Attentional Resource Demands of Traditional and Novel Instrument Landing Systems

A Thesis
Presented in Partial Fulfillment of the Requirements for the
Degree of Master of Science
With a
Major in Psychology
in the
College of Graduate Studies
By

Zachary A. Spielman

Major Professor: Brian P. Dyre, Ph. D.
Committee Members: Rajal Cohen, Ph. D., Steffen Werner, Ph. D.
Department Chair: Todd Thorsteinson, Ph. D.

December 2017

Authorization to Submit Thesis

The thesis of Zachary A. Spielman, submitted for the degree of Master of Science with a major in Psychology and titled "Attentional Resource Demands of Traditional and Novel Instrument Landing Systems" has been reviewed in final form. Permission, as indicated by the signatures and dates given below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor: ____Date:_________
Brian P. Dyre, Ph. D.

Committee Members: ____Date:_________
Steffen Werner, Ph. D.

____Date:_________
Rajal Cohen, Ph. D.

Department Administrator: ____Date:_________
Todd Thorsteinson, Ph. D.

Abstract

A peripherally-located optical flow (POF) instrument landing system (ILS) display affords greater flight performance and sensitivity to a concurrent visual signal detection task, as compared to the display specified by Mil-Std 1787B. This study examined whether such an advantage extends to auditory secondary tasks. During a simulated approach and landing task participants maneuvered an aircraft along a predetermined glide path using either a peripherally-located optical flow instrument landing system or the Mil-Std 1787B landing system. Participants concurrently performed a visual or auditory signal detection task. Participant flight performance and signal detection task sensitivity were measures of the task time-sharing efficiency afforded by the ILS display. The visual task required participants to monitor two engines and identify the simultaneous occurrence of two critical conditions. The auditory task required participants to detect the sequential presentation of an assigned call sign and command line during a stream of call signs and command lines. Sensitivity in visual and auditory monitoring was assessed using the signal detection parameter A, while glide-path deviations served as the measure of landing performance. The mean performance of the flight control task by instrument landing display type matched performance patterns from previous studies. Visual signal detection performance improved while using the POF ILS for the flight task, suggesting a lower visual resource demand in comparison to the Mil-Std ILS. Display type did not affect auditory signal detection task performance. Since the auditory task is similarly affected by both displays in the flight control task but the visual task is more impacted by the Mil-Std display, we conclude the POF ILS imposes a central resource demand like the Mil-Std display but without the same visual processing demands.

Acknowledgements

I would like to thank Dr. Brian Dyre for his continued mentorship throughout my time spent in the graduate program at the University of Idaho. I had the benefit his guidance during studies leading up to my thesis providing experience and perspective to the research process. His commitment to a high standard of work has already begun influencing current.

I would also like to thank the research associates in the Human-in-the-loop Simulation Laboratory as they provided support developing the experiment equipment and running participants while I worked from a distance. Their contribution to this project made it possible.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

ATC	Air Traffic Control
CDI	Course Direction Indicator
FCT	Flight Control Task
HUD	Head-up Display
IFR	Instrument Flight Rules
ILS	Instrument Landing System
IPL	Inferior Parietal Lobule
LGN	Lateral Geniculate Nucleus
POF	Peripheral Optic Flow
SDT	Signal Detection Task
SPL	Superior Parietal Lobule
VDI	Vertical Direction Indicators
MATB-II	The Multi-Attribute Task Battery II

Chapter 1 - Attentional Resource Demands of Traditional and Novel Instrument Landing Systems

Forty-seven percent of fatal aviation accidents occur during final approach and landing (Boeing Corporation, 2014). Nearly half these accidents transpire while piloting under instrument flight rules (IFR): at night or in poor visibility conditions (Boeing, 2002). Half of those 47% are caused by pilot error (Freissinet, 2014). The dangers inherent to this phase of flight are likely exacerbated by pilot workload and fatigue, which are often at a peak during final approach and landing (Hart and Hauser, 1987). Hence, reducing pilot workload during final approach and landing has the potential to reduce pilot error and enhance aviation safety.

Final approach and landing are the most complex and workload intensive portions of flight. Multiple tasks require attention such as monitoring meteorological conditions, plane attitude, system state indicators; communicating with air traffic control (ATC) and executing high-precision visuo-motor control to maintain proper glide path. Performing these tasks requires high-resolution focal vision to process necessary information. Using an instrument landing system (ILS) demands more of the same visual resources (Roscoe, 1980). To determine if reducing the focal visual processing demand on pilots would increase performance, Dyre and his associates proposed and tested an alternative format for the ILS (Bulkley, 2009; Bulkley, Spielman, Dyre, 2011; Spielman, Vargas, Hammack, Bulkley, Lew, & Dyre, 2013; Spielman, Evans, Holmberg, & Dyre, 2014). The peripherally-located display codes ILS information in an optical flow field (Gibson, 2008), using design elements specifically intended to reduce attentional and central visual field load under IFR conditions. This research found that the peripheral display afforded greater landing precision and reduced workload under dual visual task conditions (Bulkley et al. 2009; Spielman et al. 2013).

The Peripheral Optic Flow (POF) display improved both flight performance and sensitivity to a secondary visual task in comparison to a more traditional focal visual display (Spielman et al., 2013) suggesting reduced visual resource demand from the POF. The visual and central processing demand are investigated here by comparing the attentional demand of POF display compared to the traditional Mil-STD1787b ILS display. Using a simulated flight environment, participants followed a predetermined glide-path aided by either the POF or the Mil-Std ILS display, while concurrently monitoring a secondary task displayed either visually or aurally. Measures of attentional demand included flight performance, secondary task sensitivity, and the NASA Task Load Index (TLX) to assess subjective mental workload.

1.1 ILS Displays

Instrument landing systems are designed as guidance support when minimum visibility requirements are not met at the point of final approach and landing. ILSs provide horizontal and vertical guidance to the pilot. A typical ILS has two flight director bars, each tuned to a distinct pair of radio signals (the flight director bars as indicated in Figure 1.1). Landing strips broadcast the signals such that to receive both signals equally the aircraft must be on the correct course and glide path. A flight director bar responds to the strength of both signals. Equal signal strength is indicated by one flight director bar positioned in the middle of the other. To maintain glideslope, the pilot keeps both flight director bars at each other's center. The simulation uses invisible waypoints to replicate the function of the radio signals. The simulation replicates a traditional ILS as specified by the Mil-Std1787b (Mil-Std) head-up display (Figure 1.1). The peripherally-located optical flow display is a concept design tested in the Human-in-the-loop Simulation Laboratory (Figure 1.2).

The Mil-Std uses four display elements to represent vertical and horizontal deviations: A vertical deviation indicator (VDI), a course deviation indicator (CDI) and two flight director bars: one vertical and one horizontal. The VDI consists of five tick marks aligned vertically like rungs of a ladder; the center tick represents the target altitude. An arrow pointing to one of the tick marks indicates the aircraft's location relative to the target altitude; when at target altitude, the arrow indicates the middle tick mark; if altitude is too high, the arrow moves towards upper tick marks and vice versa. The CDI is an arrow pointing towards the center receiving point of the two radio signals, in this case, a waypoint. A straight up and down arrow indicates the waypoint is directly ahead (See Figure 1.1 for example of Mil-Std ILS). The vertical director bar indicates the magnitude of error in the aircraft's lateral position relative to the glide path by its distance from the midpoint of the horizontal director bar. The horizontal bar indicates magnitude of error in the aircraft's vertical position relative to the glide path by its distance from the vertical director bars midpoint. When both flight director bars bisect each other, forming a '+', the aircraft is on glide path (see Figure 1.3). The speed of the flight director bars' movement represents the speed the aircraft is moving in relation to the glideslope.

Bulkley et al. (2009) developed the POF head-up display by adapting a speed magnitude display developed by Cox (2000). The speed magnitude display had fields of arrows that flowed a specific speed and direction depending on the movement of the aircraft. Arrows flowing towards the pilot indicated forward motion. The repurposed display used fields of flowing arrows in four quadrants: the left, right, upper, and lower visual periphery. The fields of moving arrows served the same function as the Mil-Std flight director bars. The vertical flight director bar, coding lateral deviations, was replaced by arrows flowing in the left or right visual periphery. The horizontal flight director bar, coding vertical deviations, was replaced by arrows flowing in the upper or lower visual periphery (see Figure 1.2).

The POF communicated error magnitude in two ways: a) arrow movement speed and b) arrow size. Faster movement and larger arrows indicated greater distances from the glide slope. Slower movement and smaller arrows indicated closer proximity to the glide slope. The aircraft's rate-of-change in position can be determined by the change in arrow movement and size per unit time. The POF display employed sudden jumps in size to capture attention and the location and movement of the arrows to speed information processing. The POF ILS displayed nothing if the aircraft is on the correct glide-path, the Mil-Std's display equivalent of a '+'. Thus, no arrows displayed means no error.

1.2 How the POF Display Exploits Functions of Ambient Vision

The multiple resource model posited by Wickens (1981) discusses two visual resource pools within the visual modality: ambient and focal. When resources necessary to maintain task performance are demanded from the same pool by a concurrent task—for example, reading two spatially-separated displays of independent processes that both require focal visual processing—the model predicts a decrease in performance relative to single task performance on one or both tasks, and increased mental workload. In contrast, when two tasks draw resources from separate pools—for example, reading a display requiring ambient processing concurrently with a display requiring focal processing for two independent processes—the model predicts more efficient time-sharing of the tasks, a smaller performance decrement, and less of an increase in mental workload as compared to sharing two tasks that draw from the same pool of resources. Pilot workload can therefore be reduced by redesigning some tasks to stimulate the ambient visual processes offloading demand on the focal-visual resource pool.

Ambient visual processes have been found to support spatial orientation, motor planning, and object awareness with more direct access to action oriented cognitive systems (Bridgeman, 1992). Spatial mapping and action planning functions have been referred to as the “where” and “how” purpose of ambient processing (Goodale & Milner, 1992), in contrast to the “what” purpose of focal processing. Brain regions associated with executing arm movements and enhancing object awareness are also associated with ambient processing (Rizzolatti & Matelli, 2003). Ambient processing is stimulated by low spatial frequencies occupying large visual areas and is less susceptible to information degradation due to poor illumination or attentional narrowing (Leibowitz, Shupert, & Post, 1982) than focal processing. Ambient processing has been found to require much less attentional resources (Leibowitz, Shupert, & Post, 1982) than focal processing as well. Although many researchers have labeled ambient processing differently, all suggest generally the same functional properties.

The location, size, motion, and sudden onsets present in the symbology of the POF display were specifically designed to stimulate ambient visual processes. Located in the far periphery, the POF is displayed to visual regions known to trigger ambient processes (Leibowitz, Shupert, & Post, 1982). Cells activated by sensors in the periphery of the retina are sensitive to motion and low-spatial frequencies such as a field of flowing arrows. These cells, called M-cells for magnocellular, activate a more direct route to brain regions linked to arm movement, and spatial and object awareness (Rizzolatti & Matelli, 2003) compared to their parvocellular counter parts. One region activated by M-cells, the inferior parietal lobe, uses visual stimuli for planning and executing arm movements (i.e. shifting a throttle up or down in response to peripheral stimuli). The movement or flow of arrow fields and sudden changes in arrow size also serve to capture attention and change the urgency displayed in the POF.

The fixed location of each display and direct response mapping aid in response automaticity as well. The same four regions are always used for the same purpose in contrast to the oscillating and shifting shape of the Mil-Std flight director bars. An intuitive response mapping is employed such that arrows appearing on the right indicate the pilot needs to correct to the right. Using simple, clear, intuitive response mappings reduces the number of rules the pilot must keep in mind to effectively use information from the POF display reducing workload.

Ambient visual processes are more subconscious than conscious, which also reduces operator workload (Leibowitz, Shupert, & Post, 1982). The subconscious processing is described by Leibowitz et al. (1982) as a reflexive response much like regaining stability after losing balance from tripping on an object. Consider also the speed and response time required to regain balance; there is limited time for processing and decision making but an accurate and comprehensive response is still required. Identifying tripping hazards and obstacles entails comparing visually identified objects to knowledge of tripping hazards, an activity requiring conscious processing and attention. Compared to the focal visual system, the reflexive ambient visual processes are less susceptible to information loss under high stress situations when processing orientation information (Lebowitz et al., 1982). As a spatial orienting sense, ambient processes are connected to vestibular and somatosensory to carry out the “where” and “how” functions described by Goodale and Milner (1992). The POF display aids the “where” function with a sudden onset of a flowing field of arrows in a far region of the periphery to stimulate ambient processing. The “how” function is supported by intuitive response mapping to the location of that field.

Subconscious visual processing is also demonstrated in a study involving a patient with severe left focal visual neglect. The patient was presented with two houses with identical right sides but differing left sides. One house had a left side identical to the right, the other house’s left side was on fire. When asked to select the more livable house the patient consistently answered by pointing to the house not on fire but could not explain their reasoning (Marshal & Halligan, 1988). This study illustrates the potential to deliver meaningful information through visual processing without requiring central information processing resources.

The POF display symbology is designed to stimulate ambient visual processes by catering to its sensitivities. Stimuli are presented in the periphery as large fields of flowing arrows, visual information with low spatial frequency. A deviation from glide-path triggers the sudden, attention capturing appearance of light and motion. As deviations increase, the display increases in saliency and urgency with sudden onsets of larger arrows and increasing flow speed. The peripheral location itself indicates the direction needed to return to glide path aiding in response compatibility. Finally, as ILS functions to assist pilots during the spatial orientation of a vehicle in relation to glide path and ambient visual processes have been found to manage spatial orientation functions.

1.3 Human Performance with the POF ILS

The POF display has demonstrated an advantage over the Mil-Std display during simulated aircraft descents (Bulkley, 2009). A simulated aircraft-landing scenario using simplified lateral and vertical 1^{st}-order up/down or left/right controls (with exponential lag) to make lateral and vertical translations was used to compare the two displays' functionality as ILSs. The environment simulated a linear landing approach from initial position to the beginning of the runway during a starless and moonless night with only the runway and background city lights visible. Bulkley compared three conditions within subjects using three different displays; POF, Mil-Std, and Cruise HUD (no ILS). The Cruise HUD has all the information presented in the head-up display except for the ILS. Compared to the Mil-Std and Cruise HUD, the POF improved flight performance of participants. A test comparing participant performance when using the POF in different peripheral locations demonstrated the display remained equally effective in the far periphery (Bulkley et al., 2011). Three visual angles were tested: a) a large-format display that showed the field of flowing arrows within rectangular regions defined relative to the center of the HUD, with the lateral flight indicator (left and right regions) subtending +/-5 to 62.5° by +/-0 to 16.875° (HxV) and the vertical flight command indicator (top and bottom regions) subtending +/-0 to 45° by +/-6.875 to 16.875° (HxV); b) a near peripheral display occupying in the inner half of the large format display; and c) a far peripheral display occupying the outer half of the large format display (see Figure 1.4). Both studies used the same flight simulation configuration. All following studies utilized the far peripheral location.

The far peripheral display has two advantages: 1) it does not compete with any existing instrumentation reducing the potential for clutter and 2) it further removes information from the central visual field. Minimizing clutter reduces potentially hindering pre-attentive visual processes that guide attention to important regions (Beck, 2008). Freeing the central visual field may enhance a pilot's ability to recognize meaningful patterns, segment important information, and search for needed information in the cruise HUD using focal vision because of reduced clutter (Rosenholtz, 2007).

Although comparing flight performance during each display type is a good indicator of general processing demands, it does not provide evidence of the visual resource demand imposed by either display. To measure visual resource demand, a secondary visual task was added. Comparing the performance impact each ILS display has on the secondary visual task measures the attentional time-sharing efficiency of each display. Additionally, the dual-task method is sensitive to differences in workload due to greater difficulty of performing both tasks.

Performance on a secondary visual task provided a measure of visual resource demand imposed by each display (Spielman et al., 2013). The visual task replicated the engine monitoring task used in the Multi-Attribute Task Battery II (MATB-II) Software for human performance and workload Research (Santiago-Espada, Myer, Latorella, & Comstock, 2011). Consistent with previous studies the POF display afforded equal or less error maintaining glide path than a Mil-Std display. Additionally, participants demonstrated greater detection sensitivity to the secondary task when using the POF display versus the Mil-Std, suggesting the POF imposes less visual resource demand freeing participants to allocate more resources to the signal detection task. Perhaps the source of the advantage is POF display qualities making use of the ambient visual system while performing a secondary task requiring the focal visual system.

The POF display was designed to trigger ambient visual processes whereas the Mil-Std requires focal vision (Leibowitz et al., 1984). The POF display supported greater landing performance and secondary task sensitivity compared to the Mil-Std focal display (Spielman et al., 2013). Thus, it is reasoned that splitting visual tasks between the focal and the ambient system improves participant ability to time-share tasks. Performance on a secondary auditory task provides a measure used to compare the attentional demand of both displays. Comparing the attentional demand between displays using an auditory secondary task will offer insight to the central processing demands of the two displays.

The control configurations previously used to compare the POF and Mil-Std displays required participants to determine only the direction of the glide path in relation to the vehicle. To correct, the participants used 1^{st}-order controls holding the joystick left or right until the aircraft was realigned, then releasing the self-centering controller to halt vehicle movement. A realistic approach and landing procedure involves higher order controls, complicating the task requiring three types of information; direction of glide slope, the vehicle's rate-of-change in relation to glide slope, and the magnitude of error from glide slope.

An effective ILS provides clear information regarding three aspects of the aircraft's relative position to the glide path: a) the magnitude of error, b) the rate at which error is changing, and c) the direction of the error. The distance between the aircraft and the glide path is the magnitude of error. The speed at which the aircraft is moving towards or away from the glide path is the error rate-of-change. Increasing error and decreasing error indicate if the vehicle direction is moving away from or closer to glide path respectively. Pilots beginning their realignment with glide path must constantly refer to the ILS to make correct adjustments to their heading and position.

A common example of understanding the magnitude, rate of change, and direction of error is changing lanes on a multi-lane highway. Knowing the number of highway lanes to cross and the direction to turn the wheel is like knowing the magnitude and direction of error. Using the magnitude and direction of error a driver determines how much to turn the wheel, shifting the orientation of the vehicle out of line with the highway. The driver then judges the car's rate of approach toward the target lane. Using experience and the rate of approach the driver synchronizes entering the target lane and reorienting with the highway. However, to do so the steering wheel must rotate past center alignment then back for a full correction due to the yaw rotation that occurs when the wheel is turned; the driver must rotate the car the opposite direction to realign with the new lane. Such steering characteristics make changing lanes (lateral position) a 2^{nd}-order control task.

Where changing lateral position with an automobile requires second order control, with an airplane this maneuver requires third order controls. More control elements must be coordinated to change position: a lateral correction alone requires that bank, pitch, and yaw maneuvers be done in a coordinated manner. Excessive motions can cause rotations around the roll, yaw, and pitch axes. As control order increases, a pilot requires perception of higher-order time derivatives of position—the rate and acceleration of approach to the target position—to predict the trajectory of the airplane and make successful course corrections.

Therefore, an ILS must present information about rates of change in an efficient manner to best facilitate the pilot's precise higher-order control over the glide path.

To examine the influence of higher-order control on the effectiveness of the POF and Mil-Std displays, Spielman et al. (2014) replicated the experiments of Bulkley (2009, 2011) but with an approach path incorporating a waypoint requiring a 54-degree left turn and initiating descent to the runway. Added to the simulation were 2nd order controls allowing participants control over yaw, rather than first order control of lateral position to maneuver the aircraft through a turn. The higher order control configuration and complex flight path required participants to use all the information provided by the ILS to make effective corrections. Performance when using the POF display either equaled or surpassed that of the Mil-Std display (Spielman et al. 2014).

The results of Spielman et al. (2013; 2014) were consistent with the finding of Bulkley et al. (2009) that the POF display produced equal or superior glide path control as compared to the Mil-Std display for simulated instrument approaches and landings. Furthermore, the POF ILS also allowed increased sensitivity in detecting critical engine states, suggesting that it lowered visual processing demand. Finally, participants' raw NASA TLX scores (Hart, 2006), collected after each condition, indicated a significantly reduced subjective mental workload when using the POF ILS, during both single and dual task conditions, corroborating the detection sensitivity measure. The Spielman (2013) results suggest that the POF reduces workload because the display takes advantage of the ambient system. Pairing an ambient visual task with a focal visual task may reduce competition for visual attention, whereas pairing two focal visual tasks (Mil-Std and visual SDT) does not.

Chapter 2 Identifying Resource Demands of ILS Displays

Ambient visual processing is thought to have more direct influence on spatial organization and motor-response than focal visual processing suggesting that the POF should reduce central processing demands relative to the Mil-Std ILS when paired with any secondary task - either visual or other. The POF is hypothesized to reduce attentional demand through ambient visual processing that requires less conscious processing and more directly activates spatial orientation and motor-response to spatial stimuli.

Both flight control tasks (FCTs) demand visual processing resources. The Mil-Std ILS demands focal visual resources, while POF was designed to offload focal visual resources by taking advantage of the ambient vision’s direct pathway to spatial and motor-response centers of the brain.

Maintaining glide slope is a spatial orientation task; hence, POF display is designed to stimulate ambient visual processing to perform. Both FCTs require continuous manual/spatial responses that directly reflect the information provided within the FCT. However, if the POF functions as designed then responses will be reflexive, much like a well-rehearsed skill, whereas the Mil-Std responses are more similar to rule based decision making demanding more processing resources (Rasmussen, 1983).

The limited resource pools predicted by Wickens’ multiple resource model (Wickens, 2008) and supported by other models (Baddeley, 2003) help explain how reducing overall attentional and processing demand afforded by the POF supports better overall secondary task performance. A single resource pool has an information processing limit that, when reached, becomes selective of priority information. As a result, lower priority task performance begins to suffer. Central processing resources are also limited. We can use these processing models to categorize how information is attended to and prioritized as well as relative resource capacity of the perceptual and central processing channels by comparing task performance.

Secondary tasks function to increase the information load when a single task does not exceed a resource pool’s processing limit, forcing attentional prioritization of information. Comparing single task performance to task performance when carried out concurrently can reveal task priority as well as dual task resource demand compatibility. Typically, with visual tasks such as flying an aircraft, we expect decreased performance when paired with another visual task due to exceeding visual processing resource limits. We expect less performance decrement when minimizing resource competition such as when performing two tasks not sharing a modality and thus the same resource pool. The following study explored the attentional demand differences of pairing an ambient visual (POF display) and a focal visual task (Mil-Std display) with focal visual and auditory secondary tasks.

2.1 Flight control task

The flight control task employed the same control-order configuration and flight path structure as Spielman et al. (2014). The higher-order-control configuration along the simulated Moscow-Pullman airport flight path required participants to extrapolate higher order time derivatives to successfully execute the landing task. Flying the aircraft required a predictive strategy incorporating both magnitude and rate-of-change of error. Flight control task difficulty was adjusted to induce resource-limited performance decrements rather than data-limited performance decrements. The control configuration was adjusted after pilot testing to ensure participants could successfully land the plane but with difficulty (Spielman et al., 2014). The purpose of the more complex approach path was to increase resource demand such that the task was performed proficiently in isolation but required task prioritization with the addition of a concurrent task.

2.2 Visual signal detection task

The visual signal detection task (SDT) burdened visuo-spatial processing resources. The visual secondary task's single to dual task decrement measured the visual resource demand imposed by a given display. Comparing the decrement size difference between displays indicated which demanded more visual resources. Since the visual SDT required focal visual processing, the comparison reveals the focal visual resource cost differences between focal-focal visual task pairings and ambient-focal visual task pairings.

One meter in front of the participant, a 9”-diagonal LED screen displayed the visual signal detection task (SDT). The visual task required focal attention and the head down location made receiving visual stimulus from both displays simultaneously difficult. The visual task simulated status indicators of two engines based on the MATB-II system monitoring task (Santiago-Espada, Myer, Latorella, & Comstock, 2011). Each engine appeared as a pair of vertical rectangles representing an engine parameter; temperature or pressure. All four rectangles had “high” and “low” areas indicated by tick marks near the top and bottom of the rectangle representing “critical-areas”. A red indicator bar within each rectangle oscillated up and down in a pseudo random, seemingly independent, fashion. If a single indicator bar stopped within a critical area the respective engine was in “critical condition”. Pulling the trigger button on the joystick indicated the participant recognized a “double-critical condition” when both engines simultaneously entered a critical condition (see Figure 2.1).

2.3 Auditory signal detection task

Simulating a simplified air-traffic-control interaction, the auditory task burdened phonological loop processing resources. The auditory secondary task’s single to dual task decrement measured the central resource demand imposed by a given display. Auditory task resource demand should have less overlap with the FCT resource demands compared to the visual task. Adding an auditory task and observing the task decrement pattern across the visual and auditory tasks helped determine the resource demand of each ILS display.

The design of the auditory task was analogous to the visual task design. The two engines of the visual task display equated to the two word categories: color and command. The same pseudo random patterns used to move the engine parameter indicator bars combined words from each category. Aurally presented phrases occurred at the same rate an indicator bar entered a critical zone in the visual task. The participant pulled the joystick trigger to indicate hearing their assigned words called in sequence (Figure 2.2). Zhang and Mueller’s (2005) ‘A’ served as the SDT sensitivity measure.

This experiment used a dual-task paradigm to compare the resource demands imposed by the Mil-Std and POF displays. The implications of an ambient visual display suggest less impact on both secondary tasks by the POF display than the Mil-Std display used for the FCT. Furthermore, secondary task performance decrements should be equal when using the POF display since ambient visual processes require no attentional demand therefore interfering with neither task. Task performance decrements in the dual task condition demonstrated participants were performing at or near their resource limitations during the single FCT. Analyzing how the tasks are time-shared in each dual task condition helped specify the type of resource demand imposed by each ILS.

Chapter 3 Method

3.1 Participants

Participants were recruited by word of mouth, Sona-systems experiment recruiter software, and social media notifications. Participants were paid $8 per hour and offered a $12 completion bonus after the study totaling $60 payment for completing the study. The participants ranged from 18 to 24 years of age. Participants reported no previous experience in aviation simulations, games, personal computer simulators, previous aviation-related studies, or real aircraft. Using participants with no previous experience with any of the display formats ensures neither the current MIL-STD or POF ILS is advantaged or disadvantaged by participant experience at the beginning of testing. Fourteen participants began the study; two never returned to complete final sessions, leaving twelve that completed all four sessions.

3.2 Design

A 2x2x3 within-subjects experimental design was used in this study. Two ILS displays were used during the FCT, the peripherally located optical flow display and the Mil-Std 1787b display and two signal detection tasks, one visual and one auditory. Each trial was separated into three epochs. The first epoch was characterized by a flat approach to the first waypoint. The second began at the waypoint initiating the turn and descent towards the runway. The third was the final runway approach. The four possible combinations of the 2x2 factors required four two-hour sessions to ensure each participant received each combination. The counterbalanced factor presentation order assured equal distribution of order variations throughout participants.

Two main effects and four interactions across three flight control patterns (the epochs) were tested. Three measures were used to detect differences between factors; flight performance, signal detection sensitivity, and subjective mental workload. These measures served to determine the relative attentional demand and workload of each ILS display.

The main effects investigated included the impact of ILS display on FCT and secondary task modality on SDT. Analyzing the impact of display on FCT was to determine if the POF or Mil-Std offers an advantage in helping participants find their way to the runway. Comparing how modality impacted the SDT helped determine how comparable in difficulty the tasks were to each other. The interaction effects were used to determine how the ILS display type affected resource availability and attention allocation during the two tasks. Lastly, all effects were compared across all epochs to observe how different flight task characteristics impacted performance and dual task strategy.

The three epochs had distinct patterns (Spielman et al. 2014). The first epoch was the approach to the turn. During this epoch, participants had to correct for the error caused by the disturbances causing their vehicle to transfer vertically or laterally. The second epoch was the turn. During this epoch, participants had to re-orient the vehicle to the flight path. The change in flight path, a left turn and downward trajectory, manifested as a large and sudden injection of error into the data: he 54-degree left turn and addition of a constant downward trajectory meant participants had to correct for error caused by disturbances as well as error associated with re-orienting the vehicle. The final epoch was the landing. During this epoch, participants had to correct for disturbances and continue correcting or, in some cases, maintain heading and a downward trajectory.

3.3 Procedure

Data were collected across four sessions lasting 120 minutes each. Each session contained 11 trials or 3 trial sets: a) a set of three SDT trials, b) a set of five FCT trials, and c) a set of three trials combining the factors from the two previous trial sets. Every trial contained all three epochs. Whether the FCT or the SDT came first was counterbalanced across sessions. The dual task trials always ended the session (figure 3.1).

The training immediately preceded the relevant trial set to ensure information was recent in the participant's memory before performing the task. The Power Point training lasted about five minutes. The participants were given the "physics of flight" training Power Point before the FCT trials began in session one. The FCT also included two practice flights (trial one and two). Previous studies found participants needed two flight trials before becoming proficient in the task (Spielman et al., 2014). Finally, following each trial set the NASA-TLX was administered. Sessions one through three concluded with a short debriefing. Session four had a short debriefing as well as a final questionnaire asking participants to evaluate all tasks.

3.4 Measures

Workload measures included FCT performance, SDT sensitivity, and NASA TLX. FCT performance was measured for accuracy (Amean), precision (SD) and overall performance (RMS). RMS error captured overall flight performance by combining accuracy and precision. SDT performance was measured by ability to detect critical events (‘A’ sensitivity). Last, participants rated each task on the NASA TLX subjective mental workload questionnaire to provide a sense of the perceived workload of each task.

3.4.1 Flight control task performance. Flight control task performance was split into two categories; altitude error and course error. Altitude error was measured in meters. Course error was measured in degrees. The degree of angular deviation from the way point origin created by vehicles distance from glide path determined course error. Therefor the same distance from glide path during an earlier portion of flight results in lower course error than the same distance closer to the landing strip. Three measures explain different characteristics of the participants flight path. Absolute mean error (Amean) measures flight path accuracy or the participant’s magnitude of error from flight path during a trial. Standard Deviation (SD) error represents flight path precision or the extent of oscillation around a participant’s flight path. The root mean square (RMS) error was used to measure participant flight performance. RMS error is a vector sum of two orthogonal dimensions of error: a) Amean and b) SD of error. Using RMS error, we combine these two dimensions into a single score to represent overall flight performance.

3.4.2 Signal detection task performance. To measure signal detection task performance, we used ‘A’ as the detection measure (Zhang and Mueller, 2005). ‘A’ is a measure of participants ability to detect their assigned words in the auditory task or the “double-critical” engine condition in the visual task amidst false alarms. ‘A’ is a non-parametric version of d’, a measure of signal sensitivity. It also solves an issue in A’ by accounting for either perfect hit rates, or perfect false alarm rates. The data violated parametric assumptions so ‘A’ was used as our sensitivity measure. Also, the data contained perfect scores which ‘A’ is better suited to handle during processing (Zhang and Mueller, 2005). Bias in sensitivity was also calculated to better understand participant strategy when performing the secondary task.

3.4.3 Time-sharing efficiency measure. Since time-sharing efficiency is the main comparison of interest, a single measure was developed to reflect it. Doing so simplifies the comparison and interpretation of the results. Performance on the individual tasks was used as the benchmark for each participant’s peak task performance. However, performance of each task was measured differently. To compare between the FCT and SDT, performance needed to be normalized to set each measure equal to the other. After performance was normalized the final time-sharing efficiency measure could be calculated.

To normalize the single task measures we assumed the participant's scores on the single task trial sets represented maximum performance capability. As with flight task performance, any measure of error could be used in the time-sharing efficiency measure. To calculate this, we used the inverse of error (1/error) to represent flight control performance for the dual-task efficiency measure. To normalize dual-task flight control performance relative to single task performance, we divided the error from the dual-task trial set by the error from the single task trial set for each participant to determine a percentage score.

To normalize the dual-task signal detection sensitivity measure (A) we also divided it by the sensitivity measured during the single task condition to create a percentage score. Because task sensitivity already increases as performance increases, the detection sensitivity measure will not be inverted like the error measure used in the flight control task.

Dual task performance is then illustrated using two-dimensional plots of the performance operating characteristic (POC). A POC plot represents the normalized performance on each task on separate orthogonal axes. Points in this two-dimensional space thus represent combined dual-task performance for a task combination. We used a POC graph to record all the now equally weighted, normalized performance values. The single task normalized values will be represented by triangles marking 100 percent performance on each axis (FCT on the y-axis, SDT on the x-axis). The two component tasks of each dual task condition will act as coordinates placing them within the graph. A special coordinate on the POC graph represents the theoretically perfect concurrent task performance and is derived from each single task value or (1, 1). Each task measure was now equally weighted and was combined into a single point on the POC graph. The single time-sharing efficiency performance measure is the distance from the origin to a coordinate location. The distance value was converted into the single time-sharing efficiency measure using the same method to compute each participant's trial set point on the graph. The distance of each dual task condition divided by the distance of the theoretical perfect time-sharing point from the origin yields a single percentage score representing time-sharing efficiency. The single measure is sensitive to changes in performance and can be compared across conditions and participants regardless of how each participant prioritizes the FCT and SDT.

The time-sharing efficiency measure was designed with the expectation of finding differences between ILS displays in the FCT and differences in the SDTs. Our intent was to use it to aggregate those differences into a single measure to simplify interpretation. However, no reliable differences were found between ILS displays; hence, the differences between SDT modalities is sufficient to interpret time-sharing efficiency. The time-sharing efficiency measure is redundant and provides no additional explanation of results; therefore it was not included in the final analyses.

3.4.4 Subjective mental workload

The Raw TLX (RTLX) was used to evaluate subjective mental workload. The RTLX is more sensitive and provides better experimental validity than the weighted scales (Bustamante & Spain, 2008). The results are based on the average of all subscales.

3.5 Stimuli and Apparatus

The simulated environment consisted of a 288 x 72 (X, Z) km island. The beginning of the 30 m wide runway began at the center of the island (0, 0, 0) and extended along the z-axis 950 m ending at 0, 0, -950 (X, Y, Z). White and yellow-lit spheres, spaced at 50 m intervals, lined the first and second half respectively, of the 250 m long runway. Green and red spheres, spaced at six-meter intervals, were arranged in a line spanning the full width of the beginning and end of the runway, respectively. Great reduction in visibility was due to simulated moonless and starless nighttime conditions with fog obscuring the view of the surface and runway until the end of flight.

The visual flight environment was hosted on a single computer running ViEWER 2.25 (Dyre, Grimes, & Lew, 2013). The host computer coordinated three front-projected images rendering a spatial resolution of 1400-1050 pixels (H x V), each at a refresh rate of 60 Hz. The images were projected to three screens arranged as a partial octagon with the viewpoint at the center of the partial-octagon, 1.8 m from the center of each screen all together subtending 135 x 33.75 degrees of visual angle (H x V) effectively replicating the same displays used in Spielman et al. (2013). The environmental and monochromatic green HUD characteristics were replicated from Spielman (2013) to maintain continuity across experiments. A fourth display hosted the engine monitoring task displaying a 4.36 x 4.36 degree visual angle.

Simulated lateral and vertical wind disturbance increased task difficulty and required continuous pilot corrections. Translating the vehicle independently along the lateral (x) and vertical (y) axes produced the disturbances. Each disturbance was defined by a sum-of-sines with five prime frequencies. The maximum acceleration of the disturbances did not exceed 9.8 m-2 (1 G). The disturbance displacement amplitudes for the x-and y-axes were 0.8, 7.023, 4.116, 3.225, 2.772, 0, and 25.847, 7.9494, 2.394, 1.318, 0.753, 0 respectively. The frequencies for the x- and y-axes were 0.055, 0.085, 0.145, 0.185, 0.215, 0, and 0.035, 0.065, 0.115, 0.205, 0 respectively.

The new initial vehicle location was (-3500, 1720, 12750) oriented at 54.45 degrees with waypoints positioned at (-3500, 1720, 12750), (0, 1720, 10200), and (0, 0, 0). The flight-path required 2nd-order controls to accommodate the 54.45 degrees turn involved. The higher fidelity control configuration used left-right movements of a CF-F-16 joystick with the right hand to control lateral heading by combining a zero-order roll with gain = 20 degrees at maximum deflection and exponential lag constant = 0.5 s^{-1} with a first order yaw with gain = 5 degrees s^{-1} and exponential lag constant = 0.1 s^{-1}. This configuration slowly rolled the simulated plane congruent with direction of the command input to a maximum 20 degrees tilt simultaneously causing the plane to yaw at a maximum rate of 5 degrees s^{-1}. Forward-backward movements of the CF-Pro throttle with the left hand will affect first-order (velocity) altitude changes with gain = 15 ms^{-1} and exponential lag constant = 0.01 s^{-1}. Forward movement, increasing input, of the throttle caused upward vehicle translations (simulating increased thrust) while backward movement, decreased input, of the throttle caused downward vehicle translations (simulating less thrust). However, because our population consists of non-pilots, the controls remained simplified in comparison to actual pilot-aircraft control systems.

The visual signal detection task replicated the engine monitoring task used in MATB-II (Santiago-Espada, Myer, Latorella, & Comstock, 2011), a task designed to evaluate operator workload. The display, placed in a head down location, did not allow participants to fixate vision on both the engine monitoring and flight control task simultaneously. On the display was the status of two engines. On each parameter of each engine (four total) were red indicator bars that oscillated up and down in a pseudo random fashion occasionally resting at the top, above the ‘high’ tick mark and bottom, below the ‘low’ tick mark. Two engine parameters, temperature and pressure, were monitored for indications of a critical event. A critical event occurred when the temperature or pressure dropped below the ‘low’ tick mark or move above the ‘high’ tick mark. Only one engine parameter entering the critical area qualified as a critical event. The participant was instructed to monitor the display for a “double critical event” which occurred when each engine entered a critical condition at the same time. Participants were instructed indicate a double critical condition by pulling trigger. ViEWER 2.25 tracked clicks of the trigger button.

The secondary auditory task consisted of a bank of 10 command phrases and 10 call sign phrases of common usage. Participants were assigned a target command phrase and call sign phrase. Participants were exposed to different pseudo random pairings of vocalized command and call sign phrases to which they chose to either respond to or ignore. Participants were instructed to pull the trigger if they heard the target command and call phrase voiced in sequence. ViEWER 2.25 tracked trigger pulls. The same pseudo random algorithm that moved the red indicator bars of the visual monitoring task was used to activate different pairs of command and call phrases throughout a given trial resulting in wording pairings being uttered every two seconds. Using the same program for both signal detection tasks ensured the target pairing occurred with the same frequency as the double-critical condition of the visual monitoring task (Figure 2.2).

Chapter 4 Results and Discussion

This section discusses how data were handled after collection, how outliers treated, and how reliability was assessed. The data were cleaned of outliers and verified as a reliable data set for testing differences between the two FCT and SDT factors. Outliers were defined as trials that participant lateral flight path tracking performance exceeded 40 degrees course error; a "point-of-no-return" threshold. Previous studies have demonstrated such outliers are a result of misunderstanding of the task, lack of attention to the task at hand, or an unknown event that caused a participant to veer past a point of reasonable recovery (Spielman, 2014). It is reasonable to expect some outliers in performance due to fatigue or moments of inattention when a study requires four sessions, 44 trials, and nearly four and a half hours of simulator time.

4.1 Data Treatment

Participant trials were recorded by ViEWER 2.25 and saved in .dat files. An in-house tool analyzed the files, breaking each trial for every participant into three epochs organized in rows. Altitude error and course error were analyzed separately. Organized by column were the ILS, SDT, dual or single task, and session performed, Amean, SD, and RMS. The file format and organization allowed for GUANO (Lew & Dyre, 2005) or Graphical User ANalysis Of variance program software to perform ANOVAs on the data. The α threshold was set at 0.05 for significant determinations and all values used Greenhouse-Geisser corrections for violations of sphericity.

Determined a priori, the first two of the five single FCT trials on each day were not considered in the analyses, as the previous study involving the same control scheme and flight path found participants needed two trials to learn the flight task well enough to proceed with the experiment (Spielman et al., 2014).

Outlying trials were removed from analyses due to course error patterns demonstrating the participant was unable to complete the trial by reaching the runway. All data for a trial that exceeded 40 degrees course error at any point during flight was removed from analyses (average course error in total was 3.644 degrees). Table 1 lists the trials that did not meet the criterion. Each of the four participants had four or fewer trials cut except for participant nine. Participant nine had 14 of 44 total trials cut from analysis making it clear that participant nine was an outlier from the other participants. Due to disproportionate number of removed trials, participant nine was left out of the analysis entirely.

4.2 Learning effects

Trial and session were factors analyzed for learning effects or if participants had reached task proficiency with the training and practice allotted them. Ensuring no learning effects are present increases the reliability of any results found due to the factors of interest. The process also confirmed participants had sufficient training to perform the tasks during data collection trials. The data was also analyzed for order effects. Although all factors are counterbalanced across and within sessions, removing outliers can unbalance the design. Once the data is verified the factor effects can be analyzed.

A common practice to ensure a learning effect does not cause a type one error is to counterbalance the order of conditions across sessions such that any learning would affect all conditions equally. Participant performance within a trial set is also tested for learning effects. If there is not a learning effect then the data analyses pertaining to the hypotheses are more reliable. Ensuring no learning effects occurred also indicates training was sufficient and effects found in the data are not caused by participants still learning the task.

4.2.1 Trial

A within-subjects ANOVA was run for the single and dual task trial sets for each task. Altitude error and course error were both analyzed for course precision (SD) and overall flight performance (RMS). Statistics are reported were significant differences were found.

4.2.1.1 FCT. A within-subjects ANOVA with factors task load (single or dual) and trial (three, four, five) demonstrated no differences across trials in either the single or dual FCT condition. No difference between trials suggests the participants reached a performance plateau by trial three, four, and five in the single task condition and were proficient by the dual task condition.

4.2.1.2 SDT. A 2 x3 within-subjects ANOVA with task load (single or dual) and trial (one, two, three) as factors demonstrated no differences across trials in either the single or dual SDT condition were found. The result suggests training was effective or the task was easy to learn.

4.2.2 Session

4.2.2.1 Single FCT by session. No differences in altitude performance across sessions were found. An effect was found in course error. A within-subjects ANOVA, $F(3, 10) = 3.620$, $MSE = 2.889$, $p = 0.049$, $\eta^2 = 0.124$ demonstrated significant differences in course precision between sessions. Sessions three and four were significantly different than session one and two when measured by SD (refer to figure 4.1). All tasks were counterbalanced across session such that equal trials of both ILS were performed on day three hence the effect cannot be attributed to ILS. The results suggest that participants were still getting better at the flight control task although this did not show up in the analyses by trial.

4.2.2.2 Dual FCT by session. A within-subjects ANOVA showed no effect across session for the dual FCT condition. The pattern of means does show a slight improvement as the days progress. The pattern and result suggest that participants quickly established a strategy to perform the two tasks concurrently and used it throughout the sessions.

4.2.2.3 Single SDT by session. A within-subjects ANOVA showed no effect by session for single SDT sensitivity. Participants were performing at high success rates suggesting the training was sufficient to get performance up to proficiency or the task was simple enough to perform alone.

4.2.2.4 Dual SDT by session. An ANOVA found no effect of session on Dual SDT sensitivity. However, the trend suggests participants were still building a strategy to perform the task concurrent the FCT. The dual FCT performance matches the dual SDT performance pattern.

4.3 Single Task Display Comparison

RMS and SD altitude and course error were analyzed using 2 x 3 within-subjects ANOVA with ILS display (POF, Mil-Std) and epoch (1, 2, 3) as factors. Statistics are reported where significant differences were found.

No significant effect of display on FCT performance was found for any measure. However, as shown in table 2 the trends all favor the POF display. The difference between means in altitude RMS error was nearly twelve meters or, roughly a four-story building suggesting there may be some advantage to using the POF ILS. Course error differences remained within a degree between ILSs. The trend of the means is consistent with previous patterns of performance along the same flight path using the same control configuration (Spielman, Evans, Holmberg, & Dyre, 2014). The result suggests a potential lack of power to observe the effect of ILS.

A within-subjects ANOVA demonstrated significant differences between epochs in altitude error for SD, $F(1.089, 10.890) = 42.735$, $MSE = 246.981$, $p > 0.001$, $\eta^2 = 0.845$, and RMS, $F(1.075, 10.745) = 31.648$, $MSE = 852.593$, $p > 0.001$, $\eta^2 = 0.789$. Epoch two was significantly different from epoch one and three in altitude error (Figure 4.2). The increased error in epoch two lead to further investigation of display effect of FCT performance. The difference between ILS displays was not significant but the difference between means suggest an advantage of the POF display when recovering from large amounts of error or handling larger amounts of control inputs at a single time (Figure 4.3). Managing altitude during epoch two required participants to maintain a constant downward trajectory therefor manage a sudden introduction of continual control input. The regular disturbances still shifted the vehicles location vertically in a virtually unpredictable manner adding a second source of error to respond to. In this way, flight control during the second epoch can be characterized as error recovery and different from the first two epochs.

A within-subjects ANOVA demonstrated significant differences between epochs in course error for SD, $F(1.181, 11.813) = 147.023$, $MSE = 2.07$, $p > 0.001$, $\eta^2 = 2.473$, and RMS, $F(1.150, 11.5) = 82.704$, $MSE = 2.026$, $p > 0.001$, $\eta^2 = 2.341$. Course error in epoch one had reliably less error than epoch two and three (figure 4.4). However, ILS performance remained similar in all epochs resulting in no significant difference. The way course error is measured may explain the pattern of course error across the three epochs.

Participants are performing the same as previous studies with the same environment and control configuration which is promising for the POF display as it continues to demonstrate lower average error over the Mil-Std. Furthermore, flight precision during the second epoch was better when tracking altitude using the POF display. Demonstrating a control characteristic that may require further investigation to understand how epoch two differs from epoch one and three for altitude. Course error epoch one also suggests improved performance by those using the POF but the effect is unreliable in this study. Again, further investigation is needed to understand the difference in flight characteristics between epochs and how the POF is supporting different behaviors.

4.4 Single Task SDT by Modality

'A' sensitivity was analyzed using 2 x 3 within-subjects ANOVA with SDT modality (Auditory, Visual) and epoch (1, 2, 3) as factors. Statistics are reported where significant differences were found.

Performed alone, SDT sensitivity was similar regardless of epoch and modality. Having equal sensitivity to both tasks despite differences in modality indicates equal task difficulty, which is best for comparing the imposed workload the FCT has on each SDT. The previous study using the single visual SDT task recorded average participant sensitivity at 0.964. Participants in this study performed the single visual SDT at an average of 0.858 demonstrating the participants here performed the task worse than previous studies (Spielman, et al., 2013).

4.5 Performance Decrements

4.5.1 FCT decrement. RMS and SD altitude and course error were analyzed using 2 x 2 within-subjects ANOVA with ILS display (POF, Mil-Std) and task load (single, dual) as factors. The result indicates FCT performance was not impacted by the addition of the concurrent SDT in the dual task conditions. It is likely participants were prioritizing the FCT as they were instructed. A dual task decrement was anticipated hence the time-sharing efficiency measure was created to account for differences in the FCT and SDT. Since no FCT decrement was found differences in SDT performance are enough to draw conclusions regarding our hypotheses.

4.5.2 SDT decrement. 'A' sensitivity was analyzed using 2 x 2 within-subjects ANOVA with SDT modality (Auditory, Visual) and task load (single, dual) as factors. A within-subjects ANOVA demonstrated significant main effect of sensitivity by task load (single, dual), $F(1, 10) = 11.656$, $MSE = 0.039$, $p = 0.007$, $\eta^2 = 0.232$. The test did not demonstrate a significant main effect of SDT. However, it did reveal a significant interaction between SDT modality and task load, $F(1, 10) = 5.747$, $MSE = 0.001$, $p = 0.037$, $\eta^2 = 0.046$. The result indicates participants indeed prioritized the FCT over the SDT as instructed but the auditory SDT showed less resource conflict resulting in higher task sensitivity (Figure 4.5).

The POF display was expected to support greater time-sharing efficiency than the Mil-Std. Hence, the magnitude of the decrements were also compared to identify if one ILS had a significantly greater impact on the visual SDT than the other ILS. No significant difference in the decrement cause by the ILSs was found either. A non-significant result suggests that the POF does not improve the participants ability to perform concurrent visual tasks. As noted before, the average SDT sensitivity by ILS display suggests an advantage of the POF consistent with all previous studies. However, we do not have the power in this study to rely on the means.

Both ILS displays caused a significant decrement in auditory SDT sensitivity. The auditory task decrement was less than the visual task as is expected since visual processing is required for the FCT as well. The modality bottleneck does not explain the auditory task decrement, suggesting there is common demand for central processing resources between the auditory and FCT requiring the participant to prioritize concurrent tasks. The ILSs made the same impact on the auditory tasks as well, suggesting the source of the resource interference is the same.

4.6 Dual Task Interaction

4.6.1 Flight performance measures. A 2 x 3 within-subjects ANOVA with ILS display (POF, Mil-Std) and epoch (1, 2, 3) as factors was used to analyze altitude and course RMS and SD error. Statistics are reported where significant differences were found.

Despite the SDT condition, participants reliably performed the dual FCT the same. The pattern of flight performance across epoch matched the single FCT pattern (Figure 4.2). The within-subjects ANOVA demonstrated significant differences between epochs in altitude error for SD, $F(1.075, 10.753) = 46.234$, $MSE = 287.061$, $p < 0.001$, $\eta^2 = 0.697$, and RMS, $F(1.082, 10.824) = 22.020$, $MSE = 2783.583$, $p = 0.001$, $\eta^2 = 0.422$. Due to the significant difference of epoch two, performance was compared by display during epoch two alone using a 2 x 2 ANOVA comparing FCT displays and SDT modality performance during epoch 2 (Figure 4.6). The test demonstrated a non-significant difference in maintaining altitude precision (SD) during epoch two $F(1, 10) = 4.566$, $MSE = 495.665$, $p = 0.058$, $\eta^2 = 0.092$. Although non-significant, the evidence towards the POF supporting greater performance, corroborated by the single FCT performance during epoch two, suggests there may be an advantage to the POF display when recovering from large amounts of error or handling larger amounts of control inputs at a single time.

The course error pattern across epochs also matches the single task (Figure 4.4). Epoch 1 is significantly different from epoch two and three. As with the single task course error, no difference between ILS display was found across epochs. Note again that the way course error is measured is a possible explanation for the increasing error by epoch, especially once epoch two presents an additional task of reorienting it is difficult to recover to near perfect glide-slope upon reaching the runway.

4.6.2 Dual SDT. A 2 x 2 within-subjects ANOVA measured differences in 'A' with SDT modality (Auditory, Visual) and FCT (POF, Mil-Std) as factors. The analysis demonstrated a significant main effect of SDT $F(1, 10) = 7$, $MSE = 0.002$, $p = 0.024$, $\eta^2 = 0.036$. Secondary task sensitivity was impacted by both ILS displays equally. The analyses demonstrated no interaction effect as well.

Display impact was analyzed within each SDT modality separately using an ANOVA test. The analysis revealed the Mil-Std FCT had a significantly greater impact on the visual SDT $F(1, 10) = 6.464$, $MSE = 0.003$, $p = 0.029$, $\eta^2 = 0.084$. The same analysis within the POF condition demonstrated a non-significant effect of display on sensitivity $F(1, 10) = 7$, $MSE = 0.001$, $p = 0.577$, $\eta^2 = 0.004$. The mean sensitivity of both SDTs during the POF FCT nearly match the mean sensitivity of the auditory SDT during the Mil-Std display condition (See figure 4.7). The POF display impacting both modalities the same as the Mil-Std display's impact on the auditory but not the visual SDT suggests the resource demand for all three conditions is similar. If true, then the resource bottleneck is not modality related. This result supports the hypothesis that the POF display utilizes the more direct ambient visual system demanding only minimal central processing resources.

4.7 NASA TLX Comparisons

Subjective mental workload was analyzed using 2 x 2 within-subjects ANOVA with FCT (POF, Mil-Std) and task load (single, dual) as factors revealing no difference between the ILS displays or task load.

Subjective mental workload was analyzed using 2 x 2 within-subjects ANOVA with SDT modality (Auditory, Visual) and task load (single, dual) as factors demonstrated a main effect of SDT $F(1, 10) = 19.153$, $MSE = 55.604$, $p = 0.001$, $\eta^2 = 0.083$. and a main effect of task load $F(1, 10) = 33.254$, $MSE = 139.253$, $p < 0.001$, $\eta^2 = 0.363$ (Figure 4.8). The test did not result in an interaction effect. The auditory task had a lower rating than the visual task

(Figure 4.8). As expected, participants gave a much higher mental workload rating to the dual task condition over the single task.

The single visual SDT had a higher workload rating despite participants performing better on it compared to the auditory SDT. The visual task maintained the higher rating in the dual task condition, consistent with the poorer performance observed when the visual SDT was combined with another task. The results show the visual task imposed greater mental workload on participants than the auditory.

Chapter 5 General Discussion

We compared a POF display designed to stimulate ambient visual processes to reduce pilot workload during final approach and landing phases of flight to the Mil-Std1787b display. Participants using the POF display had less RMS and SD error across all conditions compared to the Mil-Std, particularly when the FCT difficulty increased during epoch two. Participants showed greater task-sharing efficiency between engine monitoring task and the POF display and compared to the Mil-Std display, as indicated by greater sensitivity during the visual SDT. However, the POF time-sharing advantage did not extend to the ATC monitoring task. Therefore, the POF display showed a reduction in focal visual processing demand but did not reduce non-visual attentional demand. Ambient processing of the POF therefore required at least some attentional resources.

The pattern of FCT performance suggests the POF display reduces pilot altitude error compared to the Mil-Std display. This pattern replicates findings from previous comparisons of the two displays (Bulkley, 2009; Spielman et al., 2013; Spielman et al., 2014). The most pronounced difference of altitude error between displays occurred during epoch two, a period characterized by increased control input requirements. Thus, the POF display supports greater altitude tracking performance and matches the Mil-Std display supporting course corrections when no other orientating information is present.

The POF supported greater time-sharing efficiency with the visual SDT compared to the Mil-Std display. The finding suggests the POF display’s focal visual resource demand is less than the Mil-Std visual resource demand. Therefore, the POF display may be more suited to situations in which many focal visual tasks are present. Reducing flight path error and increasing pilot visual resource availability would be an advantage in the cockpit.

However, participants time-shared the auditory SDT equally well regardless of ILS type. The finding suggests the POF imposes a non-visual demand equal to the Mil-Std thus not removing attentional demand all together. Either ambient visual processes were not the only resource used to process the POF display or ambient processing itself does requires some central processing. The pattern of means on the flight control task suggest the POF can more efficiently time-share two tasks of separate modalities in comparison to the Mil-Std however more trials are required before the FCT data is reliable source for comparison of auditory task time-sharing differences between the POF and Mil-Std displays.

Targeting ambient visual processing was intended to reduce attentional demand in both visual and central processing channels. While the POF does move resource demand away from focal-visual channels, a central processing burden equal to that imposed by the Mil-Std still exists. As found in this experiment, the POF does impose a burden, but less than the Mil-Std. After testing more participants, the FCT performance differences may prove more reliable demonstrating a reduced overall demand of the POF assuming all else remains the same.

The results of this study are based on small sample size and an uneven design due to the removing a participant for too many failed trials. Had time and resources permitted, more participants would improve the power in this study making the results more reliable when generalizing to the population. Improving training or reducing FCT difficulty may be two solutions to minimizing the learning effect found for course error across session as well. Since this experiment protocol required four two-hour sessions, perhaps the more feasible option is reducing FCT difficulty and increasing SDT difficulty. The FCT required the most training and extra trials which added approximately 15 minutes to each session. Furthermore, using angular error as the course error measure added noise to the data. Using angular deviation from the flight path does not account for lateral deviation equally across the entire flight path. As the participant approached the waypoint, smaller lateral deviations equaled the same amount of course error. Increasing angular error is then expected if a participant maintains the same lateral deviation during the entire approach to the runway. Lateral deviation from flightpath is recommended for use in future studies. Another limitation is only one instrument can be redesigned as a POF display therefore limiting the potential impact. The feasibility of the POF needs to be investigated further as well. Having fields of flowing arrows displayed about the cockpit may be difficult and/or expensive to implement.

The POF demonstrated more efficient time-sharing with other focal-visual tasks. Since most tasks require focal-visual resources this characteristic has potential to reduce pilot workload in the cockpit when landing under IFR. Furthermore, the POF display will not add additional burden to auditory tasks such as ATC communications. The location of the display can reduce clutter in the cockpit and open space to view other instruments while landing the plane. Lastly, the POF display is not limited to an ILS. Magnitude of error, rate-of-change, and direction are three types of information the POF efficiently delivers. Anywhere such information is needed, the POF has potential to deliver with low focal-visual or auditory resource competition.

References

Baddeley, A. (2003). Working Memory: Looking Back And Looking Forward. *Nature Reviews Neuroscience*, 4(10), 829-839.

Beck, M., Lohrenz, M., Trafton, J. G., & Gendron, M. (2008). The Role Of Local And Global Clutter In Visual Search. *Journal of Vision,* 8(6). Retrieved April 10, 2014, from http://www.journalofvision.org/content/8/6/1071.short

Boeing Corporation, A. (2014, August 1). Statistical Summary of Commercial Jet Airplane Accidents. Retrieved April 13, 2015, from http://www.boeing.com/news/techissues/pdf/statsum.pdf

Bridgeman, B. (1992). Conscious vs Unconscious Processes. *Theory & Psychology, 2*(1), 73-88. doi:10.1177/0959354392021004

Bulkley, N., Dyre, B. P., Lew, R., & Caufield, K. (2009). A Peripherally-Located Virtual Instrument Landing Display Affords More Precise Control Of Approach Path During Simulated Landings Than Traditional Instrument Landing Displays. *Proceedings of the 53rd Annual Meeting of the Human Factors and Ergonomics Society*, 31-35. San Antonio, TX: HFES

Bulkley, N. K., Spielman, Z., & Dyre, B. P. (2011) Peripherally-located Virtual Instrument Landing Displays. *In Proceedings of the 16th International Symposium on Aviation Psychology* (ISAP), Dayton, OH.

Bustamante, E. A., & Spain, R. D. (2008). Measurement Invariance of the NASA TLX. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, 52, 1522-1526.

Cox, E. M. (2000). The effect of a virtual indicator on speed control. (Unpublished Master's Thesis, University of Idaho).

Dyre, B. P., Grimes, J. G., & Lew, R. (2013). ViEWER 2.25: A Virtual Environment Workbench for Education and Research. http://www.webpages.uidaho.edu/~bdyre/viewer.htm.

Freissinet, S. (2014). Statistic analysis of airplane accidents. 1001 Crash. Retrieved May 12, 2014, from http://www.1001crash.com/index-page-statistique-lg-2-numpage-3.html

Gibson, R., Schvaneveldt, R., & Gray, R. (2008). Visual Misperception In Aviation: Glide Path Performance In A Black Hole Environment. *Human Factors*, 50(4), 699-711.

Goodale, M. A., & Milner, A. D. (1992, January). Separate visual pathways for perception and action. Retrieved December 05, 2017, from https://www.ncbi.nlm.nih.gov/pubmed/1374953

Hart, S. G. (2006). NASA-Task Load Index (NASA-TLX); 20 years later. In Proceedings of the *Human Factors and Ergonomics Society 50th Annual Meeting* (pp. 904-908). Santa Monica, CA: Human Factors & Ergonomics Society

Hart, S. G., & Hauser, J. R. (1987). Inflight application of three pilot workload measurement techniques. *Aviation, Space, & Environmental Medicine*, 58(5), 402-410.

Hettinger, L. J., Brickman, B. J., Roe, M. M., Nelson, W. T., & Haas, M. W. (1996). Effects Of Virtually-Augmented Fighter Cockpit Displays On Pilot Performance, Workload, And Situation Awareness. *Proceedings of the Human Factors Society 40th Meeting,* 30-33. Philadelphia, PA: HFES.

Leibowitz H. W., Post R. B. (1982). "The two modes of processing concept and some implications," in Organization and Representation in Perception, ed. Beck J., editor. (Hillsdale, NJ: Erlbaum), 343–363.

Marshall, J. C., & Halligan, P. W. (1988). Blindsight and insight in visuo-spatial neglect. *Nature*,*336*(6201), 766-767. doi:10.1038/336766a0

Rasmussen, J. (1983). Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models. IEEE Transactions On Systems, Man, and Cybernetics, 257-266.

Rizzolatti, G., & Matelli, M. (2003). Two different streams form the dorsal visual system: anatomy and functions.*Experimental Brain Research*, 153(2), 146-57. Retrieved April 10, 2014, from http://download.springer.com/static/pdf/794/art%253A10.1007%252Fs00221-003-1588-0.pdf?auth66=1400300723_5a531c167e92cf5ec6

Roscoe, S. (1980). Aviation Psychology. Ames, IA: Iowa St. University Press.

Rosenholtz, R., Li, Y., & Nakano, L. (2007). Measuring Visual Clutter. *Journal of Vision*, 7(2). Retrieved April 10, 2014, from http://171.67.113.220/content/7/2/17.full

Santiago-Espada, Y., Myer, R. R., Latorella, K. A., & Comstock, J. R., Jr. (2011, July). The Multi-Attribute Task Battery II (MATB-II) Software for ... Retrieved November 7, 2017, from https://matb.larc.nasa.gov/files/2014/03/NASA-TM-2011-217164.pdf

Schaudt, W.A., Caufield, K.J., and Dyre, B.P. (2002). Effects of a virtual speed indicator on guidance accuracy and eye movement control during simulated flight. *Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society*, 318-322. Baltimore, MD: HFES

Spielman, Z., Evans, R., Holmberg, J., & Dyre, B. (2014). Evaluation of a Peripherally-Located Instrument Landing Display with High-Order Control of a Non-Linear Approach and Landing. *Proceedings Of The Human Factors And Ergonomics Society Annual Meeting*, 58(1), 1067-1071. Retrieved April 14, 2015.

Spielman, Z., Vargas, J., Hammack, T., Bulkley, N., Lew, R., and Dyre, B. P., (2013). Evaluation Of Peripherally-Located Instrument Landing Display Under Dual-Task Conditions. In *Proceedings of the 17th International Symposium of Aviation Psychology* (ISAP), Dayton.

Wickens, C. D. (2008). Multiple Resources and Mental Workload. *Human Factors: The Journal of the Human Factors and Ergonomics*, 50(3), 449-455. Retrieved April 11, 2014, from http://hfs.sagepub.com/content/50/3/449.

Zhang, J., & Mueller, S. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. Psychometrika, 70(1), 145-154.

Tables

Table 4.1: Trials failing to meet acceptable performance criterion.

Participant	SDT	FCT	Trial	Total Trials
1	None	Mil-Std	4	2/44
	Vis	Mil-Std	3	
2	None	OF	4	4/44
	Aud	OF	1	
	None	OF	4	
	Vis	OF	1	
9	None	Mil-Std	3, 4, 5	14/44
	Aud	Mil-Std	1, 3	
	Aud	OF	1, 2, 3	
	None	OF	3, 4, 5	
	Vis	OF	1, 2, 3	
10	None	OF	3	3/44
	None	OF	3	
	None	Mil-Std	4	

Table 4.2: Mean errors of single FCT demonstrating direction of means despite lack of significance.

Comparison	Landing Display	Altitude Error (Mean)	Standard Error	Course Error (Mean)	Standard error
Single Task	Optic Flow	38.218	2.543	3.195	0.180
Landing Display (RMS)	Mil-Std	51.406	3.368	3.908	0.236
Single Task	Optic Flow	29.095	1.434	3.073	0.175
Landing Display (SD)	Mil-Std	37.945	2.062	3.646	0.201
Dual Task	Optic Flow	42.358	3.409	3.660	0.297
Landing Display (RMS)	Mil-Std	50.657	3.477	3.728	0.194
Dual Task	Optic Flow	31.380	1.781	3.414	0.238
Landing Display (SD)	Mil-Std	39.360	2.358	3.553	0.190

Figures

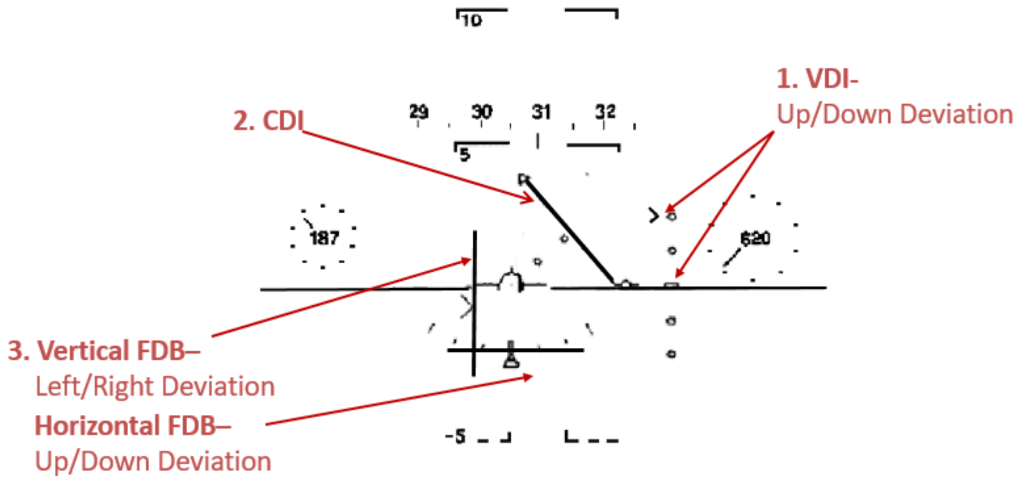

Figure 1.1: MIL-STD 1787 B Head-up Display. 'A' indicates the course deviation indicator (CDI). 'B' indicates the vertical deviation indicator (VDI).

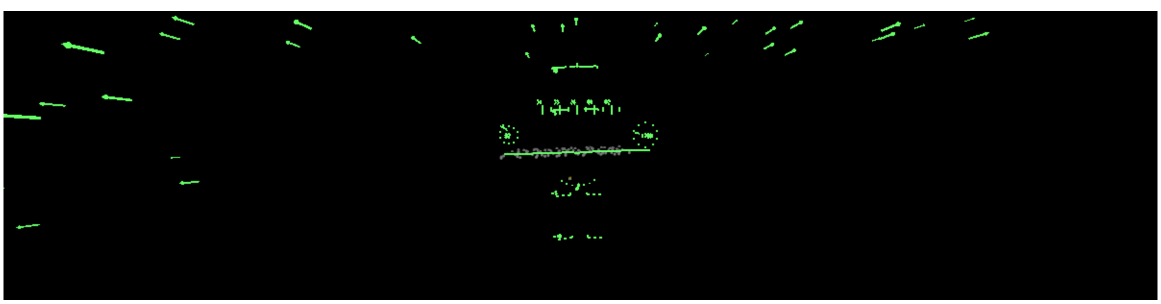

Figure 1.2: The peripheral optical flow display and cruise HUD. The peripheral OF display with the lateral flight indicator subtending +/-33.75 to 62.5° by +/-0 to 16.875° (HxV) and the vertical flight command indicator subtending +/-0 to 45° by +/-11.875 to 16.875° (HxV). Display is indicating the pilot should make upward vertical and leftward horizontal corrections.

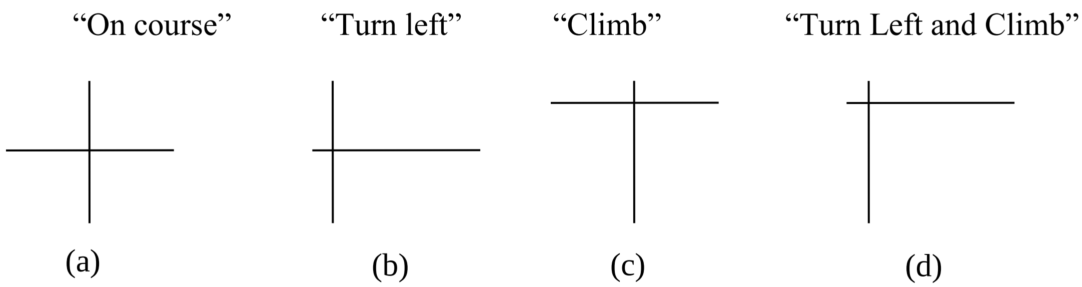

Figure 1.3: The flight director bars of the Mil-STD 1787B instrument landing system. Here is a depiction of what each basic position could mean.

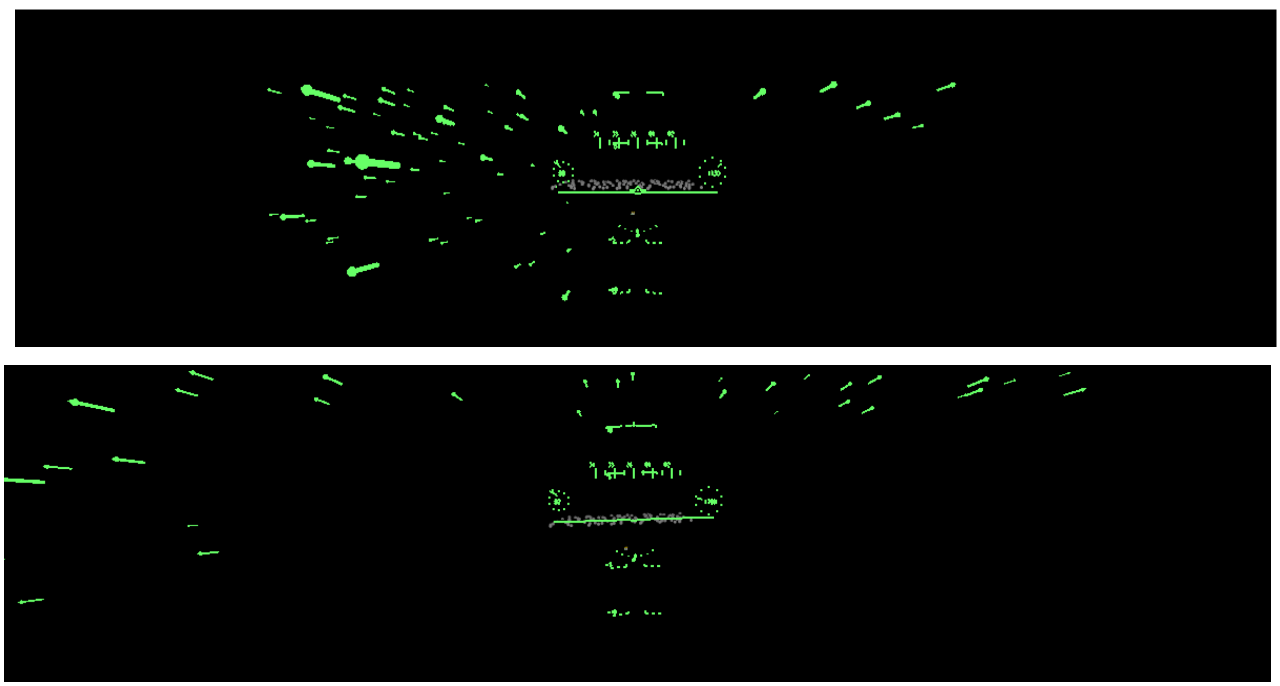

Figure 1.4: An example of the near periphery location (top) and far periphery (bottom) locations tested in Bulkley (2009).

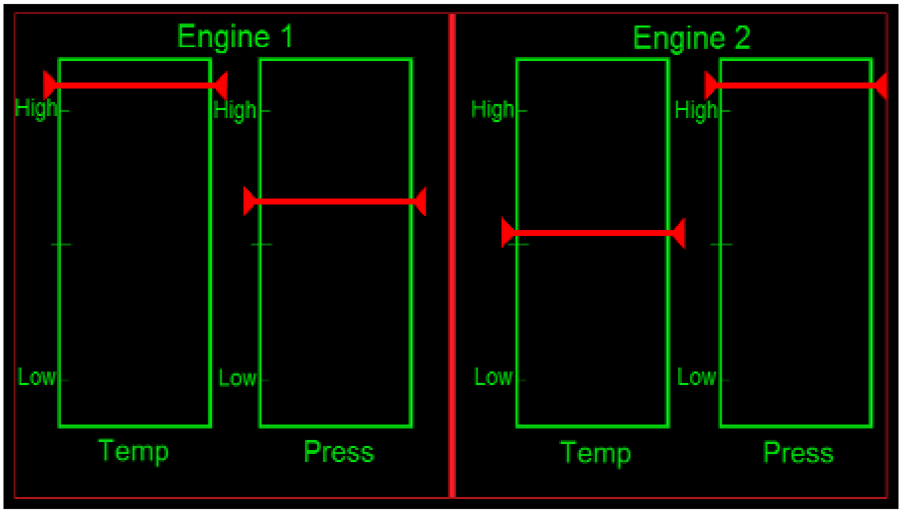

Figure 2.1: The visual engine-monitoring task. It currently displays one example of a double critical condition.

Color	Command
Black	Check
Blue	Climb
Brown	Dive
Gold	Hold
Blue	Check
Black	Open
White	Dive
White	Climb

Figure 2.2: A visual example of the auditory task. The colors and commands are not the exhaustive list of those presented during the experiment. A two second gap occurred between aurally presented pairs.

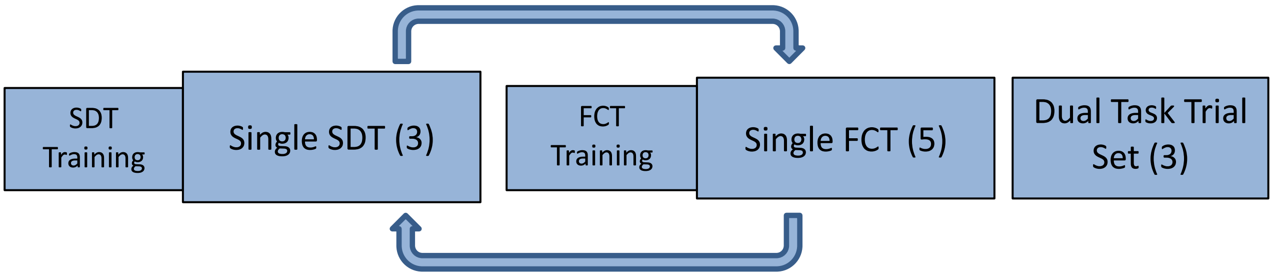

Figure 3.1: Experiment Protocol. Whether the SDT or the FCT was presented first during a session was counterbalanced across sessions within each participant. The modality and display type combination presented during a session was counterbalanced across all participants.

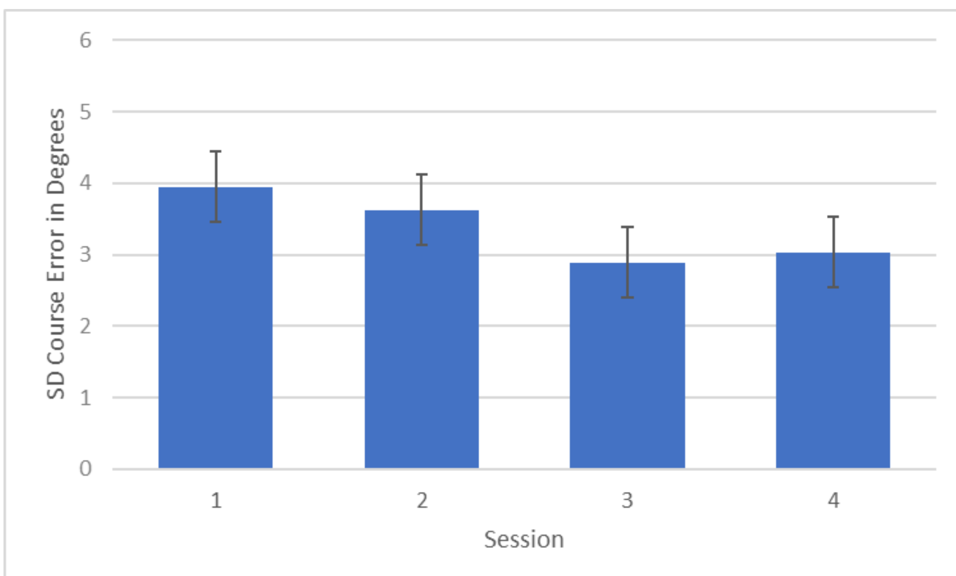

Figure 4.1: Single task SD of course error by session. A significant difference between sessions indicates participant course precision improved up until the third session. Within-subjects ANOVA, $F(3, 10) = 3.620$, $MSE = 2.889$, $p = 0.049$, $\eta^2 = 0.124$ Error bars represent 95% within-subjects confidence intervals.

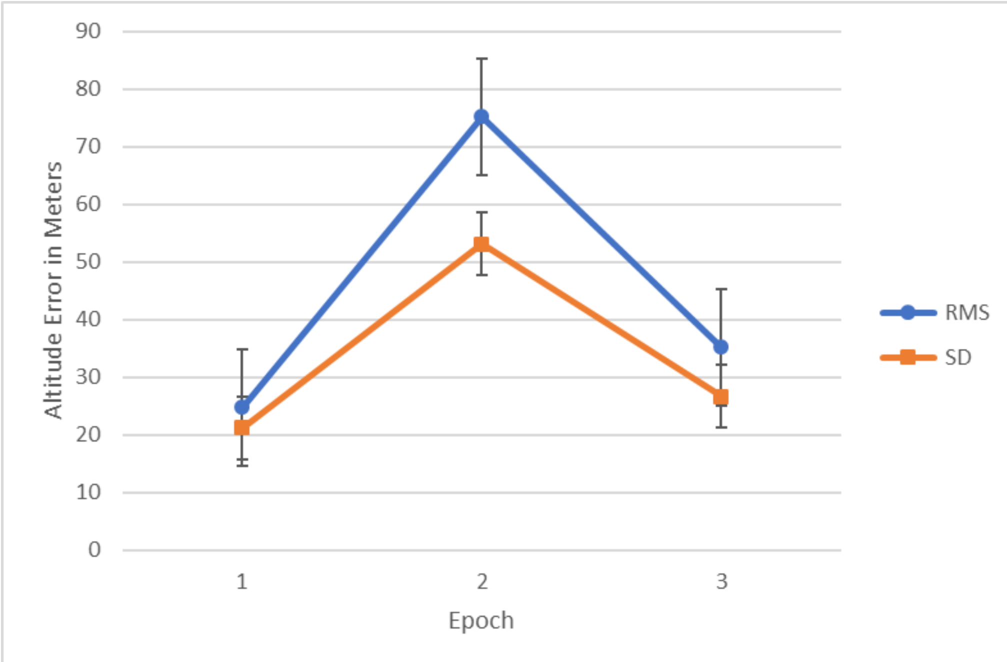

Figure 4.2: Changing altitude error by single task epoch. Epoch 2 error is significantly greater than Epoch 1 and 3. Participants manage to compensate for the introduction of a new glide-path by Epoch 3. Error bars represent 95% within-subjects confidence intervals.

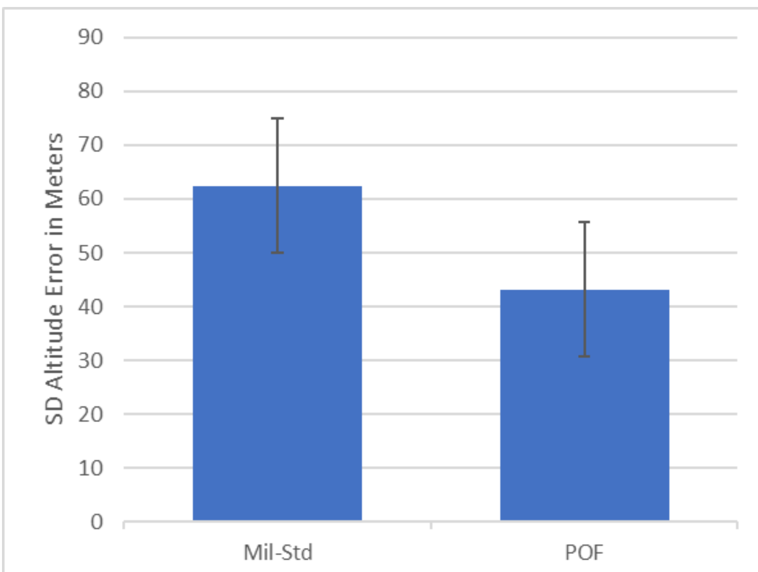

Figure 4.3: Single task comparison within Epoch 2 of display type. $F(1, 10) = 4.309$, $MSE = 343.156$, $p = 0.065$, $\eta^2 = 0.182$. Error bars represent 95% within-subjects confidence intervals.

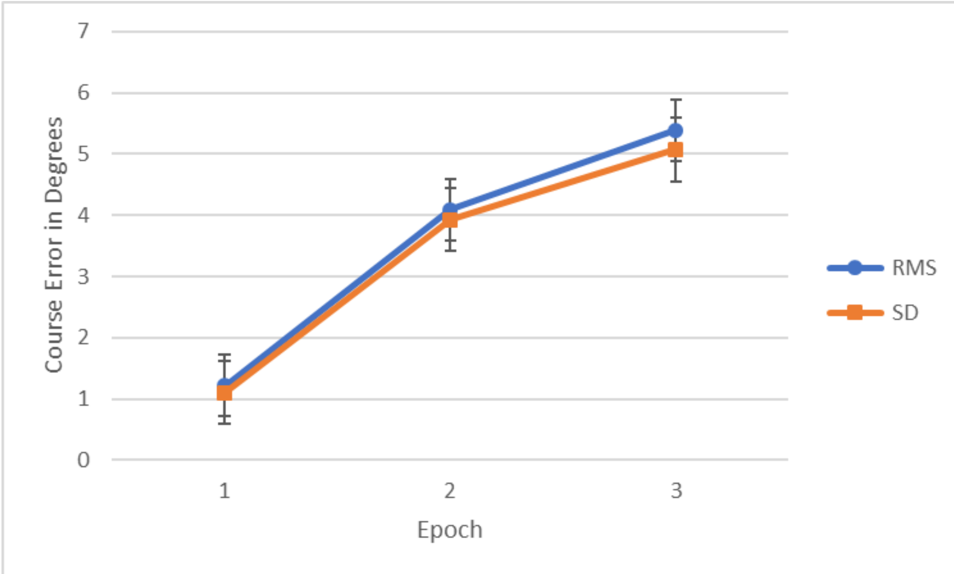

Figure 4.4: Changing course error by single task epoch. The Epoch 2 and 3 were significantly different from Epoch 1. It appears participant's overall ability to control vehicle course diminished as they approached the runway. Error bars represent 95% within-subjects confidence intervals.

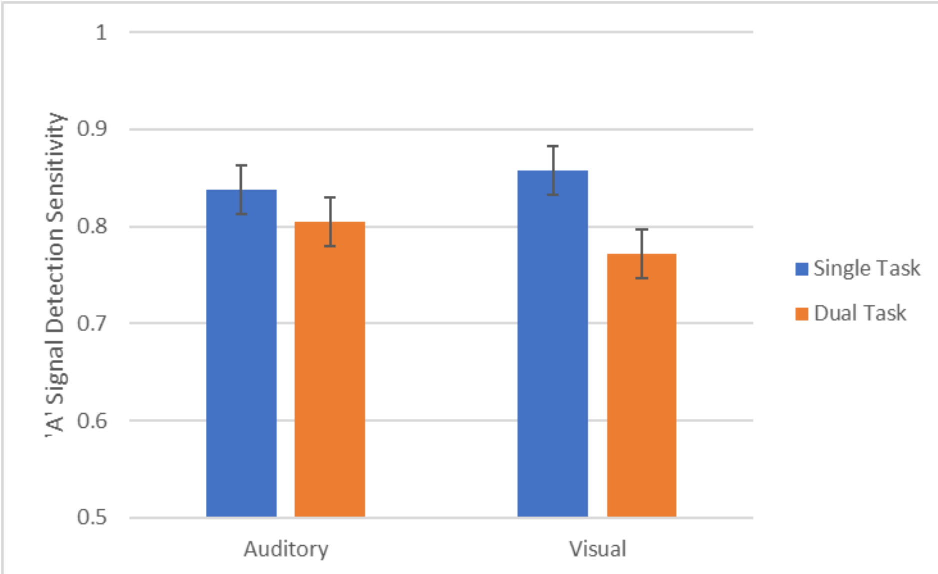

Figure 4.5: Signal detection task by task load. The SDT by task load interaction was significant $F(1, 10) = 5.747$, MSE = 0.001, $p = 0.037$, $\eta^2 = 0.046$. Error bars represent 95% within-subjects confidence intervals.

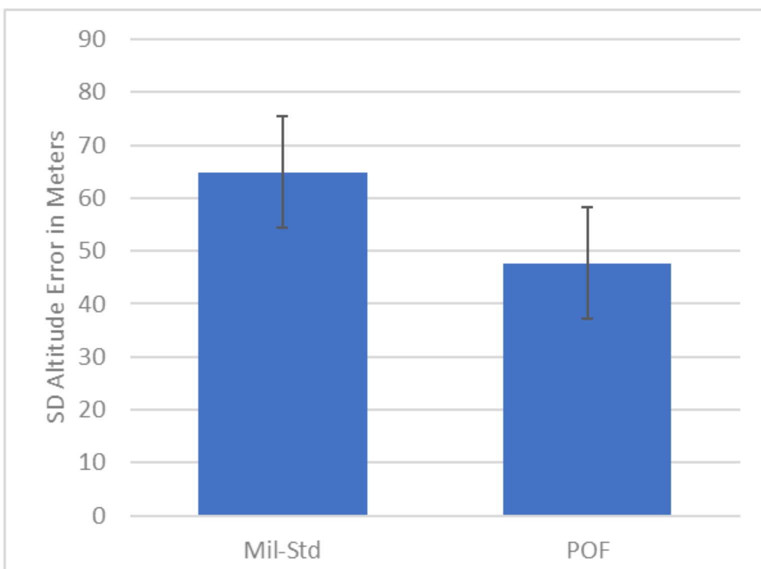

Figure 4.6: Dual task comparison within epoch 2 of display type $F(1, 10) = 4.566$, MSE = 495.665, $p = 0.058$, $\eta^2 = 0.092$. Error bars represent 95% within-subjects confidence intervals.

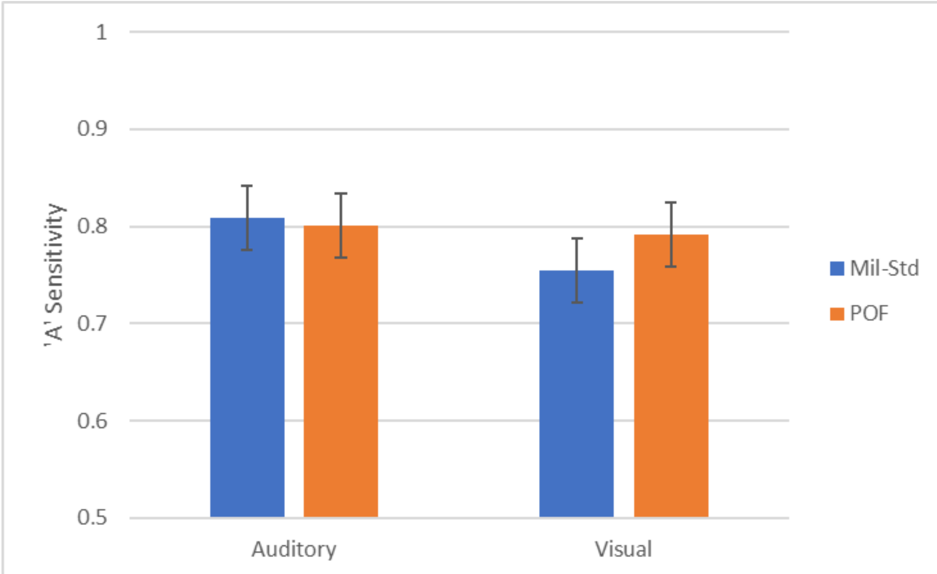

Figure 4.7: Dual task sensitivity differences by SDT in combination with display type. Significant difference between modalities within Mil-STD display condition $F(1, 10) = 6.464$, $MSE = 0.003$, $p = 0.029$, $\eta^2 = 0.084$). The POF display had the same impact on modality sensitivity $F(1, 10) = 7$, $MSE = 0.001$, $p = 0.577$, $\eta^2 = 0.004$. No interaction effects. Error bars represent 95% within-subjects confidence intervals.

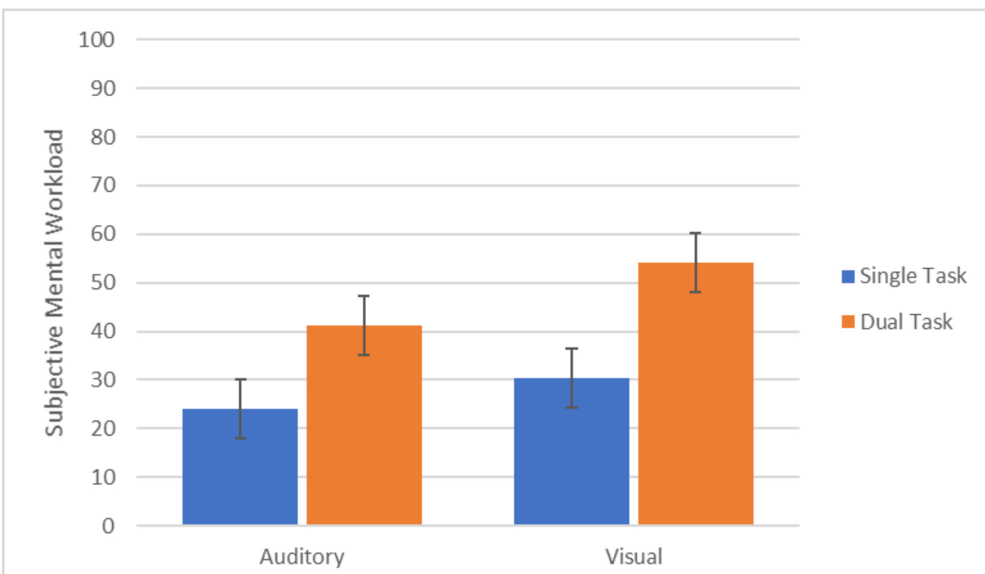

Figure 4.8: Subjective mental workload of SDT by task load. A main effect of SDT was found $F(1, 10) = 19.153$, $MSE = 55.604$, $p = 0.001$, $\eta^2 = 0.083$. A main effect of task load was found $F(1, 10) = 33.254$, $MSE = 139.253$, $p < 0.001$, $\eta^2 = 0.363$. No interaction effect was found. Error bars represent 95% within-subjects confidence intervals.