CPDS display conditions (D1 = U = 2.00, z = -8.633, p < .001; D2 = U = 9.50, z = -8.460, p < .001; D3 = U = 11.50, z = -8.601, p < .001). Operators in the integrated VS Advanced display condition (*Mdn* = 00.00) took less time to initiate an edit in the VSCS control interface than operators in the standalone CPDS Display 1 (*Mdn* = 13.00), 2 (*Mdn* = 9.00), and 3 (*Mdn* =10.00) conditions. Table 8 displays the results of all of the comparisons. Results of the Kruskal-Wallis test were also significant, $H(4) = 125.090$, p < .001. Pairwise comparisons with adjusted p-values confirmed the findings of the Mann-Whitney tests indicating that the integrated VS Advanced condition was significantly different than all three standalone CPDS display conditions (D1 = p < .001; D2 = p < .001; D3 = p < .001). Finally, the Friedman test also confirmed a significant effect, $\chi^2(4) = 73.559$, p < .001. Follow-up pairwise comparisons were conducted using a Dunn-Bonferroni test. Results of the comparisons confirmed the findings of the other tests indicating the integrated VS Advanced display was significantly faster than the three standalone CPDS displays (D1 = $p < .001$; D2 = $p < .001$; D3 = $p < .001$).

Table 8. Integrated V. Standalone Displays (Initial Edit Time)

Comparison	***U***	***z***	**Sig. (2-tailed)**
CPDS D1 / VS Basic	812.00	-1.119	.263
CPDS D1: N = 35; *M* = 18.26; *SE* = 2.52			VS Basic: N = 54; *M* = 13.63; *SE* = 1.25
CPDS D2 / VS Basic	1,055.00	1.442	.149
CPDS D2: N = 33; *M* = 10.39; *SE* = 1.09			VS Basic: N = 54; *M* = 13.63; *SE* = 1.25
CPDS D3 / VS Basic	1,134.50	1.340	.180
CPDS D3: N = 36; *M* = 10.28; *SE* = 0.92			VS Basic: N = 54; *M* = 13.63; *SE* = 1.25
CPDS D1 / VS Advanced	2.00	-8.633	< .001
CPDS D1: N = 35; *M* = 18.26; *SE* = 2.52			VS Advanced: N = 55; *M* = 0.35; *SE* = 0.13
CPDS D2 / VS Advanced	9.50	-8.460	< .001
CPDS D2: N = 33; *M* = 10.39; *SE* = 1.09			VS Advanced: N = 55; *M* = 0.35; *SE* = 0.13
CPDS D3 / VS Advanced	11.50	-8.601	< .001
CPDS D3: N = 36; *M* = 10.28; *SE* = 0.92			VS Advanced: N = 55; *M* = 0.35; *SE* = 0.13

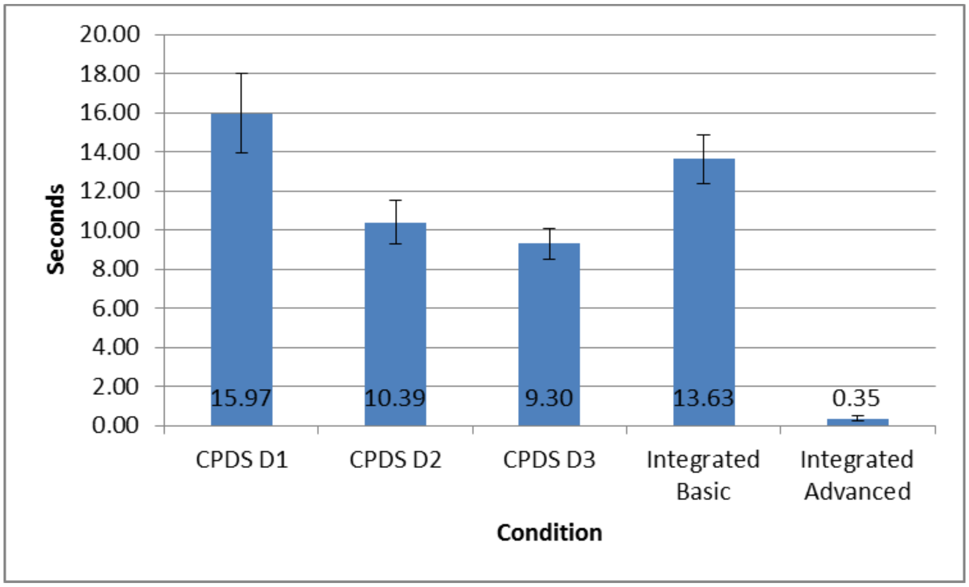


Figure 39. Initial Edit Time

Total Edit Time ($T_{4b} - T_3$). Mann-Whitney tests were conducted between each of the Part Task 4b CPDS conditions (standalone) and each of the Part Task 4 integrated displays. Results indicated the integrated VS Advanced condition was significantly different than all three standalone CPDS display conditions (D1 = $U = 72.00$, $z = -7.543$, $p < .001$; D2 = $U = 71.50$, $z = -7.547$, $p < .001$; D3 = $U = 117.00$, $z = -7.318$, $p < .001$). Operators in the integrated VS Advanced display condition (*Mdn* = 00.00) took less time to upload all edits for a particular alert than operators in the standalone CPDS Display 1 (*Mdn* = 19.00), 2 (*Mdn* = 15.00), and Display 3 (*Mdn* =13.50) conditions. Table 9 displays the results of all of the comparisons. Results of the Kruskal-Wallis test were also significant, $H(4) = 105.850$, $p < .001$. Pairwise comparisons with adjusted *p*-values confirmed that the integrated VS Advanced condition was significantly different than all three standalone CPDS display conditions (D1 = $p < .001$; D2 = $p < .001$; D3 = $p < .001$). Finally, the Friedman test also confirmed a significant effect, $\chi^2(4) = 53.525$, $p < .001$. Follow-up pairwise comparisons were conducted using a Dunn-Bonferroni test. Results of the comparisons confirmed the findings of the other tests indicating the integrated VS Advanced display was significantly faster than the three standalone CPDS displays (D1 = $p < .001$; D2 = $p < .001$; D3 = $p < .001$).

Table 9. Integrated V. Standalone Displays (Total Edit Time)

<table>
<tr><th>Comparison</th><th>U</th><th>z</th><th>Sig. (2-tailed)</th></tr>
<tr><td>CPDS D1 /
VS Basic</td><td>862.50</td><td>-.613</td><td>.540</td></tr>
<tr><td colspan="2">CPDS D1: N = 34; M = 25.71; SE = 3.37</td><td colspan="2">VS Basic: N = 55; M = 21.02; SE = 1.86</td></tr>
<tr><td>CPDS D2 /
VS Basic</td><td>867.50</td><td>-.570</td><td>.568</td></tr>
<tr><td colspan="2">CPDS D2: N = 34; M = 32.26; SE = 5.51</td><td colspan="2">VS Basic: N = 55; M = 21.02; SE = 1.86</td></tr>
<tr><td>CPDS D3 /
VS Basic</td><td>1,205.50</td><td>1.750</td><td>.080</td></tr>
<tr><td colspan="2">CPDS D3: N = 36; M = 16.67; SE = 2.09</td><td colspan="2">VS Basic: N = 55; M = 21.02; SE = 1.86</td></tr>
<tr><td>CPDS D1 /
VS Advanced</td><td>72.00</td><td>-7.543</td><td>< .001</td></tr>
<tr><td colspan="2">CPDS D1: N = 34; M = 25.71; SE = 3.37</td><td colspan="2">VS Advanced: N = 55; M = 2.0; SE = 0.45</td></tr>
<tr><td>CPDS D2 /
VS Advanced</td><td>71.50</td><td>-7.547</td><td>< .001</td></tr>
<tr><td colspan="2">CPDS D2: N = 34; M = 32.26; SE = 5.51</td><td colspan="2">VS Advanced: N = 55; M = 2.0; SE = 0.45</td></tr>
<tr><td>CPDS D3 /
VS Advanced</td><td>117.00</td><td>-7.318</td><td>< .001</td></tr>
<tr><td colspan="2">CPDS D3: N = 36; M = 16.67; SE = 2.09</td><td colspan="2">VS Advanced: N = 55; M = 2.0; SE = 0.45</td></tr>
</table>

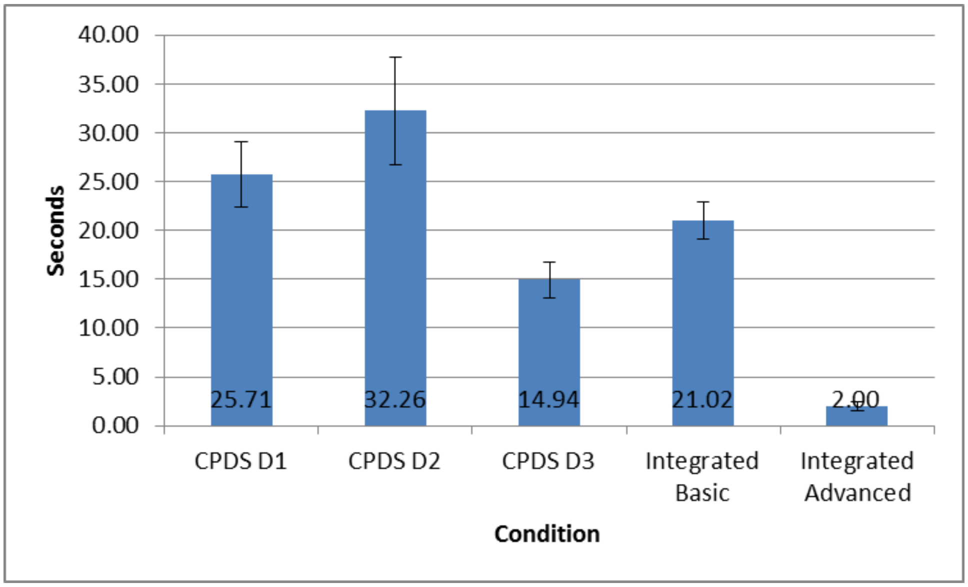


Figure 40. Total Edit Time

Total Response Time (T_{4b} – T_0). Mann-Whitney tests were conducted between each of the Part Task 4b CPDS display conditions (standalone) displays and each of the Part Task 4 integrated displays. Results indicated that the integrated VS Basic condition was significantly different than the standalone CPDS Display 1 condition (U = 628.00, z = -2.515, p = .012) and the integrated VS Advanced condition was significantly different than all three standalone CPDS display conditions (D1 = U = 495.00, z = -4.273, p < .001; D2 = U = 638.00, z = -3.359, p = .001; D3 = U = 756.00, z = -3.018, p = .003). Operators in the integrated VS Basic display condition (*Mdn* = 34.50) took less time from the initial appearance of an alert to upload all edits for a particular alert than operators in the standalone CPDS Display 1 condition (*Mdn* = 58.00). Operators in the integrated VS Advanced display condition (*Mdn* = 23.50) also took less time from the initial appearance of an alert to upload all edits for a particular alert than operators in the standalone CPDS Display 1 (*Mdn* = 58.00), 2 (*Mdn* = 50.00), and 3 (*Mdn* = 42.00) conditions. Table 10 displays the results of all of the comparisons. Results of the Kruskal-Wallis test were also significant, $H(4)$ = 26.298, p < .001. Pairwise comparisons with adjusted p-values indicated that the integrated VS Advanced condition was significantly different than all three standalone CPDS display conditions (D1 = $p < .001$; D2 = $p = .003$; D3 = $p = .013$). Finally, the Friedman test also confirmed a significant effect, $\chi^2(4) = 15.562$, $p = .004$. Follow-up pairwise comparisons were conducted using a Dunn-Bonferroni test. Follow-up analyses of the Friedman only confirmed a significant effect between the integrated VS Advanced condition and the standalone CPDS Display 1 condition ($p = .016$).

Table 10. Integrated V. Standalone Displays (Total Response Time)

Comparison	***U***	***z***	**Sig. (2-tailed)**
CPDS D1 / VS Basic	628.00	-2.515	.012
CPDS D1: N = 33; *M* = 54.09; *SE* = 4.61		VS Basic: N = 56; *M* = 40.43; *SE* = 2.93	
CPDS D2 / VS Basic	758.00	-1.615	.106
CPDS D2: N = 34; *M* = 49.85; *SE* = 5.52		VS Basic: N = 56; *M* = 40.43; *SE* = 2.93	
CPDS D3 / VS Basic	913.00	-.966	.334
CPDS D3: N = 37; *M* = 49.70; *SE* = 5.83		VS Basic: N = 56; *M* = 40.43; *SE* = 2.93	
CPDS D1 / VS Advanced	495.00	-4.273	< .001
CPDS D1: N = 33; *M* = 54.09; *SE* = 4.61		VS Advanced: N = 64; *M* = 29.38; *SE* = 2.43	
CPDS D2 / VS Advanced	638.00	-3.359	.001
CPDS D2: N = 34; *M* = 49.85; *SE* = 5.52		VS Advanced: N = 64; *M* = 29.38; *SE* = 2.43	
CPDS D3 / VS Advanced	756.00	-3.018	.003
CPDS D3: N = 37; *M* = 49.70; *SE* = 5.83		VS Advanced: N = 64; *M* = 29.38; *SE* = 2.43	

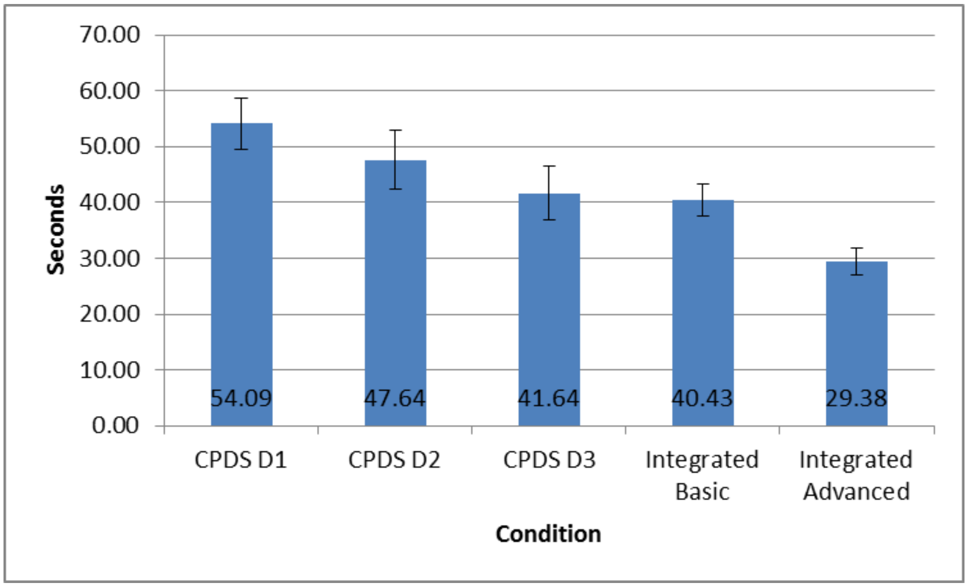


Figure 41. Total Response Time

6.3.2 Question 7

Is there a performance difference on secondary tasks between standalone and integrated displays?

Simple Secondary Tasks. The same 12 simple secondary tasks per trial presented in Part Task 4b were also presented in Part Task 4. Number of tasks responded to did not significantly deviate from normality so independent *t*-tests were conducted between each of the Part Task 4 CPDS display conditions (standalone) displays and each of the Part Task 4 integrated displays. No significant differences between standalone and integrated displays were identified. Table 11 displays the results of all of the comparisons. Since there was an overlap of operators in both experiments the data was also analyzed using a one-way ANOVA (i.e., treating the data as between-subjects and a one-way repeated measures ANOVA (i.e., treating the data as within-subjects). No significant results were identified.

Table 11. Integrated V. Standalone Displays (Simple Secondary Tasks)

Comparison	Equal Variances	Levene's Test Equality of Variances		t-test for Equality of Means				
		F	Sig.	*t*	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
CPDS D1 / VS Basic	Assumed	.001	.972	< .001	20	1.00	.00	1.579
CPDS D1: N = 10; *M* = 6.0; *SE* = 1.16				VS Basic: N = 12; *M* = 6.0; *SE* = 1.07				
CPDS D2 / VS Basic	Assumed	.117	.736	-.425	20	.675	-.70	1.647
CPDS D2: N = 10; *M* = 5.30; *SE* = 1.27				VS Basic: N = 12; *M* = 6.0; *SE* = 1.07				
CPDS D3 / VS Basic	Assumed	.376	.546	.123	20	.200	.20	1.629
CPDS D3: N = 10; *M* = 6.20; *SE* = 1.24				VS Basic: N = 12; *M* = 6.0; *SE* = 1.07				
CPDS D1 / VS Advanced	Assumed	.325	.575	.173	20	.865	.250	1.446
CPDS D1: N = 10; *M* = 6.0; *SE* = 1.16				VS Advanced: N = 12; *M* = 5.75; *SE* = .906				
CPDS D2 / VS Advanced	Assumed	.919	.349	-.296	20	.770	-.450	1.521
CPDS D2: N = 10; *M* = 5.30; *SE* = 1.27				VS Advanced: N = 12; *M* = 5.75; *SE* = .906				
CPDS D3 / VS Advanced	Assumed	1.78	.197	.300	20	.767	.450	1.501
CPDS D3: N = 10; *M* = 6.20; *SE* = 1.24				VS Advanced: N = 12; *M* = 5.75; *SE* = .906				

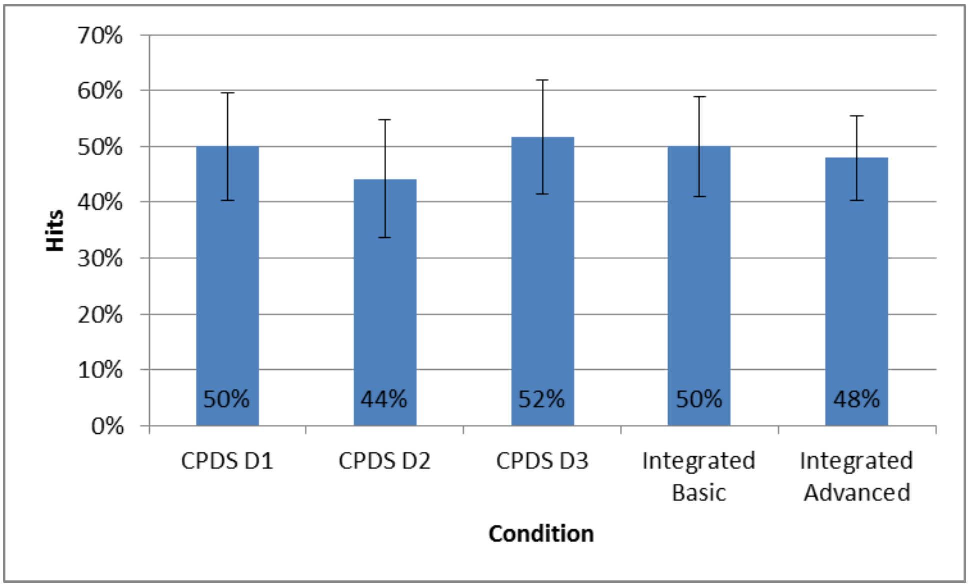


Figure 42. Simple Secondary Task % Hits

Response time to simple secondary tasks was also analyzed. Reaction time was not normally distributed, so transformations were investigated. Based on Kirk (2013) a square root transformation was chosen for analyzing response time to simple secondary tasks between standalone and integrated displays (See Appendix H for transformation comparisons).

After transforming the data 2 of the 3 display conditions still significantly deviated from normality so Mann-Whitney tests were conducted between each of the Part Task 4 CPDS display conditions (standalone) displays and each of the Part Task 4 integrated displays. No significant differences between standalone and integrated displays were identified. Table 12 displays the results of all of the comparisons. Since there was an overlap of operators in both experiments the data was also analyzed using a Kruskal-Wallis test (i.e., treating the data as between-subjects and aFriedman test (i.e., treating the data as within-subjects). No significant results were identified.

Table 12. Integrated V. Standalone Displays (Simple Secondary Task RT)

Comparison	*U*	*z*	Sig. (2-tailed)
CPDS D1 / VS Basic	2.283.00	.736	.462
CPDS D1: N = 59; *M* = 9.98; *SE* = .924		VS Basic: N = 72; *M* = 11.10; *SE* = .916	
CPDS D2 / VS Basic	1,935.00	.225	.822
CPDS D2: N = 53; *M* = 10.90; *SE* = 1.07		VS Basic: N = 72; *M* = 11.10; *SE* = .916	
CPDS D3 / VS Basic	2,038.00	-.866	.387
CPDS D3: N = 62; *M* = 12.21; *SE* = 1.02		VS Basic: N = 72; *M* = 11.10; *SE* = .916	
CPDS D1 / VS Advanced	2,235.00	.954	.340
CPDS D1: N = 59; *M* = 9.98; *SE* = .924		VS Advanced: N = 69; *M* = 11.14; *SE* = .855	
CPDS D2 / VS Advanced	1,935.00	.550	.582
CPDS D2: N = 53; *M* = 10.90; *SE* = .1.07		VS Advanced: N = 69; *M* = 11.14; *SE* = .855	
CPDS D3 / VS Advanced	1,965.00	-.820	.422
CPDS D3: N = 62; *M* = 12.21; *SE* = 1.02		VS Advanced: N = 69; *M* = 11.14; *SE* = .855	

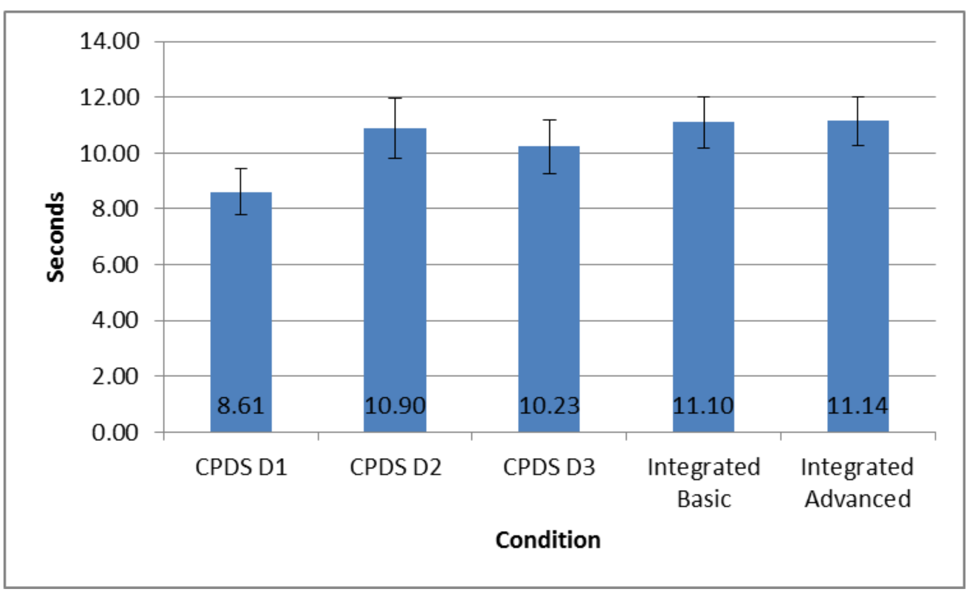


Figure 43. Simple Secondary Task Response Time

Complex Secondary Tasks. The same 12 complex secondary tasks per trial presented in Part Task 4b were also presented in Part Task 4. Number of tasks responded to significantly deviated from normality so Mann-Whitney tests were conducted between each of the Part Task 4 CPDS display conditions (standalone) displays and each of the Part Task 4 integrated displays. No significant differences between standalone and integrated displays were identified. Table 13 displays the results of all of the comparisons. Since there was an overlap of operators in both experiments the data was also analyzed using a Kruskal-Wallis test (i.e., treating the data as between-subjects and a Friedman test (i.e., treating the data as within-subjects). No significant results were identified. Accuracy of responses was not analyzed as all operators responded accurately to all tasks. This is not surprising given the tasks were very directed questions such as, fuel remaining or changing a radio to specified channel.

Table 13. Integrated V. Standalone Displays (Complex Secondary Tasks)

Comparison	*U*	*z*	Sig. (2-tailed)
CPDS D1 / VS Basic	70.50	.749	.497
CPDS D1: N = 10; *M* = 10.80; *SE* = .467		VS Basic: N = 12; *M* = 11.42; *SE* = .193	
CPDS D2 / VS Basic	60.50	.037	1.00
CPDS D2: N = 10; *M* = 11.4; *SE* = .221		VS Basic: N = 12; *M* = 11.42; *SE* = .193	
CPDS D3 / VS Basic	69.50	.684	.539
CPDS D3: N = 10; *M* = 11.00; *SE* = .394		VS Basic: N = 12; *M* = 11.42; *SE* = .193	
CPDS D1 / VS Advanced	75.00	1.093	.346
CPDS D1: N = 10; *M* = 10.80; *SE* = .467		VS Advanced: N = 12; *M* = 11.58; *SE* = .149	
CPDS D2 / VS Advanced	67.50	.563	.628
CPDS D2: N = 10; *M* = 11.4; *SE* = .221		VS Advanced: N = 12; *M* = 11.58; *SE* = .149	
CPDS D3 / VS Advanced	76.00	1.174	.314
CPDS D3: N = 10; *M* = 11.00; *SE* = .394		VS Advanced: N = 12; *M* = 11.58; *SE* = .149	

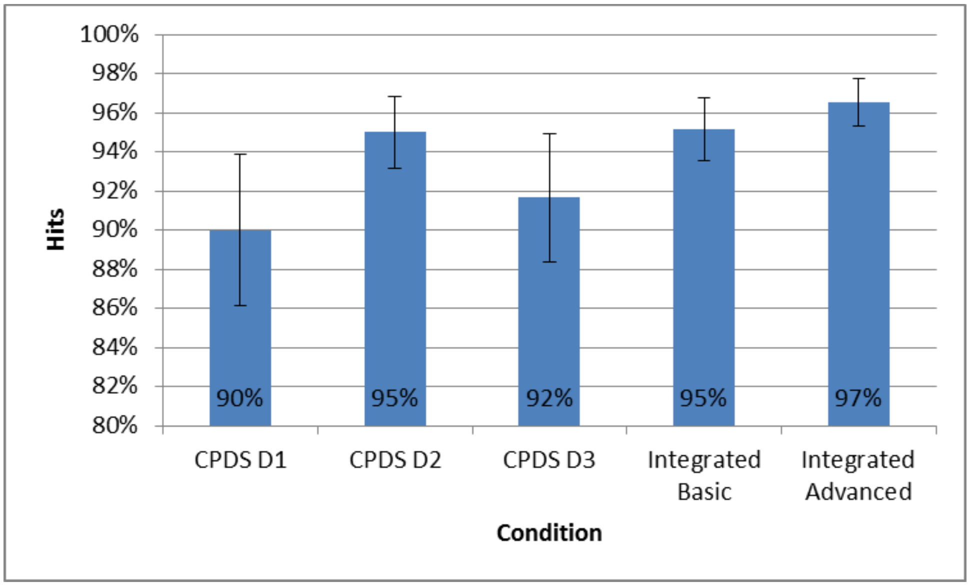


Figure 44. Complex Secondary Task % Hits

Response time to complex secondary tasks was also analyzed. Outliers were identified as specified in the results of Question 5. Reaction time was not normally distributed, so transformations were investigated. Based on Kirk (2013) a square root transformation was chosen for analyzing response time to complex secondary tasks between standalone and integrated displays (See Appendix H for transformation comparisons).

After transforming the data 2 of the 3 display conditions still significantly deviated from normality so Mann-Whitney tests were conducted between each of the Part Task 4 CPDS display conditions (standalone) displays and each of the Part Task 4 integrated displays. Results indicated that the integrated VS Advanced condition was significantly different than all three CPDS display conditions (D1 = U = 5,569.00, z = -2.002, p = .045; D2 = U = 5,459.00, z = -2.970, p = .003; D3 = U = 5,377.00, z = -2.158, p = .031). Operators in the integrated VS Advanced display condition (*Mdn* = 20.07) showed a significantly faster response time complex secondary tasks than operators in the standalone CPDS Display 1 (*Mdn* = 27.89), 2 (*Mdn* = 32.58), and 3 (*Mdn* = 28.33) conditions. Table 14 displays the results of all of the comparisons.

Since there was an overlap of operators in both experiments the data was also analyzed using a Kruskal-Wallis test (i.e., treating the data as between-subjects and a Friedman test (i.e., treating the data as within-subjects). Results of the Kruskal-Wallis test were also significant, $H(4) = 10.464$, $p = .033$. Pairwise comparisons with adjusted *p*-values only identified as significant difference between the integrated VS Advanced display condition and the standalone CPDS Display 2 condition ($p = .02$). The Friedman test also identified a significant effect, $\chi^2(4) = 18.208$, $p = .001$. Follow-up pairwise comparisons were conducted using a Dunn-Bonferroni test. Results indicated operators using the standalone CPDS Display 2 ($p = .009$) and 3 ($p = .007$) conditions showed a significantly slower response time to complex secondary tasks than operators in the integrated VS Advanced condition.

Table 14. Integrated V. Standalone Displays (Complex Secondary Task RT)

Comparison	***U***	***z***	**Sig. (2-tailed)**
CPDS D1 / VS Basic	6,450.00	-.453	.651
CPDS D1: N = 102; *M* = 39.09; *SE* = 3.09			VS Basic: N = 131; *M* = 41.56; *SE* = 3.40
CPDS D2 / VS Basic	6,206.00	-1.743	.081
CPDS D2: N = 109; *M* = 51.14; *SE* = 4.46			VS Basic: N = 131; *M* = 41.56; *SE* = 3.40
CPDS D3 / VS Basic	6,175.00	-.745	.456
CPDS D3: N = 100; *M* = 40.79; *SE* = 3.32			VS Basic: N = 131; *M* = 41.56; *SE* = 3.40
CPDS D1 / VS Advanced	5,569.00	-2.002	.045
CPDS D1: N = 102; *M* = 39.09; *SE* = 3.09			VS Advanced: N = 129; *M* = 35.92; *SE* = 3.28
CPDS D2 / VS Advanced	5,459.00	-2.970	.003
CPDS D2: N = 109; *M* = 51.14; *SE* = 4.46			VS Advanced: N = 129; *M* = 35.92; *SE* = 3.28
CPDS D3 / VS Advanced	5,377.00	-2.158	.031
CPDS D3: N = 100; *M* = 40.79; *SE* = 3.32			VS Advanced: N = 129; *M* = 35.92; *SE* = 3.28

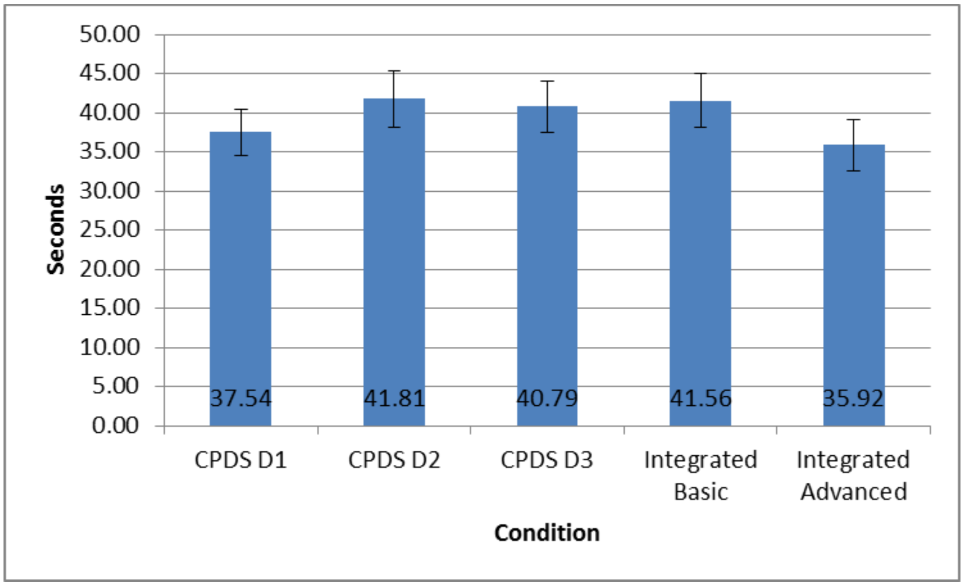


Figure 45. Complex Secondary Task Response Time

6.3.3 Question 8

Is there a workload difference between standalone and integrated displays?

Each of the six NASA-TLX scales were analyzed using the raw data on a 1 to 7 scale for differences between each of the Part Task 4 CPDS display conditions (standalone) displays and each of the Part Task 4 integrated displays.

Mental. Responses in the CPDS Display condition 1 and the VS Basic condition violated the assumption of normality so Mann-Whitney tests were used to analyze the data. No significant differences between standalone and integrated displays were identified. Table 15 displays the results of all of the comparisons. Since there was an overlap of operators in both experiments the data was also analyzed using a Kruskal-Wallis test (i.e., treating the data as between-subjects and a Friedman test (i.e., treating the data as within-subjects). No significant results were identified.

Table 15. Integrated V. Standalone Displays (NASA-TLX – Mental)

Comparison	*U*	*z*	Sig. (2-tailed)
CPDS D1 / VS Basic	55.50	-.307	.771
CPDS D1: N = 10; *M* = 4.00; *SE* = .447		VS Basic: N = 12; *M* = 3.83; *SE* = .423	
CPDS D2 / VS Basic	45.00	-1.016	.346
CPDS D2: N = 10; *M* = 4.50; *SE* = .500		VS Basic: N = 12; *M* = 3.83; *SE* = .423	
CPDS D3 / VS Basic	49.00	-.743	.497
CPDS D3: N = 10; *M* = 4.30; *SE* = .423		VS Basic: N = 12; *M* = 3.83; *SE* = .423	
CPDS D1 / VS Advanced	54.00	-.408	.722
CPDS D1: N = 10; *M* = 4.00; *SE* = .447		VS Advanced: N = 12; *M* = 3.92; *SE* = .468	
CPDS D2 / VS Advanced	47.00	-.871	.418
CPDS D2: N = 10; *M* = 4.50; *SE* = .500		VS Advanced: N = 12; *M* = 3.92; *SE* = .468	
CPDS D3 / VS Advanced	50.00	-.677	.539
CPDS D3: N = 10; *M* = 4.30; *SE* = .423		VS Advanced: N = 12; *M* = 3.92; *SE* = .468	

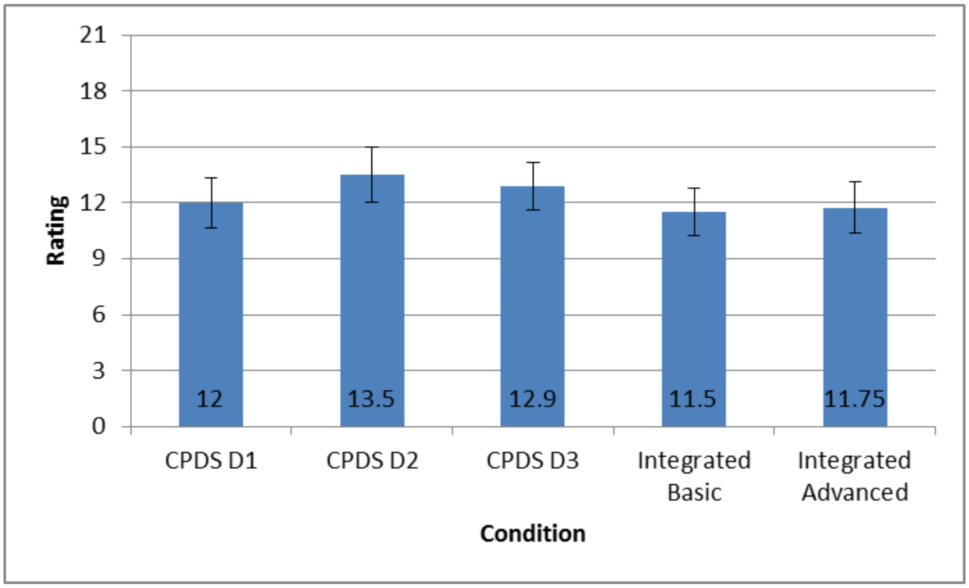


Figure 46. NASA-TLX – Mental Workload Ratings

Physical. Responses in the VS Basic condition violated the assumption of normality so Mann-Whitney tests were used to analyze the data. No significant differences between standalone and integrated displays were identified. Table 16 displays the results of all of the comparisons. Since there was an overlap of operators in both experiments the data was also analyzed using a Kruskal-Wallis test (i.e., treating the data as between-subjects and a Friedman test (i.e., treating the data as within-subjects). No significant results were identified.

Table 16. Integrated V. Standalone Displays (NASA-TLX – Physical)

Comparison	***U***	***z***	**Sig. (2-tailed)**
CPDS D1 / VS Basic	55.50	-.313	.771
CPDS D1: N = 10; *M* = 2.30; *SE* = .396		VS Basic: N = 12; *M* = 2.08; *SE* = .260	
CPDS D2 / VS Basic	48.00	-.822	.456
CPDS D2: N = 10; *M* = 2.90; *SE* = .623		VS Basic: N = 12; *M* = 2.08; *SE* = .260	
CPDS D3 / VS Basic	46.00	-.958	.381
CPDS D3: N = 10; *M* = 2.80; *SE* = .573		VS Basic: N = 12; *M* = 2.08; *SE* = .260	
CPDS D1 / VS Advanced	55.00	-.343	.771
CPDS D1: N = 10; *M* = 2.30; *SE* = .396		VS Advanced: N = 12; *M* = 3.92; *SE* = .468	
CPDS D2 / VS Advanced	48.00	-.818	.456
CPDS D2: N = 10; *M* = 2.90; *SE* = .623		VS Advanced: N = 12; *M* = 3.92; *SE* = .468	
CPDS D3 / VS Advanced	47.50	-.856	.418
CPDS D3: N = 10; *M* = 2.80; *SE* = .573		VS Advanced: N = 12; *M* = 3.92; *SE* = .468	

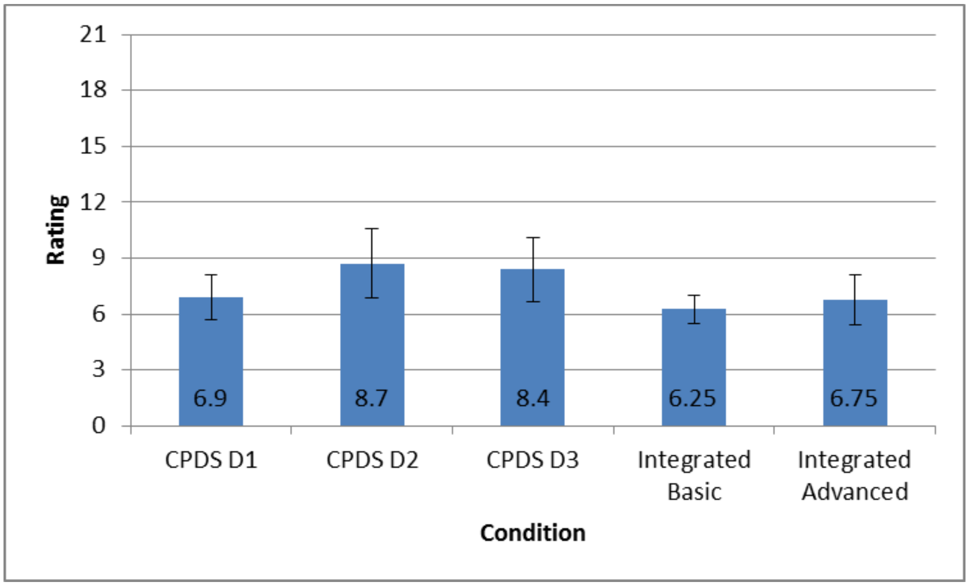


Figure 47. NASA-TLX – Physical Workload Ratings

Temporal. Responses in the CPDS Display 3 condition violated the assumption of normality so Mann-Whitney tests were used to analyze the data. No significant differences between standalone and integrated displays were identified. Table 17 displays the results of all of the comparisons. Since there was an overlap of operators in both experiments the data was also analyzed using a Kruskal-Wallis test (i.e., treating the data as between-subjects and a Friedman test (i.e., treating the data as within-subjects). No significant results were identified.

Table 17. Integrated V. Standalone Displays (NASA-TLX – Temporal)

Comparison	***U***	***z***	**Sig. (2-tailed)**
CPDS D1 / VS Basic	54.50	-.375	.722
CPDS D1: N = 10; *M* = 3.30; *SE* = .448		VS Basic: N = 12; *M* = 3.25; *SE* = .392	
CPDS D2 / VS Basic	44.00	-1.078	.314
CPDS D2: N = 10; *M* = 4.00; *SE* = .632		VS Basic: N = 12; *M* = 3.25; *SE* = .392	
CPDS D3 / VS Basic	60.00	0.00	1.00
CPDS D3: N = 10; *M* = 3.20; *SE* = .389		VS Basic: N = 12; *M* = 3.25; *SE* = .392	
CPDS D1 / VS Advanced	63.00	.202	.872
CPDS D1: N = 10; *M* = 3.30; *SE* = .448		VS Advanced: N = 12; *M* = 3.58; *SE* = .452	
CPDS D2 / VS Advanced	51.50	-.570	.582
CPDS D2: N = 10; *M* = 4.00; *SE* = .632		VS Advanced: N = 12; *M* = 3.58; *SE* = .452	
CPDS D3 / VS Advanced	65.00	.341	.771
CPDS D3: N = 10; *M* = 3.20; *SE* = .389		VS Advanced: N = 12; *M* = 3.58; *SE* = .452	

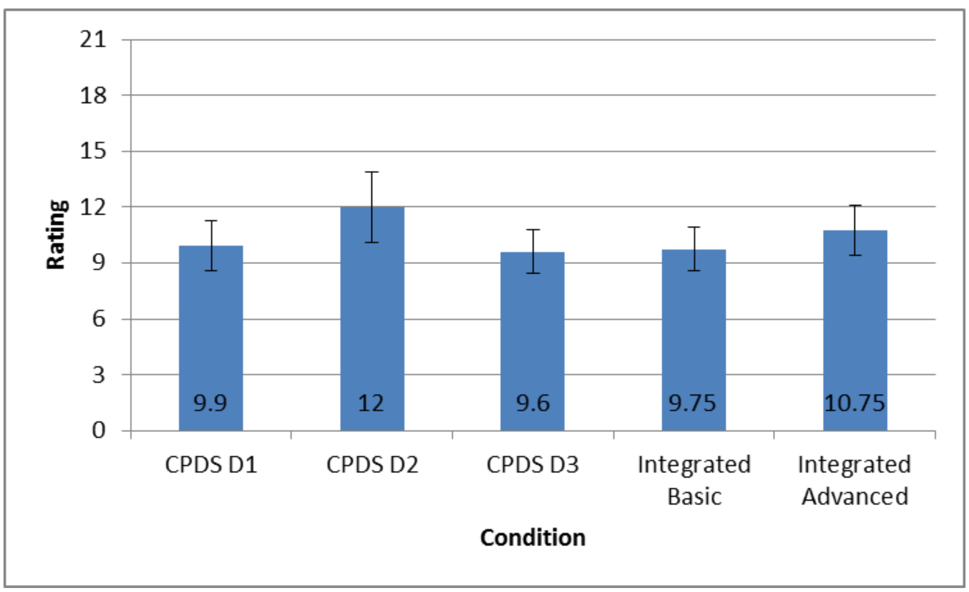


Figure 48. NASA-TLX – Temporal Workload Ratings

Performance. Responses in the CPDS Display 3 condition and the VS Advanced condition violated the assumption of normality so Mann-Whitney tests were used to analyze the data. Results indicated that the standalone CPDS Display 2 condition was significantly different than integrated VS Basic condition, $U = 28.00$, $z = -2.168$, $p = .032$. There was also a marginal effect identified between the standalone CPDS Display 2 condition and the integrated VS Advanced condition, $U = 30.00$, $z = -2.043$, $p = .050$. Operators in the standalone CPDS Display 2 condition (*Mdn* = 3.50) reported higher levels of performance degradation than operators in the integrated VS Basic (*Mdn* = 2.00) and VS Advanced (*Mdn* = 2.00) conditions. Table 18 displays the results of all of the comparisons. Results of the Kruskal-Wallis test did not confirm this finding, $H(4) = 7.518$, $p = .111$; however, the a Friedman test did identify a significant effect between integrated and standalone displays, $\chi^2(4) = 10.104$, $p = .039$. Follow-up pairwise comparisons were conducted using a Dunn-Bonferroni test. After adjusting *p*-values for multiple comparisons, no pairwise comparisons were identified as significant.

Table 18. Integrated V. Standalone Displays (NASA-TLX – Performance)

Comparison	*U*	*z*	Sig. (2-tailed)
CPDS D1 / VS Basic	58.00	-.140	.923
CPDS D1: N = 10; *M* = 2.40; *SE* = .340		VS Basic: N = 12; *M* = 2.33; *SE* = .284	
CPDS D2 / VS Basic	28.00	-2.168	.036
CPDS D2: N = 10; *M* = 3.60; *SE* = .452		VS Basic: N = 12; *M* = 2.33; *SE* = .284	
CPDS D3 / VS Basic	40.00	-1.371	.203
CPDS D3: N = 10; *M* = 3.00; *SE* = .394		VS Basic: N = 12; *M* = 2.33; *SE* = .284	
CPDS D1 / VS Advanced	54.00	-.419	.722
CPDS D1: N = 10; *M* = 2.40; *SE* = .340		VS Advanced: N = 12; *M* = 2.33; *SE* = .396	
CPDS D2 / VS Advanced	30.00	-2.043	.050
CPDS D2: N = 10; *M* = 3.60; *SE* = .452		VS Advanced: N = 12; *M* = 2.33; *SE* = .396	
CPDS D3 / VS Advanced	42.00	-1.222	.254
CPDS D3: N = 10; *M* = 3.00; *SE* = .394		VS Advanced: N = 12; *M* = 2.33; *SE* = .396	

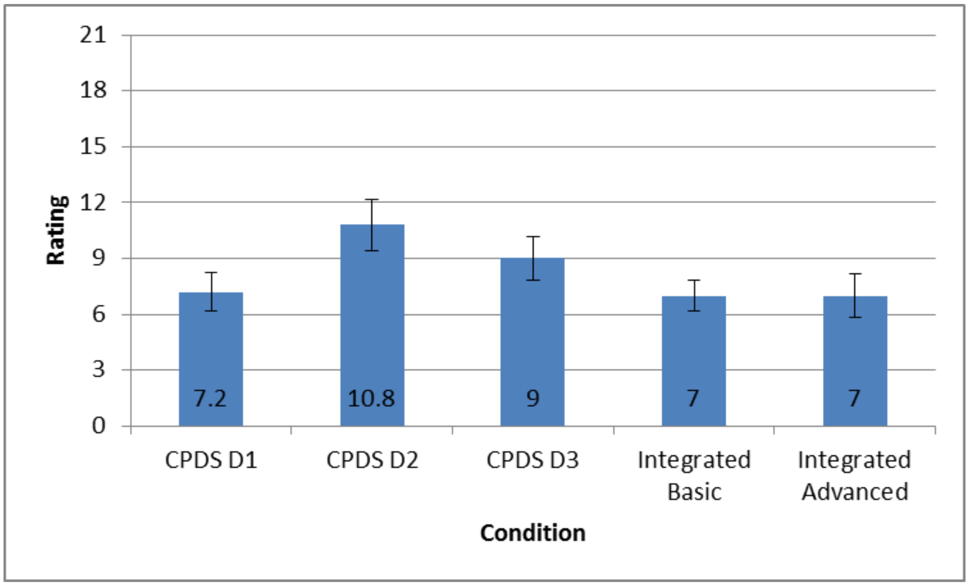


Figure 49. NASA-TLX – Performance Degradation Ratings

Effort. Responses in the CPDS Display 1 and the VS Basic conditions violated the assumption of normality so Mann-Whitney tests were used to analyze the data. No significant differences between standalone and integrated displays were identified. Table 19 displays the results of all of the comparisons. Since there was an overlap of operators in both experiments the data was also analyzed using a Kruskal-Wallis test (i.e., treating the data as between-subjects and a Friedman test (i.e., treating the data as within-subjects). No significant results were identified.

Table 19. Integrated V. Standalone Displays (NASA-TLX – Effort)

<table>
<tr><th>Comparison</th><th>U</th><th>z</th><th>Sig. (2-tailed)</th></tr>
<tr><td>CPDS D1 /
VS Basic</td><td>51.50</td><td>-.577</td><td>.582</td></tr>
<tr><td colspan="2">CPDS D1: N = 10; M = 3.60; SE = .400</td><td colspan="2">VS Basic: N = 12; M = 3.50; SE = .435</td></tr>
<tr><td>CPDS D2 /
VS Basic</td><td>39.50</td><td>-1.394</td><td>.180</td></tr>
<tr><td colspan="2">CPDS D2: N = 10; M = 4.40; SE = .499</td><td colspan="2">VS Basic: N = 12; M = 3.50; SE = .435</td></tr>
<tr><td>CPDS D3 /
VS Basic</td><td>46.00</td><td>-.976</td><td>.381</td></tr>
<tr><td colspan="2">CPDS D3: N = 10; M = 3.90; SE = .277</td><td colspan="2">VS Basic: N = 12; M = 3.50; SE = .435</td></tr>
<tr><td>CPDS D1 /
VS Advanced</td><td>51.50</td><td>-.578</td><td>.582</td></tr>
<tr><td colspan="2">CPDS D1: N = 10; M = 3.60; SE = .400</td><td colspan="2">VS Advanced: N = 12; M = 3.33; SE = .355</td></tr>
<tr><td>CPDS D2 /
VS Advanced</td><td>36.50</td><td>-1.582</td><td>.123</td></tr>
<tr><td colspan="2">CPDS D2: N = 10; M = 4.40; SE = .499</td><td colspan="2">VS Advanced: N = 12; M = 3.33; SE = .355</td></tr>
<tr><td>CPDS D3 /
VS Advanced</td><td>42.50</td><td>-1.195</td><td>.254</td></tr>
<tr><td colspan="2">CPDS D3: N = 10; M = 3.90; SE = .277</td><td colspan="2">VS Advanced: N = 12; M = 3.33; SE = .355</td></tr>
</table>

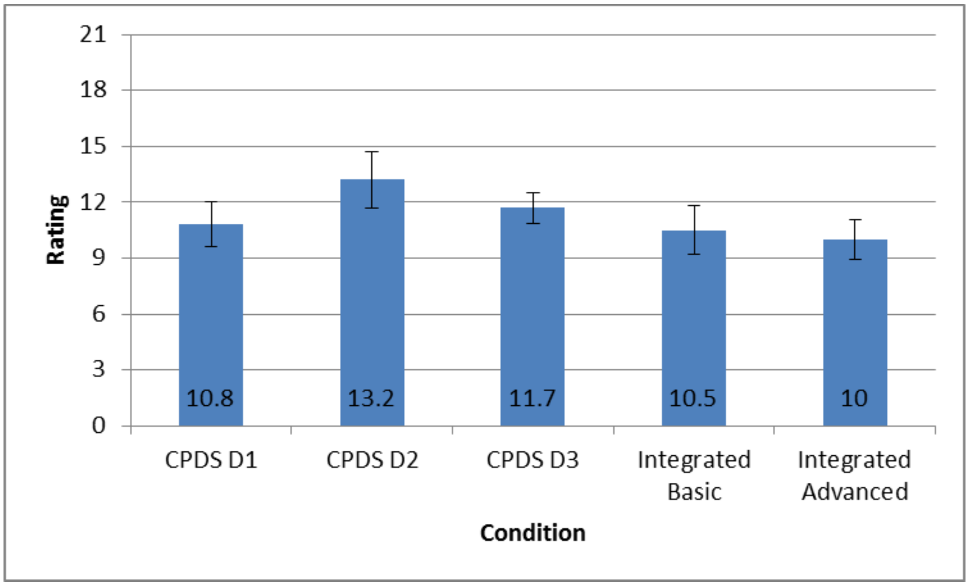


Figure 50. NASA-TLX – Effort Workload Ratings

Frustration. Responses for physical workload did not violate the assumption of normality so one-way independent *t*-tests were used to analyze the data. No significant differences between standalone and integrated displays were identified. Table 20 displays the results of all of the comparisons. Since there was an overlap of operators in both experiments the data was also analyzed using a one-way ANOVA (i.e., treating the data as between-subjects and a one-way repeated measures ANOVA (i.e., treating the data as within-subjects). No significant results were identified.

Table 20. Integrated V. Standalone Displays (NASA-TLX – Frustration)

Comparison	Equal Variances	Levene's Test Equality of Variances		t-test for Equality of Means				
		F	Sig.	*t*	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
CPDS D1 / VS Basic	Assumed	.413	.528	-.110	20	.913	-.050	.453
CPDS D1: N = 10; *M* = 2.70; *SE* = .300				VS Basic: N = 12; *M* = 2.75; *SE* = .329				
CPDS D2 / VS Basic	Assumed	2.51	.129	1.737	20	.098	1.050	.604
CPDS D2: N = 10; *M* = 3.80; *SE* = .533				VS Basic: N = 12; *M* = 2.75; *SE* = .329				
CPDS D3 / VS Basic	Assumed	.920	.349	.764	20	.454	.450	.589
CPDS D3: N = 10; *M* = 3.20; *SE* = .512				VS Basic: N = 12; *M* = 2.75; *SE* = .329				
CPDS D1 / VS Advanced	Assumed	1.13	.301	-.508	20	.617	-.300	.591
CPDS D1: N = 10; *M* = 2.70; *SE* = .300				VS Advanced: N = 12; *M* = 3.00; *SE* = .477				
CPDS D2 / VS Advanced	Assumed	.302	.589	1.121	20	.276	.800	.714
CPDS D2: N = 10; *M* = 3.80; *SE* = .533				VS Advanced: N = 12; *M* = 3.00; *SE* = .477				
CPDS D3 / VS Advanced	Assumed	.027	.872	.285	20	.778	.200	.701
CPDS D3: N = 10; *M* = 3.20; *SE* = .512				VS Advanced: N = 12; *M* = 3.00; *SE* = .477				

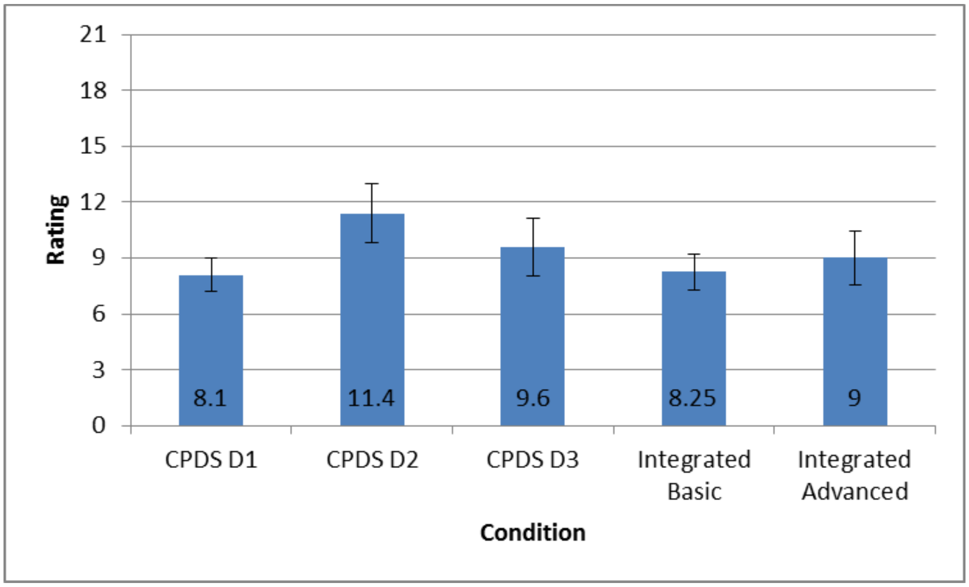


Figure 51. NASA-TLX – Frustration Ratings

Results from the analyses of the NASA-TLX across the three display configurations did not support the proposed hypothesis that operators would report higher levels of workload in the integrated display conditions.

CHAPTER 7

CONCLUSION

The primary objective of the proposed research was to attempt to identify differences in subjective perception and objective performance in detect-and-avoid tasks between three different display presentation types of the Conflict Prediction Display System (CPDS). The CPDS application is designed to provide operators with readily actionable Level 3 SA (Endsley, 1995a); however the application allows for varying how the display presents this future information as well as the level of information provided. The investigation was undertaken in order to identify an optimal configuration of the CPDS for Level 3 SA in support of operator detect-and-avoid (DAA) tasks. More specifically, the display configurations were chosen in an attempt to identify the effects of the conflict probes and information level on pilot performance in order to inform future design decisions. While the comparison of Display 1 and Display 3 allowed for the separation of the effects of self-separation information the comparison of Display 2 and Display 3 allowed for the separation of the effects of conflict probes in the plan view and vertical profile view.

A secondary objective was to utilize previously collected data to identify differences in workload and performance between integrated and standalone displays. Although intuitively it may seem optimal to integrate displays which depict related information such as ownship navigation and surrounding traffic as it agrees with the proximity compatibility principle, which states that if a task requires high processing proximity (i.e. different information sources need to be sampled), display integration should be high (Wickens & Carswell, 1995), only a few published papers (Sauer et al., 2002; Friedman-Berg & Ahlstrom, 2005; Fern et al., 2015) have investigated the real-world implications of integrated versus standalone display presentations using high fidelity simulations. In order to add to growing body of literature available, five primary performance and two secondary performance metrics were assessed.

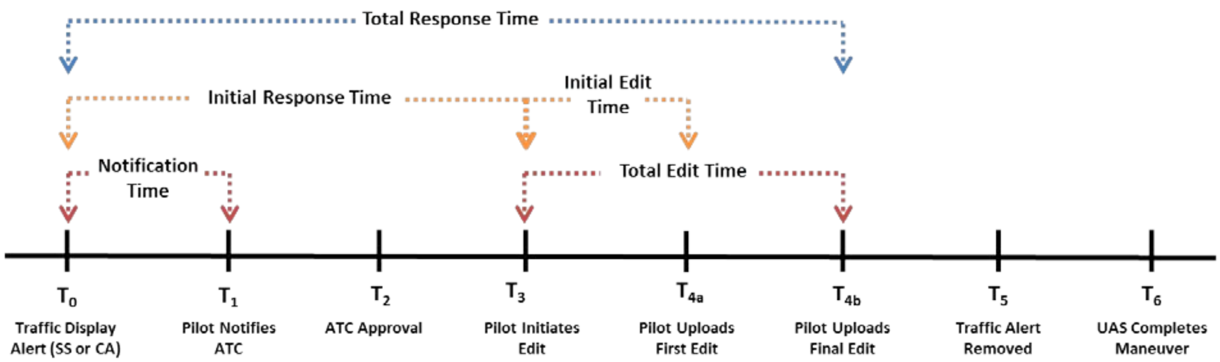


Figure 52 DAA Interaction Timeline

7.1 Notification Time

Notification time was a measure of the time it took operators to notify air-traffic control (ATC) of the appearance of a self-separation or collision avoidance alert and the need to execute a maneuver. Of the 95 total responses, 21 or 22% of them were negative values indicating operators were notifying ATC of a request to maneuver prior to actually getting into an alert state. Display condition 3 contained the largest percentage of pre-emptive notifications (27%). The Display 2 condition did not provide operators with conflict probes and the Display 1 condition which only provided collision avoidance and near mid-air collision (NMAC) alerts, could attribute to the lack of pre-emptive notifications. The uncertainty of the conflict space in Display 2 and the lack of self-separation (SS) conflict probes in Display 1 could have lead operators to become overly cautious in notifying ATC too early. The additional SS conflict probes provided by Display 3 may have lead operators to be conservative in their reactions to the probes on the display which were not actually currently active alerts thereby to contacting ATC sooner.

7.2 Initial Response Time

Initial response time was a measure of the time it took operators to initiate a maneuver response, or edit, in the simulated control interface in response to a traffic display alert. Operators in all the display conditions were also a bit overzealous when accessing the Vigilant Spirit Control (VSCS) interface for modifying the flight path of the simulated UAV. Similar to notification time, of the 102 total responses, 22 or roughly 22% were negative values indicating operators were accessing the interface required to modify the planned flight path prior to actually getting into an alert state. In this case the Display 2 condition contained the highest percentage of pre-emptive accessing of the display (24%). Unlike the Display 1 and Display 3 conditions, the Display 2 condition did not provide operators with conflict probes. This lack of probes could have lead operators to mentally project where the conflict space actually was even though they were provided sufficient information in the no-go heading and altitude bands to identify a safe flight path. This uncertainty of information and additional utilization of mental resources to project the conflict space could have lead operators in this condition to begin accessing the interface for modifying the flight path sooner than the other two conditions.

7.3 Initial Edit Time

Initial edit time is a measure of the time it took operators to input an initial edit in the VSCS interface and send a modified flight path command to the UAV. Although in all three conditions operators were notifying ATC and accessing the VSCS interface prior to actually entering an alert state there were no initial flight path modifications made prior to entering an alert state. In other words, operators did not deviate from their planned flight path until they had breached the self-separation bounds with another aircraft. Given that self-separation alerts are cautionary in nature and do not actually require operators to maneuver this was not a surprising finding. Whether an operator maneuvers their aircraft or not while in a self-separation alert is based strictly on operator preference and their level of comfort occupying airspace in such close proximity to another aircraft. On average, operators in the Display 1 condition were in a self-separation alert almost eight seconds longer than in the Display 2 and Display 3 conditions prior to commanding their first deviation from the planned flight path. The Display 1 condition was the only condition which did not provide a visual indication of self-separation zones either with conflict probes or no-go bands, therefore it is not surprising that operators in this condition waited longer for the encounter to evolve prior to deviation from their flight path.

7.4 Total Edit Time

Total edit time was a measure of the time it took operators to complete an edit in VSCS to maneuver in response to a traffic alert. If operators only made one flight path edit for an encounter this time was equal to Initial Edit Time. This measure was used to account for operators making several flight path modifications to a given alert. Operators in the Display 2 condition (14 alerts with multiple edits) uploaded multiple edits more often than either the Display 1 (9) or Display 3 (8) conditions. Since the method of modifying the flight path was the same across all conditions, this difference is most likely attributed to the difference in the CPDS display conditions (i.e, the lack of probes in the Display 2 condition). Additionally, on average operators in the Display 2 condition required a longer duration to make their final flight path edit for a given alert and in the Display 1 or Display 3 conditions. The Display 2 condition did not provide conflict probes as in the other 2 conditions; therefore operators were forced to mentally project the conflict space in which they were trying to avoid. Although they were provided with heading and altitude bands which would provide them a safe flight path alternative, they were more likely to require multiple edits and a longer duration for editing than the other two conditions. This could potentially be due to the lack of certainty in the conflict space in which they were attempting to avoid.

7.5 Total Response Time

Total response time was a measure of the time it took operators to upload a final maneuver to the aircraft in response to a traffic alert as measured from the initial onset of the alert. Total response time was the longest in Display 1 condition, but only by about four more seconds than the Display 2 and Display 3 conditions. Similar to Total Edit Time, this can probably be attributed to the fact that self-separation alerts were not visually presented to operators in the Display 1 condition. Since this time is measured from the onset of the alert to the final commanded maneuver operators in the Display 1 condition may have been completely unaware they were in a self-separation alert prior to deviating from their planned flight path. Given that self-separation alerts are cautionary in nature and do not actually require operators to maneuver the extended duration of the total response time does not necessarily entail a negative consequence. In fact, minimizing deviations from a planned flight helps air-traffic controllers obtain Level 3 SA in monitoring traffic because they are aware of the planned flight path of all aircraft and can use mental models to project where aircraft should be at some future state. One may presume that this increased duration in the Display 1 condition could have led to more instances of self-separation alerts evolving into collision avoidance (CA) alerts; however, of all three conditions the Display 1 (6 CA alerts across all operators) condition produced less CA alerts than the Display 2 (8 CA alerts across all operators) or Display 3 (8 CA alerts across all operators) conditions.

7.6 Secondary Task Performance

In addition to maintaining safe separation from other aircraft operators were asked to complete simple and complex secondary tasks. Both tasks were presented on the far right display. The simple secondary task required operators to acknowledge a number of alerts. Each trial presented the operators with six warnings and six cautions. Alerts were allowed to timeout after 25 seconds. The complex secondary task required operators to interact with the mouse and keyboard as well as the VSCS display for certain tasks. The complex task presented operators with a simulated chat window with several operators on the far right display. A total of 12 chat messages at random intervals directing them to take a particular action.

All three display conditions were relatively equal in response to both secondary and complex secondary tasks. In all three display conditions operators responded to about half of the simple secondary tasks and since these tasks timed out after 25 seconds the potential for variability was limited. Of the simple tasks responded to response time on average ranged between 10 and 12 seconds. In all three display conditions operators responded to almost all complex secondary tasks. Since these tasks were persistently available once they appeared in the chat window the variability in response time was much greater than in the simple tasks. Response times for all three display conditions were relatively equal.

The pairwise comparisons of the CPDS Displays 1 and 3 and Displays 2 and 3 were able to identify several important factors to help inform future designs. In terms of *information level* the Display 1 versus Display 3 comparisons were designed to tease out the effects of self-separation alert presentation on operator performance. Objectively, operators using Display 3 took significantly less time complete their initial edit than while using Display 1. When multiple edits were made in response to a particular alert, operators also took significantly less time to complete all edits when using Display 3 compared to Display 1. However, subjectively operators perceived Display 1 to be less cluttered than Display 3. Given both displays used conflict probes and “No-Go” bands to present information regarding alerts we should be able to attribute these differences to the addition of self-separation information in the Display 3 condition.

In terms of *information display* the Display 2 versus Display 3 comparisons were designed to tease out the effects of conflict probes on operator performance. Objectively, operators using Display 3 took significantly less time to notify ACT of a request to maneuver and less time to upload all edits for a particular alert than while using Display 2. Subjectively operators believed they were better able to respond appropriately to both self-separation and collision avoidance alerts while using Display 3 as well as better handle all pilot tasks which provides us with information regarding the perceived accuracy between the two displays. In addition, operators also believed they were better able to immediately respond to both self-separation and collision avoidance alerts while using Display 3 which provides us with information regarding the perceived efficiency between the two displays. However, it should be noted that operators also perceived Display 3 to be more cluttered than Display 2. Given both displays presented the same information level (self-separation and collision avoidance alerts) we should be able to attribute these differences to the addition of conflict probes in Display 3.

After all planned analyses were completed several other subjective measures of interest were reviewed to try and better understand operator preference between the display configurations. Across all conditions operators agreed the displays were easy to use and provided the information necessary to maintain safe separation while navigating. Regarding ease of use, ratings were almost identical between Display 1 (M = 4.40, SE = .163) and Display 3 (M = 4.20, SE = .20), however operators did not feel quite as strongly about the ease of use of Display 2 (M = 3.40, SE = .452) although their ratings were still quite high. No significant differences were found between Display 1 and 3 (p = .343) or Display 2 and 3 (p = .153) regarding ease of use.

When asked whether or not the display provided the necessary information for safe separation ratings were again almost identical between Display 1 (*M* = 4.40, *SE* = .163) and Display 3 (*M* = 4.50, *SE* = .167). Similar to ease of use, operators again did not feel as strongly about Display 2 (*M* = 3.40, *SE* = .371) with regard to information availability. No significant difference was identified between Display 1 and 3 ($p = .678$); however, the difference between Display 2 and 3 was significant ($p = .032$) indicating operators were significantly less likely to believe Display 2 provided the necessary information to maintain safe separation compared to Display 3. After all trials were complete operators where given a forced-choice alternative question comparing Display 1 versus Display 3 and Display 2 versus Display 3. Given the choice between Display 1 and Display 3 operators were fairly split on their preference with five preferring Display 1, four preferring Display 3 and one having no preference. However, given the choice between Display 2 and Display 3 operators overwhelmingly chose Display 3 with 8 out of 10 showing preference for Display 3, seven of which indicated a strong preference. From a strictly subjective perspective these forced-choice comparisons indicate operators prefer the display of the conflict probes over just "No-Go" bands; however, there was no clear preference regarding whether or not the probes display self-separation information.

Finally, in order to ensure none of the display differences could be attributed to differences in the training received, operators were asked if they believed they had sufficient training to be able to operate safely during each trial. Across all three display conditions operators strongly agreed they were provided sufficient training to safely operate each display (Display 1: *M* = 4.60, *SE* = .221; Display 2: *M* = 4.30, *SE* = .153: Display 3: *M* = 4.30, *SE* = .153).

Overall, few significant differences were identified; however, other than differences between the integrated and standalone displays in the primary task very few of the comparisons reached statistical significance. However, given the applied nature of the research it is important to go beyond just the statistical comparisons and look at the real-world implications of the data or what some call practical significance.

> To the operational community, SA is not a model or theoretical approach – it has become a foundational tenet associated with keeping your head in the game a key to ensuring operational safety.
>
> The operational community can be incredibly suspicious of academic theorizing, and a research scientist has little chance of affecting an end user unless there exists a "magic decoder ring" to translate the science to actionable components. (Byrne, 2015, p. 85)

Byrne (2015) emphasizes an important point which is not as frequently stressed by the research community which often focuses on *p*-values and effect sizes. One example of this is the data on total edit time for the CPDS display conditions. A review of the data suggests operators in the display condition which did not depict the conflict probes took more than 17 seconds longer to complete all of the edits related to an alert than operators in the display condition which showed self-separation and collision avoidance conflict probes. Seventeen seconds is a long time in a practical sense and framed in terms of aviation safety it becomes even more important. The cruising speed of the MQ-9 UAV simulated in the study is approximately 200 knots, which equates to about a mile traveled. Framing that in terms of the study, operators using the Display 3 configuration would have traveled approximately a mile further while editing their route than operators using Display 2. A 737 sharing the airspace would have traveled almost three miles in this 17 second time period. The next generation UAS which already inhabit the skies in closed airspace cruise at speeds reaching 350 knots. Table 21 provides distance traveled equivalents for the maximum average differences for each metric assessed in the current study.

Table 21. Time to Distance Calculations

Metric	Maximum Avg. Difference	Distance Traveled (MQ-9)	Distance Traveled (Next Gen UAS)	Distance Traveled (737)
Notification Time	11.39s D3 < D1	0.72 miles	1.27 miles	1.85 miles
Initial Response Time	9.24s D3 < D2	0.59 miles	1.03 miles	1.50 miles
Initial Edit Time	6.67s D3 < D1	0.43 miles	0.74 miles	1.08 miles
Total Edit Time	17.32s D3 < D2	1.11 miles	1.92 miles	2.80 miles
Total Response Time	12.46s D3 < D1	0.79 miles	1.39 miles	2.02 miles

Another important variable to consider that was not part of the original planned analysis is the total number of alerts encountered and the alert duration. There were no differences in the number of alerts encountered in the difference experimental conditions. This is not a surprise since this condition was semi-controlled by an air-traffic controller attempting to ensure a minimum of eight encounters per trial. However, a paired-samples t test did reveal a significant difference between the Display 1 condition and Display 3 condition ($p = .015$). In practical terms though, alerts in the Display 1 condition ($M = 69.61$s; $SE = 4.25$) were on average almost 16 seconds longer than alerts in the Display 3 condition ($M = 53.65$s; $SE = 5.01$). Given this comparison was designed to identify the effects of information level (i.e., the addition of self-separation alerts in Display 3) on performance one can conclude presenting self-separation alerts as conflict probes will result in lower alert duration time. An investigation of alert durations between the three CPDS conditions and the two integrated VS conditions also revealed several significant effects. Mann-Whitney tests identified significant differences in duration time between the standalone CPDS Display 1 condition and the VS Basic condition, $U = 1{,}439.00$, $z =$

-3.242, p = .001, the standalone CPDS Display 1 condition and the VS Advanced condition, U = 1,278.50, z = -4.757, p < .001, and the standalone CPDS Display 2 condition and the VS Advanced condition, U = 1,646.00, z = -2.418, p = .016. Both the Kruskal-Wallis and Friedman tests confirmed the findings between the Display 1 condition and VS Basic and Advanced conditions, but did not confirm the finding between the Display 2 and VS Advanced condition. In practical terms, alerts in the Display 1 condition (M = 69.61s; SE = 4.25) were on average almost 26 seconds longer than alerts in the VS Advanced condition (M = 43.74s; SE = 3.58).

So what does all this mean in theoretical terms? Based on Endsley's (1995s) model of SA it is not surprising that few significant differences were found between the three CPDS display conditions. Although situational awareness (SA) is a separate construct from performance and SA and performance are only related probabilistically meaning an increase in SA will increase the probability of an increase in performance, experienced operators such as those used in this study, are often better equipped to translate better SA into better performance (i.e., the probabilistic relation is increased for experienced operators). Given that the three conditions only modified the presentation of Level 3 SA information which is already the highest level SA and operators were all experienced pilots the effect sizes were expected to be small. Had we conducted the study using non-pilots who are inexperienced in dealing with congested traffic in all three dimensions while simultaneously responding to system alerts and air-traffic control requests for information it is likely more significant effects would have been identified; however this trade-off would have required more assumptions when attempting to translate the results to the real-world. (i.e., non-pilots will not be flying UAVs in the national airspace). However, given the technical maturity of the CPDS the primary goal was to attempt to maximize the utility of the Level 3 SA provided to the experienced operator by teasing out the effects of information level and conflict probes. In circumstances where potential lives are at stake it is imperative that we attempt to maximize the operator's ability to deal with an ever changing environment. Although 17 seconds may not be significant based on a Friedman test it may be there difference between a mid-air collision between a UAS and a manned aircraft when UAS are integrated in the NAS.

In order to better understand the potential contribution of the Level 3 SA provided by the CPDS it may be more appropriate to compare the results to the Basic Standalone display evaluated in Part Task 4 which did not provide readily actionable Level 3 SA information. Of the 5 response metrics the best performing CPDS display (Display 3) demonstrated faster average response time than the Basic Standalone condition of Part Task 4 on three (notification time, initial edit time, and total edit time). Regarding alert duration, alerts in the CPDS Display 3 condition (*M* = 53.65s; *SE* = 5.01) were on average almost 10 seconds longer than alerts in the Basic Standalone condition (*M* = 43.44s; *SE* = 3.37).

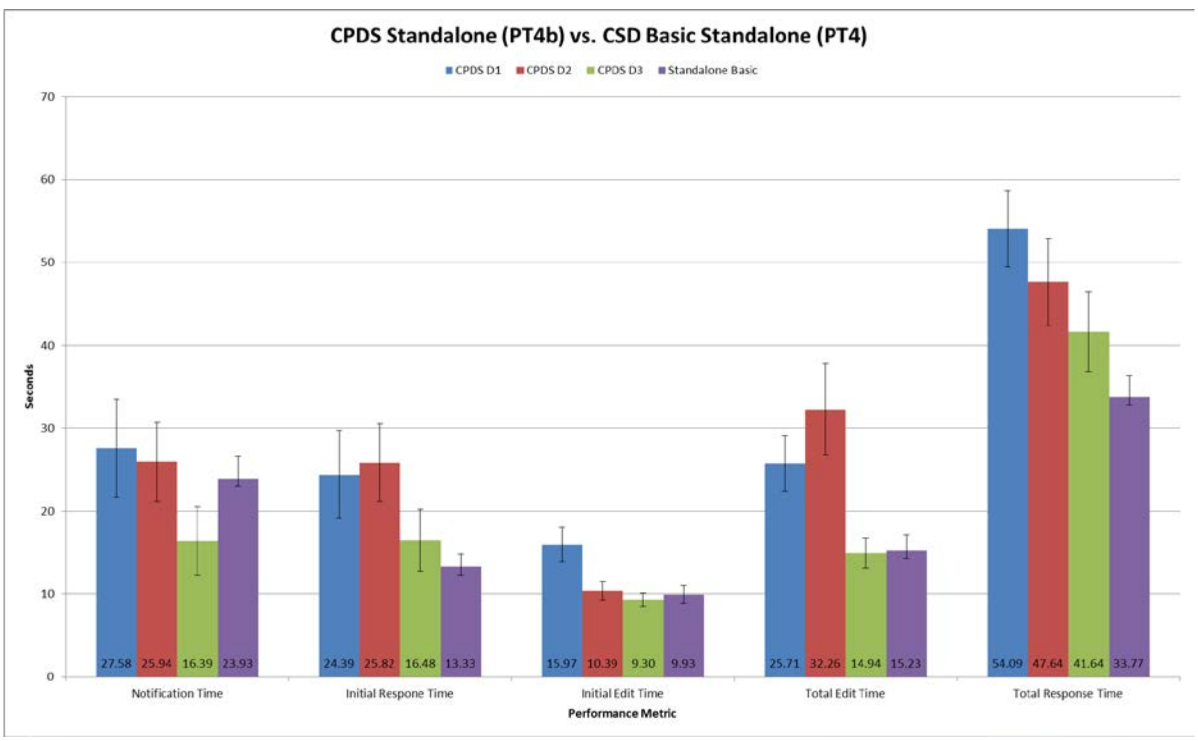


Figure 53. CPDS versus CSD Standalone displays

So why didn’t the standalone display which provided readily actionable Level 3 SA outperform the standalone display which did not provide Level 3 SA? Based on observations and limited feedback provided during the study the initial inclination is the conflict probes were potentially too complex. The conflict probes provided by the CPDS are generated by complex algorithms which take into account many variables (see Tadema, 2011) which are not necessarily intuitively thought of by a pilot. The algorithms can also monitor and take into account far more variable than the human operator which is limited by working memory capacity. Algorithms can also differentially weight certain variables in real-time and the environment changes. For example, the algorithms use a variable to account for intruder track uncertainty, which although may make the probe more accurate it may not be a variable accounted for by the operator watching the probe evolve over time. Due to the complex nature of the algorithms generating the probes, they will often evolve over time in unexpected ways. Although they are accurate they do not necessarily meet the operator’s mental model of how they *should* evolve which can have a negative effect on SA. If the operators do not trust or understand the information being provided to them it is likely they will not utilize to its full potential or take use additional mental resources to attempt to understand why it is doing what it is doing.

In summary, although no new models or theories were developed and the data did not change the landscape of the research literature, the practical implications of the study are apparent. Varying the methodology of displaying Level 3 SA to operators can have significant practical ramifications on exactly how actionable the Level 3 SA information is and whether it can actually detract from operators SA by requiring additional resources to understand the projected information being display and how it fits their mental model or what might be causing it to not fit their mental model. No conclusive evidence was found for the benefit of an integrated presentation over a standalone presentation. Although the Advanced Integrated display displayed significantly better performance than the standalone CPDS, much of this can be attributed to the navigation tool provided in the Advanced display. When an alert appeared in the Advanced display a window would appear notifying the operator of the alert level and giving a suggested heading or altitude. The operator was able to simply accept the suggestion and bypass the normal input methods required for modifying the flight path of the UAV in all other conditions. Given the Basic Integrated display did not outperform the standalone CPDS displays on as many metrics nor to the same magnitude as the Advanced Integrated it would be difficult to attribute the performance differences to the integrated versus standalone variable. Going forward, this information will need to be combined with the additional constraints of the operational community such as level of rigor of testing, software stability, cost and duration of integration when deciding which alternative best balances their global needs of safety, cost and schedule.

REFERENCES

REFERENCES

Association for Unmanned Vehicle Systems International. (2013, March). *The Economic Impact of Unmanned Aircraft Systems Integration in the United States*.

Bell, S., Drury, J., Estes, S., & Reynolds, C. (2012). GDTI: A ground station display of traffic information for use in sense and avoid operations. *2012 IEEE/AIAA 31st Digital Avionics Systems Conference (DASC), Williamsburg, VA,* 8A2-1 – 8A2-13.

Bridges, W. (2014). UAS-NAS Live Virtual Constructive – Distributed Environment (LVC-DE) Part Task 4B Encounters (Unpublished report). NASA Ames Research Center, Moffett Field, CA.

Byrne, E. (2015). Commentary on Endsley's "Situation awareness misconceptions and misunderstandings". *Journal of Cognitive Engineering and Decision Making, 9,* 84-56.

Chiappe, D.L., Rorie, R.C., Moran, C.A., & Vu, K.P.L. (2012). A situated approach to the acquisition of shared SA in team contexts. *Theoretical Issues in Ergonomics Science, 15,* 69-87.

Chiappe, D.L., Strybel, T.Z., & Vu, K.P.L. (2011). Mechanisms for the acquisition of situation awareness in situated agents. *Theoretical Issues in Ergonomics Science, 13,* 625-647.

Chiappe, D.L., Vu, K.P.L., & Strybel, T.Z. (2012). Situation awareness in the NextGen air traffic management system. *International Journal of Human-Computer Interaction, 28,* 140-151.

Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their mutual constraints within the human information-processing system. *Psychological Bulletin, 104,* 163-191.

Dekker, S.W.A. (2015). The danger of losing situation awareness. *Cognition, Technology & Work, 17,* 159-161.

Dekker, S.W.A. & Hollnagel, E. (2004). Human factors and folk models. *Cognition, Technology & Work, 6,* 79-86.

Dekker, S., Hummerdal, D.H., & Smith, K. (2010). Situation awareness: Some remaining questions. *Theoretical Issues in Ergonomics Science, 11,* 131-135.

Draper, D.H., Pack, J.S., Darrah, S.J., Moulton, S.N., & Calhoun, G.L. (2014). Human-machine interface development for common airborne sense and avoid program. *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting, Chicago, IL,* 44-48.

Duggan, G.B., Banbury, S., Howes, A., Patrick, J., & Waldron, S.M. (2004). Too much, too little or just right: Designing data fusion for situation awareness. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, New Orleans, LA,* 528-532.

Efron, B. & Tibshirani, T. (1993). *An introduction to the bootstrap*. New York, NY: Chapman & Hall.

Endsley, M.R. (1988). Design and evaluation for situation awareness enhancement. *Proceedings of the Human Factors and Ergonomics Society 32nd Annual Meeting, Santa Monica, CA,* 97-101.

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

Endsley, M.R. (1995b). Measurement of situation awareness in dynamic systems. *Human Factors, 37,* 65-84.

Endsley, M.R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M.R. Endsley & D.J. Garland (Eds.), *Situation Awareness Analysis and Measurement* (3-32). Mahwah, NJ: Lawrence Erlbaum.

Endsley, M.R. (2015a). Situation awareness: Operationally necessary and scientifically grounded. *Cognition, Technology & Work, 17,* 163-167.

Endsley, M.R. (2015b). Situation awareness misconceptions and misunderstandings. *Journal of Cognitive Engineering and Decision Making, 9,* 4-32.

Endsley, M.R. (2015c). Final reflections: Situation awareness models and measures. *Journal of Cognitive Engineering and Decision Making, 9,* 101-111.

Endsley, M.R. & Jones, D.G. (2004). Designing for Situation Awareness: An approach to user-centered design. Boca Raton, FL: CRC Press.

Endsley, M.R. & Rodgers, M.D. (1994). Situation awareness information requirements analysis for en route air traffic control. *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting, Nashville, TN,* 71-75.

Endsley, M.R. & Smith, R.P. (1996). Attention distribution and decision making in tactical air combat. *Human Factors, 38,* 232-249.

Feitshans, G.L., Rowe, A.J., Davis, J.E., Holland, M., & Berger, L. (2008). *Vigilant spirit control station (VSCS)—'The face of COUNTER'* (AFRL-RH-WP-TR-2012-0015). Air Force Research Laboratory, 711 Human Performance Wing, Wright-Patterson Air Force Base, OH.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Field, A. (20013). *Discovering statistics using SPSS* (4th ed.). Thousand Oaks, CA: SAGE Publications.

Fern, L., Rorie, R.C., Pack, J., Shively, R.J., & Draper, M. (2015). An evaluation of detect and avoid (DAA) displays for unmanned aircraft systems: The effect of information level and display location on pilot performance. *Proceedings of the 15th AIAA Aviation Technology, Integration, and Operations Conference, Dallas, TX*.

Fern, L., Shively, R.J., Johnson, W., Trujillo, A., Pestana, M., & Hobbs, A. (2011). Human factors challenges facing UAS integration into the NAS. *Proceedings of the 16th International Symposium on Aviation Psychology, Dayton, OH*.

Friedman-Berg, F. & Ahlstrom, U. (2005). Evaluating controller use of advanced weather products by evaluating user interaction patterns. *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting, Orlando, FL,* 30-34.

Friedman-Berg, F., Rein, J. & Racine, R. (2014). Minimum visual information requirements for detect and avoid in unmanned aircraft systems. *Proceedings of the Human Factors and Ergonomics Society 58th Annual Meeting, Chicago, IL,* 54-58.

Gillie, T. & Berry, D. (1994). Object displays and control of dynamic systems. *Ergonomics, 37*, 1885-1903.

Granada, S., Dao, A.Q., Wong, D., Johnson, W.W., & Battiste, V. (2005). Development and integration of a human-centered volumetric cockpit display for distributed air-ground operations. *Proceedings of the 13th International Symposium on Aviation Psychology, Dayton, OH*, 279-284.

Gugerty, L. (1998). Evidence from a partial report task for forgetting in dynamic spatial memory. *Human Factors*, *40,* 498-508.

Gugerty, L. & Tirre, W. (1997). Situation awareness: A validation study and investigation of individual differences. *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting, Philadelphia, PA,* 564-568.

Hansen, J.P. (1995). An experimental investigation of configural, digital, and temporal information on process displays. *Human Factors, 37,* 539-552.

Hoaglin, D.C. & Iglewicz, B. (1987). Fine-Tuning some resistant rules for outlier labeling. *Journal of the American Statistical Association, 82,* 1147-1149.

Hockey, G.R. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology, 45,* 73-93.

Jones, D.G. & Endsley, M.R. (1996). Sources of situation awareness errors in aviation. *Aviation, Space and Environmental Medicine, 67,* 507-512.

Kirk, R.E. (2013). *Experimental design: Procedures for the behavioral sciences* (4th ed.). Thousand Oaks, CA: SAGE Publications.

Klein, G.A., Moon, B., & Hoffman, R.T. (2006a). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems, 21,* 70-73.

Klein, G.A., Moon, B., & Hoffman, R.T. (2006b). Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems, 21,* 88-92.

Lambregts, A.A., Tadema, J., Rademaker, R.M., & Theunissen, E. (2009). Defining maximum safe maneuvering authority in 3D space required for autonomous integrated conflict resolution. *2009 IEEE/AIAA 28th Digital Avionics Systems Conference (DASC), Orlando, FL,* 5.C.1-1 – 5.C.1-17.

Lee, J.D. & Sanquist, T.F. (1996). Maritime automation. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (365-384). Mahwah, NJ: Lawrence Erlbaum.

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology, 49,* 746-766.

Lichacz, F. M.J. (2008). Augmenting understanding of the relationship between situation awareness and confidence using calibration analysis. *Ergonomics, 51,* 1489-1502.

Logan, G.D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492-527.

Monk, K., Shively, R.J., Fern, L., & Rorie, R.C. (2015). Effects of display location and information level on UAS pilot assessments of a detect and avoid system. *Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting, Los Angeles, CA,* 50-54.

Parasuraman, R., Sheridan, T.B., & Wickens, C.D. (2008). Situation awareness, mental workload and trust in automation: Viable empirically supported cognitive engineering constructs. *Journal of Cognitive Engineering and Decision Making, 2,* 140-160.

Press, M. (1986). Situation awareness: Let's get serious about the clue-bird. (unpublished manuscript.

Rorie, R.C., & Fern, L. (2015). The impact of integrated maneuver guidance information on UAS pilots performing the detect and avoid task. *Proceedings of the Human Factors and Ergonomics Society 59th Annual Meeting, Los Angeles, CA,* 55-59.

Radio Technical Commission for Aeronautics, Inc. Special Committee 228 (2013). *Detect and avoid (DAA) Whitepaper*. RTCA-SC228-WG1-WP01-01

Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin, 114,* 510-532.

Salmon, P., Stanton, N., & Young, M. (2012). Situation awareness on the road: Review, theoretical and methodological issues and future directions. *Theoretical Issues in Ergonomics Science, 13,* 472-492.

Santiago, C. & Mueller, E.R. (2015). Pilot evaluation of a UAS Detect-and-Avoid (DAA) system's effectiveness in remaining well clear. *11th USA/Europe Air Traffic Management Research and Development Seminar (ATM2015)*, Lisbon, Portugal.

Sauer, J., Wastell, D.G., Hockey, R.J., Crawshaw, C.M., Ishak, M., & Downing, J.C. (2002). Effects of display design on performance in a simulated ship navigation environment. *Ergonomics, 45,* 329-347.

Smeaton, G.P., Dineley, W.O., & Tucker, S.M. (1995). Display requirements for ECDIS/ARPA overlay systems. *Journal of Navigation, 48,* 13-28.

Sorensen, L.G., Stanton, N., & Banks, A.P. (2010). Back to SA school: Contrasting three approaches to situation awareness in the cockpit. *Theoretical Issues in Ergonomics Science, 12,* 451-471.

Tadema, J. (2011). Unmanned aircraft systems HMI & automation: Tackling control, integrity and integration issues. The Netherlands: Shaker Publishing B.V.

Tadema, J. & Theunissen, E. (2008). A concept for UAV operator involvement in airborne conflict detection and resolution. *2008 IEEE/AIAA 27st Digital Avionics Systems Conference (DASC), St. Paul, MN,* 4.C.1-1 – 4.C.1-12.

Tadema, J. & Theunissen, E. (2008). Design and evaluation of a GUI for operator involvement in airborne conflict detection and resolution. *2008 IEEE/AIAA 27st Digital Avionics Systems Conference (DASC), St. Paul, MN,* 5.A.2-1 – 5.A.2-12.

Tadema, J. & Theunissen, E. (2009). Exploring the potential of energy and terrain awareness information in a Synthetic Vision (SV) display for UAV control. *Proceedings of SPIE: Enhanced and Synthetic Vision 2009, Orlando, FL,* 9.1 – 9.13.

Tadema, J., Theunissen, E., & Kirk, K.M. (2010). Beyond traffic depiction: Conformally integrating the conflict space to support level 3 situation awareness. *Proceedings of SPIE: Enhanced and Synthetic Vision 2010, Orlando, FL,* 768902-1 – 768902-16.

Taylor, R.M., Endsley, M.R., & Henderson, S. (1996). In B.J. Hayward & A.R. Lowe (Eds.), *Applied Aviation Psychology: Achievement, Change and Challenge* (447-454). Aldershot, UK: Ashgate Publishing.

Tukey, J. W. (1977) *Exploratory data analysis*. Reading, PA: Addison-Wesley.

Uhlarik, J. & Comerford, D.A. (2002). A review of situation awareness literature relevant to pilot surveillance functions (DOT/FAA/AM-02/3). Federal Aviation Administration.

Ververs, P.M. & Wickens, C.D. (1998). Head-up displays: Effects of clutter, display intensity, and display location on pilot performance. *The International Journal of Aviation Psychology, 8,* 377-403.

Wickens, C.D. (1992). *Engineering psychology and human performance* (2nd ed.). New York, NY: Harper Collins.

Wickens, C.D. (2000). Human factors in vector map design: The importance of task-display dependence. *Journal of Navigation, 53,* 54-67.

Wickens, C.D. (2002).Multiple resources and performance prediction. *Theoretical Issues in Ergonomics Science, 3,* 159-177.

Wickens, C.D. (2005). Attentional tunneling and task management (AHFD-05-23/NASA-05-10). NASA Ames Research Center.

Wickens, C.D. (2008). Situation awareness: Review of Mica Endsley's 1995 articles on situation awareness theory and measurement. *Human Factors, 50,* 397-403.

Wickens, C.D. & Carswell, C.M. (1995). The proximity compatibility principle: Its psychological foundation and relevance to display design. *Human Factors, 37,* 473-494.

Yeh, M. & Wickens, C.D. (2000). Attention filtering in the design of electronic map displays: A comparison of color-coding, intensity coding, and decluttering techniques (ARL-00-4/FED-LAB-00-2). U.S. Army Research Laboratory.

APPENDICES

APPENDIX A

NASA AMES CATEGORY II HUMAN RESEARCH MINIMAL RISK CONSENT

Ames
Research
Center

CATEGORY II - HUMAN RESEARCH
MINIMAL RISK CONSENT

To the Research Participant: Please read this consent form and the attached protocol and/or subject instructions carefully.
Make sure all your questions have been answered to your satisfaction before signing.

A. I agree to participate in the ______ research experiment as described in the attached protocol or subject instructions.

I understand that I am employed by ______ who can be contacted at ______

B. I understand that my participation could cause me minimal risk*, inconvenience, or discomfort. The purpose and procedures have been explained to me and I understand the risks and discomforts as described in the attached research protocol.

C. To my knowledge, I have no medical conditions, including pregnancy, that will prevent my participation in this study. I understand that if my medical status should change while I am a participant in the research experiment there may be unforeseeable risks to me (or the embryo or fetus if applicable). I agree to notify the Principal Investigator (PI) or medical monitor of any known changes in my condition for safety purposes.

D. My consent to participate has been freely given. I may withdraw my consent, and thereby withdraw from the study at any time without penalty or loss of benefits to which I am entitled. I understand that the PI may request my withdrawal or the study may be terminated for any reason. I agree to follow procedures for orderly and safe termination.

E. I am not releasing NASA or any other organization or person from liability for any injury arising as a result of my participation in this study.

F. In the event of injury or illness resulting from this study and calling for immediate action or attention, NASA will provide, or cause to be provided, the necessary emergency treatment. If I am eligible for and receive workers' compensation benefits while participating in this study, I cannot sue my employer because the law makes workers' compensation my only remedy against my employer. I may have other remedies against other persons or organizations, depending on the circumstances of the injury. The United States Government will pay for any claims of injury or loss of life to the extent required by the Federal Employees Compensation Act or the Federal Tort Claims Act.

G. I hereby agree that all records collected by NASA in the course of this study are available to the research study investigators, support staff, and any duly authorized research review committee. I grant NASA permission to reproduce and publish all records, notes, or data collected from my participation, provided there will be no association of my name with the collected data and that confidentiality is maintained, unless specifically waived by me. All stated precautions will be taken to protect anonymity, but there is a small risk that some or all of the participants' data could become identifiable.

H. I have had an opportunity to ask questions and have received satisfactory answers to all my questions. I understand that the PI for the study is the person responsible for this activity and that any questions regarding the research will be addressed to him/her during the course of the study. I have read the above agreement, the attached protocol and/or subject instructions prior to signing this form and I understand the contents.

* ***Minimal Risk*** means that the probability and magnitude of harm or discomfort anticipated in the research are not greater, in and of themselves, than those ordinarily encountered in daily life or during the performance of routine physical or psychological examinations or tests.

Signature of Research Participant — Date	Signature of Principal Investigator — Date
Printed / Typed Name of Research Participant	Printed / Typed Name of Principal Investigator
Telephone Number of Research Participant	Telephone Number of Principal Investigator
Address	Subject Signature: Authorization for Videotaping
City, State, Zip Code	

ARC 475 (JUN 2015)

Previous editions are obsolete

APPENDIX B

PILOT DEMOGRAPHICS FORM

*Please fill in the blanks or circle your response to **each question** below*

PART I - Pilot Experience

1. Age: ________________

2. Do you have manned pilot flying experience: **Yes** **No**

If Yes, please complete the following:

a) Military**:** **Yes** **No**

b) Flight Hours:

Civilian ________ **Military Non-Combat** _________ **Military Combat** _________

Approximate Hours in Civil Airspace (i.e. not restricted or special use) _____________

c) IFR rated**:** **Yes** **No**

c) Other Ratings:

____

____

____

d) Aircraft Types:

____

____

____

____

3. How would you rate your familiarity with flying using the Traffic Collision Avoidance System (TCAS)?

Not Familiar	Somewhat Familiar	Familiar	Very Familiar	Expert

3.a. How many years experience have you had with TCAS? ________________

4. How would you rate your familiarity with flying using other traffic displays?

Not Familiar	Somewhat Familiar	Familiar	Very Familiar	Expert

4.a. How many years experience have you had with other traffic displays?

____

4.b. Which other traffic displays have you used?

____

____

____

5. Do you have UAS flying experience: Yes No

If Yes, please complete the following:

a) Training: **18X Undergraduate Pilot Training Other:** ____________

b) Military: **Yes No**

b) Flight Hours:

Civilian ________ **Military Non-Combat** _________ **Military Combat** _________

Approximate Hours in Civil Airspace (i.e. not restricted or special use)

____

d) Aircraft Types:

____

____

____

____

PART II - Flight Simulation

1. Do you have any desktop flight simulation experience on programs such as MS Flight Sim?

Yes No

If Yes, Please Specify:

a) Number of hours: __________

b) Type:

____

____

2. Do you have any flight simulation experience on rated flight training simulators?

Yes **No**

If Yes, Please Specify:

a) Number of hours: __________

b) Type:

____

____

____

APPENDIX C

PT4B SUBJECTIVE QUESTIONNAIRE – POST TRIAL (CPDS DISPLAY 1)

1. Rate your ability to maintain safe flight operation **without conflicting with intruders:**

Ability was:	Unacceptable ☐	Bad ☐	Satisfactory ☐	Good ☐	Excellent ☐

Comments: ____
____

2. Rate your ability to minimize deviations from the planned path:

Ability was:	Unacceptable ☐	Bad ☐	Satisfactory ☐	Good ☐	Excellent ☐

Comments: ____
____

3. Rate your ability to **handle all pilot tasks** (including chat, radio changes, alerts, range/bearing)?

Ability was:	Unacceptable ☐	Bad ☐	Satisfactory ☐	Good ☐	Excellent ☐

Comments: ____
____

4. Rate the acceptability of the **amount of time provided** to you to adequately assess any Collision Avoidance alerts after they appeared?

Ability was:	Unacceptable ☐	Bad ☐	Satisfactory ☐	Good ☐	Excellent ☐

Comments: ____
____

Rate the extent to which you agree with the following statements about the **AIRSPACE DISPLAY**:

	Strongly Disagree	***Somewhat Disagree***	***Neither Agree nor Disagree***	***Somewhat Agree***	***Strongly Agree***
5. This display was easy to use					
6. This display provided the necessary information to maintain well clear					
7. I did not find this display overly cluttered					
8. I felt I had enough training with this display to be able to operate it safely during this trial					
9. This display allowed me to respond immediately to self-separation threats					
10. This display allowed me to respond appropriately to self-separation threats					
11. This display allowed me to respond immediately to collision avoidance threats					
12. This display allowed me to respond appropriately to collision avoidance threats					
13. The vertical profile display was easy to understand					
14. The conflict probes in plan view were easy to understand					
15. The conflict probes in the vertical profile display view were easy to understand					
16. The “No-Go” heading band was easy to understand					
17. The “No-Go” altitude band was easy to understand					

18. Rate the degree to which each feature contributed to your ability to most appropriately assess and avoid conflicts:

Plan View Conflict Probes	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Vertical Profile Conflict Probes	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Plan View "No-Go" Heading Band	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Vertical Profile "No-Go" Altitude Band	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Vertical Profile Display	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Intruder CPA	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Time to CPA	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐

19. From the following list, please rate the display features in order of preference that offer the **highest potential** to assist you in your ability to remain well clear. **(#1 = highest potential, #2 = second, etc.)**

___Plan View Conflict Probes

___Plan View "No-Go" Heading Band

___Vertical Profile Display Conflict Probes

___Vertical Profile Display "No-Go" Altitude Band

20. Did the **alerting thresholds** used in this trial allow you sufficient time to assess the situation and maintain well-clear?

_____Yes _____No

Comments:

____

____

____

21. Were any information elements **unnecessary** or **confusing**?

22. Were any information elements **missing** that you might need?

23. Please discuss any suggestions for **improving** this display or **any other comments, issues, concerns.**

APPENDIX D

PT4B SUBJECTIVE QUESTIONNAIRE – POST TRIAL (CPDS DISPLAY 2)

1. Rate your ability to maintain safe flight operation **without conflicting with intruders**:

Ability was:	Unacceptable ☐	Bad ☐	Satisfactory ☐	Good ☐	Excellent ☐

Comments: ____

____

2. Rate your ability to minimize deviations from the planned path:

Ability was:	Unacceptable ☐	Bad ☐	Satisfactory ☐	Good ☐	Excellent ☐

Comments: ____

____

3. Rate your ability to **handle all pilot tasks** (including chat, radio changes, alerts, range/bearing)?

Ability was:	Unacceptable ☐	Bad ☐	Satisfactory ☐	Good ☐	Excellent ☐

Comments: ____

____

4. Rate the acceptability of the **amount of time provided** to you to adequately assess any Self Separation alerts after they appeared?

Ability was:	Unacceptable ☐	Bad ☐	Satisfactory ☐	Good ☐	Excellent ☐

Comments: ____

____

5. Rate the acceptability of the **amount of time provided** to you to adequately assess any Collision Avoidance alerts after they appeared?

Ability was:	Unacceptable ☐	Bad ☐	Satisfactory ☐	Good ☐	Excellent ☐

Comments: ____

____

Rate the extent to which you agree with the following statements about the **AIRSPACE DISPLAY**:

	Strongly Disagree	***Somewhat Disagree***	***Neither Agree nor Disagree***	***Somewhat Agree***	***Strongly Agree***
6. This display was easy to use					
7. This display provided the necessary information to maintain well clear					
8. I did not find this display overly cluttered					
9. I felt I had enough training with this display to be able to operate it safely during this trial					
10. This display allowed me to respond immediately to self-separation threats					
11. This display allowed me to respond appropriately to self-separation threats					
12. This display allowed me to respond immediately to collision avoidance threats					
13. This display allowed me to respond appropriately to collision avoidance threats					
14. The vertical profile display was easy to understand					
15. The “No-Go” heading band was easy to understand					
16. The “No-Go” altitude band was easy to understand					

17. Rate the degree to which each feature contributed to your ability to most appropriately assess and avoid conflicts:

Plan View "No-Go" Heading Band	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
VPD View "No-Go" Altitude Band	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Vertical Profile Display	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Intruder CPA	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Time to CPA	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐

18. From the following list, please rate the display features in order of preference that offer the **highest potential** to assist you in your ability to remain well clear. **(#1 = highest potential, #2 = second, etc.)**

___Plan View "No-Go" Heading Band

___Vertical Profile Display "No-Go" Altitude Band

19. Did the **<u>alerting thresholds</u>** used in this trial allow you sufficient time to assess the situation and maintain well-clear?

_____Yes _____No

Comments:

____

____

____

20. Were any information elements **unnecessary** or **confusing**?

____

____

____

21. Were any information elements **missing** that you might need?

22. Please discuss any suggestions for **improving** this display or **any other comments, issues, concerns.**

APPENDIX E

PT4B SUBJECTIVE QUESTIONNAIRE – POST TRIAL (CPDS DISPLAY 3)

1. Rate your ability to maintain safe flight operation **without conflicting with intruders:**

Ability was:	Unacceptable	Bad	Satisfactory	Good	Excellent
	☐	☐	☐	☐	☐

Comments: ____
____

2. Rate your ability to minimize deviations from the planned path:

Ability was:	Unacceptable	Bad	Satisfactory	Good	Excellent
	☐	☐	☐	☐	☐

Comments: ____
____

3. Rate your ability to **handle all pilot tasks** (including chat, radio changes, alerts, range/bearing)?

Ability was:	Unacceptable	Bad	Satisfactory	Good	Excellent
	☐	☐	☐	☐	☐

Comments: ____
____

4. Rate the acceptability of the **amount of time provided** to you to adequately assess any Self Separation alerts after they appeared?

Ability was:	Unacceptable	Bad	Satisfactory	Good	Excellent
	☐	☐	☐	☐	☐

Comments: ____
____

5. Rate the acceptability of the **amount of time provided** to you to adequately assess any Collision Avoidance alerts after they appeared?

Ability was:	Unacceptable	Bad	Satisfactory	Good	Excellent
	☐	☐	☐	☐	☐

Comments: ____
____

Rate the extent to which you agree with the following statements about the **<u>AIRSPACE DISPLAY</u>**:

	Strongly Disagree	***Somewhat Disagree***	***Neither Agree nor Disagree***	***Somewhat Agree***	***Strongly Agree***
6. This display was <u>easy</u> to use					
7. This display provided the <u>necessary information</u> to maintain well clear					
8. I did not find this display overly cluttered					
9. I felt I had enough <u>training</u> with this display to be able to operate it safely during this trial					
10. This display allowed me to respond <u>immediately</u> to <u>self-separation</u> threats					
11. This display allowed me to respond <u>appropriately</u> to <u>self-separation</u> threats					
12. This display allowed me to respond <u>immediately</u> to <u>collision avoidance</u> threats					
13. This display allowed me to respond <u>appropriately</u> to <u>collision avoidance</u> threats					
14. The <u>vertical profile display</u> was easy to understand					
15. The <u>conflict probes</u> in <u>plan view</u> were easy to understand					
16. The conflict probes in the vertical profile display view were easy to understand					
17. The <u>“No-Go” heading band</u> was easy to understand					
18. The <u>“No-Go” altitude band</u> was easy to understand					

19. Rate the degree to which each feature contributed to your ability to most appropriately assess and avoid conflicts:

Plan View Conflict Probes	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Vertical Profile Conflict Probes	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Plan View "No-Go" Heading Band	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Vertical Profile "No-Go" Altitude Band	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Vertical Profile Display	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Intruder CPA	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐
Time to CPA	**Not At All** ☐	**Very Little** ☐	**Some** ☐	**Quite a Bit** ☐

20. From the following list, please rate the display features in order of preference that offer the **highest potential** to assist you in your ability to remain well clear. **(#1 = highest potential, #2 = second, etc.)**

21.

___Plan View Conflict Probes

___Plan View "No-Go" Heading Band

___Vertical Profile Display Conflict Probes

___Vertical Profile Display "No-Go" Altitude Band

22. Did the **alerting thresholds** used in this trial allow you sufficient time to assess the situation and maintain well-clear?

_____Yes   _____No

Comments:

____

____

23. Were any information elements **unnecessary** or **confusing**?

24. Were any information elements **missing** that you might need?

25. Please discuss any suggestions for **improving** this display or **any other comments, issues, concerns.**

APPENDIX F

NASA-TLX

<u>NASA TLX Workload Ratings</u>

Please circle the number that best describes your opinion for each of the questions below.

Mental Demand: *How mentally demanding was the task?*	1 Low	2	3	4	5	6	7 Hig h
Physical Demand: *How physically demanding was the task?*	1 Low	2	3	4	5	6	7 Hig h
Temporal Demand: *How hurried or rushed was the pace of the task?*	1 Low	2	3	4	5	6	7 Hig h
Effort: *How hard did you have to work to accomplish your level of performance?*	1 Low	2	3	4	5	6	7 Hig h
Frustration: *How insecure, discouraged, irritated, stressed, and annoyed were you?*	1 Low	2	3	4	5	6	7 Hig h
Performance Degradation: *How degraded was your ability to meet task goals?*	1 Low	2	3	4	5	6	7 Hig h

APPENDIX G

OUTLIER IDENTIFICATION METHODS

Experimental Condition	*N*	1.5 IQR			2.2 IQR			3 SD			MAD		
		n	***n* deleted**	**% deleted**	***n***	***n* deleted**	**% deleted**	***n***	***n* deleted**	**% deleted**	***n***	***n* deleted**	**% deleted**
Notify Time(sec) - D1	33	31	2	6.06	32	1	3.03	32	1	3.03	31	2	6.06
Notify Time(sec) - D2	34	31	3	8.82	31	3	8.82	33	1	2.94	31	3	8.82
Notify Time(sec) - D3	36	33	3	8.33	34	2	5.56	34	2	5.56	33	3	8.33
Initial RT(sec) - D1	36	33	3	8.33	34	2	5.56	36	0	0.00	32	4	11.11
Initial RT(sec) - D2	37	33	4	10.81	35	2	5.41	35	2	5.41	33	4	10.81
Initial RT(sec) - D3	40	36	4	10.00	38	2	5.00	38	2	5.00	36	4	10.00
Initial Edit(sec) - D1	36	35	1	2.78	35	1	2.78	35	1	2.78	32	4	11.11
Initial Edit(sec) - D2	37	33	4	10.81	34	3	8.11	36	1	2.70	33	4	10.81
Initial Edit(sec) - D3	40	36	4	10.00	38	2	5.00	39	1	2.50	36	4	10.00
Total Edit(sec) - D1	36	34	2	5.56	34	2	5.56	35	1	2.78	34	2	5.56
Total Edit(sec) - D2	37	34	3	8.11	35	2	5.41	36	1	2.70	28	9	24.32
Total Edit(sec) - D3	40	36	4	10.00	36	4	10.00	39	1	2.50	34	6	15.00
Total Response(sec) - D1	36	33	3	8.33	34	2	5.56	35	1	2.78	33	3	8.33
Total Response(sec) - D2	37	34	3	8.11	35	2	5.41	36	1	2.70	34	3	8.11
Total Response(sec) - D3	40	37	3	7.50	38	2	5.00	39	1	2.50	36	4	10.00
Alert Duration(sec) - D1	74	74	0	0.00	74	0	0.00	74	0	0.00	70	4	5.41
Alert Duration(sec) - D2	68	67	1	1.47	68	0	0.00	67	1	1.47	67	1	1.47
Alert Duration(sec) - D3	67	66	1	1.49	66	1	1.49	66	1	1.49	64	3	4.48
Simple RT(sec)- D1	59	59	0	0.00	59	0	0.00	59	0	0.00	59	0	0.00
Simple RT(sec) - D2	53	53	0	0.00	53	0	0.00	53	0	0.00	53	0	0.00
Simple RT(sec) - D3	62	62	0	0.00	62	0	0.00	62	0	0.00	62	0	0.00
Complex RT(sec) - D1	108	102	6	5.56	103	5	4.63	104	4	3.70	97	11	10.19
Complex RT(sec) - D2	114	109	5	4.39	112	2	1.75	112	2	1.75	96	18	15.79
Complex RT(sec) - D3	107	100	7	6.54	102	5	4.67	102	5	4.67	93	14	13.08
Total	**1267**	**1201**	**66**	**5.21**	**1222**	**45**	**3.5**	**1237**	**30**	**2.37**	**1157**	**110**	**8.68**

APPENDIX H

SECONDARY TASK RESPONSE TIME TRANSFORMATIONS

Table 22. Reaction Time (RT) Transformation Comparison (Simple RT)

Transformation Type	Condition	Min	Max	Range	Ratio (Largest/Smallest)
Log(Y)	Simple RT D1	0.037426	1.281942	1.244515	1.349971
	Simple RT D2	0.152288	1.400365	1.248077	
	Simple RT D3	-0.33724	1.342817	1.680059	
Log(Y+1)	Simple RT D1	0.320146	1.304059	0.983913	1.217336
	Simple RT D2	0.383815	1.417306	1.03349	
	Simple RT D3	0.164353	1.362105	1.197752	
Reciprocal(1/Y)	Simple RT D1	0.917431	0.052247	0.865185	3.20341
	Simple RT D2	0.704225	0.039777	0.664448	
	Simple RT D3	2.173913	0.045413	2.1285	
SQRT(Y)	Simple RT D1	1.044031	4.374929	3.330898	1.205175
	Simple RT D2	1.191638	5.01398	3.822343	
	Simple RT D3	0.678233	4.692547	4.014314	

Table 23. Reaction Time (RT) Transformation Comparison (Simple RT)

Transformation Type	Condition	Min	Max	Range	Ratio (Largest/Smallest)
Log(Y)	Simple RT D1	0.037426	1.281942	1.244515	1.349971
	Simple RT D2	0.152288	1.400365	1.248077	
	Simple RT D3	-0.33724	1.342817	1.680059	
Log(Y+1)	Simple RT D1	0.320146	1.304059	0.983913	1.217336
	Simple RT D2	0.383815	1.417306	1.03349	
	Simple RT D3	0.164353	1.362105	1.197752	
Reciprocal(1/Y)	Simple RT D1	0.917431	0.052247	0.865185	3.20341
	Simple RT D2	0.704225	0.039777	0.664448	
	Simple RT D3	2.173913	0.045413	2.1285	
SQRT(Y)	Simple RT D1	1.044031	4.374929	3.330898	1.205175
	Simple RT D2	1.191638	5.01398	3.822343	
	Simple RT D3	0.678233	4.692547	4.014314	

Table 24. Reaction Time Transformation Comparisons (Simple RT)

Transformation Type	Condition	Min	Max	Range	Ratio (Largest/Smallest)
Log(Y)	Simple RT D1	0.037426	1.384353	1.346927	1.389169
	Simple RT D2	0.152288	1.400365	1.248077	
	Simple RT D3	-0.33724	1.396548	1.73379	
	Simple RT VS_Basic	0.117271	1.38952	1.272249	
	Simple RT VS_Advanced	0.127105	1.384891	1.257786	
Log(Y+1)	Simple RT D1	0.320146	1.401917	1.081771	1.209118
	Simple RT D2	0.383815	1.417306	1.03349	
	Simple RT D3	0.164353	1.413635	1.249282	
	Simple RT VS_Basic	0.363612	1.406881	1.043269	
	Simple RT VS_Advanced	0.369216	1.402433	1.033217	
Reciprocal	Simple RT D1	0.917431	0.041271	0.87616	3.211364
	Simple RT D2	0.704225	0.039777	0.664448	
	Simple RT D3	2.173913	0.040128	2.133785	
	Simple RT VS_Basic	0.917431	0.040783	0.876648	
	Simple RT VS_Advanced	0.746269	0.04122	0.705049	
SQRT	Simple RT D1	1.044031	4.922398	3.878367	1.144883
	Simple RT D2	1.191638	5.01398	3.822343	
	Simple RT D3	0.678233	4.991994	4.313761	
	Simple RT VS_Basic	1.144552	4.951767	3.807215	
	Simple RT VS_Advanced	1.157584	4.925444	3.76786	

Table 25. Reaction Time Transformation Comparisons (Complex RT)

Transformation Type	Condition	Min	Max	Range	Ratio (Largest/Smallest)
Log(Y)	Simple RT D1	0.7024	2.0696	1.3672	1.356275398
	Simple RT D2	0.6042	2.2398	1.6355	
	Simple RT D3	0.6734	2.1423	1.4689	
	Simple RT VS_Basic	0.3385	2.1927	1.8543	
	Simple RT VS_Advanced	0.6513	2.1528	1.5015	
Log(Y+1)	Simple RT D1	0.7810	2.0733	1.2922	1.310171155
	Simple RT D2	0.7007	2.2423	1.5416	
	Simple RT D3	0.7569	2.1454	1.3884	
	Simple RT VS_Basic	0.5024	2.1955	1.6931	
	Simple RT VS_Advanced	0.7388	2.1558	1.4170	
Reciprocal	Simple RT D1	0.1984	0.0085	0.1899	2.381858805
	Simple RT D2	0.2488	0.0058	0.2430	
	Simple RT D3	0.2121	0.0072	0.2049	
	Simple RT VS_Basic	0.4587	0.0064	0.4523	
	Simple RT VS_Advanced	0.2232	0.0070	0.2162	
SQRT	Simple RT D1	2.2450	10.8342	8.5892	1.300952657
	Simple RT D2	2.0050	13.1792	11.1742	
	Simple RT D3	2.1712	11.7797	9.6085	
	Simple RT VS_Basic	1.4765	12.4840	11.0075	
	Simple RT VS_Advanced	2.1166	11.9231	9.8065	