University of Iowa
Iowa Research Online

Theses and Dissertations

Summer 2014

# Flight deck crew coordination indices of workload and situation awareness in terminal operations

Kyle Kent Edward Ellis
*University of Iowa*

Follow this and additional works at: https://ir.uiowa.edu/etd

Part of the Industrial Engineering Commons

Copyright 2014 Kyle Kent Edward Ellis

This dissertation is available at Iowa Research Online: https://ir.uiowa.edu/etd/1313

## Recommended Citation

Ellis, Kyle Kent Edward. "Flight deck crew coordination indices of workload and situation awareness in terminal operations." PhD (Doctor of Philosophy) thesis, University of Iowa, 2014.
https://doi.org/10.17077/etd.y8zux206

Follow this and additional works at: https://ir.uiowa.edu/etd
Part of the Industrial Engineering Commons

المنارة للاستشارات
www.manaraa.com

# FLIGHT DECK CREW COORDINATION INDICES OF WORKLOAD AND SITUATION AWARENESS IN TERMINAL OPERATIONS

by

Kyle Kent Edward Ellis

A thesis submitted in partial fulfillment of the
requirements for the Doctor of Philosophy degree
in Industrial Engineering in the Graduate College of
The University of Iowa

August 2014

Thesis Supervisor: Associate Professor Thomas Schnell

المنارة للاستشارات

www.manaraa.com

Copyright by

KYLE KENT EDWARD ELLIS

2014

All Rights Reserved

www.manaraa.com

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

---

PH.D. THESIS

---

This is to certify that the Ph.D. thesis of

Kyle Kent Edward Ellis

has been approved by the Examining Committee for the thesis requirement for the Doctor of Philosophy of Industrial Engineering degree in Industrial Engineering at the August 2014 graduation.

Thesis Committee: ___________________________________

Thomas Schnell, Dissertation Supervisor

___________________________________

Andrew Kusiak

___________________________________

Geb Thomas

___________________________________

Yong Chen

___________________________________

Jonathan T Mordkoff

المنارة للاستشارات

www.manaraa.com

To Dr. Robert M Norman: F-14 pilot, NASA researcher, musician, and friend. Your dedication to aviation will never be forgotten.

www.manaraa.com

“When once you have tasted flight, you will forever walk the earth with your eyes turned skyward, for there you have been, and there you will always long to return.”

Leonardo Da Vinci

المنارة للاستشارات

www.manaraa.com

# ACKNOWLEDGEMENTS

This research is the combined effort of several individuals, to whom I am forever grateful. The impetus for the research, the countless hours of coordination and planning, and selfless mentorship have bolstered not only the advancement of aviation safety and the pursuit of novel science, but myself as a researcher.

I want to thank my advisor, Dr. Thomas Schnell, for taking a chance on a freshman with nothing more than a dream to work with aircraft and make a difference. I would also like to thank my colleagues at NASA Langley Research Center and the Operator Performance Laboratory at the University of Iowa for providing me with endless opportunities to work with the best-of-the-best, and for providing the world class facilities and technology that made this research possible. Additionally, I would like to thank Dr. Kara Latorella of NASA Langley, for without her I would have never had the opportunity to step foot in the door at NASA Langley and become a part of the aviation research conducted there. I would like to thank my committee members, Dr. Andrew Kusiak, Dr. Geb Thomas, Dr. Yong Chen, and Dr. Jonathan T Mordkoff for their participation and leadership.

I especially want to thank my family and friends. Their continuous support, encouragement, belief, and love have led me to where I am today. I particularly want to thank Sasha Congiu, who’s never ending support on all levels and bottomless knowledge of grammar and technical writing was invaluable to this research. Lastly, I want to thank my mother and father for instilling in me the confidence and freedom to believe that anything is possible.

المنارة للاستشارات
www.manaraa.com

## TABLE OF CONTENTS

المنارة للاستشارات
www.manaraa.com

المنارة للاستشارات
www.manaraa.com

المنارة للاستشارات

www.manaraa.com

## LIST OF TABLES

www.manaraa.com

www.manaraa.com

TABLE OF FIGURES

المنارة للاستشارات
www.manaraa.com

www.manaraa.com

المنارة للاستشارات

www.manaraa.com

المنارة للاستشارات

www.manaraa.com

www.manaraa.com

# LIST OF EQUATIONS

المنارة للاستشارات
www.manaraa.com

# CHAPTER 1.

# INTRODUCTION

## 1.1 Statement of the Problem

Air travel has proved to be one of the safest means of transportation. However, accidents still occur. Specifically in the field of aviation, when an accident occurs, the results are often catastrophic. Major advancements in aircraft design and maintenance have reduced mechanical errors and greatly improved the overall safety of aircraft operation. The leading cause of modern aviation accidents is not mechanical failure, but operator error. Pilot-related causes accounted for 70% of non-commercial accidents, and 60% of commercial accidents in 2010 (AOPA, 2010). Therefore, reduction of pilot error is necessary to improve aviation safety. A pilot’s ability to quickly interpret complex amounts of information and execute the appropriate response is critical to minimize errors inherent to a demanding environment.

## 1.2 Proposed Solution Approach

Eye-tracking quantitatively captures pilot eye-scan data, which can be used as a means to evaluate pilot tasking. Additionally, understanding trained crew behaviors enables researchers to determine expected task demands that must be met across each phase of flight and evaluate differences from observed pilot behavior. The basic metrics within eye-tracking provide clear, measurable elements that can be used to reliably characterize pilot task load through exhibited eye-scan behavior.  In addition, evaluating the task load and eye-scan behavior of the pilot flying (PF) and pilot monitoring (PM) allows for cross-comparison, which can be used to evaluate coordinated efforts. This research effort utilizes both new and existing quantitative methods to characterize individual pilot state behavior and pilot crew coordination using eye-tracking metrics.

المنارة للاستشارات

www.manaraa.com

Eye-tracking metrics and derived measures of coordination provide a correlation between self-reported, subjective workload ratings of varying flight deck experiment scenarios. Throughout this research initiative, two simulation studies were conducted. The first study assessed individual pilot eye-scan behavior, performance, and workload. The second study assessed behavioral indicators of crew coordination while utilizing the findings from the initial single pilot study. This research improves existing knowledge of eye-tracking in flight deck operations and provides further advancement in the quest for characterization of pilot state and crew coordination.

Assessment of crew resource management (CRM) literature and training techniques reveals several key components of aircrew behavior to help identify when and why errors are likely to occur in multi-personnel flight deck operations. CRM theory was used to develop normative models of expected workload for each pilot. Additionally, normative models of eye-scan behavior are developed to assess the difference between expected versus observed eye-tracking behavior for each pilot. Poorly coordinated crews are represented by significant differences between expected and observed eye-tracking behavior.

### 1.2.3 Central Hypotheses and Specific Aims

Crew coordination in the context of aviation is a specifically choreographed set of tasks performed by each pilot, defined for each phase of flight. The means to accomplish required tasks are prescribed by the flight deck interface, standard operating procedures (SOPs) and CRM, which is designed to balance the workload between crewmembers. Based on the constructs of CRM task load balancing and SOPs for each phase of flight, a shared understanding of crew workload is considered representative of well-coordinated crews.

المنارة للاستشارات
www.manaraa.com

Research has shown that pilot tasking is identifiable through pilot eye-scan (Ellis & Schnell, 2009). Therefore, eye-scan models developed for each pilot crewmember can characterize normative tasking and expected workload. Nominal behavior is defined by SOPs and CRM theory. Differences between expected versus observed eye-scan of the pilot reveals a departure from nominal behavior. A departure from expected behavior is therefore indicative of reduced crew coordination. While unexpected behavior is common, it is nevertheless indicative of a reduction in shared crew situation awareness. This research effort investigates the relationship between eye-scan exhibited by each pilot and the level of coordination between crewmembers. However, the relationship between crew coordination and crew performance is not evaluated in this research effort. Characterization of crew coordination using eye-scan behavior metrics must first be established before evaluating the effect of coordination on crew performance.

## 1.3 Contributions

### 1.3.1 Practical Contribution

Pilots on current flight decks find themselves in various situations that dynamically change their cognitive workload. Situations vary from mundane monitoring tasks commonly experienced during cruise flight, to intense, dynamic task sets during takeoff and landing. Research to improve the interaction between the pilot and the aircraft interface would benefit from quantitative analysis as opposed to qualitative analysis. Quantitative feedback eliminates the subjective bias across participants and standardizes results to provide more accurate analysis of pilot eye-scan in varying flight deck operation scenarios. Quantitative assessment of pilot eye-scan is flight deck specific due to areas of interest (AOIs) that are unique to flight deck configuration. AOIs are defined by the individual sources of information available across the flight deck interface. AOIs are useful for defining and developing models of eye-tracking data and are often

المنارة للاستشارات
www.manaraa.com

dependent on flight tasking during specific phases of flight. While eye-tracking models produced by this research may not apply globally across interfaces, the concepts to derive, evaluate, and classify individual pilot attention and crew coordination may be useful in other environments or applications.

### 1.3.2 Theoretical Contributions

Current avionics are inept at recognizing pilot capabilities and limitations that result from dynamic workload levels. Potential information overload in flight deck operations exists across all phases of flight. Several systems within the aircraft, such as the flight management system (FMS) and autopilot (A/P), are often effective at making flight tasks easier. Conversely, the FMS and A/P systems can also make difficult flight tasks harder when the systems fail to function as anticipated. Additionally, unexpected occurrences may happen in flight that lead to unusual aircraft attitude or the loss of energy state awareness, as demonstrated by the fatal accidents of Flash Air 604 and Aeroflot 821 (Egyptian Ministry of Civil Aviation, 2004), (Interstate Aviation Committee, 2009).

Flight decks on current-generation aircraft have several systems onboard to monitor the health state of the aircraft; however, there are no systems that monitor the state of the pilot. If aircraft were able to actively monitor pilot state, it would be possible to provide dynamic displays with all pertinent information in the proper context based upon the pilot’s current in-situ abilities. The flight deck interface could be tailored to rapidly increase pilot awareness to the critical task and relevant source(s) of information. Methods to increase pilot awareness are particularly useful in non-normal flight scenarios when pilot attention is channelized on a single instrument during high stress.

To achieve active monitoring of the pilot crew state, it is necessary to address the flight deck environment and flight operations tasks as a complex adaptive system,

www.manaraa.com

composed of PF, PM, and flight automation. The pilot-automation interaction is optimized by an increased awareness of current state through active feedback and modification of display information saliency. As a result, aviation safety and human-flight deck interface design is improved through active monitoring of the pilot crew, ultimately leading to a reduction in pilot error.

Additionally, eye-tracking research to evaluate crew state and crew coordination supports NASA aviation safety project research objectives. The National Aeronautics and Space Administration (NASA) is currently developing and testing the NextGen flight deck concept. Part of the NextGen flight deck concept is crew state monitoring, which is a research effort focused on characterizing pilot state in flight deck operations through the use of psychophysiological measures. Informed estimation of situation awareness, workload, and crew coordination can be derived from eye-scan behavior models and used to support crew state monitoring research. Eye-tracking methods that characterize pilot state directly benefit NASA's crew state monitoring research effort. The agency is ultimately pursuing the use of eye-tracking systems as remote sensors that provide a non-invasive solution deployable in NextGen flight decks.

المنارة للاستشارات
www.manaraa.com

# CHAPTER 2.

# REVIEW OF TECHNICAL LITERATURE

## 2.1 Mechanism of Visual Search

Basic visual search is composed of two components: fixations and saccadic movements. A *fixation* is a set of look-points or a series of eye-gaze vector data points that is focused on a stationary target in the person's visual field (Applied Science Laboratories, 2007). A fixation is the duration of time for which an individual is visually collecting and interpreting information available within the foveal range of the eye. When the fixation is made on a point close to the individual, such as on a flight deck, visual angle decreases significantly depending on the distance from the eye. The central 1.5 degrees of visual field have a visual resolution many times greater than that of the peripheral vision (Rao, Zelinsky, Hayhoe, & Ballard, 1997). The foveal range is the only field in which the eye is capable of interpreting fine resolution information, such as reading words in a book. The highest resolution necessary of any eye-tracker needs to be at least within two degrees of visual angle when converting the reading information analogy to that of heads down displays on a flight deck (Rayner & Bertera, 1979). The foundational components of eye fixations are the duration, frequency, and location.

www.manaraa.com

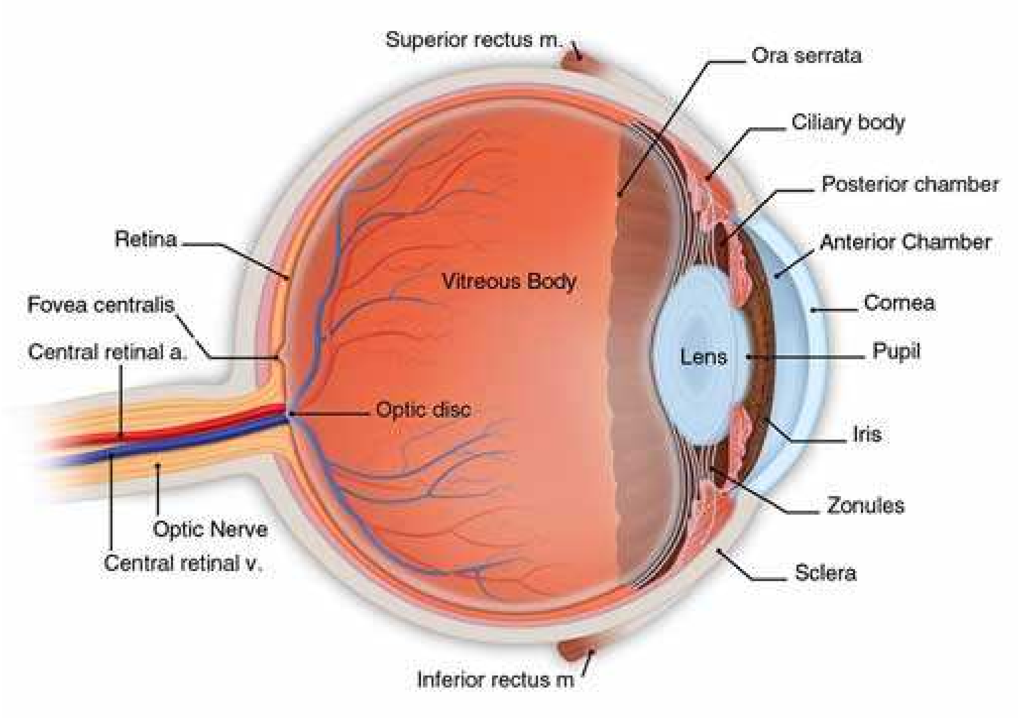


**Figure 1. Anatomy of the Human Eye (Eye and Eyesight, n.d.)**

The eye movement from one fixation to the next is called a *saccade*. A saccade connects one fixation to the next and can be measured in terms of radial degrees. Different components of a saccade include the length of the saccade (angular degrees), the speed of the saccade in degrees per second, and the direction of the saccade. When reading, the eye makes rapid movements, as many as four to five per second, moving from one fixation to the next, focusing on a few words each time (Rayner & Bertera, 1979). The eye does not transmit visual signals to the brain when making a saccade. Therefore, a saccade is made each time information is obtained from one fixation, and another fixation is necessary to observe further information elsewhere.

A scan pattern or scan path emerges when combining saccadic movements and their associated fixations. Since fixations only cover a finite space filled with information, saccadic movements trace the area of desired information so fixations can collect all the information necessary for the brain to interpret the overall image. Previous

المنارة للاستشارات
www.manaraa.com

research provides two theories connecting saccadic movements and fixations, concluding on the meaning of scan paths. One proposal indicates that specific scan paths and their associated fixations allow for the formation of visual-motor memory to encode objects and scenes (Noton & Stark, 1971). Another proposal suggests that changes in scan paths are most commonly associated with the dynamic demands of a given task (Yarbus, 1967).

These two theories relating to scan path are to a degree complementary; the scan path of a given person is strictly dependent on the individual's physical and cognitive state given the task set and current environment. In other words, certain task sets and their associated demands will vary based upon trained techniques and the psychophysiological state. Training results in the formation of visual scan techniques to encode scenes and sources of information. Evaluation of visual scan techniques makes it possible to correlate a specific scan pattern with an individual's specific state or at least begin to infer a state based upon the scan path behavior. However, characterizing pilot state requires explicit consideration of the individual differences across people — behaviors elicited by psychophysiological states, and the environment in which eye-tracking data is collected.

### 2.1.1 Impact of Individual Differences to Eye-Tracking Data Quality

There are several factors that impact eye-tracking performance, including physical features of the participant (e.g. blue versus brown eyes) and environmental conditions (e.g. lighting variation and vibration). Screening participants to fit the optimal specifications that work well with the eye-tracking system is effective in reducing physical-related differences. However, often times controlling physical features of the

المنارة للاستشارات
www.manaraa.com

participant set is impossible; therefore, care should be given to select an eye-tracking system that is capable of collecting data across a wide variety of participants.

Participants likely to cause issues for researchers attempting to obtain consistently high-quality eye-tracking data usually portray at least one of the following: a history of ocular trauma, ophthalmological diseases (previous or current), lazy eyes, pathologic nystagmus or other ocular disorders, and different forms of corrective lenses, including both eyeglasses and contact lenses. Consequently, it is highly recommended that researchers screen their participants prior to participation in any eye-tracking study to ease the effort required to collect quality eye-tracking data from the eye-tracking system. Pupil color greatly affects the quality of eye-tracking data for many eye-tracking systems. High precision eye-tracking systems require a sharp contrast between the pupil and the iris in order to calculate pilot gaze. Bright-pupil eye-tracking systems require direct infrared reflection off the retina; therefore, participants with blue eyes are often easier to track. Blue eyes contain less infrared reflective melanin in the iris. In contrast, brown or hazel eyes are usually ideal for dark-pupil eye-tracking systems that utilize a dark pupil contrast (Boyce, Ross, Monaco, Hornak, & Xin, 2006), (Wang, Lin, Liu, & Kang, 2005).

Pilots who are sleep deprived also pose another problem, particularly relevant to research investigating long-duration flight and the under-engagement of pilots in flight deck operations. The partial closure of the eyelid can become an issue when the eyelid begins to cover portions of the pupil. Many remote eye-tracking systems can operate with some portion of the pupil covered, but a majority of the pupil must still be shown in order to calculate the circular center of the pupil. Pupil center is often used as an integral part of gaze vector calculation. In the event the eyelid covers the pupil, data collection of blink rate and eyelid closure is most relevant to determine the overall reduction of pilot visual attention.

Glasses can create reflections that pose the biggest threat to eye-tracking quality. Hard-edged bi- or tri-focal lenses pose the largest problem due to distortion of the eye

المنارة للاستشارات

www.manaraa.com

image as seen from the perspective of the eye-tracking cameras. Distortions typically occur because of lens shape, which cause problems with systems using corneal reflection, bright retinal reflection, dark pupil circle, limbus or iris features, etc. Soft contact lenses usually do not cause problems; however, hard contact lenses can cause edge problems in bright pupil systems caused by dirt or dust trapped beneath the lens. Generally, single vision corrective eyeglasses do not cause problems unless they have an anti-reflective coating. However, lenses with curved front surfaces can create issues because of problems caused by reflecting the infrared source back into the camera.

Several factors exist that can become problematic to the testing environment of an eye-tracking system. Many issues are observed in environments affected by dynamic locations, such as that of an in-flight aircraft that experiences varying light conditions from the sun, turbulence, and other vibration effects that move the pilot’s head relative to the eye-tracking camera. Since eye-tracking systems are based upon visual contrast, extreme lighting conditions pose the greatest threat to eye-tracking data. Thus, in the context of flight simulators, extreme lighting from the projection system is often the largest problem to resolve.

Problems associated with extreme ambient lighting include small pupil diameter, squinting that places the eyelid over the pupil, glare that causes the pilot to change eye-tracking behavior, and degradation of the eye-tracking systems ability to detect features of the face for head-tracking purposes. Fortunately, in most simulators, ambient light levels are easily controllable. It is important to adjust lighting to be sufficient for normal operations on the flight deck as well as conditioning the light level enough to optimize eye-tracking data. Controlling ambient lighting is not as easily achieved on in-flight aircraft, requiring other forms of light mitigation that do not impede the pilot’s ability to fly under normal conditions.

المنارة للاستشارات

## 2.2 Eye-tracking on the Flight Deck

Researchers have used several different metrics to study pilot eye-scan behavior. One challenge is to identify which set of metrics strongly correlate with subjective and objective measures of performance and workload. Jacob and Karn (2003) identify three theories of eye-tracking data analysis:

- Top-down based on cognitive theory: “Longer fixations on a control element in the interface reflect a participant’s difficulty interpreting the proper use of that control.”
- Top-down based on a design hypothesis: “People will look at a banner advertisement on a web page more frequently if we place it lower on the page.”
- Bottom-up: “Participants are taking much longer than anticipated making selection on this screen. We wonder where they are looking.”

Top-down theories allow researchers to have a general idea of how a participant will react. In addition, researchers can look for trends that prove their hypothesis. The third theory, the bottom-up approach, can result in new methods of analysis. For example, post-run analysis may question why a pilot would spend more time on the attitude direction indicator (ADI) over the airspeed indicator (AI). The answers to such questions can provide further understanding of what consumes pilots’ cognitive capacity and estimate workload and performance.

Understanding how the research team attempts to interpret the data is important in determining not only what metrics to use but also how they will be used. For example, past research has demonstrated the difference between novice and more experienced participants in various usability studies (Fitts, Jones, & Milton, 1950), (Crosby & Peterson, 1991), (Card, 1984), (Altonen, Hyrskykari, & Raiha, 1998). Common effects have been observed in fixation duration, fixation frequency, saccadic movement, and scan-pattern changes. Experienced pilots are typically more comfortable performing

www.manaraa.com

flight tasks with a knowledge-based strategy of where to look on the flight deck in order to gain situation awareness. Differences in novice versus expert eye-tracking metrics indicate efficient eye-scan behavior. The following review of literature follows a top-down cognitive theory approach to identify eye-tracking metrics that quantify pilot eye-scan.

### 2.2.1 Eye-tracking Metrics

Visual scanning requirements change frequently as a function of the flight maneuver task (Hankins & Wilson, 1998), (Itoh, Hayashi, Tsukui, & Saito, 1990). Chosen eye-tracking metrics must either be task specific or applicable to all forms of workload scenarios. NASA researchers have conducted several eye-tracking studies, which have resulted in a set of eye-tracking metrics for use in future research initiatives. Eye-tracking metric definitions include:

- Average Dwell Time – The total time spent looking at an instrument divided by the total number of individual dwells on that instrument.
- Dwell Percentage – Dwell time on a particular instrument as a percent of total scanning time.
- Dwell Time – The time spent looking within the boundary of an instrument.
- Fixation – A series of continuous lookpoints that stay within a pre-defined radius of visual degrees.
- Fixations Per Dwell – The number of individual fixations during an instrument dwell.
- Glance – A "subconscious" (i.e., non-recallable) verification of information with a duration histogram peaking at 0.1 seconds. (also referred to as an "orphan").

المنارة للاستشارات

- Lookpoint – The current gaze location of where the pilot is looking, frequency of data points depending on the eye-tracking system used.
- One-way Transition – The sum of all transitions from one instrument to another (one direction only) in a specified instrument pair.
- Out of Track – A state in which the eye-tracking system cannot determine where the pilot is looking, such as during a blink or when the participant's head movement has exceeded the tracking capabilities of the system setup.
- Saccade – The movements of the eye from one fixation to the next. Also considered to be the spatial change in fixations.
- Scan – Eye movement technique used to accomplish a given task. Measures used to quantify a scan include (but are not limited to) transitions, dwell percentages, and average dwell times.
- Transition – The change of a dwell from one instrument to another.
- Transition Rate – The number of transitions per second.
- Two-way Transition – The sum of all transitions between an instrument pair, regardless of direction of the transition. (Harris, Glover, & Spady, 1986)

Defining AOIs is the critical first step in analyzing eye-tracking data. AOIs are three-dimensional regions that classify spatial information sources and may be intersected by a pilot's gaze. It is important to specify what AOIs represent in order to compile meaningful results. Additionally, AOIs may be task specific; therefore, proper definition helps determine the eye-scan behavior necessary to complete the task, which is based on the interface used (Jacob & Karn, 2003). AOIs common to flight deck interfaces often include the head-down displays (HDDs), which are broken down into smaller regions, such as the AI, altimeter (ALT), ADI, and horizontal situation indicator (HSI). However, limitations to the size of AOIs based on eye-tracking system accuracy must be considered.

المنارة للاستشارات
www.manaraa.com

The limiting factors of the definition of AOIs rest solely on the performance capabilities of the eye-tracking system being used. The size of AOIs is dependent on the accuracy of the eye-tracking system. The smaller an AOI, the more accurate an eye-tracking system must be in order to identify a person's eye gaze within that region. AOIs must be specific enough to define the important regions of an interface, but must not be so specific that meaningful data is unattainable due to the noise or inaccuracy of eye-tracking system.

Metrics within AOIs, such as dwell time and number of transitions, help researchers describe a pilot's eye-scan behavior. NASA research has utilized large AOIs, such as the instrument panel (IP) and the forward view out the window (OTW) to frame pilot eye-scan behavior in the context of head-up and head-down. NASA research was effective in describing pilot attention across the peripheral range of a pilot's eyes and proved it is capable of yielding significant results (Ellis, Kramer, Shelton, Arthur, Prinzel, & Norman, 2011). However, the results from Ellis et al. (2011) suggest the more AOIs there are, the more accurate the definition of pilot eye-scan behavior becomes.

### 2.2.2 Eye-scan behavior Relative to Pilot Tasking

Ellis and Schnell (2009) evaluated 12 pilots in a realistic, single pilot approach to land simulation. The simulation was designed to present a set of complex flight tasks to yield a wide variety of task sets and induce a range of physical and cognitive workload levels (Ellis & Schnell, 2009). The variation in cognitive tasking levels was achieved through the level of automation used to perform the approach to land flight operation. Automation varied between full A/P and auto-throttle, no A/P with auto-throttle, and no

www.manaraa.com

A/P and no auto-throttle. The variation in task set requirements induced a significantly wide range of pilot workload levels across automation conditions (Ellis & Schnell, 2009). Eye-tracking data was used to determine the effect on pilot eye-scan behavior across automation conditions. Ellis and Schnell (2009) show pilot eye-scan behavior is significantly affected by task loading, and specific scan patterns are identifiable through the use of common eye-tracking metrics, such as fixations, dwell times, and gaze dispersion. Ellis and Schnell (2009) conclude eye-tracking data is useful in modeling pilot workload and performance across a set of defined tasks.

## 2.3 Characterizing Pilot Eye-Scan and Existing Models

Researchers have utilized eye-tracking metrics to characterize pilot eye-scan behavior (Jacob & Karn, 2003). The trends of the various measures listed above in Section 2.2.1, are fundamental in the development of behavioral models to evaluate operator workload. Eye-scan models created to characterize human behavior have attempted to quantitatively assess operator situation awareness, workload, engagement, and probability of error; however, the complex nature of the flight deck interface poses a significant challenge to precise measurement.

Models are simplified to address key fundamentals of eye-scan behavior necessary to complete a certain task. In the human-machine interface, the flight deck provides sources of information that serve three primary functions the pilot must address: aviate, navigate, and communicate (Schutte & Trujillo, 1996). Each function is addressed by specific interface locations, such as the primary flight display (PFD), the multi-function display or navigation display (MFD or ND), and the control display unit (CDU). Models can be derived based upon expected behaviors specific to required tasking, and quantified deviations from that behavior can then be used as a measure to infer pilot workload.

www.manaraa.com

### 2.3.1 Pilot Scan

When flying without reference to external visual cues, pilots have to rely on instruments to safely control the aircraft. Proper instrument-scan is a skill that needs to be acquired through instruction, training and experience. In addition, a pilot's instrument-scan needs to be maintained through currency of use. Cross-checking instruments is the first fundamental skill of pilot instrument scan. Cross-checking is defined as the logical and continuous observation of instruments for attitude and performance information (Federal Aviation Administration, 2012). Cross-checking instruments is necessary to maintain attitude, performance, and navigational goals. Instruments that are pertinent to the pilot to perform a given maneuver vary depending on the executed maneuver. Maneuvers are often specific to the aircraft's phase of flight. Therefore, a pilot's scan must gaze upon instruments (also referred to as AOIs in the context of eye-tracking) specific to the task maneuver required by a specific phase of flight or operational objective. The ability of pilots to perform task maneuvers is developed through training attitude instrument flying skills.

Airplane attitude instrument flying is an aeronautical skill trained to all instrument rated pilots to safely aviate and navigate an aircraft. The two methods used to train instrument flight are "control and performance" and "primary and supporting" (Federal Aviation Administration, 2012). Both methods rely on flight instrumentation, differing only in the manner attention is given to each instrument, relying primarily on the ADI.

The control and performance method breaks instruments into three groups: control instruments, power instruments, and navigation instruments. Control instruments provide information on immediate power and attitude changes. Control instruments include the ADI and the Engine Information and Crew Alerting System (EICAS) (Federal Aviation Administration, 2012). The EICAS, found on commercial turbojets, displays N1 and N2 fan speed, exhaust gas temperature, and the engine pressure ratio, providing

المنارة للاستشارات
www.manaraa.com

engine power output information. Performance instruments provide information on aircraft performance. Performance instruments include the AI providing aircraft speed, the ALT providing aircraft altitude, vertical speed indicator (VSI) providing aircraft climb performance, the HSI, pitch attitude indicator, and the slip/skid indicator (Federal Aviation Administration, 2012).

Performance indicators reference speed, heading, and altitude information, describing the horizontal, vertical, or lateral direction an aircraft is heading. Navigation instruments provide information on aircraft position relative to a fix or navigation facility. Navigation instruments include displays that provide global positioning system (GPS) information, moving map displays (MMD), very-high omnidirectional range/nondirectional radio beacon (VOR/NDB information, and localizer and glideslope information) (Federal Aviation Administration, 2012).

A four-step process was developed by the FAA to aid the process of attitude instrument flying: establish, trim, cross-check, and adjust (Federal Aviation Administration, 2012). The four-step process describes the control loop between the pilot, the instrumentation and the aircraft controls. To adjust the attitude of the aircraft, the pitch and/or bank of the aircraft must be manipulated in coordination with the power settings to establish the desired performance. The pilot must then trim the aircraft so constant control pressure is not required and the aircraft will fly its current trajectory. Trimming the aircraft allows the pilot to relieve pressure on the controls and momentarily divert attention to other tasks (Federal Aviation Administration, 2012). Once the aircraft is established and trimmed, the pilot is trained to perform a cross-check of the instruments. Cross-checking the instruments reveals any deviations from desired performance and control, and informs the pilot on how much change is required to the control inputs. Pilots must then make adjustments to the control inputs based on any required changes observed from the cross-check (Federal Aviation Administration, 2012).

المنارة للاستشارات
www.manaraa.com

The primary and supporting method is an extension of the power and performance method (Federal Aviation Administration, 2012). Utilizing primary and supporting flight instruments in coordination with power and control instruments allows for fine adjustment of attitude instrument flight control. The primary and supporting method focuses on scanning between instruments that most accurately depict the aspect of aircraft attitude being controlled. The four key elements to aircraft attitude control for both methods are pitch, bank, roll, and trim.

Pitch control is achieved by viewing the ADI, ALT, AI, and VSI, shown in Figure 2. To maintain straight-and-level flight, a pilot maintains constant altitude, airspeed, and typically heading (Federal Aviation Administration, 2012). To achieve straight-and-level flight three primary instruments must be monitored: the ALT, AI, and HSI. Altitude should remain constant as long as aircraft speed and pitch are held constant, which maintains primary pitch control. Primary pitch control may be affected by two factors: turbulence and momentary distractions away from the instruments (Federal Aviation Administration, 2012). Deviations in pitch require the pilot to make control input corrections; small deviations require small corrections, and large deviations require large corrections. Large corrections that result in rapid attitude changes should be avoided, as rapid changes in aircraft attitude may lead to spatial disorientation and unsafe aircraft attitude.

www.manaraa.com

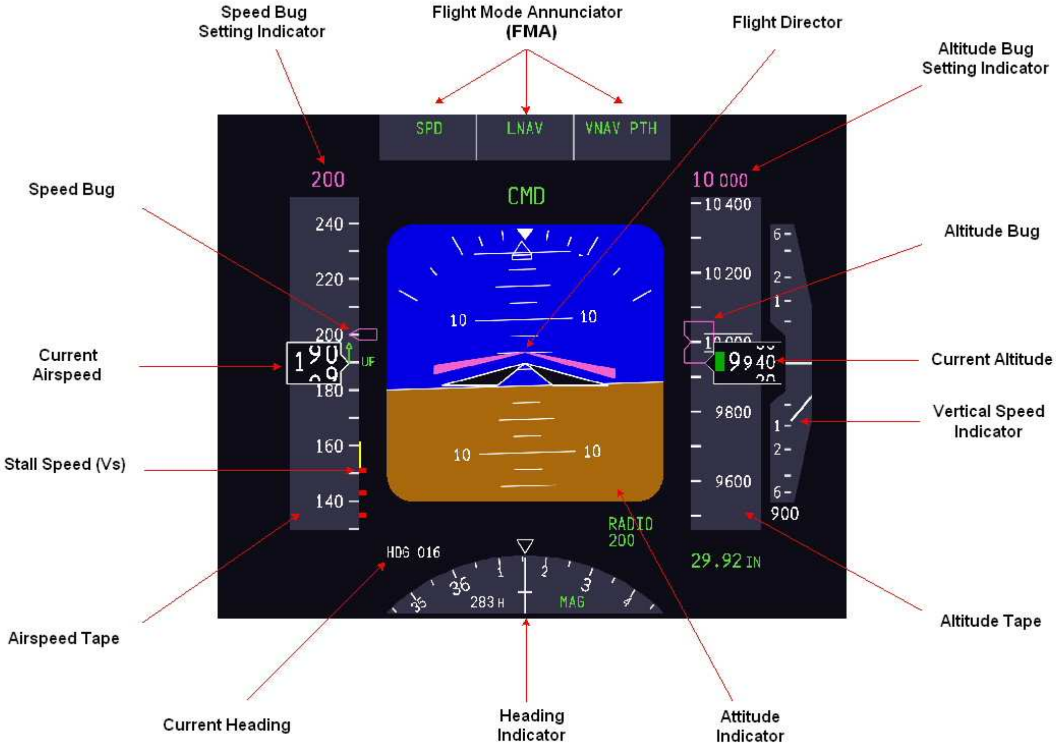


**Figure 2. EFIS Display**

Supporting instruments aid the pilot in monitoring the trend of attitude state for fine control of the aircraft. Supporting instruments of primary pitch control are the VSI, AI trend, and ALT trend (Federal Aviation Administration, 2012). AI or ALT trend is shown either by trend tapes, available on electronic flight instrumentation system (EFIS) type displays available on the PFD, or is visible by observing the rate the AI or ALT needle is moving on standard round gauges. The VSI and trend information affords pilots the ability to observe changes in the airspeed and altitude of the aircraft and make small corrections before larger corrections are required.

Primary bank control allows pilots to maintain preplanned or assigned headings (Federal Aviation Administration, 2012). The HSI is the only instrument that provides current heading information, and is therefore the primary instrument for bank control. Supporting bank control instruments include the ADI and turn rate indicator, providing aircraft roll information and the effect on changing heading. Only the slip/skid indicator displays primary yaw control information. The slip/skid indicator is the only instrument

www.manaraa.com

capable of indicating if the aircraft is moving through the air with the longitudinal axis of the aircraft aligned with the relative wind direction (Federal Aviation Administration, 2012). The AI is the primary power instrument for straight-and-level flight. Power is used to control desired airspeed, and no instrument other than the AI delivers instantaneous indication of airspeed.

Understanding the methods to scan primary and supporting instruments is essential to attitude instrument flying. The attitude indicator remains the most important instrument to attitude instrument flying, but all instruments for primary pitch, bank, and yaw must be utilized appropriately to safely maneuver the aircraft. The FAA recommends several scan pattern types that are common in instrument flight instruction; for example, the selected radial cross-check is designed so the eyes of a pilot remain on the attitude indicator 80-90 percent of the time, with the remaining time spent transitioning to other primary and supporting instruments, shown in Figure 3 (Federal Aviation Administration, 2012). Trained scan patterns help a pilot monitor all necessary flight instrumentation with an appropriate frequency to maintain a pilot’s situation awareness aircraft attitude and energy state of the aircraft.

المنارة للاستشارات

www.manaraa.com

SCAN PATTERN

Scan pattern should start with left (airspeed indicator), then right (altitude indicator/VSI), then up (slip/skid indicator), then down (heading indicator). The pilot should return attention back to the center (pitch attitude) before proceeding to the next direction. For example: left, center, right, center, up, center, down, center.

**Figure 3. EFIS Radial Cross-Check (Federal Aviation Administration, 2012)**

Scan patterns differ based upon the flight mode (automation) of the aircraft and the flight tasks required of the pilot during any given phase of flight. Pilot crew procedures are explicitly defined with regard to who is responsible for each task during each phase of flight (Federal Aviation Administration, 2003). A pilot’s scan must visit each pertinent instrument, or significant errors in the flight operation will occur. A common error is instrument fixation, which is observed when a pilot channelizes attention on a single instrument at the expense of cross-checking other instruments. Pilot scan errors are often related to poor training or unusual circumstances atypical of common flight operations that are accompanied with unusual pilot workloads and cognitive demands (Federal Aviation Administration, 2012).

Pilot scan and cross-checking, as described above, are techniques that train pilots to obtain the necessary information to complete safe flight maneuvers. Evaluation of effective pilot eye-scan and cross-check is implemented in general aviation instrument pilot training and prevalent in line operations training (Norman, 2010). Through the use

المنارة للاستشارات
www.manaraa.com

of eye-tracking metrics, models of pilot eye-scan behavior can be developed to differentiate good visual scan from bad visual scan.

### 2.3.1 The A-SA Model

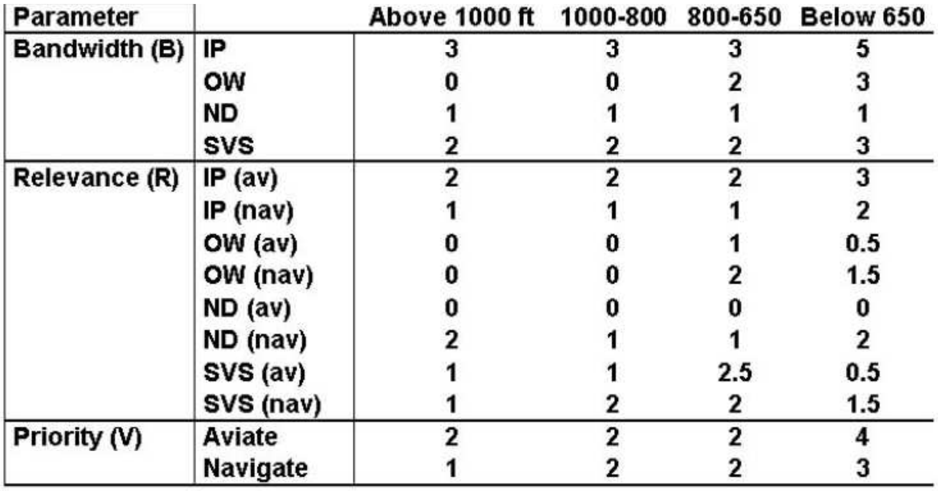
The A-SA (Attention-Situation Awareness) model is structured around two modules. The first module addresses the allocation of attention with regard to events and data channels in an environment (Wickens, McCarley, & Thomas, 2003). The second module addresses expected behavior based upon pilot tasking relative to current phase of flight and the future state of the aircraft in a given environment. The attention module is based upon the Saliency, Effort, Expectancy, and Value (SEEV) model of attention allocation (Wickens, Helleberg, Goh, Xu, & Horrey, 2001). Saliency and effort are driven by bottom-up attention allocation to salient events and the effort to allocate attention, which is affected by the contemporaneous cognitive activity. Expectancy and value of the events are dependent on the location of AOIs and their relative value to a given task. The second module is based upon Hogarth and Einhorn’s (1992) belief updating model, which places importance on defining salience, effort, and relevance of events. Relevance is defined as the corresponding value of an event relative to the pilot task of maintaining situation awareness (Wickens et al., 2003). The corresponding output of the A-SA model is designed to predict the time spent looking at an AOI. Predicting the time spent looking at an AOI can determine causal factors of errors in a task, such as loss of situation awareness.

When utilizing the A-SA model with eye-scan data, the salience parameter is replaced with the actual distribution of attention. Given the static physical dimension of a flight deck, the effort parameter is also predefined across AOIs based on the physical distance from one AOI to another. The value parameter for each AOI is based upon the well-researched hierarchy of aviate>navigate>communicate. Expectancy is calculated

www.manaraa.com

based upon the bandwidth of information contained within a given AOI (defined as the frequency of change in information). Predetermined values are established for each effort, value, and expectancy parameter, providing the necessary components to model how frequently an AOI should optimally be visited as a function bandwidth and relevance. Specifically, the model utilizes estimated coefficients of bandwidth, relevance and task priority, shown in Table 1.

**Table 1. SEEV Model Parameters Sample (Wickens, McCarley, & Thomas, 2003)**



| Parameter | | Above 1000 ft | 1000-800 | 800-650 | Below 650 |
|---|---|---|---|---|---|
| Bandwidth (B) | IP | 3 | 3 | 3 | 5 |
| | OW | 0 | 0 | 2 | 3 |
| | ND | 1 | 1 | 1 | 1 |
| | SVS | 2 | 2 | 2 | 3 |
| Relevance (R) | IP (av) | 2 | 2 | 2 | 3 |
| | IP (nav) | 1 | 1 | 1 | 2 |
| | OW (av) | 0 | 0 | 1 | 0.5 |
| | OW (nav) | 0 | 0 | 2 | 1.5 |
| | ND (av) | 0 | 0 | 0 | 0 |
| | ND (nav) | 2 | 1 | 1 | 2 |
| | SVS (av) | 1 | 1 | 2.5 | 0.5 |
| | SVS (nav) | 1 | 2 | 2 | 1.5 |
| Priority (V) | Aviate | 2 | 2 | 2 | 4 |
| | Navigate | 1 | 2 | 2 | 3 |

The A-SA model gauges how a pilot samples an AOI by quantifying the bandwidth, relevance, and priority of a given AOI (Wickens, Goh, Helleberg, Horrey, & Talleur, 2003). Output is then given as the predicted percentage dwell time of a given AOI based on flight deck configuration and task. Predicted percentage dwell time per AOI can be used to identify if a pilot is performing an optimal scan given the information required to complete a task.

## 2.4 Crew Coordination

Crew coordination is defined as a set of principles, attitudes, procedures and techniques that create a better performing crew (Thornton, Kaempf, Zeller, & McAnulty, 1992). A crew's ability to coordinate efforts effectively results in reduced human error

www.manaraa.com

and improved flight safety. The complex interaction of pilot crew and flight deck interface creates a potential for chaos and failure. When the complexity of flight systems increases, the complexity of potential failures increases. Accidents in aviation are often labeled a result of human error, where the interaction between the pilot crew and flight deck is subject to the “swiss cheese” failure model. The swiss cheese failure model describes how human interactions on the flight deck, coupled with complex modalities and established procedures, are capable of aligning in a way that result in an accident. Accidents happen when an initial error causes additional errors to occur in such a way that current safety systems fail to be effective.

Understanding crew coordination procedures and complex failure modes is useful in developing safety measures. Therefore, most safety measures are only effective against known possible scenarios. Systems and procedures intended to mitigate failures and human errors are limited to the information available. Any failure or error can lead to the onset of a string of events that develop into unfamiliar scenarios that push pilots to the limits of their expertise, often requiring them to react quickly.

Redundancy is a significant element to ensure continued safety of flight. Mechanically, aircraft are designed and manufactured with several safety factors to cope with rare events, such as backup hydraulics systems in the event of primary, or even secondary hydraulic failure. The airframe is designed to handle stress and strain loads that far exceed the operational envelope pilots are trained to operate within. As a result, if the circumstance arises when the aircraft is pushed beyond normal limits, such as extreme turbulent conditions, the aircraft maintains its structural integrity. In some cases, even the control inputs are limited so as to prevent the pilot from operating outside the operational envelope of the aircraft, such as the envelope protection used by Airbus (AirbusDriver.net, 2014).

Similar to the mechanical systems of an aircraft, the pilot is subject to several factors capable of prompting mistakes in procedures and degrading performance. Factors

المنارة للاستشارات
www.manaraa.com

include inattention caused by lack of sleep or distraction, confusion of automation modality, control input confusion, stress, task saturation, and overall high or low workload. As a result, two pilots are on the flight deck, coordinating efforts to delegate and balance individual task load, provide concurrence, and reduce the possibility of single human-in-the-loop errors. Having a redundant human-in-the-loop system in place to coordinate, execute, and monitor basic flight task actions, allows for safer and more reliable air transportation.

Factors that affect a pilot’s ability to perform tasks are addressed as individual components that describe pilot state. Analyzing flight operations as a complex adaptive system, the interacting entities that make up that system are addressed with a focus on the flight crew itself. The flight crew is made up of the PF and PM, each with abilities affected by their individual pilot state. The PF and PM must interact in a coordinated effort with the flight deck interface and its flight automation to achieve specific flight tasks. The following subchapters review the phyiscal, physiological, psychological, emotional, and environmental factors that affect functionality of the principal entitites of the man-machine flight deck system, the pilots.

### 2.4.1 Crew Resource Management (CRM) Behavioral Indicators

Situation awareness is fundamental to crew coordination. A pilot’s ability to effectively communicate and coordinate as a member of a crew improves task performance. By definition, CRM behaviors are those that improve the overall situation recognition of the crew, such as crew adaptability, flexibility, mission analysis, and situation awareness (Barker, Clothier, Woody, Mckinney, & Brown, 1996). CRM training helps pilots improve the decision-making process on the flight deck by maintaining CRM behaviors and utilizing all available resources, regardless of the

المنارة للاستشارات
www.manaraa.com

situation presented. Training crews to communicate, be assertive, advocate, and understand their individual roles and responsibilities pertaining to safe operation of the aircraft may seem trivial. CRM training enables effective execution of processes, especially during problem situations limited by time. CRM training varies by air-carrier; however, each address the same fundamental aspects of crew coordination and resource management:

- Communication Processes and Decision Behavior
    - Briefings
        - Safety
        - Security
    - Inquiry/Advocacy/Assertion
    - Crew Self-Critique (Decisions and Actions)
    - Conflict Resolution
    - Communications and Decision Making
- Team Building and Maintenance
    - Leadership/Followership/Concern for Task
    - Interpersonal Relationships/Group Climate
    - Workload Management and Situation Awareness
        - Preparation/Planning/Vigilance
        - Workload Distribution/Distraction Avoidance
    - Individual Factors/Stress Reduction

(Federal Aviation Administration, 2004)

CRM training provides pilots with standardized behaviors that enable effective crew coordination. The FAA has identified several behavioral indicators that stem from CRM training topics; core behavioral indicators include shared situation awareness and workload balancing. When a crew exhibits core behavioral indicators, effective crew

www.manaraa.com

coordination is likely to occur. A shared understanding of crewmember responsibilities allows for effective crew coordination. Additionally, effective communication is a tool to share the situation awareness possessed by each member. Crewmembers are encouraged to inquire, advocate, and assert any observations or suggestions on the flight deck, regardless of rank. Crew awareness of personal factors, such as stress, fatigue and task overload, is critical in managing available resources and coordination efforts. Furthermore, the ability to assess the linkage between behavioral indicators and crew coordination has become an increasingly important tool on the flight deck.

Eye-tracking enables the ability to assess behavioral indicators and as a result, effective crew coordination can be evaluated. Shared awareness of aircraft state, automation mode, and task responsibilities allow crews to be an effectively coordinated team. Historically, subjective response was the only feedback available to report pilot state and is prone to inherent participant bias. Today, eye-tracking can be used as a tool to monitor the eye-scan behavior of the crew on the flight deck, making it possible to better understand the interaction between crew and avionics. Understanding crew awareness, aircraft state, and phase of flight tasking makes it possible to determine the coordination of the system as a whole. Evaluation of behaviors that are indicative of crew coordination requires analysis of the factors that impact the behavior of the individual pilot and crew. Factors that impact the behaviors of the individual pilot and crew are discussed below.

### 2.4.2 Impacts of Personality and Attitude

The attitude of crewmembers is a large component of effective crew coordination and CRM. Historically, personality and associative personality traits focused mainly on understanding pilots and their behavior. Attempts were made to modify personality traits in pilots to better suit the various tasks of piloting an aircraft (Helmreich, 1984).

المنارة للاستشارات
www.manaraa.com

However, efforts to modify personality traits have had very low success rates (Helmreich, 1984). Since pilots are recruited from the general population, a more effective approach is the assessment of pilot attitude. Attitude is trainable across a variety of personality types, and yields greater improvement of pilot adaptation. However, personality traits are still perceived as important; in fact, there are observable personality traits, such as achievement motivation and interpersonal sensitivity, that correlate to measures of performance (Helmreich, 1982).

Helmreich (1984) presents two possible linkages between personality traits: attitude and resource management. One theory portrays how observed behavior is a result of pilot attitude, and pilot attitude is a result of personality given a particular situation. The second theory suggests attitudes and personality traits are independent, and that each trait affects the flight deck individually. Therefore, if personality is a significant factor in flight deck behavior and crew coordination, training would be ineffective, and focus should be spent on pilot selection. However, if both attitude and personality independently affect crew behavior and coordination, then effort toward both pilot selection and training would demonstrate success in developing coordinated crews.

Pilot opinions about crew management, regardless of personality or role, are generally in agreement (Helmreich, 1984). Contrary to the findings on crew management, pilot views of crew attitude vary significantly (Helmreich, 1984). A difference in the view of crew attitude suggests that attitudes are in fact independent of personality traits. Attitudes independent of personality traits signify CRM training is capable of adjusting or even circumventing attitudes that lead to negative behavior. CRM can therefore be taught to pilots in order to improve crew coordination.

المنارة للاستشارات

www.manaraa.com

### 2.4.4 Team Process Behavior

Team process behavior and its correlation to task performance is a significant factor that affects crew coordination and performance. In order to model the correlation, it is vital to properly identify the variables that define team process. Additionally, measures of performance are necessary to evaluate the effectiveness of team process behavior. Research has shown that it is possible to characterize team process effectiveness through examination of interactive processes (Stout, Salas, & Carson, 1994). Interactive processes such as planning, delivery of shared information, task delegation, and task load balancing are variables indicative of effective team process behavior (Stout et al., 1994).

The ability to quantify interactive processes allows for models to evaluate the level of crew coordination and in turn, become increasingly effective. The processes identified by Stout et al. (1994), such as task load balancing and shared information, are detectable through observation of pilot attention. The ability to assess pilots' eye-scan behavior allows for determination of shared attention. Shared attention of common instruments is therefore a means to evaluate crew coordination. Shared attention is a novel eye-scan behavior metric developed for use in crew coordination research. The shared attention metric is defined as a percentage of time both pilots gaze on a common AOI within a specified time frame.

### 2.4.5 Crew Coordination and Time/Risk/Complexity Management

When confronted with a time-critical, decision-making problem, pilots are forced to rapidly understand the problem, determine a corrective option, and execute the appropriate tasks. The complexity of time-critical problems, coupled with several limiting

www.manaraa.com

factors in both physical and cognitive dimensions, is incredibly influential to pilot decision-making. Factors that influence pilot decision-making include phase of flight, aircraft type, problem type, time and resources available, complexity, and the level of risk (Fischer, Orasanu, & Wich, 1995). Specifically, the factors of phase of flight, aircraft type, and problem type allow for several improvement strategies, such as training and pilot experience. On the contrary, risk, complexity, and time pressure are more difficult to manage.

The ability of a pilot to manage risk and time pressure varies based upon experience and training, attesting to the value of experience and training as error mitigation strategies. When a pilot is in an unfamiliar situation, he or she relies on professional problem-solving skills developed by training and experience. Pilots are often faced with short time and high-risk situations that push pilot problem-solving skills to the limit. Understanding the impact of risk and time pressure to pilot cognitive state is important in determining a pilot's problem-solving ability. Furthermore, the effects of time and risk to the cognitive state of the pilot affect how a crew adapts during simple or complex situations.

Research shows that the pilot flying first addresses risk and then time pressure to delegate task action on the flight deck (Fischer, Orasanu, & Wich, 1995). The pilot monitoring responds to time pressure and situational complexity, which affects the order of task execution. Situational complexity has boundaries that vary from simple to complex and even unknown. Situational complexity determines whether a pilot knows how to handle the situation or must revert to problem solving (Fischer et al., 1995). Risk, time pressure, and situational complexity are unique to each crewmember and their specific role on the flight deck.

The pilot flying assesses the risk and time pressure for a given situation and then delegates tasks. Once tasks are delegated, the pilot monitoring assesses the time pressure and complexity of the required task set and prioritizes accordingly. Understanding the

www.manaraa.com

way both pilot roles handle time, risk, and complexity reveals a deterministic approach with regard to how a crew interacts. In other words, if a crew fails to address tasks in a deterministic manner, it is logical to infer that risk, time, and complexity factors are affecting the crew. Crew interaction behaviors are identifiable through eye-tracking, capturing the eye-scan and cross-check used by each pilot to address tasks. Therefore, deviation from the way a crew is trained to interact and address tasks is indicative of reduced coordination.

### 2.4.6 Intra-team Communication and Crew Coordination “Schema”

Communication is a significant tool used to coordinate efforts in a flight crew. Variants of communication that occur on the flight deck include observation, physical gesture, and verbal communication. Due to the dependence on communication to successfully coordinate efforts, detection of factors limiting communication is necessary to determine the level of crew coordination. Factors that affect the level of communication include, noise, internal and external audio communication, improper terminology, and divided attention (Katz, Kambe, Kline, & Grubb, 2006).

Pilots are trained to perform callouts and cross-checks for several reasons. One reason is to alert another crewmember when a certain task is or should be completed. A second reason is to perform a verbal cross-check to confirm task execution. For example, when extending the landing gear, the PF calls out, “landing gear” and the PM sets the gear lever down and responds, “landing gear down.” Both pilots then visually confirm the landing gear is down with a final cross-check of the flight instrumentation. When a PM calls out, “three green” to confirm the landing gear is fully extended and locked, the task is complete. Detection of verbal and visual cross-checks is a simple way to determine if a crew is coordinated

المنارة للاستشارات

www.manaraa.com

Pilots perform visual cross-checks by looking at common sources of information, which increases shared awareness between crew members. Observing the eye-scan of both crewmembers is a means to evaluate shared awareness. For example, shared awareness can include eye-scan metrics of both pilots' AOI dwell time and transition count. Evaluating the difference between the pilots' AOI dwell time can reveal attention differences. Differences in attention measures can reveal variance in shared situation awareness and therefore, may be used as an indicator of crew coordination. Additionally, pilots failing to perform cross-checks may be due to high workload or stress (Katz et al., 2006).

### 2.4.7 Effect of Crew Formation on Team Processes and Familiarity Decline

Formation of the flight crew can have significant effects on the susceptibility to committing operational errors. Crews that disband after a short number of flights are referred to as formed crews. In addition to a formed crew, concept fixed crew continues to work as a team for an indefinite amount of time. Formed crews are less likely to commit minor in-flight errors than fixed crews with extended personal exposure to the same individuals (Barker et al., 1996).

Differences in crew formation have not increased the likelihood of major operational errors (Barker et al., 1996). However, research suggests that formed crews reduce the probability of a crew experiencing familiarity decline. Familiarity decline refers to reduced crew performance due to a higher level of familiarity between crewmembers. Familiarity decline suggests pilots who are more familiar with each other are less likely to adhere to procedural protocol. Failure to adhere to procedural protocol leads to poor coordination and is an observable crew behavior (Barker et al., 1996). Therefore, crew formation has a significant effect on crew coordination.

www.manaraa.com

### 2.4.8 Stress and Emotion

Stress is described as the interaction between dynamic elements, including perceived demand, perceived ability to cope, and the perceived importance of being able to cope with the demand (McGrath, 1976). The definition by McGrath (1976) leads to a more perceptual theory of stress and emotion and a critical interpretation of cognitive appraisal theory; however, it lacks a specific link to human performance (Staal, 2004). Human behavior models must consider the mechanisms that affect stress in order to establish a link between human performance and stress. Extending the concept of stress to a crew coordination framework adds increased complexity due to interaction effects that influence the individual crewmembers.

The Yerkes-Dodson curve is an attempt to bridge human arousal to human performance as shown in Figure 4. Arousal is defined as the level of central nervous system activity along a behavioral continuum ranging from sleep to alertness, otherwise referred to as a continuum of engagement (Razmjou, 1996). It is clear that stress and emotion are closely tied to the level of arousal and therefore, impact performance. Understanding stress and emotion are impactful factors to human performance; it is important to consider the current state of each pilot in terms of ability to perform coordinated functions. The Yerkes-Dodson theory revolves around the centralized thought that arousal can be modeled by psychophysiological states. Psychophysiological states are affected by stress and emotion and are compounding factors that impact cognition (Neiss, 1988). Therefore, the ability to accurately measure stress and emotion becomes increasingly valuable to the characterization psychophysiological state and the impact on crew coordination and relative performance.

المنارة للاستشارات
www.manaraa.com

Stressors are of particular interest due to their impact on performance. Two negative influences are associate with stressors, including a distracting influence taking attention away from primary tasks, and stress overload, which increases arousal past optimal levels suggested by the Yerkes-Dodson curve (Teichner, Arees, and Reilly, 1963). Identification of expected stressors in an environment is critical to developing human behavior models. Stressors affect the cognitive state of a crewmember. Stressors specific to the flight deck environment are often task-related events, such as interruptive events, unexpected attitude or loss of awareness. Stressors may also develop from crew interaction factors, such as the impact of personality, attitude, and familiarity, as described in Section 2.4.2 and 2.4.7.

Arousal and cognition are dynamic components of the creation of emotional states. Arousal and cognition are codependent; in other words, when both arousal and explanatory cognition are present, individuals experience emotional states congruent with those around them experiencing the same environmental stimulus. Emotional response is therefore a preparatory step to formulate action based upon the acting stimulus. Emotions are seen as managing both motivational resources and regulating behavioral and cognitive activation (Frijda, 1987), (Panksepp, 1996).

المنارة للاستشارات
www.manaraa.com

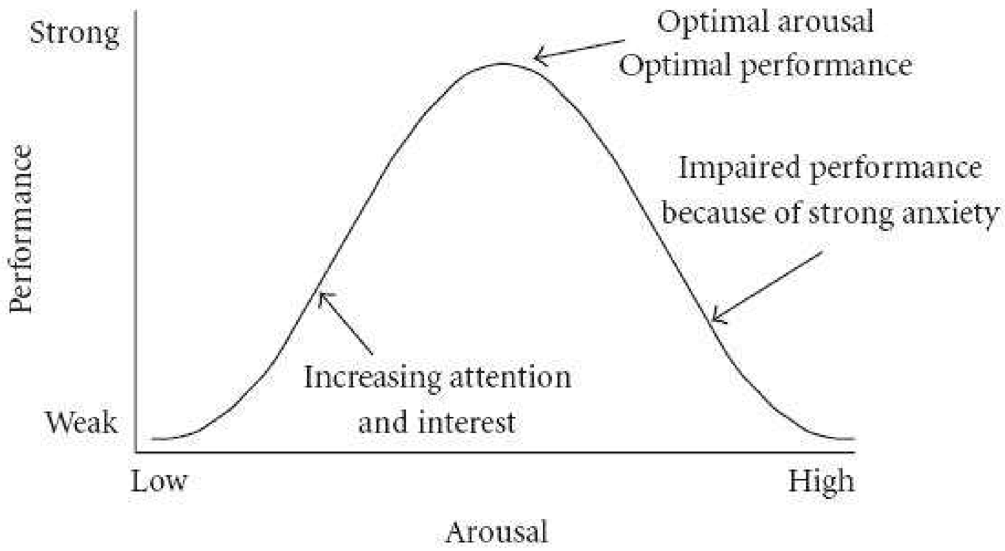


**Figure 4. Yerkes-Dodson Curve**

Aircraft are controlled by pilots and flight automation systems that are capable of acting independently, and may directly or indirectly have an effect on the actions and behavior of the crew. A basic theory of crew coordination performance is built upon the ability of pilots interacting effectively to coordinate efforts. The current state of each pilot is indeed affected by stress and emotion and can be characterized by changes in their psychophysiological state as measured through pilot eye-scan. It is therefore reasonable to conclude that eye-scan models describing the arousal state of each pilot can be used to develop a model of crew coordination. The individual pilot state and reaction to team dynamics are identified through CRM behavioral indicators as discussed in Section 2.4.1. As described earlier, the flight deck is a performing entity composed of the PF, the PM and the flight automation system, all of which are state definable. Arousal theory is an appropriate method to characterize pilot state and the effects of arousal on crew coordinated performance. The relationship between pilot engagement and performance can be used to define hazardous pilot states in a manner similar to the Yerkes-Dodson curve, shown in Figure 5 (Pope & Bogart, 1992).

المنارة للاستشارات

www.manaraa.com

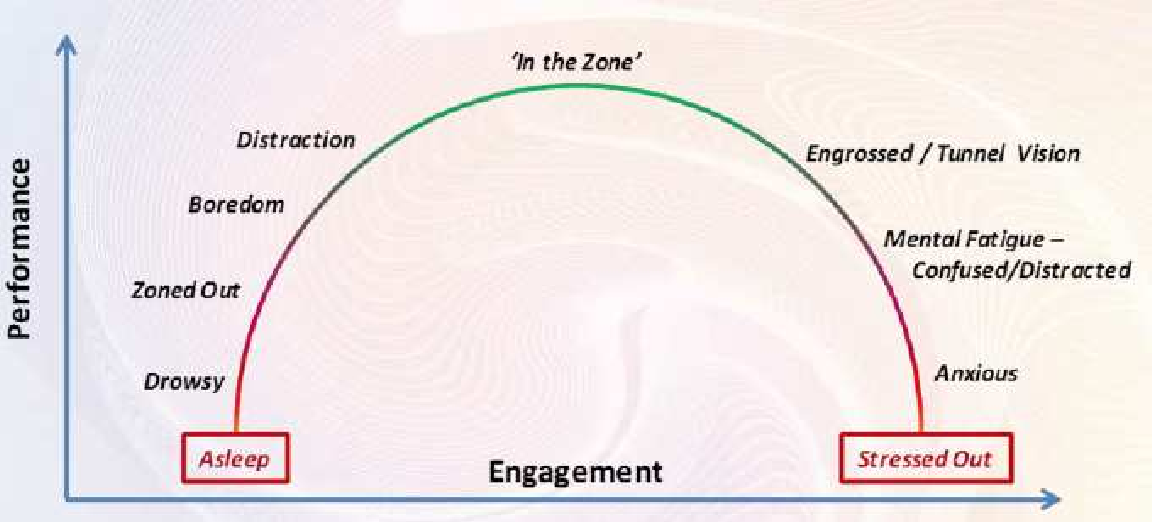


**Figure 5. Modified Yerkes-Dodson, Engagement Relative to Performance (Pope & Bogart, 1992)**

Resource theory is another approach that explains the way individuals cognitively handle tasks and associated workload. Resource theory suggests that there are a limited number of resources available to an individual useful to processing task demands. Resources are defined by capacity, attention, and effort (Wickens, 1984). Resources are described as "… a small set of scarce commodities within the human information processing system, which is associated with a distinct physiological structure, and with physiological arousal changes as increased demands are placed on it." (Wickens, 1991).

Wickens' (1991) resource theory suggests that there are three possible factors engaged in concurrent task management and performance outcomes: confusion, cooperation, and competition. Confusion of task elements explains similar tasks interfere with overall performance, while more distinct tasks have a less frequent degradation to performance. Cooperation in task elements is defined as similarities in task processes that yield combined results, such as monitoring a display and manipulating a flight control.

www.manaraa.com

Competition of resources is observable when allocating attention to one AOI versus another. Competition of resources results in decreasing supply of resources to manage other tasks. Wickens (1991) suggests cooperation between tasks improves when pilots use separate resources as opposed to shared resources. Resource theory draws a parallel to mental effort and resources, concluding that pilot performance is equal to the resources available relative to task difficulty. The relationship between available resources and task difficulty attempts to describe factors that affect performance rather than predict performance (Staal, 2004).

Research contends there are multiple groups of resources as opposed to a single generic set, such as visual, cognitive, and external aides (Marsh, Hicks, & Cook, 2004). However, regardless of the resource group, one can assume that several resources are continually utilized to accomplish pilot tasks. It is difficult to develop tools that validate models based on resource theory, as there are no means to determine the available resources a pilot may have. However, operator state affects the resources available to a pilot. Therefore, resource theory provides a descriptive logic to correlate task difficulty and operator state to performance.

A single tool, such as eye-tracking, is not capable of determining all resources utilized by pilots. Therefore, a single tool is not capable of fully characterizing pilot state. However, combining tools to measure all forms of pilot resource utilization will better define overall pilot state. Additionally, eye-tracking is a useful means to evaluate a pilot’s visual resources by observing which AOIs are viewed. Moreover, understanding what visual resources a pilot is using is necessary in estimating pilot state and its affect on pilot performance.

المنارة للاستشارات

www.manaraa.com

### 2.4.9 Individual and Crew Coping Strategies

Stressors affect emotional state, which forces pilots to resort to coping mechanisms. Coping mechanisms may be trained or instinctual when acting out an emotional response. Emotional responses are evident in nearly all psychophysiological measures, for example heart rate, which has been studied in great depth in the context of aviation (Bonner & Wilson, 2002). Understanding how pilots react to stress with particular coping strategies leads to an increased knowledge of pilot state. Coping strategies are categorized into three primary types: problem- or task-focused coping, emotion-focused coping, and avoidance coping. (Staal, 2004)

Problem or task-focused coping is the most desirable strategy when under stress. Pilots who exhibit problem- or task-focused coping strategies have sufficient situation awareness, technical skill, strategy, and experience to handle the stress of the task load. Effective pilots maximize the use of problem- and task-focused strategies to proactively accomplish tasks, thereby alleviating stress. Emotional-focused and avoidance coping strategies are considered negative responses to stressors on the flight deck. Emotional-focused and avoidance coping strategies are a result of a deficiency in situation awareness, technical skill, strategy, and experience. Pilots may handle stressors with an emotional response or may avoid the stressors altogether, depending on personality and assertiveness. The ability to identify which coping strategy is being exhibited provides further insight into pilot cognitive state.

In the context of crew coordination, negative coping strategies may cause increased stress on other crewmembers. Increased stress on both crewmembers may be evident even if one pilot is exhibiting a problem- or task-focused strategy and the other pilot is not. In a worst-case scenario, both pilots exhibiting one or both of the negative coping strategies results in undesired operational performance. Poor performance due to an emotionally reactive response or a lack of response altogether is often detected in the

المنارة للاستشارات

www.manaraa.com

control inputs made by the crew. Negative responses may reduce attention to instrumentation, ultimately affecting performance. Coping strategies help describe emotional state and how emotional state affects pilot action. Understanding coping strategies also helps provide reasoning for differences in pilot eye-scan behavior. Eye-scan behaviors, such as channelized attention, are the result of inappropriate coping strategies.

### 2.4.10 Cognitive Appraisal, Attentional Bias, and Channelized Attention

Crew eye-scan patterns are indicative of the information pilots seek to attain the situation awareness necessary to complete a task. However, pilots with heightened anxiety, stress, or high workload display an attentional bias toward any threatening stimuli producing the high workload. Attentional bias cognitively blocks the saliency of other sources of information leading to a decrease in situation awareness. Another form of attentional bias is channelized attention. Channelized attention is an eye-scan restricted to a small set of information due to high-cognitive workload. Channelized attention reduces the situation awareness of other sources of information that may be necessary to a task. Therefore, limited eye-scan is indicative of high workload required to complete a task set. If both pilots are exhibiting attentional bias or channelized attention, one can assume the task load of the crew is cognitively demanding and situation awareness is decreasing.

Another approach to describe the impacts of stress and emotion is network theory. Network theory suggests a state of high emotion and stress can lead to activation of memory representations congruent to pilot cognitive state, resulting in selective processing of information (Bower, 1981). Therefore, the current state of the individual pilot is dependent on the cognitive appraisal of task load, which can be evaluated through

المنارة للاستشارات

www.manaraa.com

observing the way information is visually processed or scanned. Determining how and when the pilot views information sources can be used to evaluate information processing. Determining when and how information should be processed can be modeled as expected behavior, also known as normative behavior. Observing pilot eye-scan behavior that is different than expected can reveal attentional bias or channelized attention. Therefore, eye-scans that differ from normative behavior can reveal reduced cognitive states.

Crew behavior can result in some attentional bias. When task-saturated, a well-coordinated crew is expected to reappraise their tasks and balance their workload through the use of proper CRM techniques. Task balancing requires a temporary shift in eye-scan behavior to address the tasks of the task-saturated crewmember. Ultimately, the crew should exhibit a reduction in attentional bias once the crew is properly coordinated. If both pilots exhibit attentional bias, current task loads or confusion of tasks may be overwhelming to the whole crew. Evaluating both pilots' eye-scan behavior is a viable means to assess situation awareness and infer cognitive demand. Evaluation of cognitive demand between pilots leads to an estimation of the level of crew coordination.

### 2.4.11 Effects of Training

The development of schemas, prototypes, scripts, attitudes, and stereotypes help reduce cognitive loading (Neuberg & Newsome, 1993). Training promotes development of effective cognition, allowing pilots to react efficiently in situations for which SOPs can be developed. Helmreich (1982) discussed training as an effective means to appropriately respond with coordinated effort to changes in tasks. Effective coordination requires task specific schemas, prototypes, scripts, attitudes, and stereotypes, which are developed by training and experience. However, training is only capable of presenting expected scenarios and foreseeable failure modes. Training highlights the significance of

المنارة للاستشارات

www.manaraa.com

understanding the flight deck interface, as well as understanding the difference between experienced and inexperienced pilots.

Flight time is experiential training that exists in addition to the recurrent training pilots are required to participate in to maintain active transport pilot licenses. Effects of training are detectable in scan patterns, revealing allocation of attention. Trained and experienced pilots in familiar and unfamiliar situations revert to the priority tasking hierarchy - aviate > navigate > communicate (Schutte & Trujillo, 1996). Observations specific to variation of pilot eye-scan begin to emerge from the priority tasking hierarchy. An untasked pilot does not have predictable eye-scan behavior, and therefore, a completely random eye-scan scan is expected. Pilots with a single task exhibit a specific scan pattern, allocating attention to required sources of information with little variance. Multi-tasked pilots exhibit priority-specific scan patterns learned from training and experience.  Trained, multi-tasked pilots, direct their attention to AOIs in order to complete tasks by following the hierarchy of aviate>navigate>communicate. Therefore, multi-tasked pilots exhibit predictable eye-scans specific to task hierarchy.

The effects of training, including effects of CRM training, are most identifiable in multi-tasked pilots. Multi-tasked pilots with experience and training will exhibit a specific scan pattern, often in coordination with fellow crewmembers. A well-coordinated crew exhibits specific scan patterns to accomplish delegated tasks. Untrained crews exhibit a more unpredictable scan pattern. The multi-tasked, untrained crew possesses fewer techniques and knowledge to complete tasks and appropriately delegate responsibilities to balance workload.

CRM training trains pilots to delegate tasks to all available resources in order to address all necessary tasking.  The PF, the PM, and the flight automation are available resources on the flight deck to accomplish flight tasks. Airlines train schemas and task scripts in the form of SOPs. SOPs, which are developed for each phase of flight, guide flight crew interaction in order to minimize errors and confusion. Normative behavior

www.manaraa.com

models can be developed from SOPs in order to predict optimal PF and PM attention allocation. Pilots may deviate from SOPs and the normative behavior model; in fact, it is expected. However, under the assumption that SOP guided task procedures are developed to allow pilots to perform optimally, deviations from normative behaviors are indicative of reduced coordination.

### 2.5 Flight Deck Interaction Model: Factors Affecting Visual Search

A flight deck interaction model was developed to highlight significant factors of pilot state and crew coordination, shown in Figure 6. The flight deck interaction model is structured on the system and operator factors model of Hilburn and Jorna (2001). The flight deck interaction model is expanded to include inputs to system task loading and its affect on pilot crew and flight automation system interactions. The model includes factors affecting pilot situation awareness along with a crew-coordination construct. The crew coordination construct leverages shared situation awareness between pilots to define the level of coordination in the system.

المنارة للاستشارات
www.manaraa.com

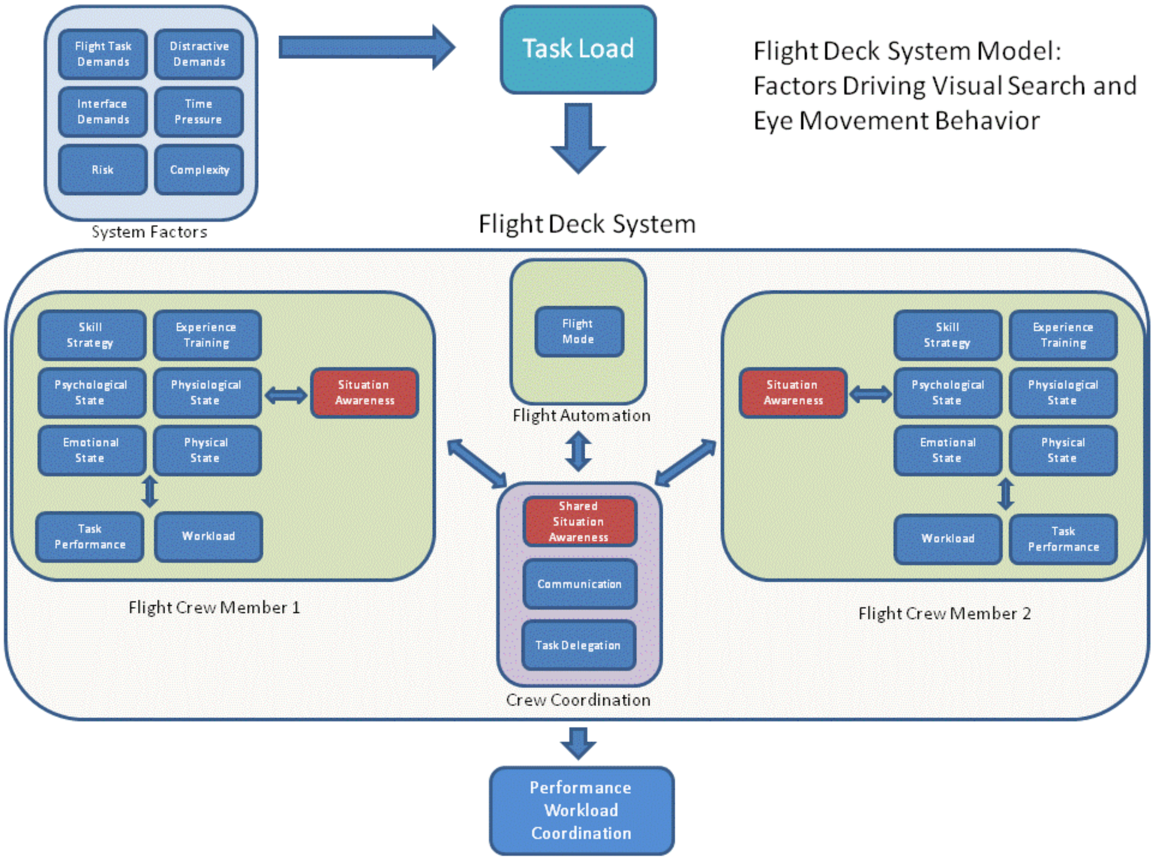


**Figure 6. Flight Deck System Factors Model**

The flight deck interaction model is a complex system of interdependent factors reduced to a framework of three components: system task load factors, flight deck system factors, and system output. System task load factors are defined by the current environment and desired future state of the aircraft. Flight deck system factors include the adaptive features of the flight deck system, represented by each individual pilot and flight automation subcomponents. Pilot subcomponents identify the factors that affect the relationship between performance/workload and situation awareness. The flight automation subcomponent is defined by flight mode. The flight deck system subcomponents are linked through a crew coordination construct, defined by shared situation awareness. Lastly, the system output component describes the performance,

www.manaraa.com

workload and crew coordination of the flight deck system as a function of shared situation awareness.

The flight deck interaction model closely follows Hancock and Warm’s (1989) arousal replacement framework model, which incorporates the system/operator state factors model presented by Hilburn and Jorna (2001). The flight deck interaction model identifies the internal and external factors that affect pilot situation awareness. The coordination component describes the relationship between individual and shared situation awareness. Therefore, the flight deck interaction model explains how eye-scan measures of situation awareness may quantitatively infer the level of crew coordination, workload, and performance.

## 2.6 Crew Coordination and Eye Movement Behavior

Eye-scan behavior in this thesis is defined by the percentage dwell time (PDT) values for each AOI and is used to develop normative models of attention. Comparing observed eye-scan behavior with normative eye-scan behavior can reveal variability of pilot attention, caused either by task overload or task underload. Evaluating whether a pilot looks at an AOI more or less when compared to the normative model implies a deviation from normative phase of flight workload. Therefore, deviations from the normative eye-scan behavior model are indicative of a decrease in crew coordination and increase in probability of pilot error.

### 2.6.1 Pilot Scan and Normative Tasked Behavior

Changes in eye scanning behavior are indicative of a variance in pilot task loading and therefore pilot workload. The effort of shifting attention and performing visual cross-check is identifiable through comparing the difference between pilot AOI PDT and AOI

المنارة للاستشارات
www.manaraa.com

shared attention, described in Section 2.4.4. AOIs are grouped into three categories: aviate, navigate, and communicate (Wickens et al., 2003), (Schutte & Trujillo, 1996). Aviate and navigate task AOIs include the PFD, AI, ALT and OTW. The communicate task AOIs include only the CDUs.

The driving constructs behind the normative eye-scan behavior are CRM and SOPs. The underlying core of CRM is shared situation awareness to efficiently delegate roles and responsibilities to balance the workload across the crew. Normative eye-scan behavior models of each pilot can be used to compare observed eye-scan behavior to identify deviations from optimal coordinated task behaviors. Departures from expected CRM and SOP behaviors are posited to be an indication of reduced crew coordination.

Differences between observed PDT and normative model PDT could be assessed for statistical significance to identify the level of crew coordination. However, to develop a normative behavior model, it is necessary to analyze normative task load, identifying expected variations in pilot workload and coordination. Pilot scan based on the normative task behavior must be evaluated to analyze differences from observed PDT values.

For a given task set, a normative model can be developed as the basis for optimal eye-scan. The normative eye-scan behavior model defines the nominal attention given to each AOI in order to obtain the necessary information to perform a set of tasks. Therefore, deviations from the normative model are indicative of reduced situation awareness. The flight deck interaction model suggests situation awareness is limited by factors affecting pilot state, such as increased workload.

### 2.6.2 Eye Movement Characterization of Crew Coordination

The PF, PM, and flight automation interact in a coordinated effort to perform tasks. During flight operations, the PF has task delegation authority. The PF utilizes all

www.manaraa.com

available resources, including the PM and the flight automation, to complete all phases of flight operations. SOPs define the interaction process of both pilots and how flight automation is used. The flight automation is capable of aviating, navigating, and communicating, and must be monitored when it is engaged. Shared situation awareness is essential to understanding which tasks are executed and who is performing them. Eye-tracking provides a quantitative measure of where the pilot is looking in order to gain situation awareness. Pilot eye-scan is driven by the need for situation awareness, which is necessary to accomplish a specific task. For example, a pilot manually flying an aircraft on short final in low visibility is focused on aviating and navigating the aircraft by following navigational guidance. In order to perform the flight operation tasks, the pilot provides input to the flight controls and measures the response feedback from the primary flight display providing guidance and adjusts the control inputs accordingly. The PM cross-checks the flight instruments while also looking OTW to visually acquire and call out the runway environment necessary to land (Ellis, Kramer, Shelton, Arthur, Prinzel, & Norman, 2011). The necessary scan to complete the flight operation is observable and predictable. Therefore, measurable comparison of observed pilot eye-scan to an optimal task eye-scan is possible.

A normative eye-scan model for both the PF and PM must incorporate the optimal coordination behavior between each pilot and the flight automation, based upon phase of flight tasking. Deviations from the normative eye-scan behavior model can be quantified and used to contrast the situation awareness of each pilot and estimate the level of crew coordination. Highly coordinated crews are expected to perform the tasks necessary to execute the flight operation as defined by the normative eye-scan model for each pilot. Poorly coordinated crews are not expected to execute tasks according to SOPs and proper CRM training and therefore will not exhibit an eye-scan behavior similar to the normative eye-scan model. Therefore, the difference in the output of the model for each pilot is expected to be a measure of crew coordination, shown in Equation 1.

المنارة للاستشارات

www.manaraa.com

### Equation 1 . Crew Coordination Index

$$\text{Crew Coordination} = \text{Absolute Variance (pilot1)} + \text{Absolute Variance (pilot2)}$$

Normative eye-scan model approaches to quantifying crew coordination are only viable for phases of flight with trained SOP and CRM behaviors. Deviations from the normative eye-scan model do not necessarily describe poor coordination, but may be indicative of unexpected events that affect crew coordination. Unexpected events may occur during any phase of flight and may lead to non-standard flight conditions, such as unusual attitude or loss of energy state awareness. Unusual attitude or loss of energy state awareness is an unsafe flight condition that may lead to loss of the aircraft, passengers, and crew.

In the event of unusual attitude or loss of energy state awareness, eye-tracking is a significant source of information to ascertain the onset of a loss of energy state awareness and deduce the type of spatial disorientation a pilot experiences. There are three types of spatial disorientation: unrecognized (type I), recognized (type II), or incapacitating (type III) (Australian Transport Safety Bureau, 2007). Monitoring pilot attention distribution and aircraft control inputs make it possible to determine if a pilot is spatially disorientated. Reduced situation awareness is the causal factor leading to loss of energy state and is detectable using attention characterization metrics.

Currently, there are no means to determine expected pilot response and task delegation during unexpected events and unsafe flight conditions. However, eye-tracking information can determine which AOIs pilots are viewing. The relationship between situation awareness and the factors that define the state of a pilot are described by the flight deck interaction model, shown in Figure 6. Situation awareness, defined by attention characterization metrics, is affected by the factors of pilot state and therefore

المنارة للاستشارات

capable of inferring the state of the pilot. The crew coordination characterization model, shown in Figure 7, replaces the pilot state factors of the flight deck interaction model with attention characterization metrics to describe the level of coordination between pilots. Combining the knowledge of crew and aircraft state allows one to predict the onset of unsafe flight conditions and provide countermeasures to prevent such conditions.

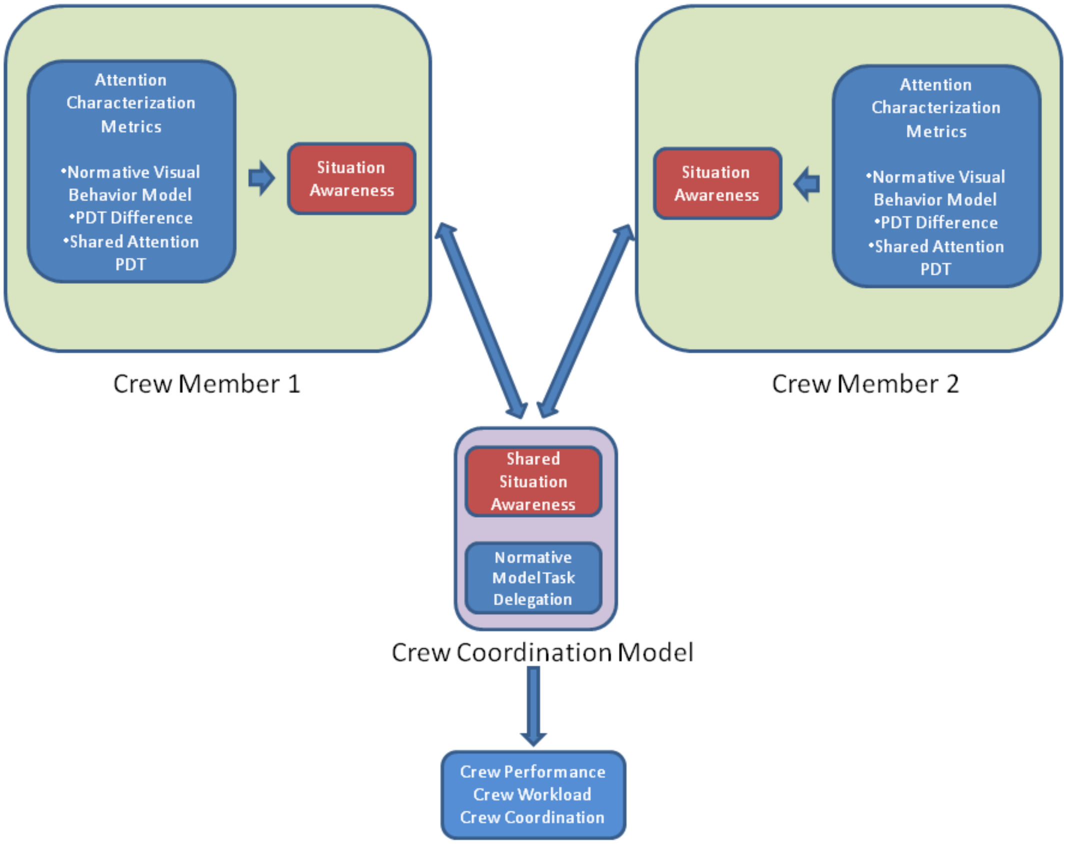


**Figure 7. Crew Coordination Characterization Model Diagram**

المنارة للاستشارات

www.manaraa.com

# CHAPTER 3.

# 757 SIMULATION CREWED DATALINK STUDY

## 3.1 Methodology

The NASA/FAA DataComm experiment was a human-in-the-loop (HITL) experiment originally designed around the utilization of new developments in the delivery of terminal operation clearances (Norman R. M., Baxley, Ellis, Adams, Latorella, & Comstock, 2010). In current flight terminal operations, clearances are made from air traffic control (ATC) to the flight crew via voice communication and a callback from the crew to ATC to confirm and monitor for errors in accepted clearances. DataComm is an accurate, persistent, auto-gate to flight management system onboard the aircraft (Norman et al., 2010). The effect on operator situation awareness, workload, and crew coordination was uncertain. To supplement the DataComm system, a moving map display was implemented to show proposed clearances, and upon acceptance by the flight crew, accepted clearances. The moving map display conditions added visual confirmation of what the crew received via the DataComm with respect to taxi route information. Display variations were found to have little effect on pilot attention across the AOIs during the arrival scenarios (Norman et al., 2010).

The DataComm experiment was designed to determine the effect of DataComm on pilot crew on the flight deck during operations in the terminal area. Leveraging the DataComm experiment conducted at NASA Langley, additional objectives were embedded into the design of the experiment to identify the following:

- The effects of DataComm modality on flight crew workload, situation awareness, and coordination during arrival operations
- Variation in crew visual behavior across phase of flight and additional task loading imposed by DataComm

المنارة للاستشارات

The eye gaze of each pilot was monitored using a remote eye-tracking system in order to assess deviations in crew eye-scan behavior. A normative workload model based on variable task loads for each scenario was developed to evaluate appropriate flight segments for data analysis. The normative workload model was used to determine expected crew behavior when executing tasks as defined by SOPs and CRM. Attention allocation metrics were used to evaluate changes to the situation awareness and workload of each pilot. The crew coordination characterization model was then used to assess the level of crew coordination.

## 3.2 Crew Coordination Hypotheses

Research Hypothesis 1: Shared understanding of crew workload is indicative of good crew coordination. Additional tasking through the use of the DataComm communication interface was expected to force crews to adapt to new tasking apart from standard terminal approach procedures typically experienced in traditional terminal area operations. Non-normal tasking between pilots decreases the shared understanding of workload between pilots. Additionally, the increase in non-normal tasking was expected to create an imbalance in crew workload and therefore reduced crew coordination.

Research Hypothesis 2: Attention allocation metrics can characterize normative tasking and expected workload for each crewmember role. The normative eye-scan model defines the average PDT in each AOI when crews exhibited excellent coordination during baseline communication and display conditions (Voice/Paper).

Research Hypothesis 3: Variations in attention allocation metrics correlate with a reduction in crew coordination. Variations in attention allocation metrics are

www.manaraa.com

representative of a shift in crew attention from normative behavior and will be indicated quantitatively in the output from the crew coordination model analysis approach described in Section 2.6. There are three attention allocation metrics evaluated: Deviation between observed and normative PDT values, Variation in the difference between PF and PM PDT values for each AOI, and the shared attention metric. Deviation of observed PDT values from the normative eye-scan model correlates with a reduction in crew coordination. Differences in the PDT between the PF and the PM correlate with a reduction in crew coordination. Reduction in the shared awareness measure correlates with a reduction in crew coordination.

## 3.3 Design of Experiment

The DataComm experiment used two independent variables to drive individual pilot and crew workload. The first independent variable was communication modality. Two communication modalities existed, varying between voice-only and DataComm operation. The voice communication modality was used as the baseline communication modality against which the DataComm modality was compared. The DataComm modality introduced the use of the DataComm system, providing clearance information from controller to flight crew, with voice modality still available for time-critical or safety-related information.

The second independent variable of the DataComm study was display methodology. Three variations of displays were used: paper map, moving map display (MMD), and moving map display with route (MMD+R). Combining the two independent variables together created the experiment test conditions. The voice/paper condition was used as the baseline condition, requiring pilots to utilize paper approach charts and airport maps, consistent with most current operations. The DataComm/MMD and condition included taxiways, runways, signage, and ownship position, providing real-time location

المنارة للاستشارات

www.manaraa.com

information to the crew. The third condition, DataComm/MMD+R, included all of the same information as the MMD condition, but added a graphical display of the expected and actual DataComm ownship route clearance. Display methodology had no effect in the context of this dissertation, as the display conditions presented in the DataComm experiment only affected pilot attention during taxi operations (Norman et al., 2010). Communication modality was significant with respect to its effect on pilot task load in the arrival scenarios of the DataComm experiment.

The DataComm experiment included 18 runs of flight crews in a flight simulator, testing eight scenario conditions. Testing included inbound standard terminal arrival routes (STARs) and instrument approach procedures (IAPs) to runway 27 and runway 33L. The STAR and IAP terminal operations were crossed with the two-communication modality and three display methodology independent variables. A total of 16 DataComm runs were randomized among the eight scenario conditions, with two replications per crew. Two additional runs to evaluate crew trust were tested but not utilized in the evaluation of crew coordination.

### 3.3.1 Scenario Descriptions

The DataComm experiment included arrival and departure scenarios at Boston Logan International airport (KBOS) and utilized a combination of current published instrument arrival procedures and clearances given by controllers. The evaluation of crew coordination investigated arrival scenarios only and did not utilize the DP scenarios. Arrivals to Runways 33L and 27 were created to provide realistic profiles and workload in the terminal area from an initial starting altitude of 18,000ft. and continued to landing and rollout. The Norwich Three and Scupp Four arrivals were chosen and shown below in Figure 8.

المنارة للاستشارات

www.manaraa.com

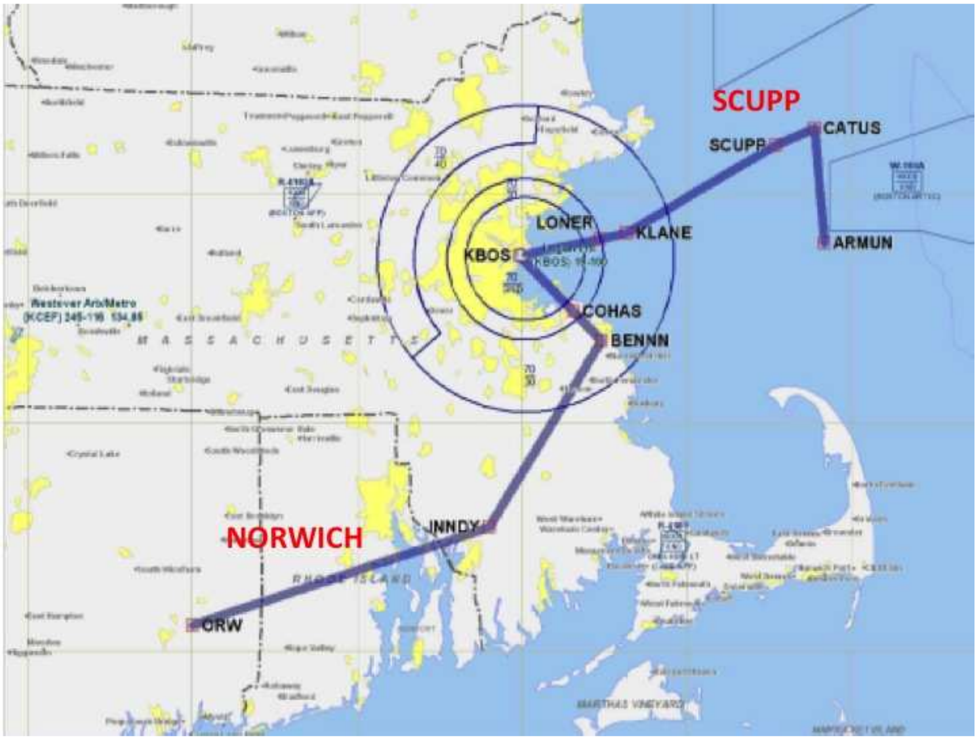


**Figure 8. KBOS NORWICH 3 and SCUPP 4 arrival routes**

Portions of the Norwich Three and Scupp Four Standard Terminal Arrival Routes (STAR) were connected to the appropriate initial approach fix (IAF) and final approach fix (FAF) for each runway approach end. Pilots were required to arm the approach and capture the localizer and glideslope of the Instrument Landing System (ILS) to continue the approach. ILS approaches to runway 33L and runway 27 are shown below in Figure 9 and Figure 10.

www.manaraa.com

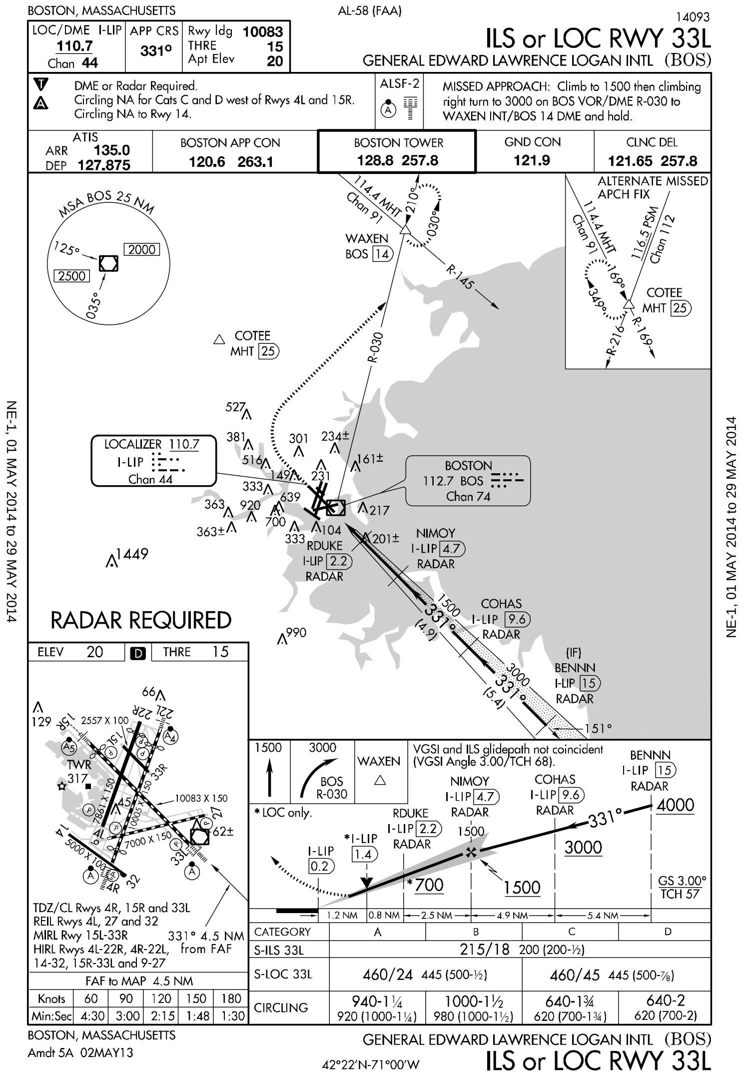
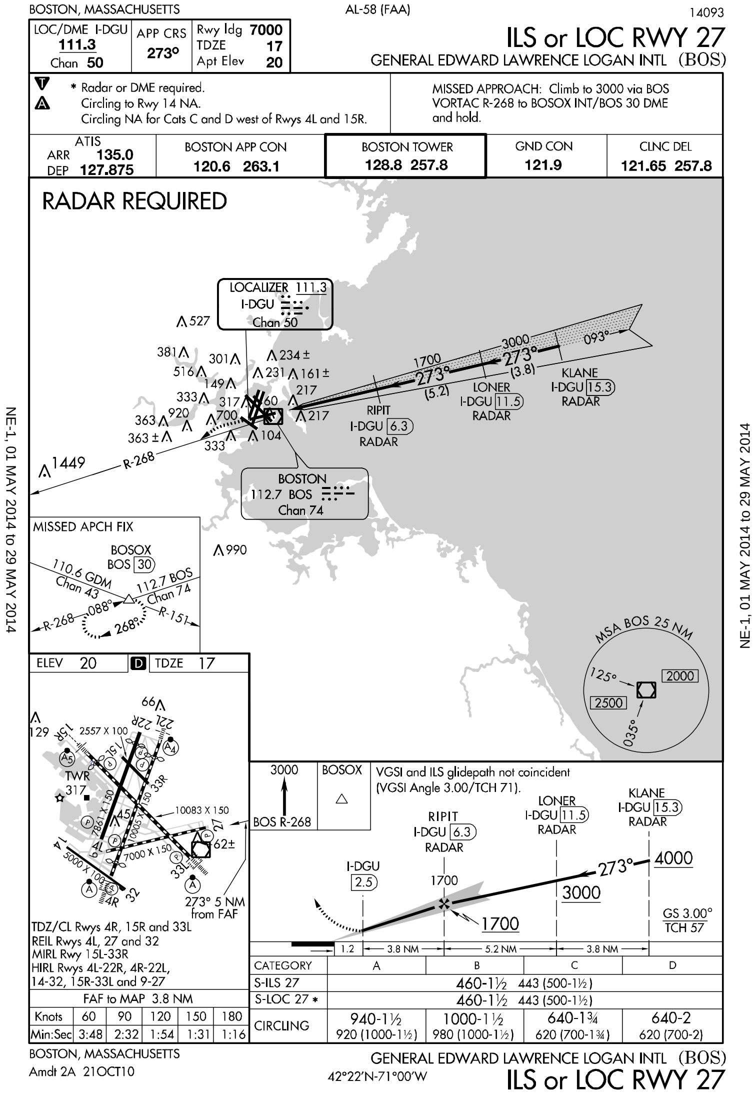


Figure 9. ILS RWY 33L

www.manaraa.com



| CATEGORY | A | B | C | D |
|---|---|---|---|---|
| S-ILS 27 | 460-1½ 443 (500-1½) | | | |
| S-LOC 27 * | 460-1½ 443 (500-1½) | | | |
| CIRCLING | 940-1½ 920 (1000-1½) | 1000-1½ 980 (1000-1½) | 640-1¾ 620 (700-1¾) | 640-2 620 (700-2) |

| Knots | 60 | 90 | 120 | 150 | 180 |
|---|---|---|---|---|---|
| Min:Sec | 3:48 | 2:32 | 1:54 | 1:31 | 1:16 |

**Figure 10. ILS RWY 27**

- Norwich Three Arrival, Kennedy Transition (Figure 11)
    - Procedure starts southwest of the airport. The scenario itself started overhead Norwich and proceeded East to INNDY, then direct to the Initial Approach Fix (BENN) for the ILS to Runway 33L.

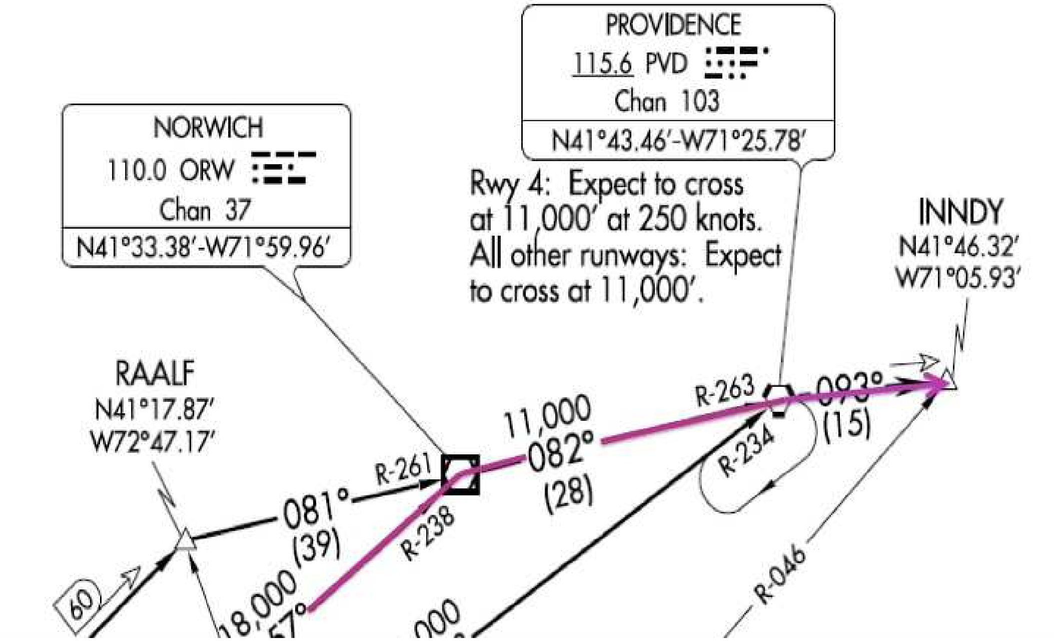


**Figure 11. NORWICH THREE arrival excerpt**

- Scupp Four Arrival, Kennedy Transition (Figure 12)
    - Procedure starts east of the airport. The scenario started overhead ARMUN and proceeded west to KBOS. Approximately 20 miles from KBOS, a clearance was given to KLANE for the ILS to Runway 27.

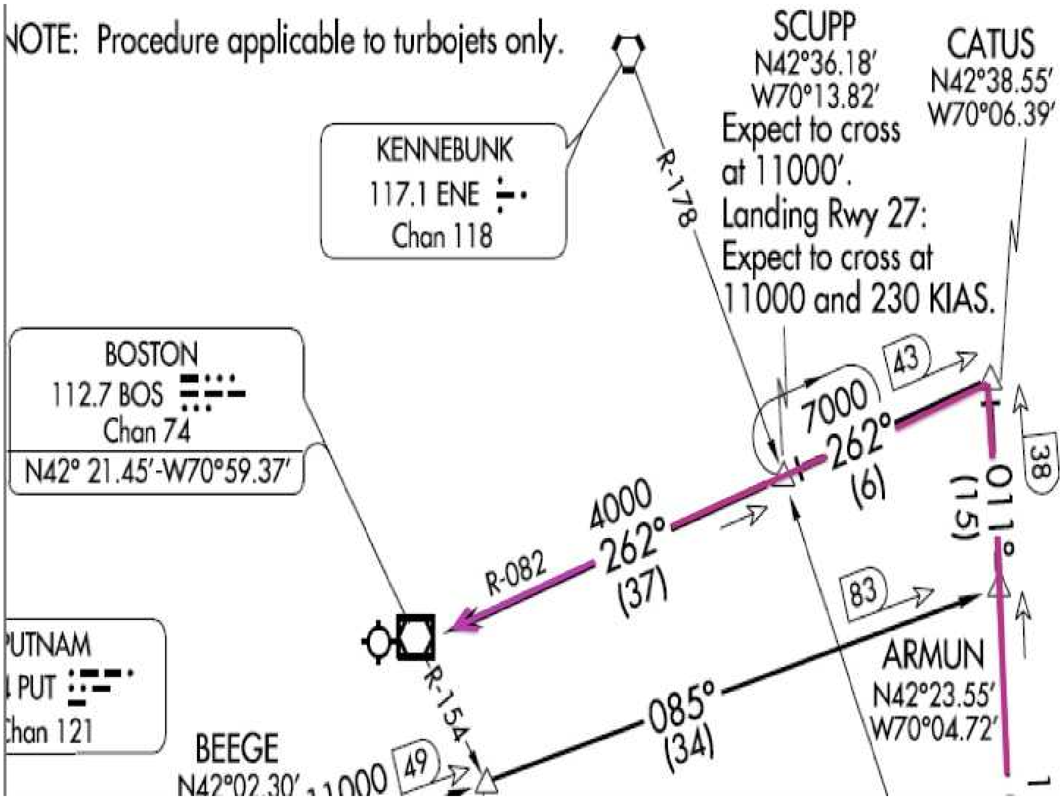


Figure 12. SCUPP FOUR arrival excerpt

### 3.3.2 DataComm Messages and Displays

Each of the 11 crews received 42 DataComm uplink messages (462 total for all crews). Crews were required to respond with a downlink message to each uplink message received. The aggregate count of these messages per crew and over the entire experiment is shown in Table 2.

**Table 2. DataComm Messages Per Crew for All Arrival Scenarios**

| **Arrival DataComm messages** | **Per Crew** | **Total** |
|---|---|---|
| • Expected D-TAXI | 12 | 132 |
| • ATIS | 6 | 66 |
| • Altimeter | 6 | 66 |
| • Frequency change | 6 | 66 |
| • D-TAXI | 6 | 66 |
| • Amended D-TAXI | 6 | 66 |
| • **TOTAL** | 42 | 462 |

www.manaraa.com

DataComm message format and page architecture were modeled after the Boeing 747-400 Future Air Navigation System 1/A (FANS-1/A) implementation. Display shapes, sizes, and colors on the ND were based on ongoing research at NASA Langley, the proposed DataComm standards, and discussions between members of the FAA and NASA DataComm team (Norman et al., 2013).

The crew accessed DataComm messages by depressing the CDU button labeled "ATC" (located on the top row of which caused the "ATC Index" page to be displayed on the CDU screen (left side of Figure 13). The "Prev Page" and "Next Page" CDU buttons were used by the flight crew to access the different pages of the controller pilot data link communication (CPDLC) message, with the ability to send a CPDLC response always on the last page of the message (right side of Figure 14). The FANS-1/A ATC index page is shown on the left of Figure 13. ATC Index (left) and ATC Request (right) pages. Depressing the button labeled "Request" accesses the ATC Request page. Depressing the fourth button on the left side accesses the ATC Log page. Depressing any button on the right side of the ATC Log (left side of Figure 14) brings up the respective DataComm message, such as the D-TAXI messages. A separate button on the CDU panel must be depressed to reach the second page of the DataComm message where the downlink response can be sent by the crew (right side of Figure 14.)

www.manaraa.com

```
ATC INDEX
<EMERGENCY      POS REPORT>
<REQUEST      WHEN CAN WE>
<REPORT
<LOG            CLEARANCE>
<LOGON/STATUS       VOICE>
- - - - - - - - - - - - -
<PRINT LOG
```

```
ATC REQUEST
ALTITUDE
<- - -             STARTUP>
SPEED
<- - -            PUSHBACK>
OFFSET
<- - -     TAXI CLEARANCE>
<ROUTE REQUEST
<ERASE REQUEST
- - - - - - - - - - - - -
<INDEX              VERIFY>
```

**Figure 13. ATC Index (left) and ATC Request (right) pages**

```
ATC LOG
 1340Z BOS            ROGER
↑ CLEARED TO START>
 1340Z BOS            WILCO
↑ PUSHBACK AT 1345Z>
 1338Z      RESPONSE RCVD
↓ REQUEST START CLEAR...>
 1338Z      RESPONSE RCVD
↓ REQUEST PUSHBACK CL...>

<INDEX           ERASE LOG>
```

```
1420Z  ATC UPLINK   2/2
(KBOS)




<STANDBY
<REJECT              WILCO>
- - - - - - - - - - - - -
<PRINT              REPORT>
```

**Figure 14. ATC Log (left) and Downlink Response (right) pages**

### 3.3.3 Experiment Protocol

Prior to the experiment, the participant pilots were scheduled two at a time over the course of several weeks. Pairs of pilots were required to be from the same flight organization in order to minimize adverse effects from differing SOPs or CRM principles specific to different airlines. All pairs used standardized, pre-briefed procedures to the maximum extent possible. During the experiment, the pilot qualified as captain performed the role of the PF in the left crew station and was responsible for control of the simulated aircraft throughout the experiment. The pilot qualified as first officer performed the role of the PM in the right crew station and was responsible for

المنارة للاستشارات

www.manaraa.com

DataComm messages for the duration of the simulation. The first officer was the PM throughout the entire experiment to increase the statistical significance of collected data.

Crews arrived at the research facility by 0745 on the first day to complete the required paperwork. At 0800, the formal briefing began, which included completing the informed consent form required by NASA's Institutional Review Board (IRB), followed by a two-hour training program. Training included: the purpose of the experiment, an interactive practice session to familiarize the crew with sending and responding to DataComm messages, a walk-through of each scenario, and a practice session to familiarize the crew with the electronic questionnaires. The pilots were scheduled from approximately 1000 to 1230 to complete part-task training and four training scenarios in the Instrument Flight Deck (IFD), the simulator used for the DataComm experiment. Prior to beginning each scenario, the crew was given a verbal briefing about the upcoming scenario. After each scenario, five to ten minutes were allotted for the crew to answer the electronic post-run questionnaire and to reconfigure the flight deck for the next scenario. After every third or fourth scenario, a break was given to ensure the crew was well rested. Following the last scenario, the crew was brought back to the briefing room to complete the post-experiment questionnaire, which generally took 20 to 30 minutes. Following the post-flight questionnaire, a semi-structured verbal debrief was given to assess the overall effect of DataComm on flight deck operations.

### 3.3.4 Participants

NASA recruited the participant pilots in support of the IFD simulation experiment, complying with all applicable procedures and laws relating to the protection of human participants as specified by the IRB. The following were specific requirements for all participant pilots:

- U.S. citizen or Permanent Resident status.

المنارة للاستشارات
www.manaraa.com

- Valid FAA Airline Transport Pilot certificate.
- Current employment by a Part 121 air carrier.
- Preference was given to participants without hard edge bi-focal or tri-focal glasses.
- Preference was given to pilots who held a Boeing 757 or 767 type-rating; however, other type ratings with CDU/FMS incorporation similar to the 757 / 767 were considered.
- Preference was given to pilots who were familiar with the FANS-1/A CDU controls, displays, and functionality through flight experience.
- Current or recent flight experience in the assigned crew role for the experiment (i.e. Captain or First Officer).

The DataComm experiment consisted of 11 crews of two pilots each, both from the same airline. Based on information collected using the biographical questionnaire, all pilots were male with an average age of 48.6 years and total flying time ranging from 6,000 to 24,000 hours with a mean of 13,832.5 hours. Pilot time in a Boeing 757 or comparable aircraft type ranged from 1,000 hours to 15,000 hours with a mean of 7,768.6 hours. Approximately half of the pilots had prior experience with DataComm, and 19 of the 22 pilots had conducted flight operations into and out of KBOS. Data on participant pilot experience levels is shown below in Table 3.

**Table 3. Subject Pilot Experience Level in Years and Hours**

| | **Mean Age** | Low Age | High Age | Std Dev Age | **Mean Years Flying** | Low Years Flying | High Years Flying | Std Dev Years Flying |
|---|---|---|---|---|---|---|---|---|
| **Captain** | 52.5 | 46.0 | 58.0 | 4.0 | 23.9 | 19.0 | 33.0 | 3.9 |
| **FO** | 44.2 | 37.0 | 56.0 | 5.6 | 15.0 | 10.0 | 26.0 | 4.8 |
| | **Mean Total Hours** | Low Total Hours | High Total Hours | Std Dev Total Hours | **Mean B757 Hours** | Low B757 Hours | High B757 Hours | Std Dev B757 Hours |
| **Captain** | 17614 | 13750 | 25000 | 3784 | 7255 | 1100 | 10000 | 3139 |
| **FO** | 11242 | 6600 | 19460 | 3391 | 5036 | 1100 | 10000 | 3032 |

www.manaraa.com

### 3.3.5 Dependent Variables

Dependent variables included workload, situation awareness, and crew coordination, assessed individually for the PF and PM. Workload was assessed by administering the Bedford Workload Scale and analysis of DataComm message response times. Measurement of the time to respond to DataComm messages was useful in determining the effect of DataComm messages on crew workload with respect to additional tasking relative to normative workload. DataComm message response times were calculated as the difference in seconds from the time a message was initially received (chime annunciated and "ATC MESSAGE" shown on the upper EICAS display) and the time that a response button ("Wilco," "Roger," or "Unable") was depressed on the message ATC uplink on page two. Message response times were averaged for all crews by communication modality and message type. Situation awareness was assessed by administering the SART and analysis of the eye-scan behavior of each pilot. Crew coordination was assessed by administering questions developed by NASA researchers developed from FAA guidance on CRM training found in AC120-51E to each pilot (Federal Aviation Administration, 2004).

Dependent variables:

- Workload
- Situation Awareness
- Crew Coordination

Metrics used to quantify the dependent variables included the following:

- Eye-Scan Behavior: AOI PDT
- Workload: Bedford Workload Scale, DataComm message response time

المنارة للاستشارات

www.manaraa.com

- Situation Awareness: Situation Awareness Rating Technique (SART)
- Crew Coordination: Crew Coordination Index, Questions developed by NASA researchers from CRM FAA guidance found in AC120-51E (Federal Aviation Administration, 2004).

The Bedford Workload Scale, the SART, and crew coordination questions are discussed below in Section 3.3.7. Additionally, crew coordination was assessed using a novel approach that cross-references the workload rating reported by each pilot and evaluates the understanding each pilot has of their crew member's workload. The novel approach to evaluate crew coordination is called the crew coordination index, described in greater detail in Section 3.3.8.

Pilot eye-scan behavior was used to assess the attention pilots gave to each instrument across the flight deck and OTW. Each instrument and the forward view OTW were defined as individual AOIs. Pilot eye-scan behavior was determined using the PDT metric for each AOI. AOIs included the PFD, the ND, the CDU, the IP, the AI, the ALT, and OTW. The PFD, ND, AI, and ALT AOIs were labeled with reference to which pilot side there were located (i.e. PFD-PF, PFD-PM). AOIs are shown below in Figure 14.

www.manaraa.com

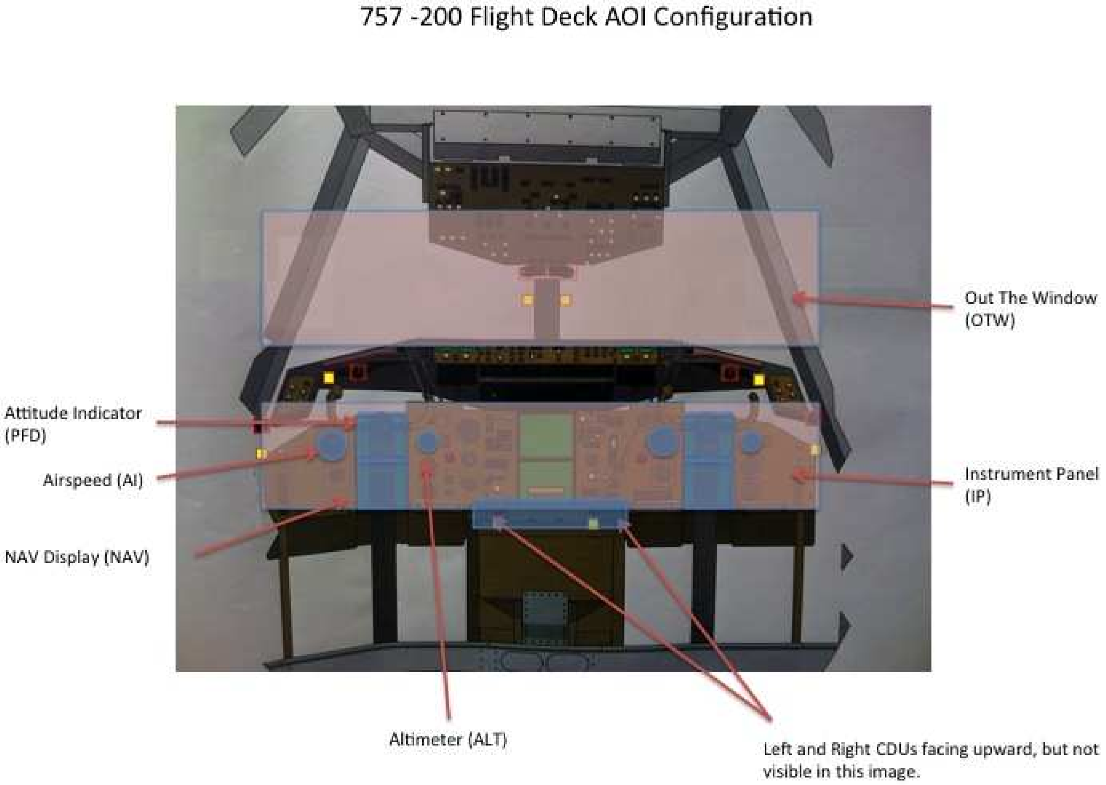


**Figure 15. 757-200 Flight Deck AOI Configuration**

### 3.3.7 Post-Scenario Questionnaire

The post-scenario questionnaire was given to both pilots after each scenario. The post-scenario questionnaire consisted of the Bedford Workload Scale (Roscoe, 1987), the SART (Taylor, 1990), and questions addressing crew coordination, acceptability and trust. The questionnaire was administered electronically on a Hewlett-Packard personal tablet computer while the participant pilot was seated in the simulator. The Bedford Workload Scale is a uni-dimensional rating scale designed to identify an operator’s spare mental capacity while completing a task, shown below in Figure 16. The Bedford scale represents a simple methodology to assess subjective workload in an operational context.

www.manaraa.com

The single dimension was assessed using a hierarchical decision tree (always completely visible to the participant) that guided the operator through a ten-point rating scale. Each point of the ten-point rating scale was accompanied by a descriptor of the associated level of workload.

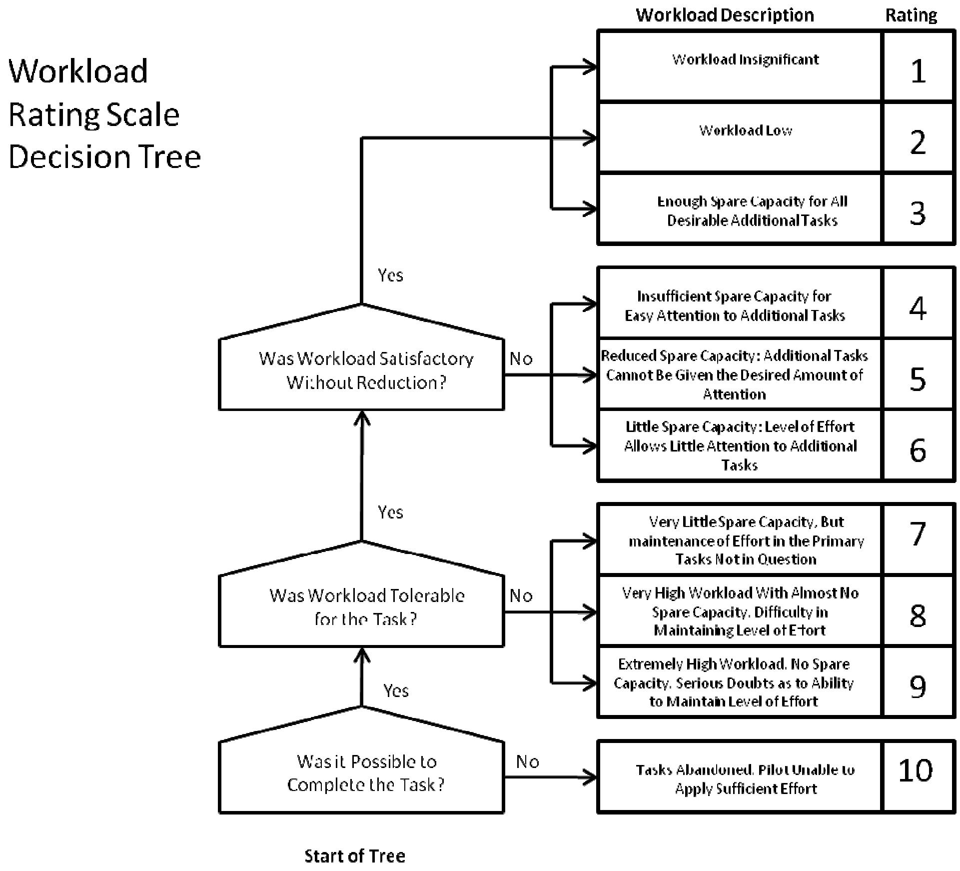


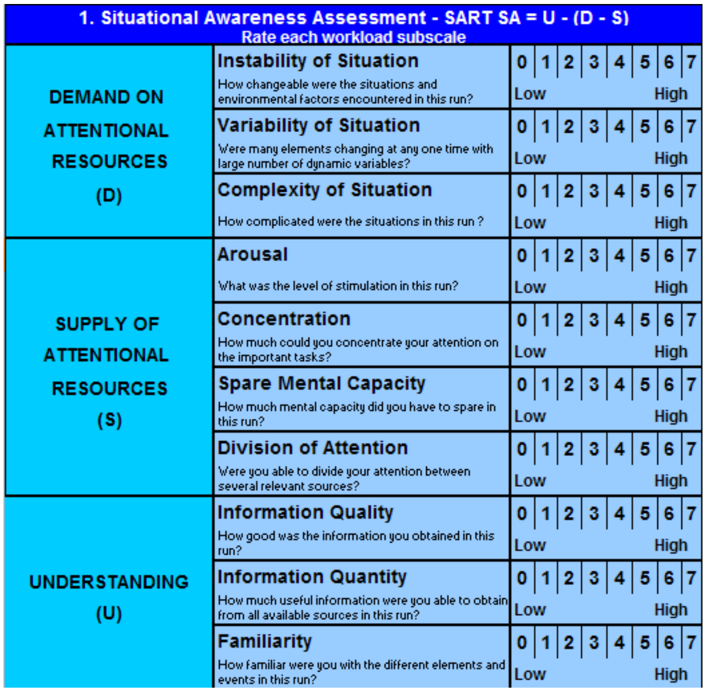
**Figure 16. Bedford Workload Scale**

The SART provides an assessment of the situation awareness based on a pilot's subjective rating. The SART incorporates the three dominant components of situation awareness: demand on pilot resources, supply of resources, and understanding of the situation. Pilots rated their perception of the impact of these components on situation awareness using bipolar scales from zero to seven as shown in Figure 17. SART scales were then transformed to provide an overall SART score for the pilot, using the formula shown in Equation 2. Scores from the application of the formula range from negative five

المنارة للاستشارات

www.manaraa.com

for extremely low situation awareness to 13 for extremely high situation awareness. The SART has been found to correlate with pilot task load and subjective measures of workload (Ellis & Schnell, 2009).



**1. Situational Awareness Assessment - SART SA = U - (D - S)**
**Rate each workload subscale**

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **DEMAND ON ATTENTIONAL RESOURCES (D)** | **Instability of Situation**<br>How changeable were the situations and environmental factors encountered in this run? | 0<br>Low | 1 | 2 | 3 | 4 | 5 | 6<br>High | 7 |
| | **Variability of Situation**<br>Were many elements changing at any one time with large number of dynamic variables? | 0<br>Low | 1 | 2 | 3 | 4 | 5 | 6<br>High | 7 |
| | **Complexity of Situation**<br>How complicated were the situations in this run ? | 0<br>Low | 1 | 2 | 3 | 4 | 5 | 6<br>High | 7 |
| **SUPPLY OF ATTENTIONAL RESOURCES (S)** | **Arousal**<br>What was the level of stimulation in this run? | 0<br>Low | 1 | 2 | 3 | 4 | 5 | 6<br>High | 7 |
| | **Concentration**<br>How much could you concentrate your attention on the important tasks? | 0<br>Low | 1 | 2 | 3 | 4 | 5 | 6<br>High | 7 |
| | **Spare Mental Capacity**<br>How much mental capacity did you have to spare in this run? | 0<br>Low | 1 | 2 | 3 | 4 | 5 | 6<br>High | 7 |
| | **Division of Attention**<br>Were you able to divide your attention between several relevant sources? | 0<br>Low | 1 | 2 | 3 | 4 | 5 | 6<br>High | 7 |
| **UNDERSTANDING (U)** | **Information Quality**<br>How good was the information you obtained in this run? | 0<br>Low | 1 | 2 | 3 | 4 | 5 | 6<br>High | 7 |
| | **Information Quantity**<br>How much useful information were you able to obtain from all available sources in this run? | 0<br>Low | 1 | 2 | 3 | 4 | 5 | 6<br>High | 7 |
| | **Familiarity**<br>How familiar were you with the different elements and events in this run? | 0<br>Low | 1 | 2 | 3 | 4 | 5 | 6<br>High | 7 |

**Figure 17. SART Assessment Card (Ellis & Schnell, 2009)**

**Equation 2. SART Evaluation SA Calculation**

$$SA = Understanding - (Demand - Supply)$$

### 3.3.8 Indices of Crew Coordination

Additional probing questions were administered in regard to crew interaction and coordination using the post-scenario questionnaire . Pilots were asked to assess their own workload, as well as estimate their crewmember’s workload during the scenario using the Bedford Workload Scale. Evaluating the perception of crewmember workload for each pilot enables the assessment of crew coordination using the CRM construct of shared

www.manaraa.com

awareness and task load balancing. Contrasting the differences in perceived versus self-reported workload values results in a pilot workload-understanding value for each pilot. Pilot workload-understanding was used to determine crew coordination as shown below in Figure 18.

Several variables are presented below in Figure 18. PFr(PM) represents the PF's rating of the PM's workload. PMr represents the PM's self reported workload. PFu represents the PF's understanding of the PM's workload. PMr(PF) represents the PM's rating of the PF's workload. PFr represents the PF's self reported workload. PMu represents the PM's understanding of the PF's workload. CCwl represents crew-coordinated workload, herein referred to as crew coordination.

**Pilot Workload Understanding**
PF Workload Understanding
PFr(PM) – PMr = PFu (-9-9)
PM Workload Understanding
PMr(PF) – PFr = PMu (-9-9)

No difference = Good Coordination

**Crew Coordination Index– *Crew Workload Understanding***
|PFu| + |PMu| = CCwl (0-18)

**Figure 18. Subjective Indices of Crew Coordination**

The crew coordination index represents the level of crew coordination. The crew coordination index is calculated as the absolute difference in pilot workload-understanding as shown in the equation at the bottom of Figure 18. The crew coordination index does not take into account a negative or positive difference in pilot workload-understanding. A negative or positive difference in pilot workload-understanding indicates overemphasis or underemphasis of crewmember workload. The emphasis of pilot workload-understanding may potentially be used to evaluate the pre-

www.manaraa.com

established schema or paradigm that exists in the way a pilot perceives their crewmember’s workload.

### 3.3.9 Boston Logan International Terminal Operations Task

The DataComm experiment was developed to simulate real-world operations and procedures in the KBOS terminal airspace. Simulated traffic was generated using actual KBOS airspace and surface data collected during periods of high-density terminal operations. Arrivals were set up to begin at FL180 and continued through landing to conclude upon completion of taxi to the terminal gate.

Scenarios were designed to vary arrival procedure tasks so pilots would not become familiar with any particular test scenario, yet remain operationally consistent to maintain a valid basis from which to assess each test condition. The experiment simulated day visual meteorological conditions (VMC) operations under Instrument Flight Rules (IFR). Pilots were expected to hand-fly the aircraft (autopilot disconnected), with flight director guidance on vertical navigation/lateral navigation (VNAV/LNAV) mode. A scenario specific STAR was programed into the FMS prior to each run, and the Auto throttle was engaged in Speed Hold Mode.

## 3.4 Research Facilities

### 3.4.1 NASA B757-200 Integrated Flight Deck (IFD)

As previously stated, the DataComm experiment was performed in NASA’s Integrated Flight Deck (IFD). The IFD simulator is a high-fidelity, 757 flight deck interface with fully operational instrumentation, hydraulically tethered flight controls,

المنارة للاستشارات
www.manaraa.com

rudder pedals, and mechanized throttle, shown in Figure 19. The simulator is based on a Boeing 757-200 aircraft flight dynamics mathematical model. The flight deck also includes a fully functional MCP, CDU with FMS, and communications unit. Outside visuals are driven by five high-definition, collimated projection displays that yield a full panorama, which provides 200 degrees of horizontal field-of-view and 40 degrees vertical field-of-view. The projection system is driven by a high-end computer generation imaging system tethered to aircraft location and state information.

**Figure 19. Integrated Flight Deck (IFD)**

In support of the DataComm experiment, the following hardware and software additions were incorporated to the IFD baseline configuration:

- Moving Map Displays (MMDs), presentable on the NDs at both crew stations, with the capability to display ownship route.

المنارة للاستشارات
www.manaraa.com

- Electronic Display Control Panels (EDCP) at both crew stations, to control scale and display mode for the NDs. Display mode selection allowed crews to see an airport depiction, with expected taxi route, while airborne during the simulated approach.
- The capability to trigger the playing of researcher-provided audio wave files based on simulated aircraft position, range to traffic, and/or specified flight deck control actuation (such as microphone transmit release).
- Additional selectable pages on both FMS CDUs, to support a hierarchical DataComm uplink and downlink capability, as well as the capability to selectively load expected or cleared routes into the MMDs.
- The capability to simulate (visually OTW) push-back from the terminal gate.

A Rockwell Collins EP-1000 KBOS database was used for OTW projection of the airport surface, taxiways, runways, buildings, obstructions, signs, and airport terrain and cultural features. Additionally, the DataComm simulation used the appropriate database to provide accurate location and frequency of navigation aids; in particular, the ILS was used for RWY 27 and RWY 33L. All frequencies aligned with published charts, and pre-recorded Automatic Terminal Information Service (ATIS) messages were used based on environmental conditions and airport status for the particular scenario. The IFD employed a navigation and communications simulation that permitted realistic voice communication, as well as accurate navigation and flight crew position awareness during standard arrivals, appropriate to each scenario. The simulator is capable of full motion testing for increased simulation fidelity; however, the DataComm experiment was conducted while the IFD was fix-based, with no motion element made available.

### 3.4.2 Dual Crew Smart-Eye Oculometer System

Eye-tracking data was collected for both crewmembers using a state-of-the-art eye-tracking system, developed by Smart-Eye Inc. Ten cameras were integrated into the IFD, using five cameras per pilot on two separate eye-tracking systems (one eye-tracking system per side) to capture the gaze vectors of both pilots simultaneously. To synchronize

المنارة للاستشارات

www.manaraa.com

the systems, Smart-Eye Inc. created a modified eye-tracking system network, tethering two systems together using a primary-secondary relationship. Each system is time stamped synchronously with GPS time so eye-gaze vector data from one pilot can be compared to that of another pilot, which is critical to interpret the data in post-analysis.

In order to collect robust eye-tracking data across the flight deck under normal flight deck operations, the system had to be capable of covering +/-45 degrees of center and +10 degrees from horizon and to the base of the CDU for each pilot. The coverage requirement had to be met while still maintaining a high level of simulator fidelity by making the cameras as inconspicuous as possible on the flight deck. Due to finite free space on the flight deck, camera placement options were limited.

A mockup of the IFD was created using 80/20™ aluminum to test which available locations for camera installation provided the greatest coverage capability on the flight deck. Tests resulted in five locations per side being chosen (mirrored locations between left and right seat) that yielded sufficient coverage to collect meaningful data for the experiment while remaining minimally obtrusive in the flight deck. Spatial accuracy and precision of the system was tested to be no greater than two degrees gaze angle for any calibration point on the display panels, example shown in Figure 23.

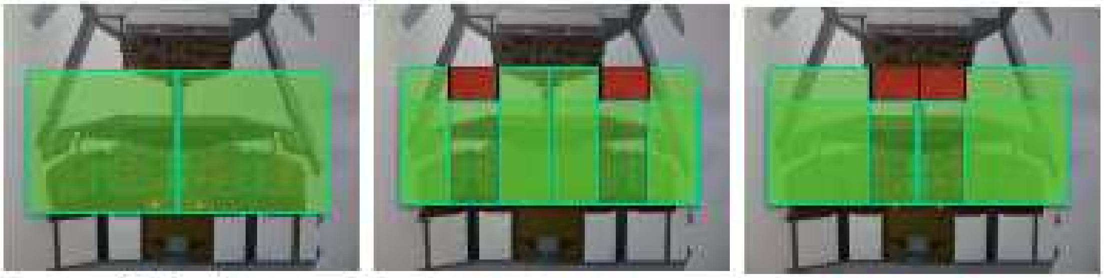
Head tracking and gaze quality across the instrument panel and OTW field of view varied between 50 and 100 percent, shown below in Figure 21 and Figure 22. Green represents values greater than 75 percent gaze quality as defined by the system gaze quality value; yellow represents 50-75 percent; and red represents <50 percent. Gaze quality values greater than 50 percent are considered satisfactory. The quality differences were due to participants wearing glasses - results shown below in Figure 22. The first two participants shown on the left were not wearing glasses, and participant three, shown on the right, was wearing glasses. The pilots in the DataComm experiment were mixed in their usage of glasses, with half of the participant population requiring reading glasses

المنارة للاستشارات

www.manaraa.com

when reviewing paper charts on the flight deck. Pilots were instructed to remove reading glasses whenever possible.

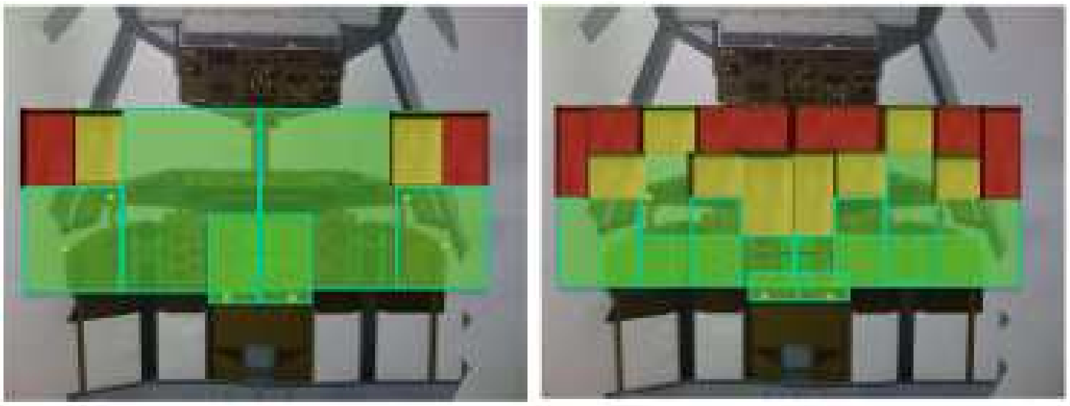
Figure 20. IFD SE Camera and Illuminator Locations



Subject 1 Subject 2 Subject 3, Glasses

Figure 21. Head Direction Quality Coverage

المنارة للاستشارات
www.manaraa.com


Subjects 1,2

Subject 3, Glasses

Figure 22. Field of View Gaze Quality

Figure 23. Spatial Accuracy

www.manaraa.com

# CHAPTER 4.

# NORMATIVE PILOT WORKLOAD MODEL DEVLOPMENT

A normative workload model was developed to provide high-resolution, dynamic PF and PM tasks and baseline workload information to be used as a comparator to pilot eye-scan behavior data. Additionally, the normative workload model was utilized to determine appropriate flight segments to analyze the eye-scan behavior for the PF and PM. The normative workload model was constructed by layering several workload factors, including tasks, distribution of attention by AOI, and subjective workload information for each minute of each scenario. The model information was acquired from expert NASA research pilots executing the DataComm scenarios and recording the expected behaviors when executing tasks.

Interruptive tasks, such as DataComm messages, increase task load and resultant workload. DataComm messages require experiment specific, flight deck interaction, to complete DataComm scenarios, described in Section 3.3.2. Scenario scripts utilized to trigger all DataComm ATC message and traffic movement were helpful to develop the normative workload model. Scenario scripts were written to be specific to each arrival scenario; therefore, the normative workload model is specific to each arrival scenario. Furthermore, interruptive tasks are specific to the type of response action required, either informative messages requiring a “Roger” or directive messages requiring a “Wilco”. The duration of each interruptive task type is taken from the analysis of the DataComm message response time, discussed in Section 4.1.

## 4.1 Influence of DataComm Message Response Time

Results of the DataComm experiment show the majority of analyzed response times were well under a minute (Mean = 20.7 seconds, SD = 17.6 seconds across all conditions) (Norman et al., 2013). There were few occasions when crews reviewed a

www.manaraa.com

message and agreed to its content but did not respond to the message within two minutes (5 of 369 (~1 percent) directive DataComm messages, and 27 of 660 (~4 percent) informative DataComm messages). Video review, researcher experience, and verbal debrief with participant pilots suggests that long response or no response events were due to the crew forgetting to respond to the message (Norman et al., 2010).

The response time data analyzed by NASA to determine acceptability of DataComm was utilized in this research for use in the development of the normative model to aid in definition of task loading and associated workload. Results from the response time analysis was important to determine the impact to task load caused by DataComm on the flight deck and the resultant task load impact to crew workload and coordination.  DataComm was used as an interruptive task, not included in any airline SOP, to disrupt crew coordination.

## 4.2 DataComm Post-Scenario Questionnaire Results

Pilots used the Bedford Workload Scale to rate the workload associated with in-flight and surface operations across communication and display modalities. The results from the DataComm experiment were used to determine the effect on each pilot's workload to develop the normative workload model. Results from the post-scenario questionnaire indicate a perception of relatively low workload for all conditions, presented in Figure 24 (Norman et al., 2010). Along the x-axis, a rating of one indicates "workload insignificant," five is "reduced spare capacity," and 10 is "task abandoned."

المنارة للاستشارات

www.manaraa.com

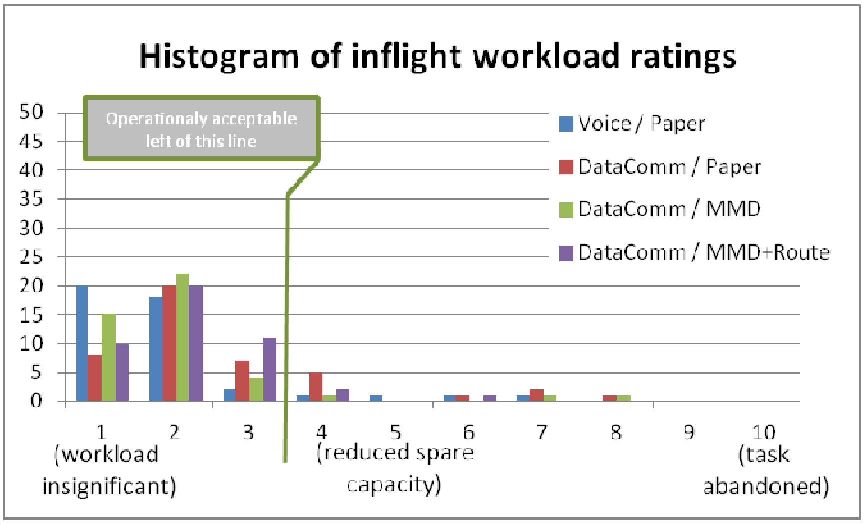


**Figure 24. Histogram of Workload Ratings for In-Flight Operations**

PF ratings of workload were significantly higher than the PM ratings of workload ($\chi^2_1$= 7.794, $p$=0. 0.005). A binomial test, with a cut point of three (“Enough spare capacity for all desirable additional tasks”) and test proportion of 75 percent, showed that most responses indicated significantly low workload for both the PF and PM (Norman et al., 2010). Figure 25 shows the mean responses for the PF and PM workload ratings for flight portions of the arrival from the DataComm study (Norman et al., 2010). The PF rated workload significantly different across the display conditions ($\chi^2_3$= 8.145, $p$=0.038); however from an operational perspective, the difference was not significant. Display conditions did not differentially affect PM ratings ($\chi^2_3$=5.749, $p$=0.125) (Norman et al., 2010).

www.manaraa.com

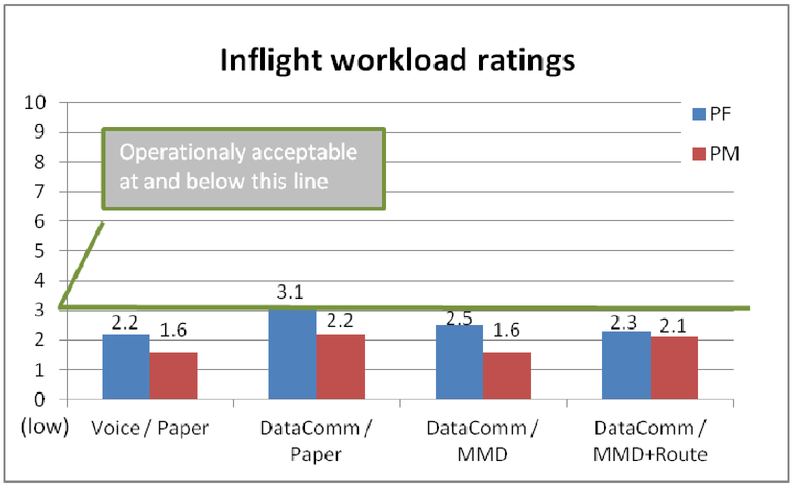


**Figure 25. In-flight workload rating by position and by condition**

The PF experienced higher workload during arrival scenarios than did the PM. The PF and PM experienced higher workload with the DataComm conditions than the Voice/Paper condition. It was posited by NASA researchers that the voice modality may have an advantage over the DataComm condition due to familiarity of using voice communication. The CDU interface for receiving DataComm messages appears to impose additional workload to the pilots (Norman et al., 2010). The additional workload due to DataComm to both the PF and PM suggests there was an effect on crew coordination.

### 4.3 Normative Workload Model for DataComm

An in-depth task analysis was completed for each pilot utilizing the results of the post-run workload analysis from the DataComm experiment. Review of the DataComm experiment results suggest the DataComm communication modality introduced tasks that were not common to current day operations. Therefore, the DataComm communication modality was an experiment variable that potentially disrupted crew coordination from the baseline voice communication condition. Each scenario was evaluated based upon task analysis, associated workload, and the DataComm interruptive tasking. The task analysis was performed using experienced NASA pilots operating as a crew to identify

www.manaraa.com

all required task responsibilities minute by minute for all DataComm scenario conditions. The task analysis also included subjective assessment of commonly visited AOIs over the duration of each scenario. The task analysis was used to develop the normative workload model based upon task demand and the number of tasks at a given time.

The normative workload model was developed using a baseline workload value based on tasking required by SOPs for each phase of the arrival and landing. Expected increases to the baseline workload were determined based upon expected increases in tasking and difficulty of the procedure. Figure 28 depicts the workload value for the pilot increasing with increased task load as the aircraft continues its approach to land. Nominal workload increases the closer an aircraft gets to touchdown as guidance becomes more difficult to follow the closer the aircraft gets to the runway. Workload remains high during the landing, rollout, and taxi phase as pilots are tasked with braking the aircraft, finding the expected taxiway exit, and monitoring for other aircraft (Norman, 2010).

Interruptive tasks are additive to workload, as they impose an increase in tasking. Interruptive tasks in the DataComm experiment are presented in the form of the DataComm messages on the CDU. Communication tasks are less important on the hierarchy of tasking for each pilot and are delegated to the PM. The impact of data link messages depends on the message content requiring flight crews to either acknowledge (Roger) or comply (Wilco). Messages requiring a “Roger” response are additive to the baseline workload by half on the Bedford Workload Scale the model is built upon. Messages requiring a “Wilco” response are additive to the baseline workload by one on the Bedford Workload Scale. Multiple messages are additive in regard to workload increase, and the duration for each interruptive task is based upon the average response times from the DataComm experiment, varying from 10.8 to 33.1 seconds. A list of the various data link messages from the DataComm experiment are listed below in Table 4 (Norman et al., 2010). Normative workload models for the Norwich arrival scenarios are

المنارة للاستشارات
www.manaraa.com

shown below in Figure 26 and Figure 27. Normative workload models for the Scupp arrivals are shown below in Figure 30 and Figure 31.

**Table 4. DataComm Messages and Response Times**

- Engine Start/Pushback – 19.5s
- Expected Taxi Out – 33.1s
- Amended Taxi – 10.8s
- Expected Taxi – 20.1s
- Taxi Out – 14.1s
- Taxi In – 25.1s
- Frequency Change – 16.0s
- ATIS (AIR) – 15.0s
- ATIS (Surface) – 23s

المنارة للاستشارات
www.manaraa.com

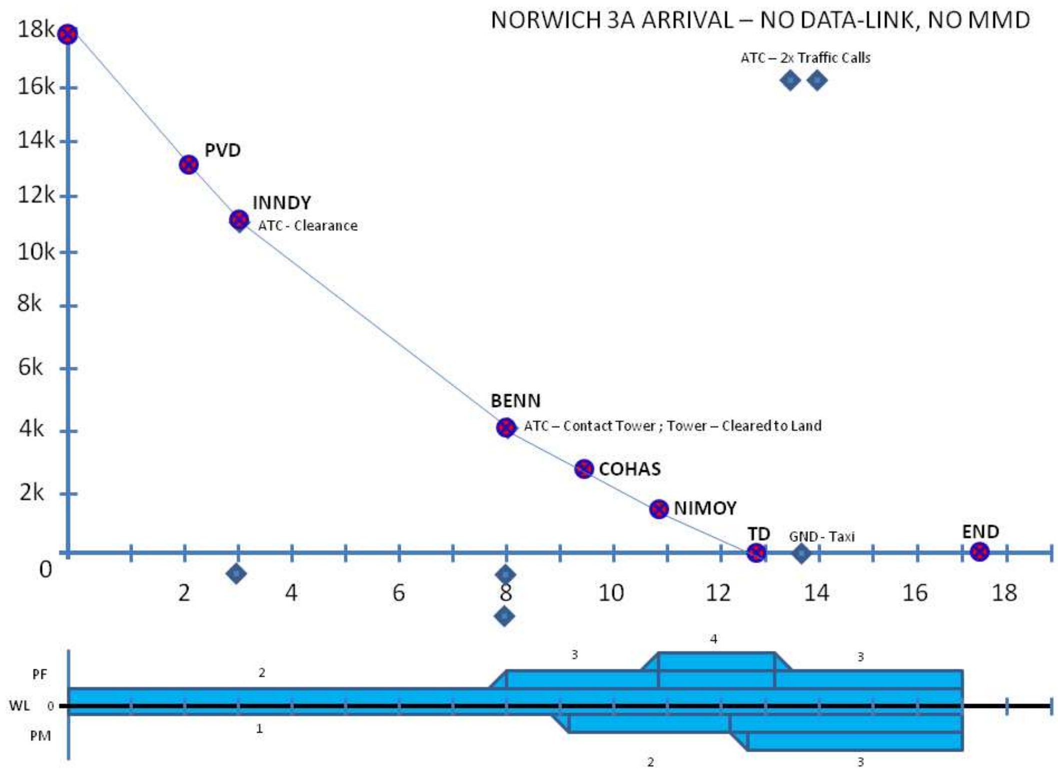


Figure 26. Norwich 3A Arrival - No DATA-LINK, No MMD

www.manaraa.com

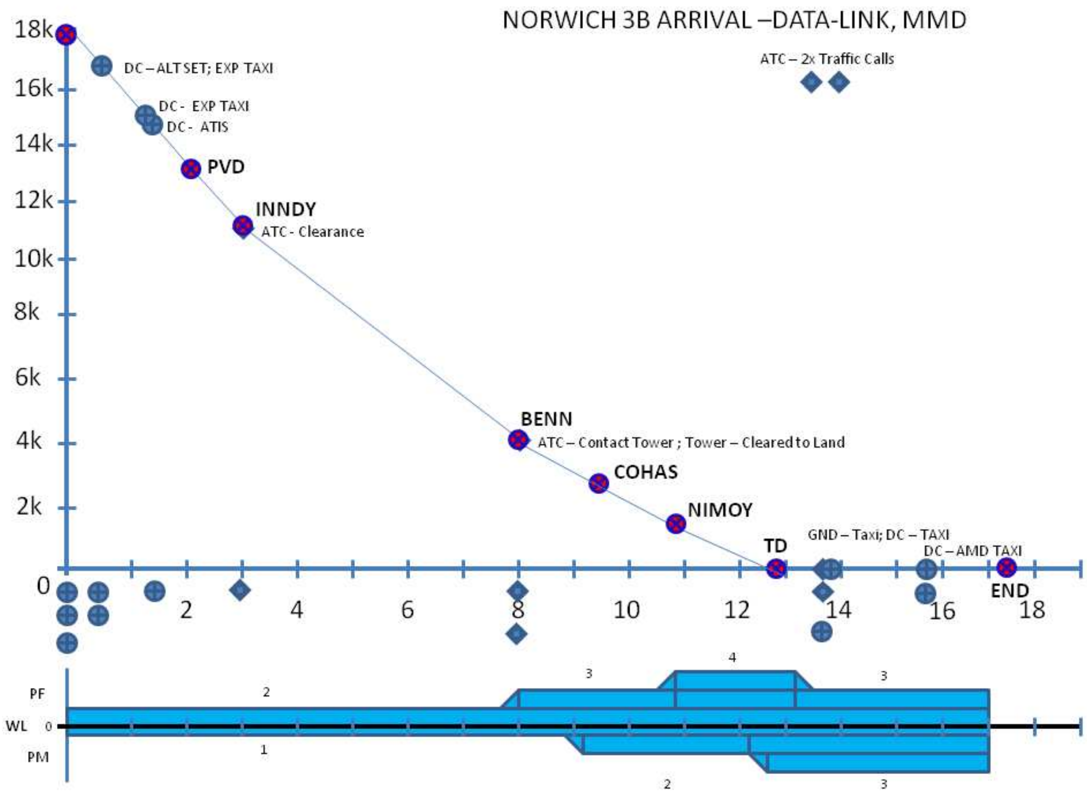


Figure 27. Norwich 3B Arrival - Data - Link, MMD

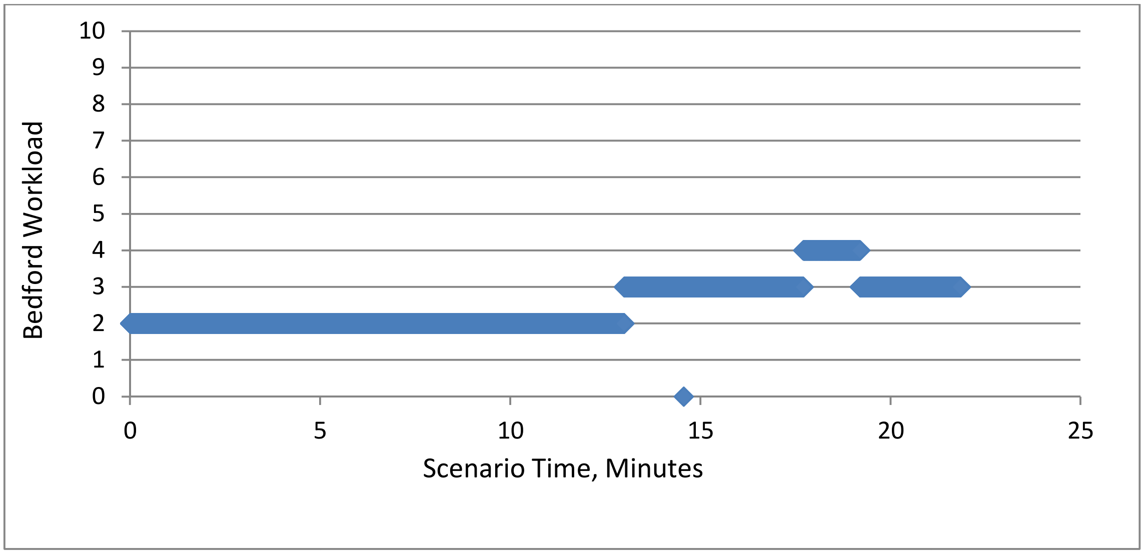


Figure 28. Pilot Flying Normative Workload Without Additional Tasks

www.manaraa.com

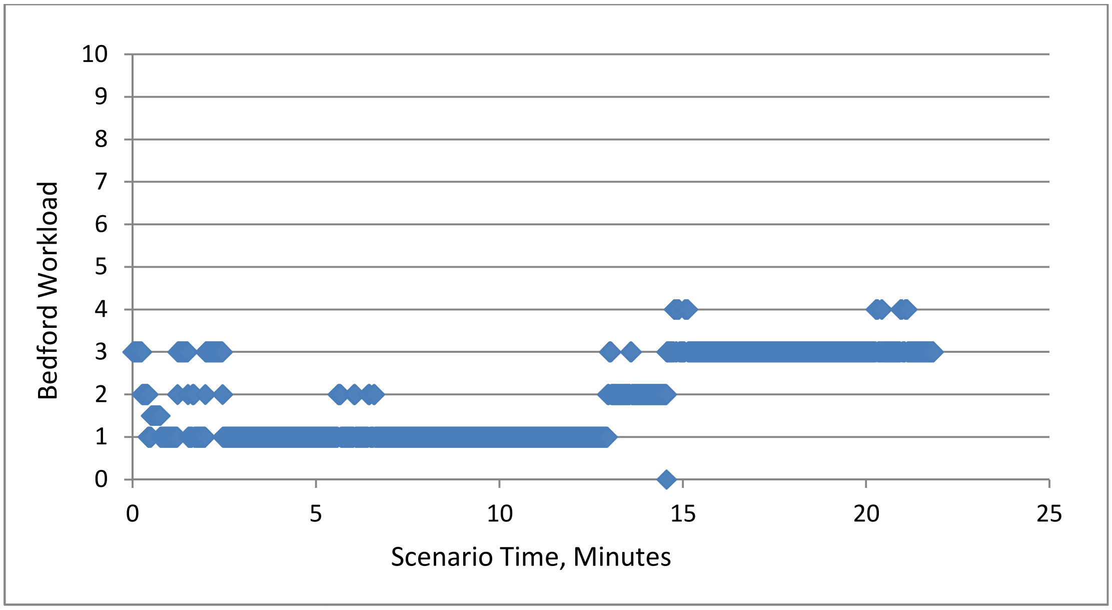


Figure 29. Pilot Monitoring Normative Workload With Additional Tasks

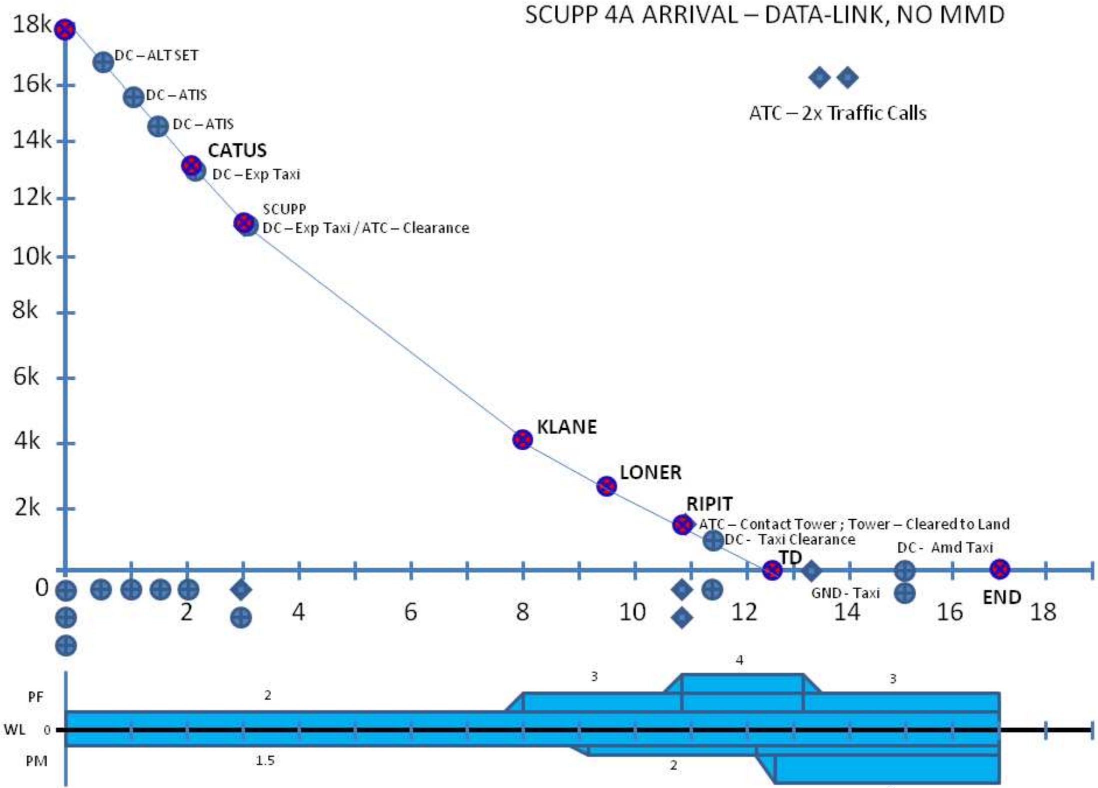


Figure 30. SCUPP 4A Arrival - Data-Link, No MMD

www.manaraa.com

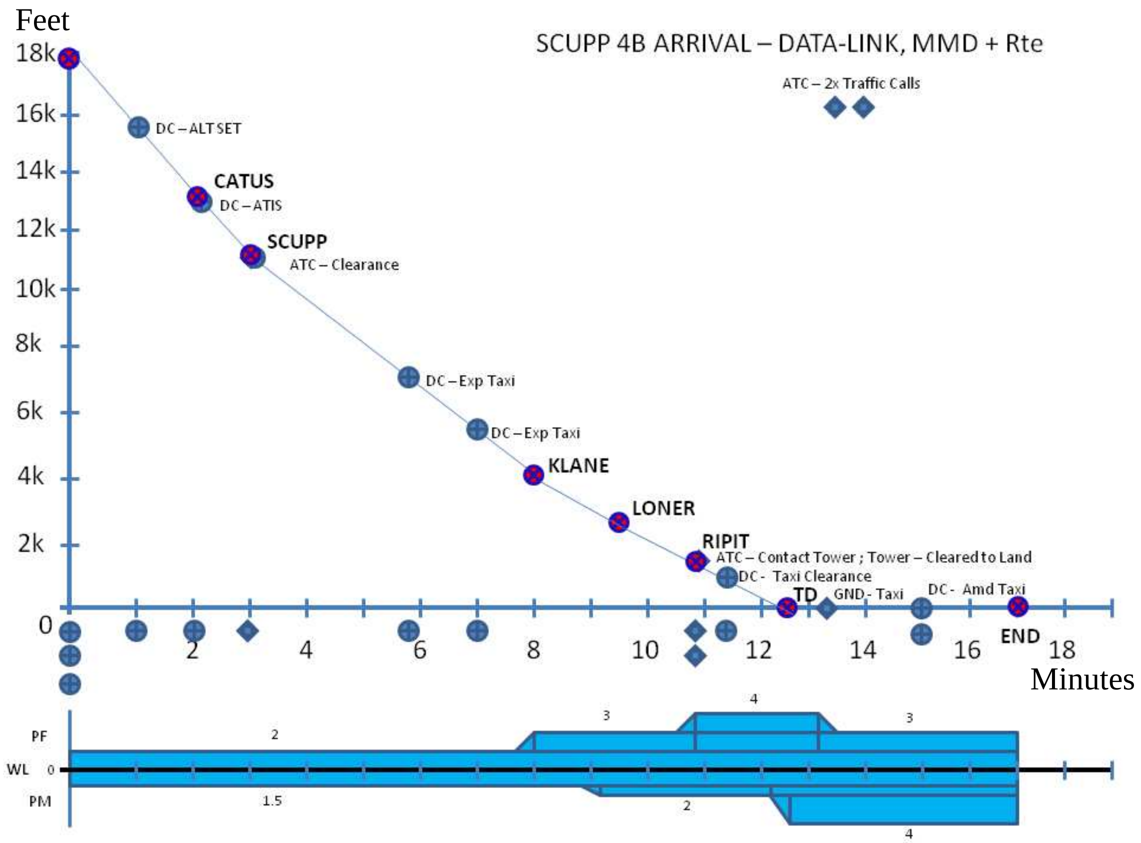


Figure 31. SCUPP 4B Arrival - Data - Link, MMD + Route

### 4.4 In-Flight Altitude Bands

Arrivals scenarios were selected to develop normative behavior models due to the well-defined SOPs that address the tasks required of the arrival and landing. Data was segmented into three different altitude bands to capture task variations over the course of the entire data run. Results from the post-scenario and post experiment workload questionnaires did not indicate operationally significant differences across communication or display conditions (Norman et al., 2013). Analyses were conducted by collapsing all communication and display modality conditions into one composite group, which were evaluated by the variation of crew coordination index scores.

المنارة للاستشارات

The altitude band segments are broken into three parts: high, middle, and low. The high segment is defined as the beginning of the scenario (starting at ~18,000ft) to 10,000ft. above ground level (AGL). The middle segment is defined as less than 10,000 ft. AGL to the final approach fix altitude of 1,700 ft. AGL. The low altitude band is defined as less than 1,700 ft. AGL to touchdown at 0 ft. AGL. The altitude bands are shown below in Figure 32.

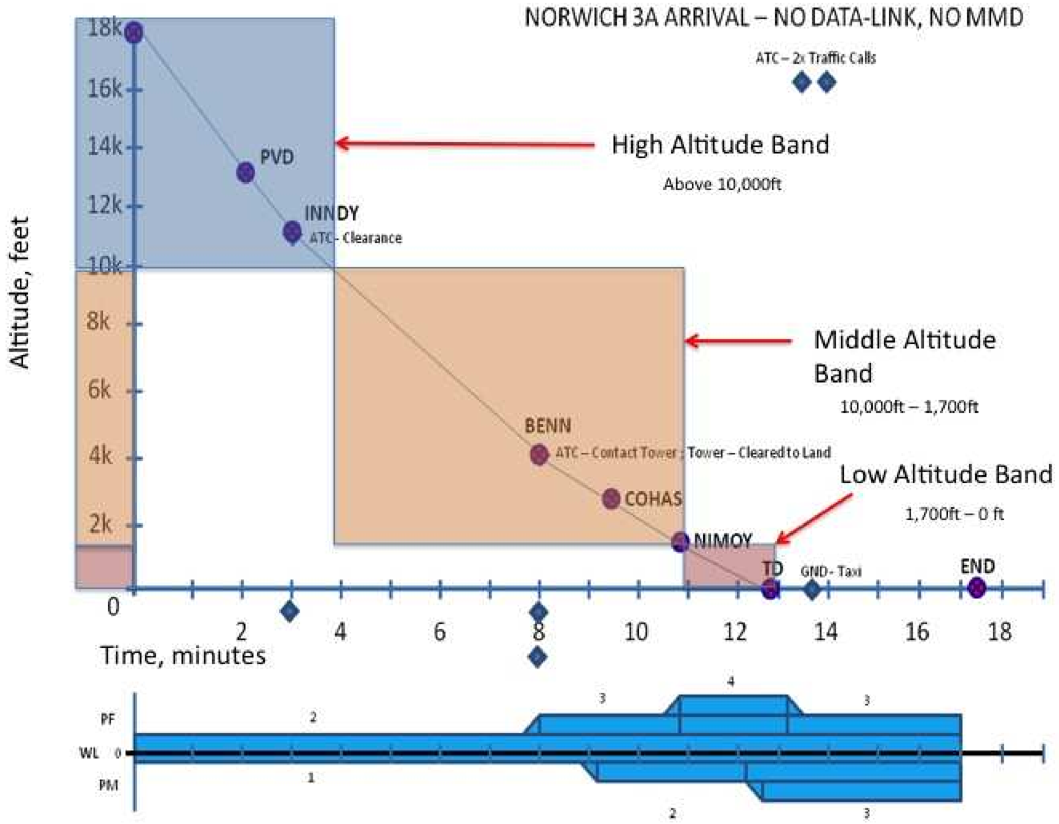


**Figure 32. Altitude band Segments**

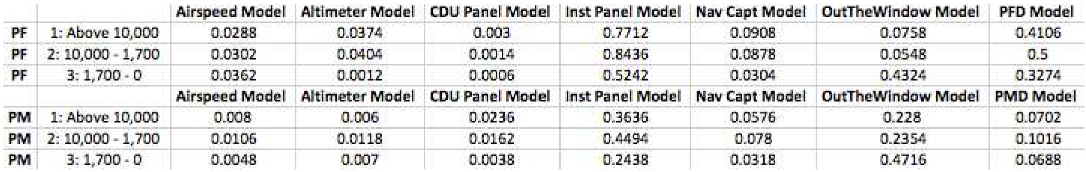
### 4.5 Normative Eye-Scan Behavior Model

The normative eye-scan behavior model was developed using eye-scan data collected in the DataComm experiment. The normative eye-scan behavior model was developed by averaging the PDT across all participants when the crew coordination index

www.manaraa.com

was less than two, indicating excellent coordination for the baseline condition (voice/paper). A normative eye-scan behavior model value was created for each AOI, altitude band, and for the PF and PM. AOI PDT values for the normative eye-scan behavior model are shown below in Table 5.

**Table 5. Altitude band Normative Visual Behavior Model Values (Percentage Dwell Time)**



| | | Airspeed Model | Altimeter Model | CDU Panel Model | Inst Panel Model | Nav Capt Model | OutTheWindow Model | PFD Model |
|---|---|---|---|---|---|---|---|---|
| PF | 1: Above 10,000 | 0.0288 | 0.0374 | 0.003 | 0.7712 | 0.0908 | 0.0758 | 0.4106 |
| PF | 2: 10,000 - 1,700 | 0.0302 | 0.0404 | 0.0014 | 0.8436 | 0.0878 | 0.0548 | 0.5 |
| PF | 3: 1,700 - 0 | 0.0362 | 0.0012 | 0.0006 | 0.5242 | 0.0304 | 0.4324 | 0.3274 |
| | | Airspeed Model | Altimeter Model | CDU Panel Model | Inst Panel Model | Nav Capt Model | OutTheWindow Model | PMD Model |
| PM | 1: Above 10,000 | 0.008 | 0.006 | 0.0236 | 0.3636 | 0.0576 | 0.228 | 0.0702 |
| PM | 2: 10,000 - 1,700 | 0.0106 | 0.0118 | 0.0162 | 0.4494 | 0.078 | 0.2354 | 0.1016 |
| PM | 3: 1,700 - 0 | 0.0048 | 0.007 | 0.0038 | 0.2438 | 0.0318 | 0.4716 | 0.0688 |

المنارة للاستشارات

www.manaraa.com

# CHAPTER 5.
# DATA ANALYSIS

## 5.1 Data Analysis Approach

The DataComm experiment data was used to evaluate the following:

- Pilot situation awareness and workload through objective measures of attention in terminal area operations.

- Crew coordination, shared situation awareness, and crew workload in terminal area operations.

Subjective workload response values from the simulation data runs were analyzed to evaluate the following:

- The PF and PM workload during each scenario.
- Each pilot’s perception of crewmember workload during each scenario.
- Scenario task analysis.
- Normative workload model development.

Eye-tracking data from the simulation experiment was analyzed to evaluate the following:

- PF and PM eye-scan behavior (PDT) by altitude band.
- Difference between normative model PDT and observed PDT for each pilot by altitude band.
- Difference between the PF and PM PDT for each common AOI by altitude band.

المنارة للاستشارات
www.manaraa.com

- Level of shared attention, defined as each pilot viewing the same or common type AOI (i.e. PF-PFD and PM-PFD) within five seconds. Shared attention is output as the percentage of time common type AOIs were viewed during each altitude band.

Level of crew coordination was correlated with eye-tracking measures of attention as described below:

- The correlation between the crew coordination index and the difference between observed and normative model PDT, which was evaluated for each AOI across each altitude band for both PF and PM by:
    - The difference between normative model PDT values and the observed PDT of the PF and PM.
        - Deviation of observed PDT from normative model PDT corresponds with non-normative behavior, which is evaluated as a quantitative measure of reduced situation awareness.
        - Deviation of observed PDT from normative model PDT indicates a deviation from expected behavior and is indicative of decreased crew coordination.
        - Difference between the normative eye-scan model PDT and the observed PDT is evaluated for correlation with crew coordination index.

- The correlation between crew coordination index and the difference between the PF and PM PDT, evaluated for each AOI across each altitude band by:
    - The difference between the PF and PM PDT as a measure of visual behavior strategy shifts between crewmembers.

www.manaraa.com

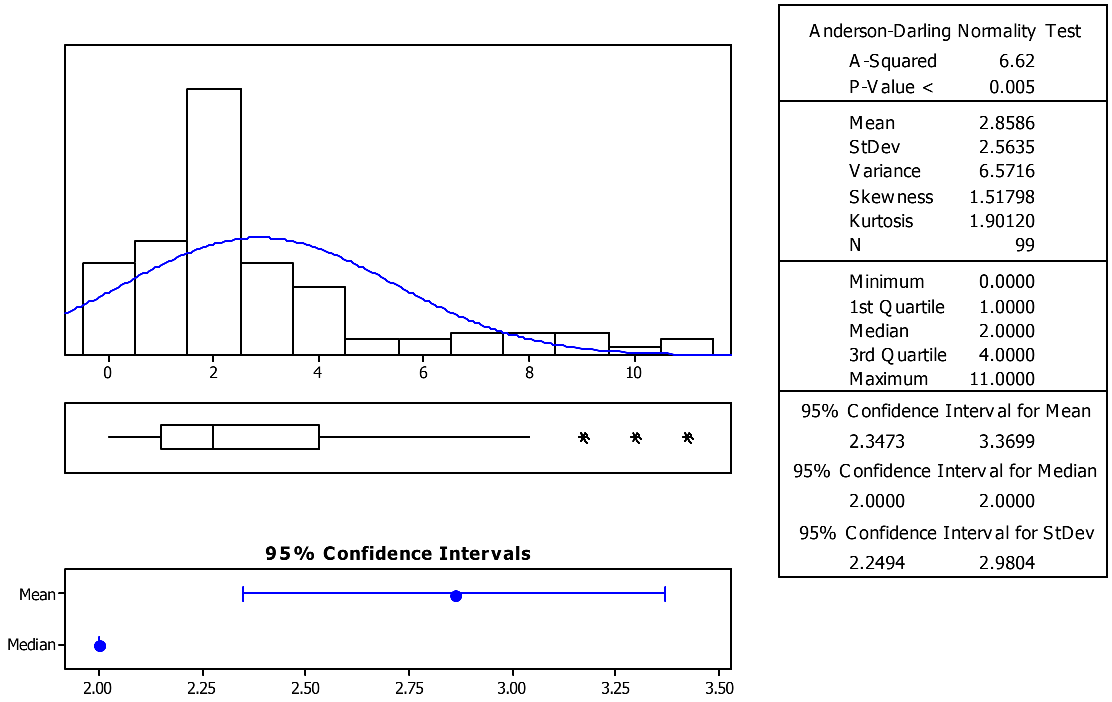
        - A zero difference in AOI PDT indicates common awareness between crewmembers (This may represent zero attention given to an AOI for both pilots).
        - A positive difference in PDT indicates increased attention given to the AOI for the PF.
        - A negative difference in PDT indicates increased attention given to the AOI for the PM.

- Correlation between crew coordination index and the measure of shared attention, evaluated for each AOI across each altitude band by:
    - The shared attention metric, which indicates increased task sharing and visual cross-check between crewmembers.
        - Values closer to one indicate increased shared awareness for a particular AOI.
        - Values closer to zero indicate less shared awareness for a particular AOI.

### 5.2 Crew Coordination Subjective Workload Analysis

Analysis of each pilots' understanding of workload was used to determine an index of crew coordination as described in Section 3.3.8. Results from the crew coordination subjective workload analysis are shown below, broken into four index ratings of crew coordination, Excellent Coordination to Poor Coordination. The DataComm experiment never yielded a scenario that proved beyond the capabilities of the crew. However, analysis of workload-understanding values shows multiple instances of poor understanding between crewmembers; referred to as poor coordination between pilots. The distribution of crew workload-understanding is shown below in Figure 33.

www.manaraa.com



| Anderson-Darling Normality Test | |
|---|---|
| A-Squared | 6.62 |
| P-Value < | 0.005 |
| Mean | 2.8586 |
| StDev | 2.5635 |
| Variance | 6.5716 |
| Skewness | 1.51798 |
| Kurtosis | 1.90120 |
| N | 99 |
| Minimum | 0.0000 |
| 1st Quartile | 1.0000 |
| Median | 2.0000 |
| 3rd Quartile | 4.0000 |
| Maximum | 11.0000 |
| 95% Confidence Interval for Mean | |
| 2.3473 | 3.3699 |
| 95% Confidence Interval for Median | |
| 2.0000 | 2.0000 |
| 95% Confidence Interval for StDev | |
| 2.2494 | 2.9804 |

**Figure 33. Crew Workload-Understanding Values Distribution**

The range of crew workload-understanding values and the corresponding index ratings is shown below in Table 6. The crew coordination index ranges were selected based upon the distribution of pilot workload-understanding values and the definition of the Bedford Workload Scale workload ratings, shown in Section 3.3.7.

**Table 6. Crew Coordination Index Range**

| Crew Coordination Index | Crew Workload-Understanding Range |
|---|---|
| Excellent Coordination (1) | 0 < 2 |
| Good Coordination (2) | 2 < 5 |
| Fair Coordination (3) | 5 < 8 |
| Poor Coordination (4) | > 8 |

Figure 34 shows the distribution of crew coordination index values for the DataComm experiment. As discussed previously, the crew coordination index is a measure of shared understanding of crew workload between pilots. Based on the constructs of SOPs and

www.manaraa.com

CRM task load balancing, the concept of shared understanding of crew workload is representative of good crew coordination. Subsequent analyses evaluate the metrics of eye-scan behavior for each pilot with respect to crew coordination index across the DataComm altitude bands.

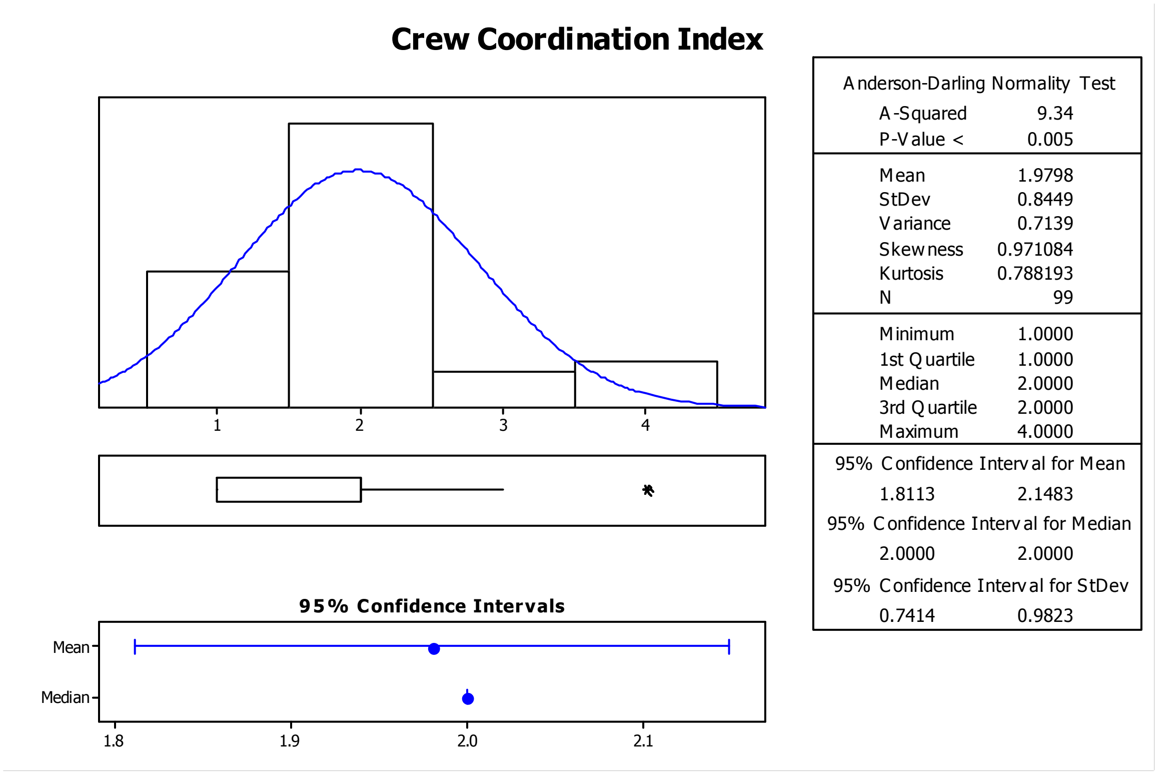


**Figure 34. Crew Coordination Index Distribution**

### 5.3 Altitude band Statistical Analyses

All statistical analyses were performed with a general linear model (GLM) ANOVA, modeling altitude band, crew coordination index, and an interaction term for both altitude band and crew coordination index. Tukey pairwise T-test comparisons were made to determine individual statistical differences between each element within the two components of the GLM model.

المنارة للاستشارات
www.manaraa.com

### 5.3.1 Pilot Flying and Pilot Monitoring Normative Visual Behavior Model Difference

Statistical analyses for the normative eye-scan behavior model difference data were performed across all AOIs for each pilot. This section evaluates the difference between the normative eye-scan behavior model PDT values (shown previously in Table 5) and the observed PDT values for each AOI across crew coordination index ratings (Excellent through Poor). Deviations from the normative model are shown as any positive or negative PDT difference from zero. Deviations from zero indicate pilot attention allocation was different from what was expected as defined by the normative eye-scan behavior model. AOIs evaluated include the PFD, ND, CDU, AI, ALT, IP, and OTW. Observations of AOIs with statistically significant differences are included in this analysis section and discussed with regard to flight operational impact. The comprehensive analysis of all AOI statistics for the PF and PM normative model differences is included in Appendix D.

#### *5.3.1.1 Instrument Panel AOI, Pilot Flying*

Analysis of the IP normative eye-scan behavior model difference for the PF, as shown in Figure 35, revealed there were statistically significant differences across altitude band ($F(2,282) = 4.46$, $p = 0.012$) and crew coordination index ($F(3,282)=7.38$, $p < 0.001$). Figure 36 shows the main effects plot for altitude band and crew coordination index. Pairwise comparisons between altitude bands show there was no statistically significant difference between the high altitude band and the middle or low altitude bands. The significant difference occurred between the middle and low altitude bands ($t=2.939$, $p=0.0092$).

Pairwise comparisons between crew coordination index ratings revealed that there were no statistically significant differences across crew coordination index ratings Excellent, Good, and Fair. However, the crew coordination index rating of Poor was statistically different from all other ratings: Excellent vs. Poor: (t=-4.389, *p*<0.001), Good vs. Poor: (t=-4.330, *p*<0.001), and Fair vs. Poor: (t=-3.711, *p*=0.001). The pairwise comparisons are graphically visible in the interaction plot shown in Figure 37.

Findings reveal that with reduced crew coordination, the PF visual behavior with regard to the IP deviated from the normative model. Additionally, there appeared to be a threshold after the crew coordination index of Fair, when the eye-scan behavior began to significantly deviate from the normative model. Findings suggest the departure from expected attentional behavior given to the instrument panel by the PF was a significant indicator of poor crew coordination.

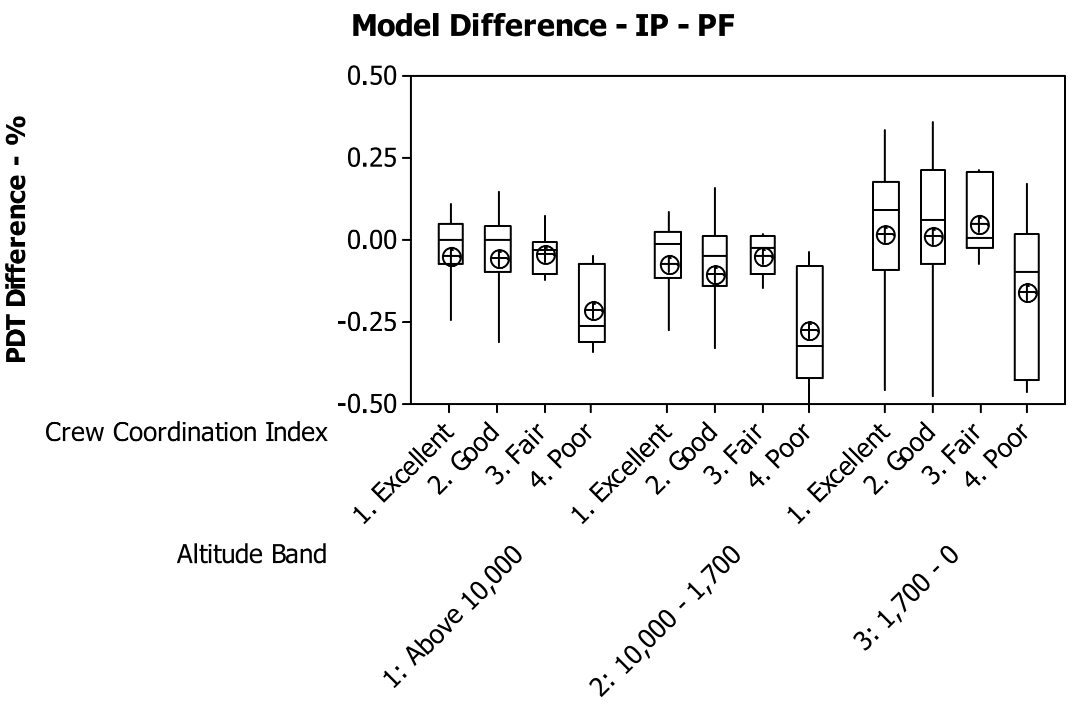


**Figure 35. Instrument Panel Model Difference - PF**

المنارة للاستشارات
www.manaraa.com

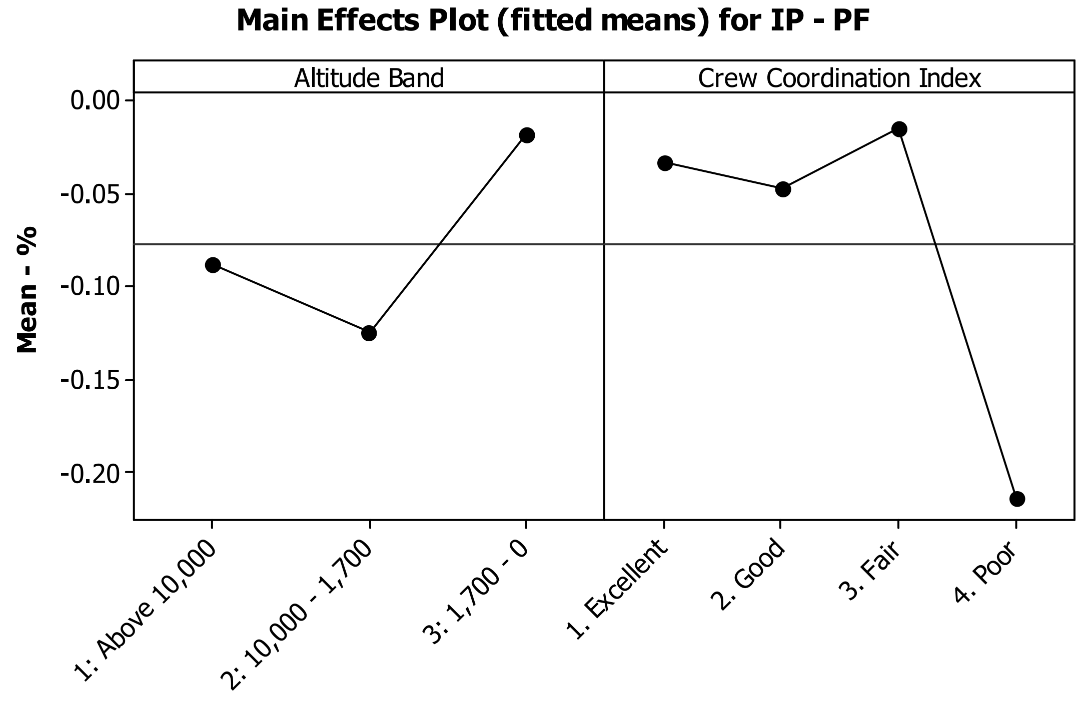


Figure 36. Instrument Panel Model Difference Main Effects Plot - PF

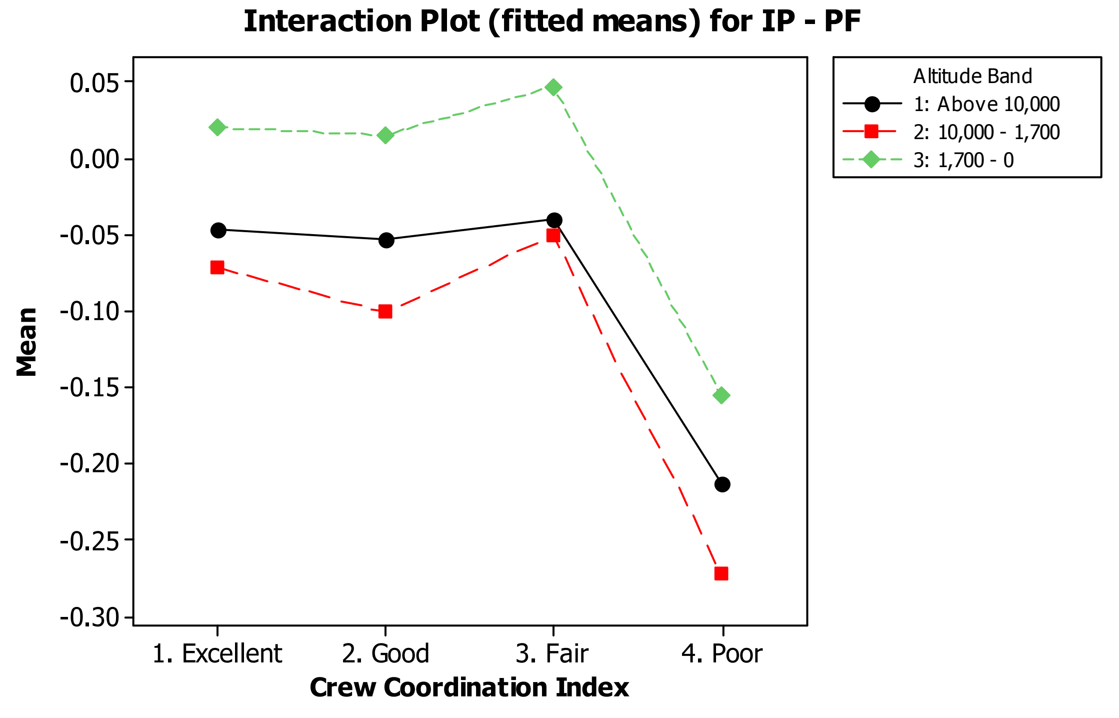


Figure 37. Instrument Panel Model Difference Interaction Plot - PF

www.manaraa.com

### *5.3.1.2 Out the Window AOI, Pilot Flying*

Analysis of the OTW normative eye-scan behavior model difference for the PF revealed there were statistically significant differences between altitude bands ($F(2,282)=61.47$, $p<0.000$). Figure 38 and Figure 39 show the deviation from the normative eye-scan behavior model across altitude bands and the crew coordination index ratings. Pairwise comparisons across the altitude bands show that there was no significant difference between the high and the middle altitude bands. The low altitude band was significantly different from both the high ($t=-9.443$, $p<0.000$) and middle ($t=-9.724$, $p<0.000$) altitude bands, with pairwise comparisons visible in Figure 40. There was no statistically significant difference across crew coordination index ratings.

Deviations from normative eye-scan behavior observed in the low altitude band did not indicate any statistical difference across crew coordination index ratings. There were no operationally significant differences observed in the eye-scan behaviors with regard to the OTW AOI beyond statistically significant difference from the normative model in the low altitude band. Findings suggest that the PF attention to the OTW AOI did not provide a significantly identifiable behavior that corresponds with reduced crew coordination. However, Figure 40 shows a visible trend suggesting a minor deviation from the normative eye-scan behavior model occurred with reduced crew coordination, a finding similar to that of the IP findings above.

www.manaraa.com

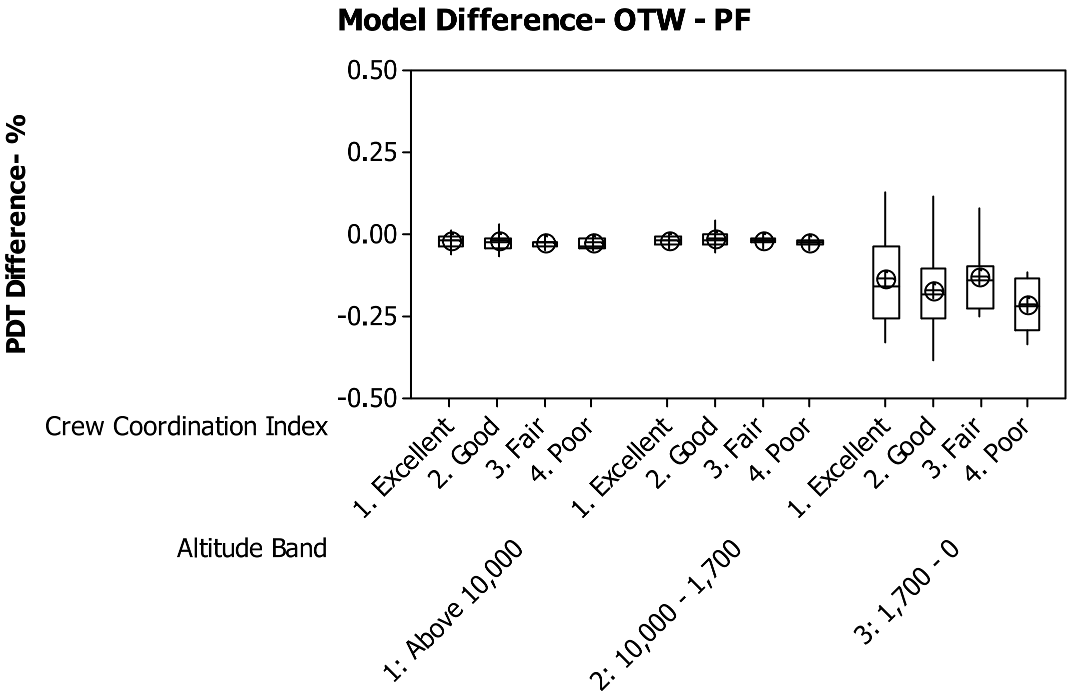


Figure 38. Out the Window Model Difference - PF

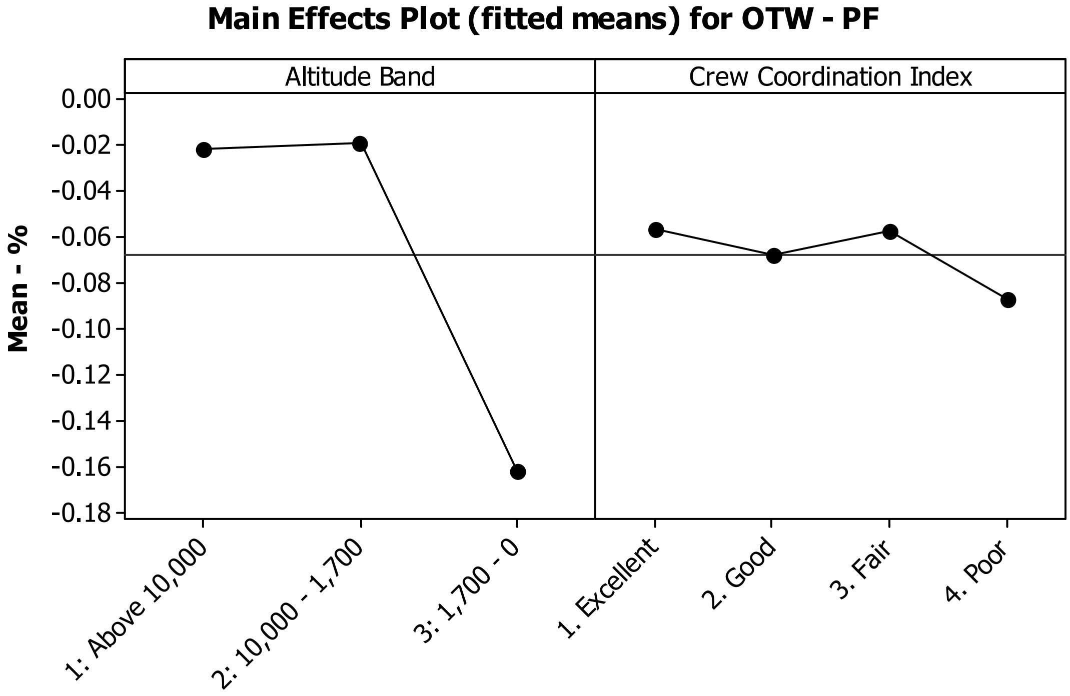


Figure 39. Out the Window Model Difference Main Effects Plot - PF

المنارة للاستشارات

www.manaraa.com

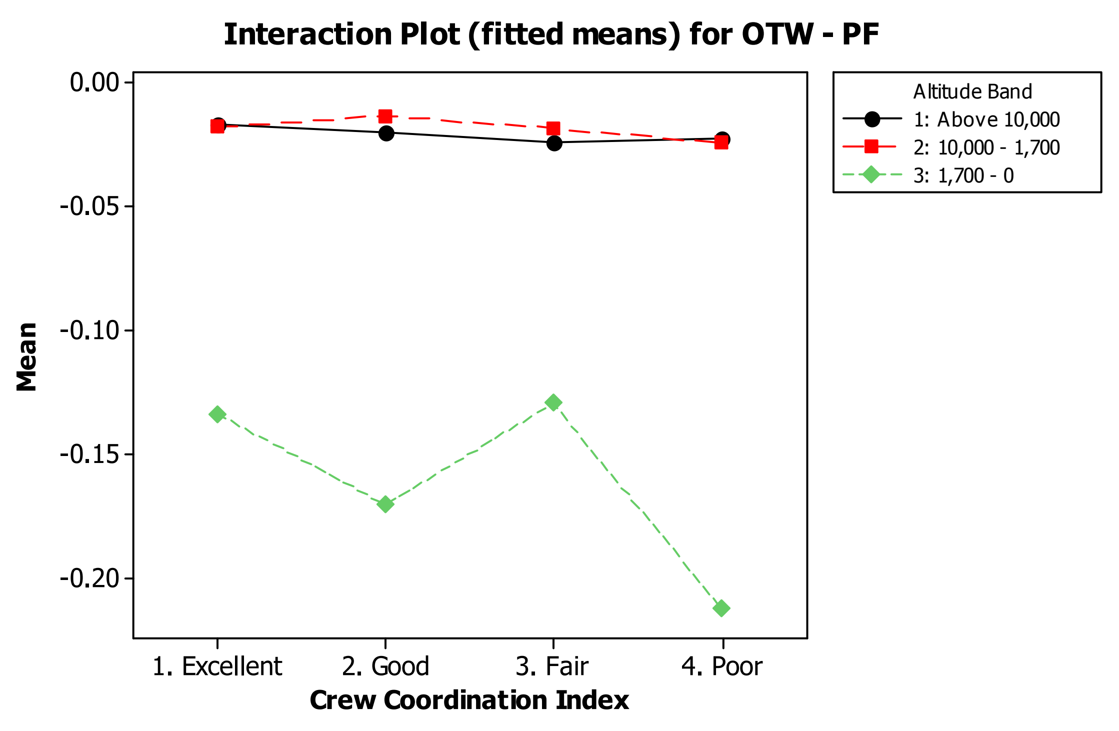


**Figure 40. Out the Window Model Difference Interaction Plot - PF**

*5.3.1.3 PFD AOI, Pilot Flying*

Analysis of the PFD normative eye-scan behavior model difference for the PF indicated statistically significant differences for both altitude band (*F*(2,282)=7.77, *p*=0.001) and crew coordination index (*F*(3,282)=15.60, *p*<0.000). Figure 41 and Figure 42 show the deviation of observed PDT from the normative model across the altitude bands and crew coordination index ratings. Pairwise comparisons across altitude bands show that the low altitude band was significantly different from the middle altitude band (t=3.930, *p*<0..001). Pairwise comparisons across the crew coordination index show a crew coordination index rating of Poor was significantly different than all other crew coordination index ratings; Excellent vs. Poor (t=-6.118, *p*<0.001), Good vs. Poor (t = -3.160, *p*<0.009), Fair vs. Poor (t=-2.930, *p*<0.018). Pairwise comparisons are shown graphically in Figure 43.

www.manaraa.com

Results show similar findings to the instrument panel AOI analysis, suggesting that with reduced crew coordination, there was a departure of visual attention from what was expected by the normative eye-scan behavior model. The PFD was an AOI defined inside the larger IP AOI, and therefore similar results were expected between IP and PFD AOIs. Findings are shown to be common across all altitude bands, suggesting that the PF eye-scan behavior differences that were indicative of reduced crew coordination were common across flight tasking.

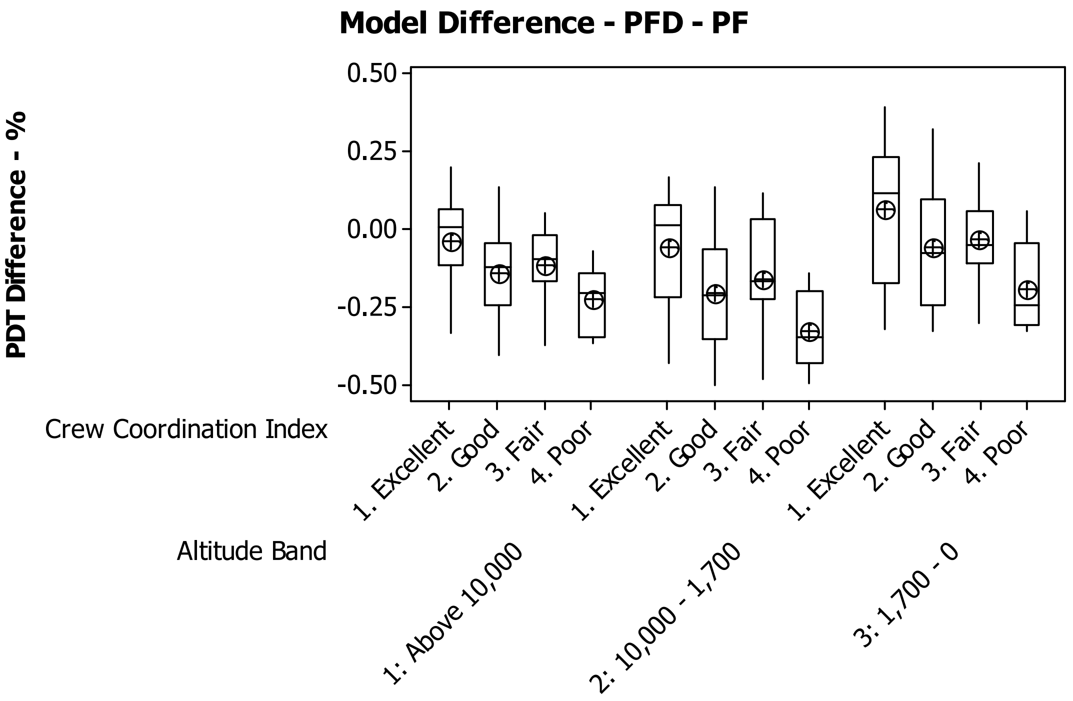


**Figure 41. PFD Model Difference - PF**

www.manaraa.com

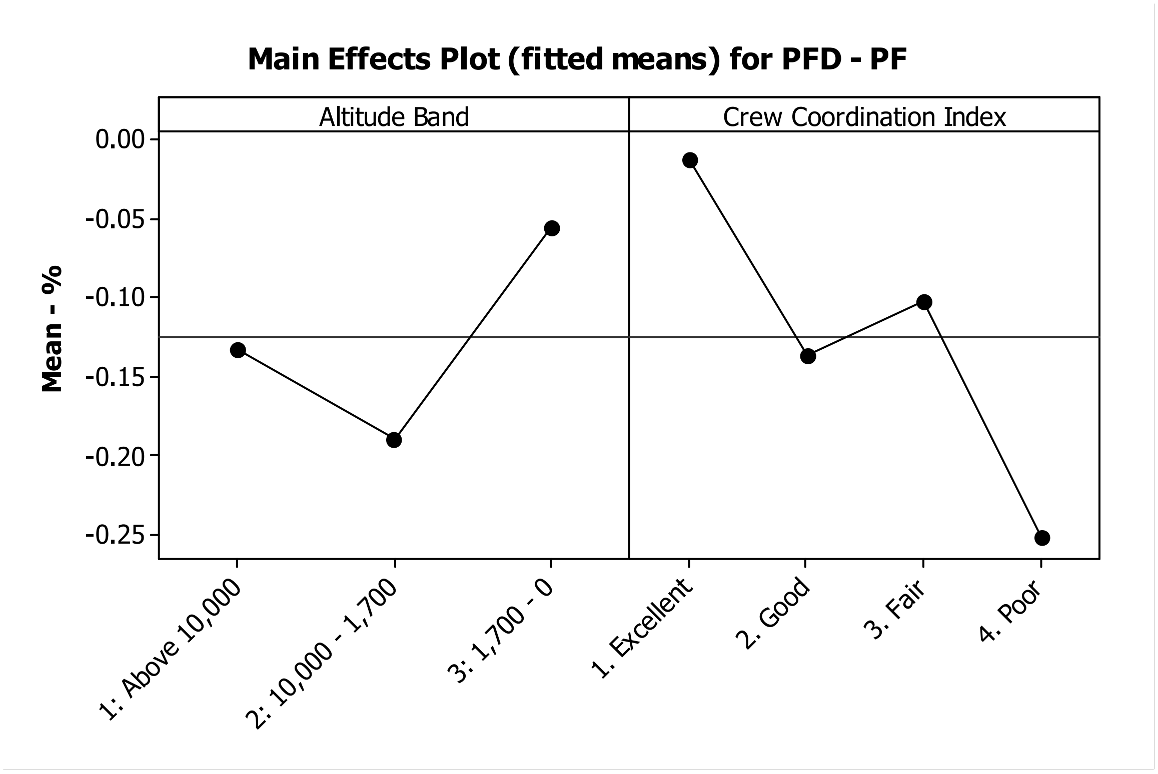


Figure 42. PFD Model Difference Main Effects Plot - PF

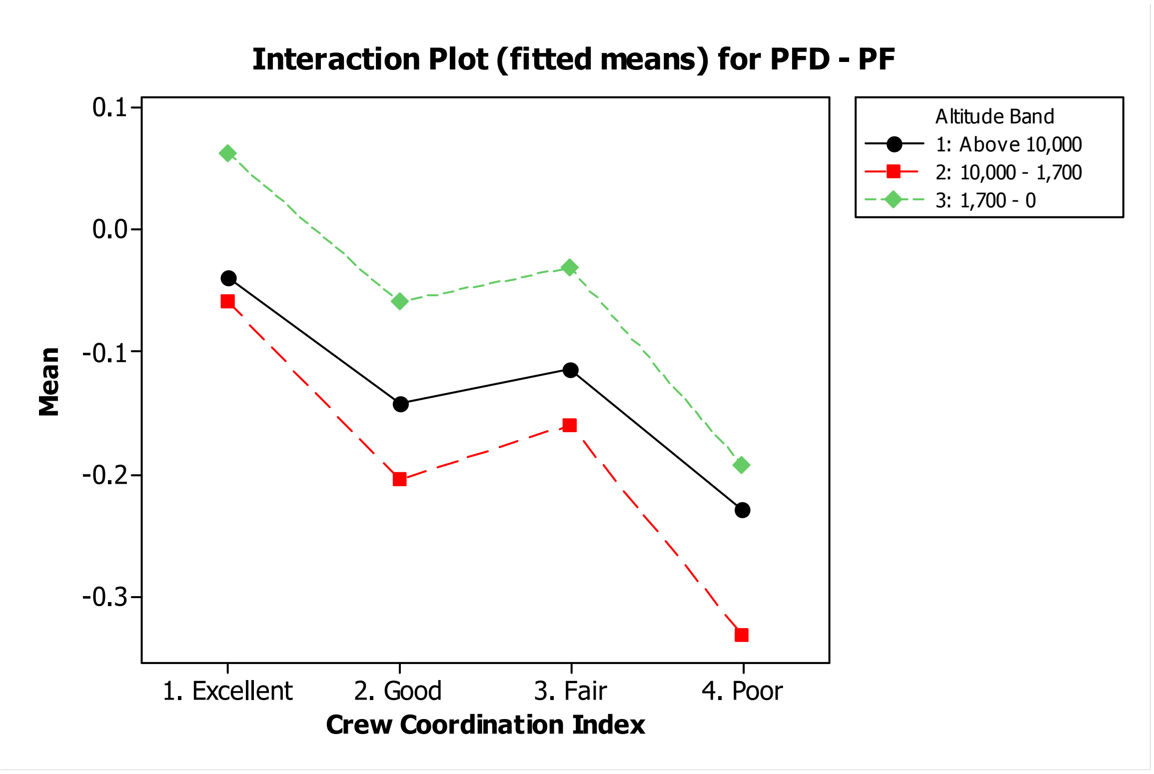


Figure 43. PFD Model Difference Interaction Plot - PF

المنارة للاستشارات

www.manaraa.com

### *5.3.1.4 Instrument Panel AOI, Pilot Monitoring*

Analysis of the IP normative eye-scan behavior model difference for the PM indicated statistically significant differences across the altitude bands ($F(2,282)=12.29$, $p<0.000$), Figure 44 and Figure 45 show the deviation of observed PDT from the normative model across the altitude bands and crew coordination index ratings. The analysis reported no statistically significant differences across crew coordination index ratings. Pairwise comparisons between altitude band segments showed no statistically significant difference between the high and the middle altitude bands. The significant difference was identified between the low altitude band when compared to both the high altitude band ($t=4.222$, $p<0.000$) and middle altitude band ($t=4.348$, $p<0.000$). Pairwise comparison findings are graphically shown in Figure 46.

Results suggest the eye-scan behavior of the PM varied significantly depending on phase of flight. The difference across altitude bands was expected as the primary tasking for each pilot changes from the final approach fix to touchdown. The FAF to touchdown segment was captured in the low altitude band where crew role tasks are well defined and attention is driven to specific AOIs. Lack of significant differences across the crew coordination index suggests that reduced crew coordination was not apparent in the eye-scan behavior of the PM with respect to the IP AOI. Findings of the PM with respect to the IP were in contrast to the results of PF eye-scan behavior that showed statistically significant differences across the crew coordination index.

www.manaraa.com

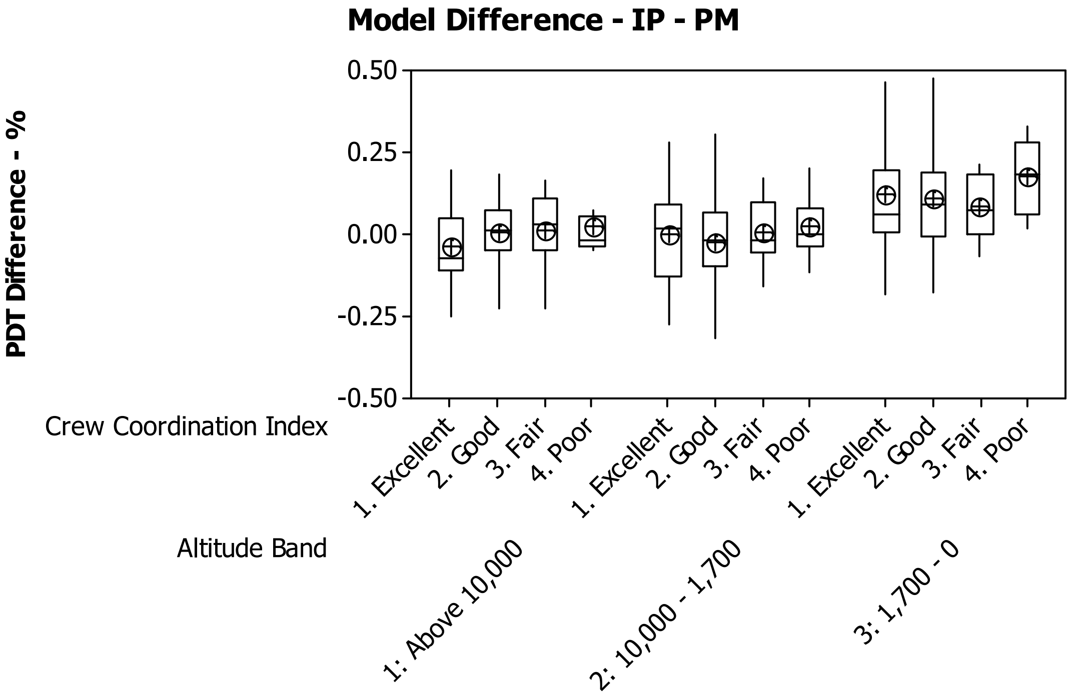


Figure 44. Instrument Panel Model Difference - PM

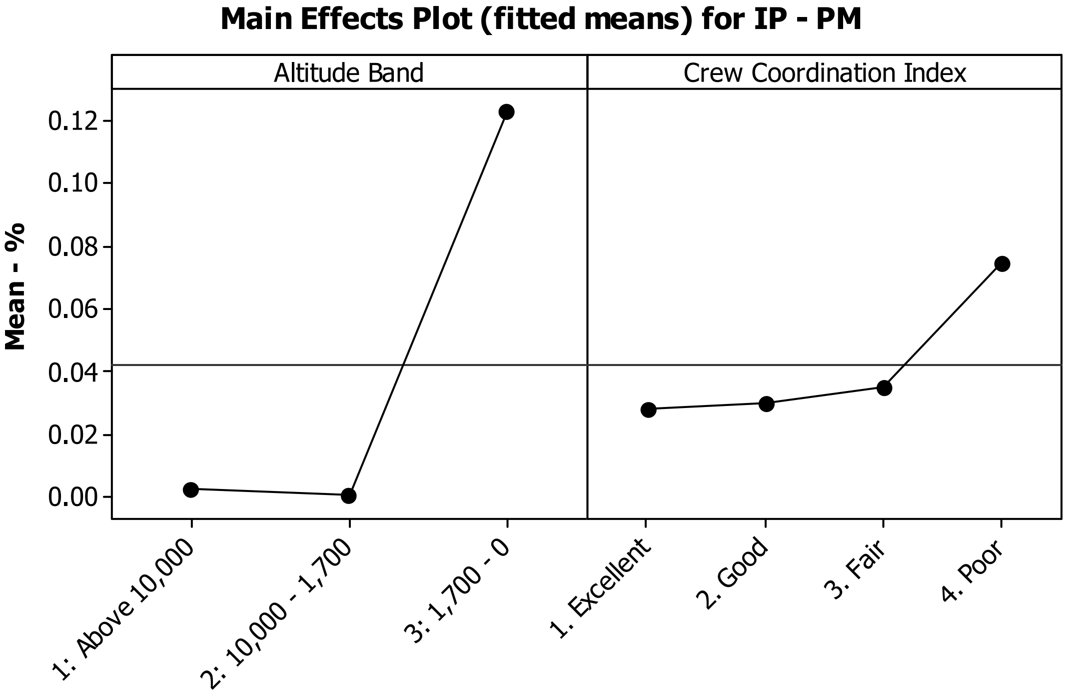


Figure 45. Instrument Panel Model Difference Main Effects - PM

المنارة للاستشارات

www.manaraa.com

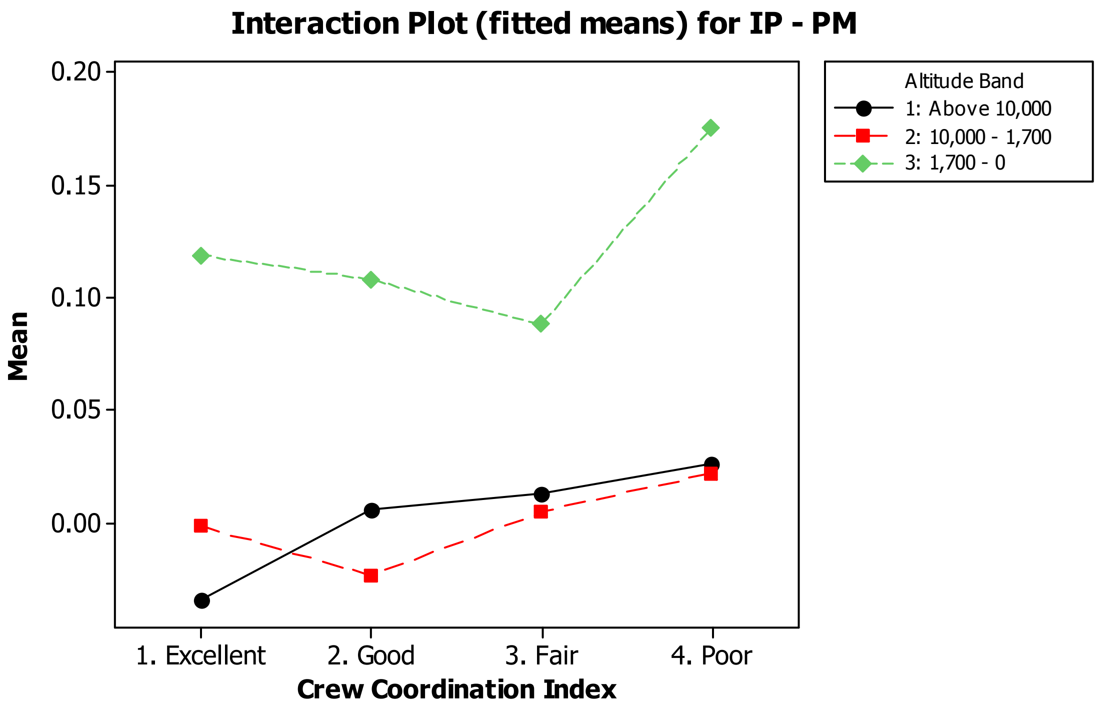


**Figure 46. Instrument Panel Model Difference Interaction Plot - PM**

*5.3.1.5 Out the Window AOI, Pilot Monitoring*

Analysis of the OTW AOI normative eye-scan behavior model difference for the PM indicated no statistically significant differences across the altitude band segments or crew coordination index (*F*(2,282)=2.44, *p*=0.065). Normative eye-scan model difference results are shown in Figure 47. Pairwise comparisons were not performed across the altitude bands or the crew coordination index ratings. Comparison data of the main effects and the interaction of effects are shown in Figure 48 and Figure 49.

Lack of significant differences across the crew coordination index suggests that the eye-scan behavior, in regard to the OTW AOI of the PM, was unaffected by reduced crew coordination. Findings suggest the information available OTW may not be as important as the flight technical information available on other AOIs. There was an observable trend in the data showing a crew coordination index of Poor induced a greater

المنارة للاستشارات
www.manaraa.com

deviation from the normative eye-scan behavior model than the other crew coordination index ratings. However, the differences from the normative eye-scan behavior model were not statistically significant. Results suggest there was some effect present with eye-scan behavior OTW with regard to reduced crew coordination.

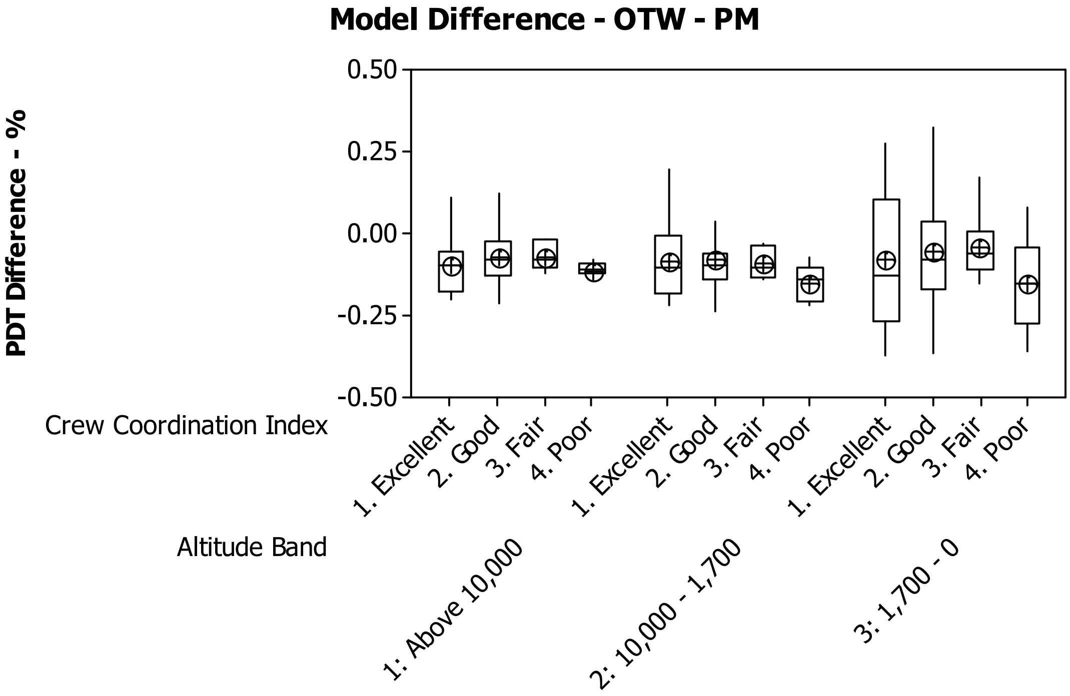


**Figure 47. Out the Window Model Difference - PM**

www.manaraa.com

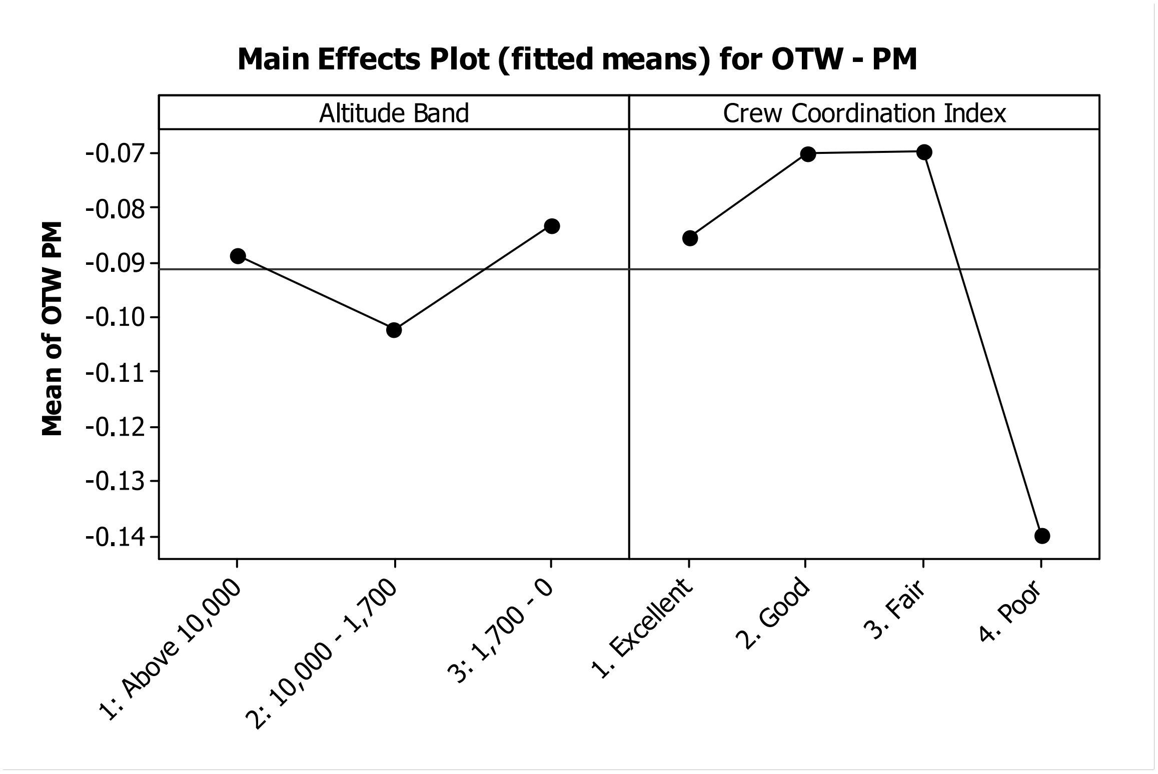


Figure 48. Out the Window Model Difference Main Effects - PM

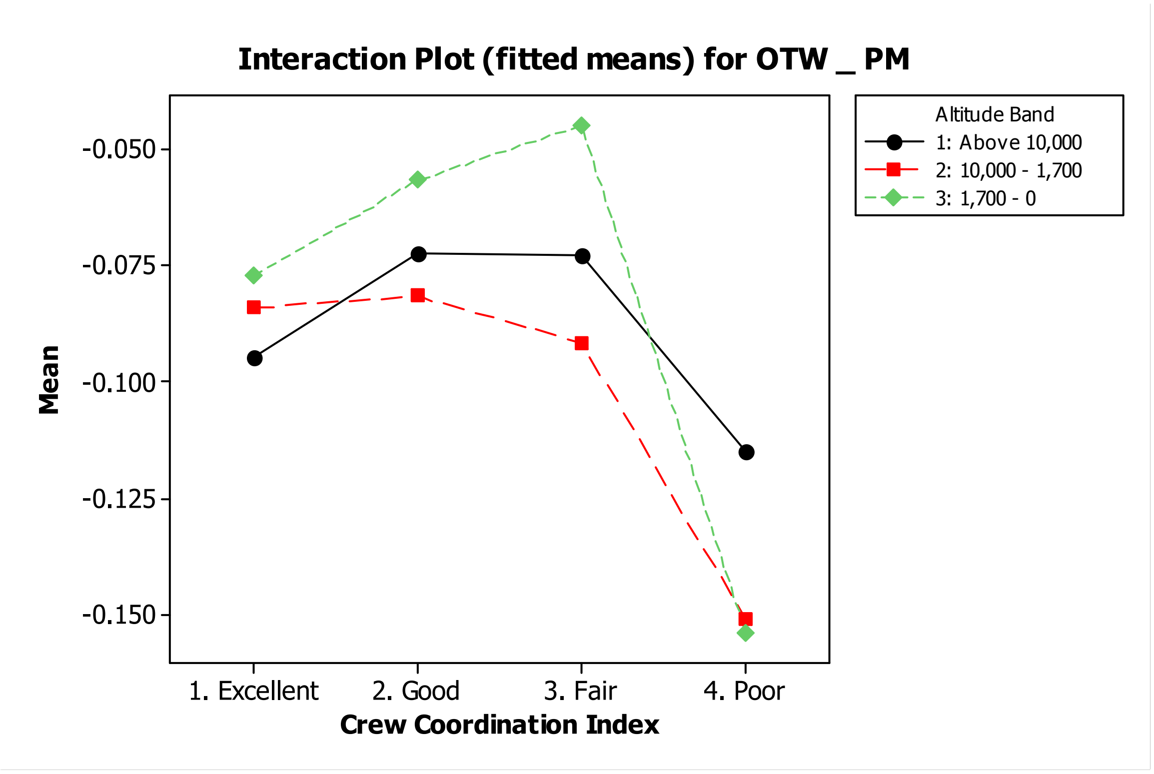


Figure 49. Out the Window Model Difference Interaction Plot - PM

المنارة للاستشارات

www.manaraa.com

#### *5.3.1.6 PFD AOI, Pilot Monitoring*

Analysis of the PFD normative eye-scan behavior model difference for the PM revealed statistically significant differences across the altitude bands ($F(2,282)=10.12$, $p<0.000$), data shown in Figure 50. Analysis revealed no statistically significant differences across crew coordination index ratings. Pairwise comparisons across the altitude bands showed no statistically significant difference between the high and the middle altitude bands. Pairwise comparison findings showed that a significant difference existed between the low altitude band and the high altitude ($t=4.287$, $p=0.0001$) and middle altitude ($t=3.387$, $p=0.0020$) bands. Pairwise comparison findings are graphically shown in Figure 51 and Figure 52.

The similarities between the IP AOI findings and the PFD AOI findings were expected, since the PFD AOI was defined inside the larger IP AOI. Findings of similar behavior indicate that the eye-scan behavior observed in regard to the IP was likely a result of the eye-scan behavior in regard to the PFD. The PFD provides both aircraft attitude state and flight path guidance, both of which are critical to the aviate and navigate tasks of the PF.

An inverse relationship in the eye-scan behavior between the PF and PM was observed when there was reduced crew coordination, particularly prevalent in the low altitude band, below 1,700 ft. It is posited that the inverse relationship in eye-scan behavior was due to visual cross-check between crewmembers and was identifiable in the PF-PM PDT difference analysis performed below. As the crew becomes less coordinated, the PM eye-scan behavior continued as expected, which was defined by the normative model. Figure 52 indicated the PM maintained visual attention to the PFD when the crew coordination worsened (crew coordination index equal to Poor). The PM behavior suggests that the reduction in crew coordination was more identifiable in the PF eye-scan behavior and less identifiable in the PM eye-scan behavior. The differences in

www.manaraa.com

effectiveness to characterize crew coordination between PF and PM eye-scan behaviors may be due to the PF's responsibility to hand-fly the aircraft. Additionally, the PF was the captain on the flight deck for the DataComm experiment. The captain under normal CRM procedures is responsible for task delegation to the PM (first officer), and failing to do so appropriately in the effort to aid in PM tasks resulted in the PF directing attention to those PM tasks.

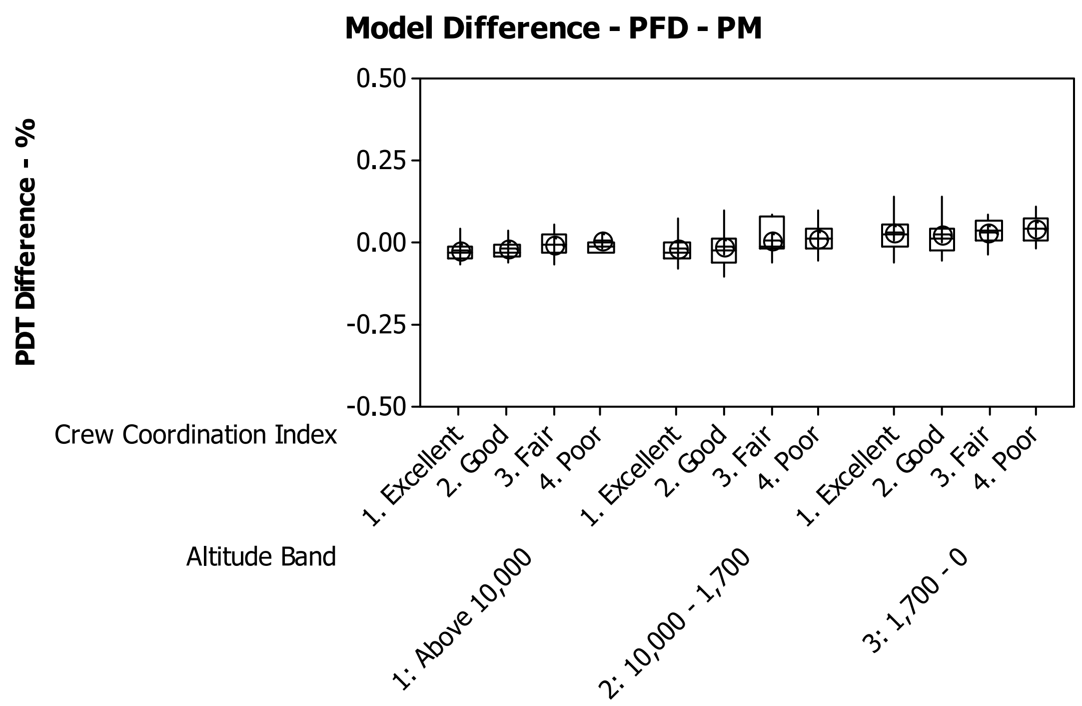


**Figure 50. PFD Model Differences - PM**

المنارة للاستشارات

www.manaraa.com

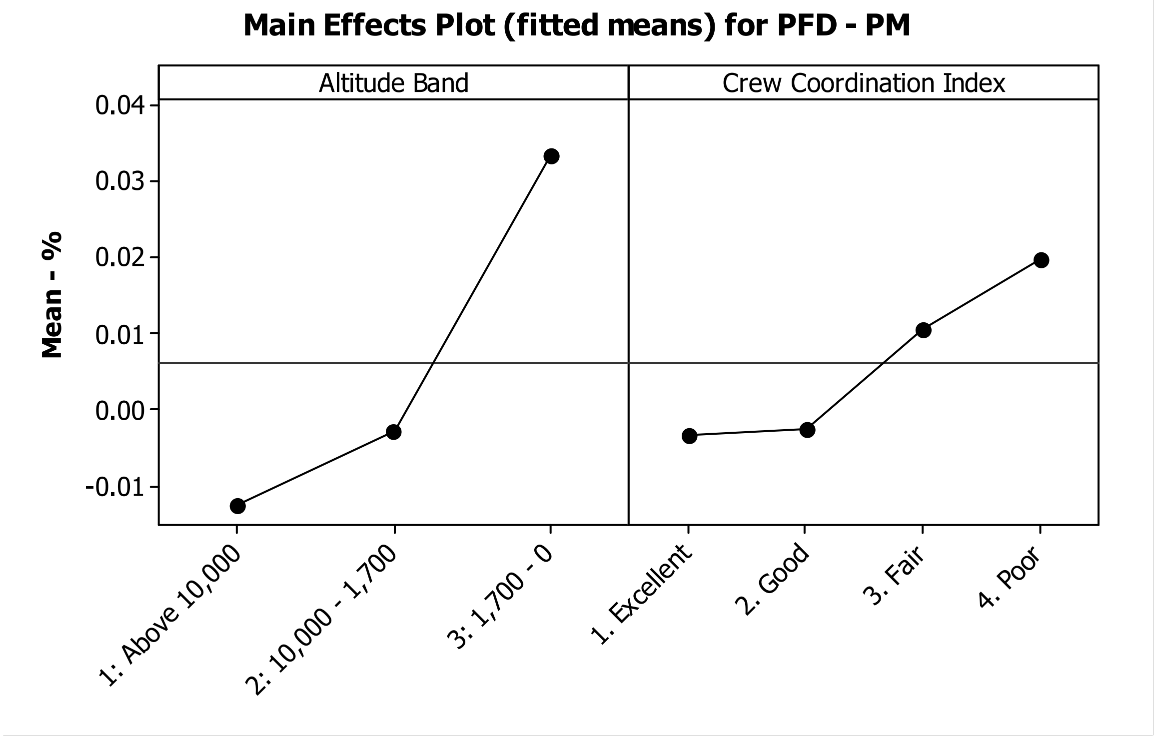


Figure 51. PFD Model Difference Main Effects - PM

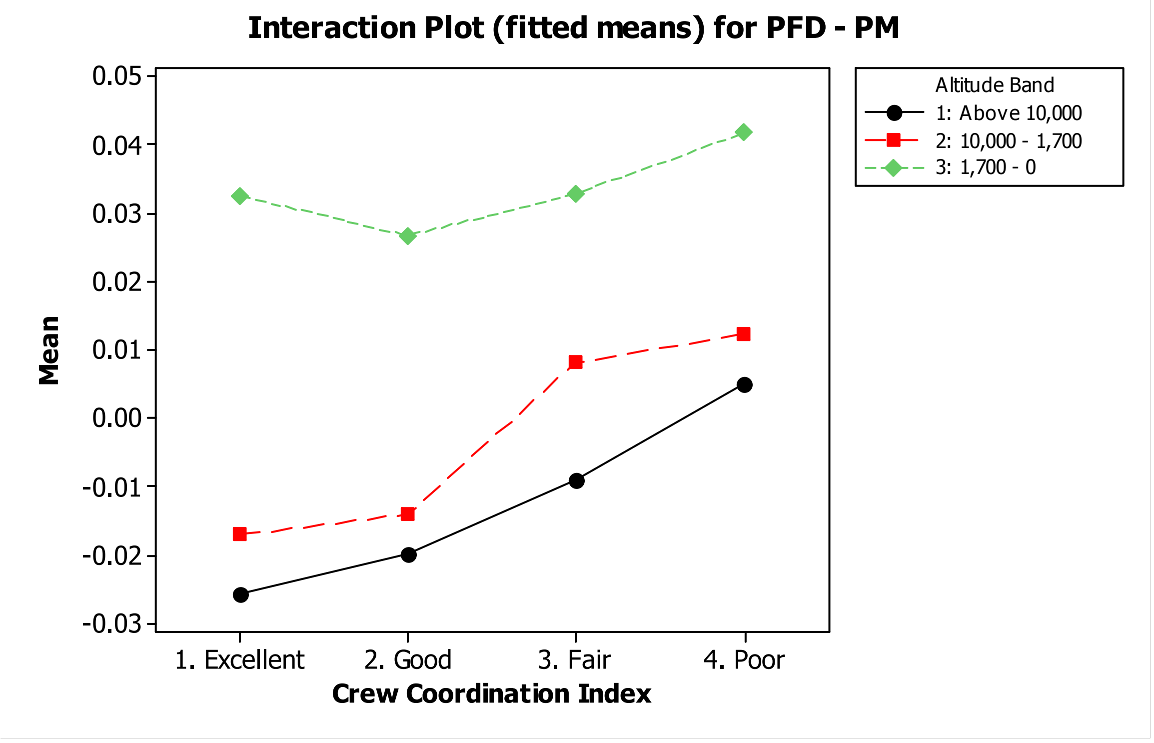


Figure 52. PFD Model Difference Interaction Plot - PM

المنارة للاستشارات

www.manaraa.com

### 5.3.2 Pilot Flying PDT Difference from Pilot Monitoring PDT

Statistical analyses were performed to evaluate the significant differences in observed PDT between the PF and the PM for all AOIs across all altitude bands and crew coordination index ratings. Positive differences indicate that the PF viewed that particular AOI more frequently, and negative values indicate the PM reviewed that particular AOI more frequently. A zero difference in PDT indicates that both pilots attended that particular AOI with the same PDT.  AOIs evaluated include the PFD, ND, CDU, AI, ALT, IP, and OTW. Observations of AOIs with statistically significant observations are included in this section and described with regard to flight operational impact. The comprehensive analysis of all PF PDT differences from PM PDT statistics is included in Appendix D.

#### *5.3.2.1 Instrument Panel PDT Difference*

Analysis of the PDT difference between the PF and PM on the IP AOI indicated statistically significant differences across the altitude band segments ($F(2,282)=9.56$, $p<0.0001$), data shown in Figure 53. The analysis reported statistically significant differences across the crew coordination index ratings, ($F(3,282)=8.49$, $p<0.0001$). Pairwise comparisons across the altitude band segments showed no statistically significant difference between the high and the middle altitude bands. Pairwise comparison analyses revealed the significant difference exists between the low altitude band and both the high altitude band ($t=-4.204$, $p<0.001$) and middle altitude band ($t=-3.124$, $p=0.005$). Pairwise comparison findings are graphically shown in Figure 54 and

المنارة للاستشارات

Figure 55. Pairwise comparisons across crew coordination index showed no significant differences between ratings of Excellent, Good, and Fair coordination. Significant differences appeared between the crew coordination index rating of Poor and all other ratings: Excellent vs. Poor (t=-4.746, $p<0.001$), Good vs. Poor (t=-4.709, $p<0.001$), Fair vs. Poor (t=-3.818, $p=0.001$).

Analyses indicated when there was a reduction in crew coordination there was a significant difference between the PF and PM cross-check of the IP displays. Crew coordination index ratings Excellent through Fair show attention given to the IP was greater for the PF. As the altitude bands change from high to low, the eye-scan behavior shifted to be increasingly equal as tasking allowed and was required of the PM. Shifts in eye-scan behavior can be explained by the specific tasking to each pilot being specific to their role, and cross-checking was not as prevalent until the aircraft was closer to the ground and on the ILS inside the FAF.

Pairwise comparison analyses revealed that the eye-scan behavior associated with a reduction in crew coordination was observed across all altitude bands. With a reduction in crew coordination, visual attention was increasingly balanced between the PF and PM. The less coordinated the crew was, the more the crew’s attentional behavior pattern shifted between the PF and PM. Findings suggest there was a threshold to differences in IP PDT between the PF and PM to identify reduced crew coordination, indicated by the pairwise comparisons between a crew coordination index rating of Poor and all other index ratings.

المنارة للاستشارات

www.manaraa.com

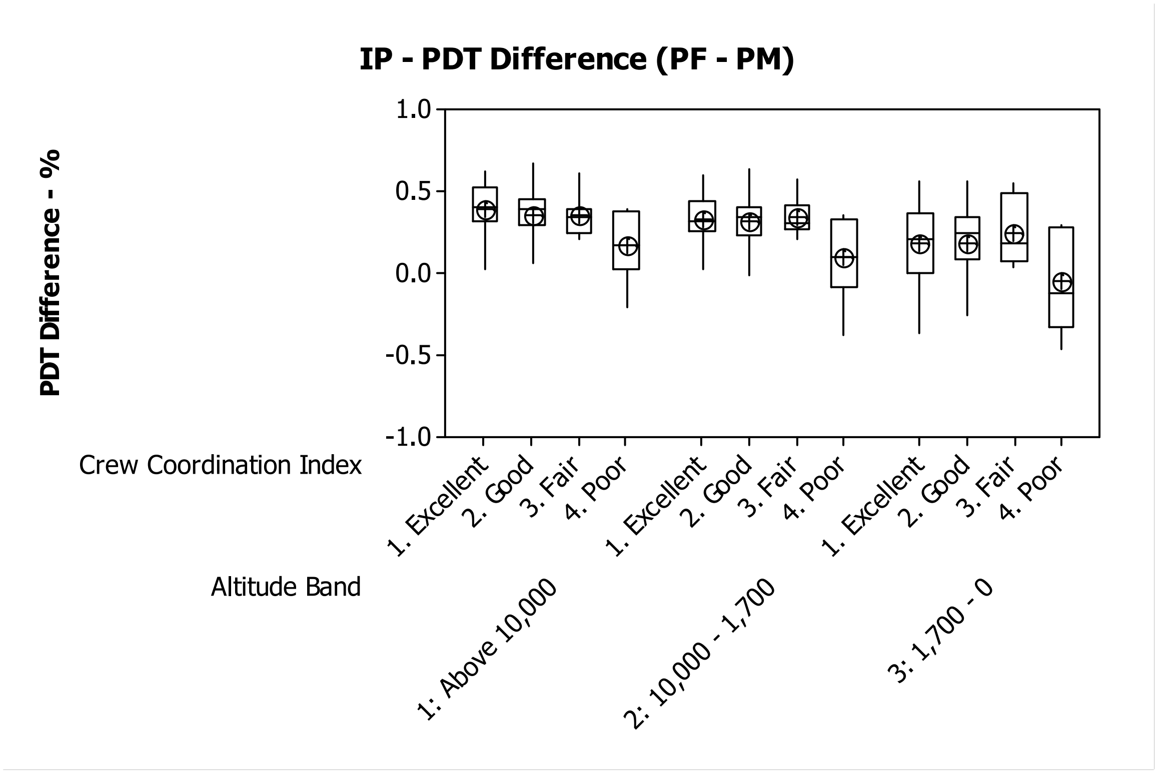


**Figure 53. Instrument Panel PDT Difference - (PF - PM)**

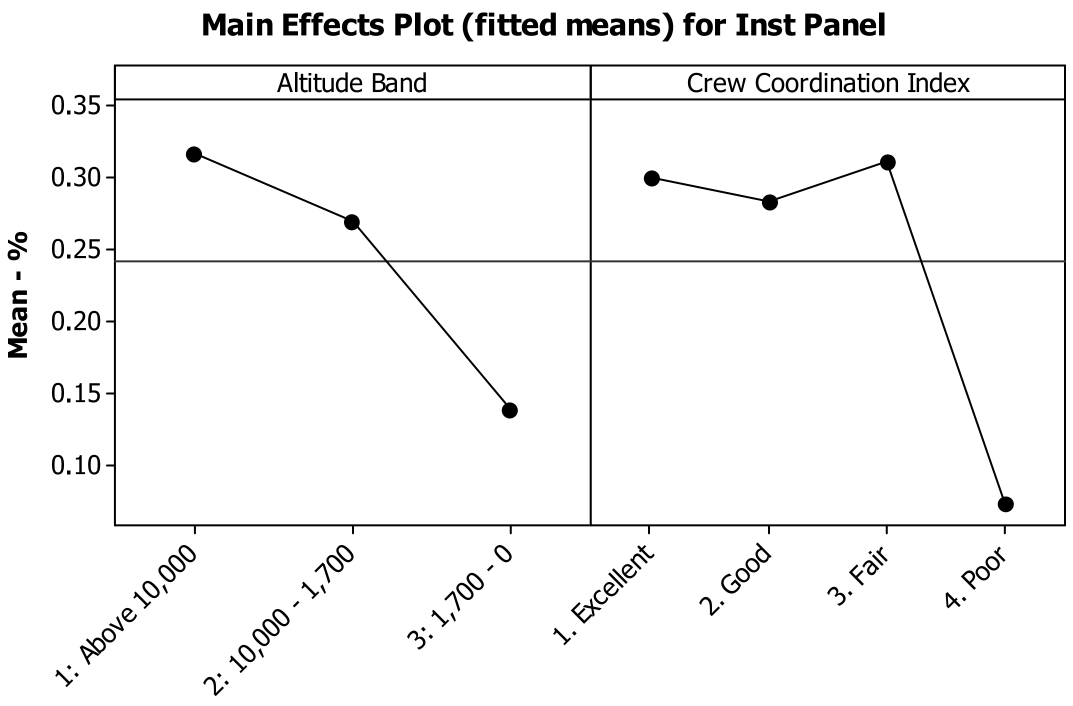


Figure 54. Instrument Panel PDT Difference Mean Effects - (PF - PM)

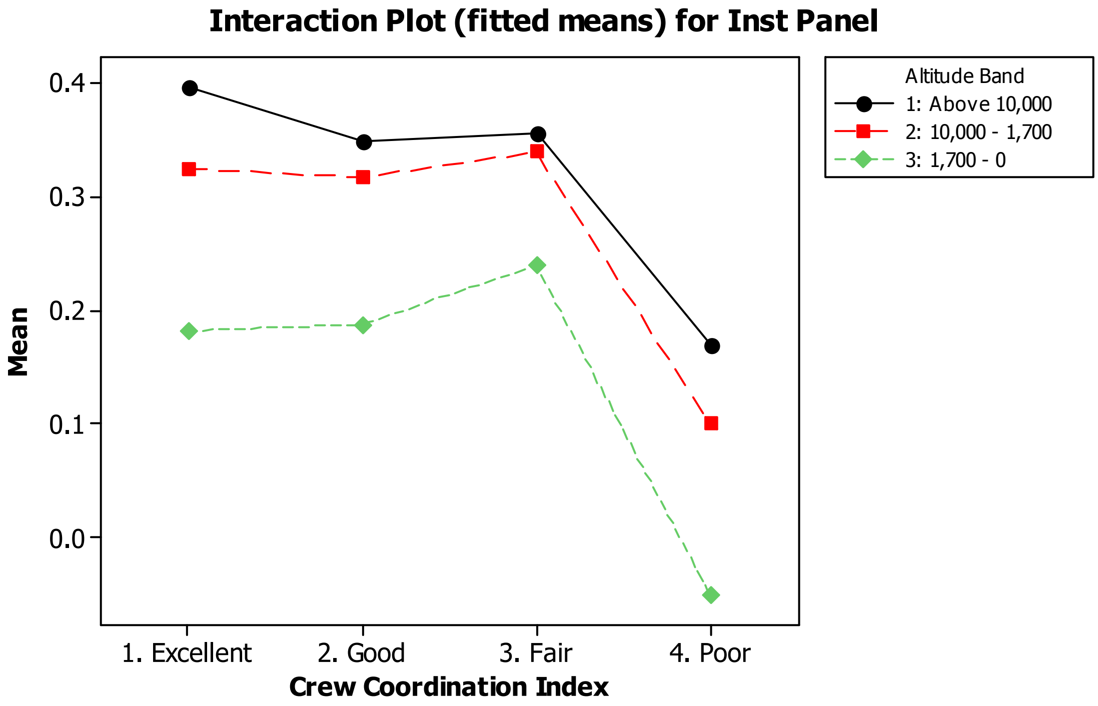


Figure 55. Instrument Panel PDT Difference Interaction Plot - (PF - PM)

www.manaraa.com

### *5.3.2.2 PFD PDT Difference*

Evaluation of the PDT difference between the PF and PM on the PFD AOI indicated statistically significant differences across the crew coordination index ratings ($F(2,282)=16.55$, $p<0.0000$). There were no statistical differences observed across altitude bands. The data comparisons are shown in Figure 56 and Figure 57. Pairwise comparisons across crew coordination index showed no significant differences across ratings of Excellent, Good, and Fair. Significant differences appeared between a crew coordination index rating of Poor and all other ratings: Excellent vs. Poor (t=-4.204, $p<0.001$), Good vs. Poor (t=-3.124, $p=0.005$), Fair vs. Poor (t=-4.746, $p<0.001$). Pairwise comparison findings are graphically shown in Figure 58.

The results of the PFD PDT difference analyses showed that there were no significant differences across the altitude bands, suggesting the visual attention given to the PFD maintains a common behavior between the PF and PM. Significant differences across the crew coordination index ratings with regard to the PFD follows the same findings to the IP. A common trend of diminishing difference between the PF and PM PDT of the PFD with reduced crew coordination was observed. Findings suggest that with reduced crew coordination there was a shift in eye-scan strategy between the PF and PM. The shift in eye-scan strategy was likely indicative of non-standard cross-checking. Contrasting the PDT difference findings to the results of the normative eye-scan behavior model differences indicated that when crew coordination index was Poor, the PF reduced the attention given to the PFD. Additionally, the PM appears to have compensated by cross-checking the PF instruments and increasing the attention given to the PFD.

المنارة للاستشارات

www.manaraa.com

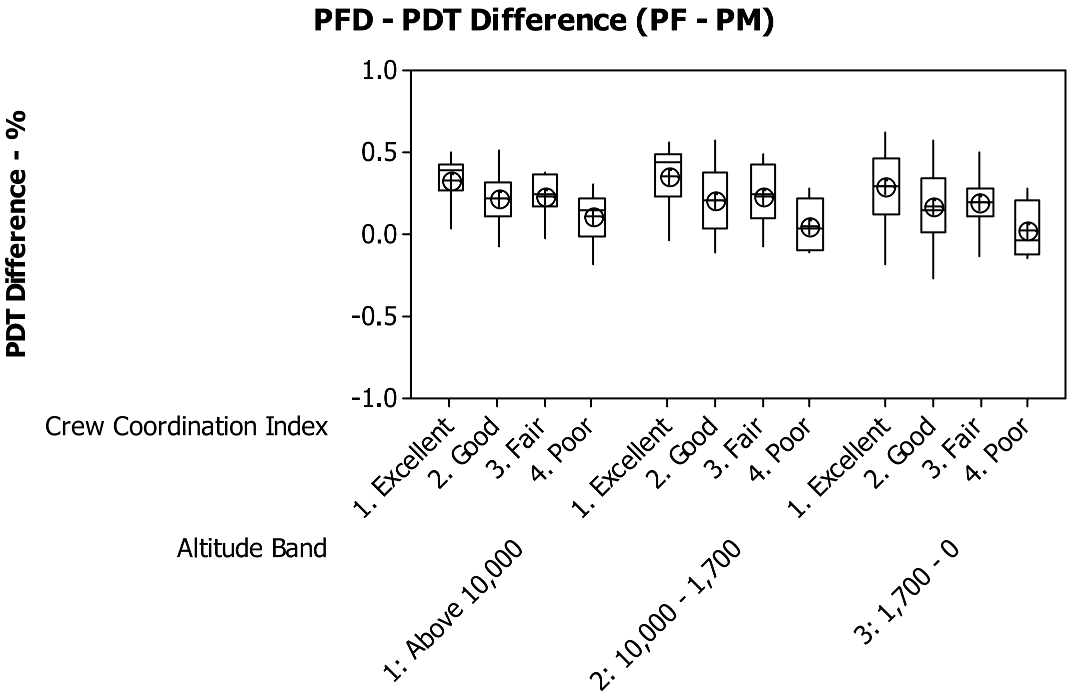


Figure 56. PFD PDT Difference - (PF - PM)

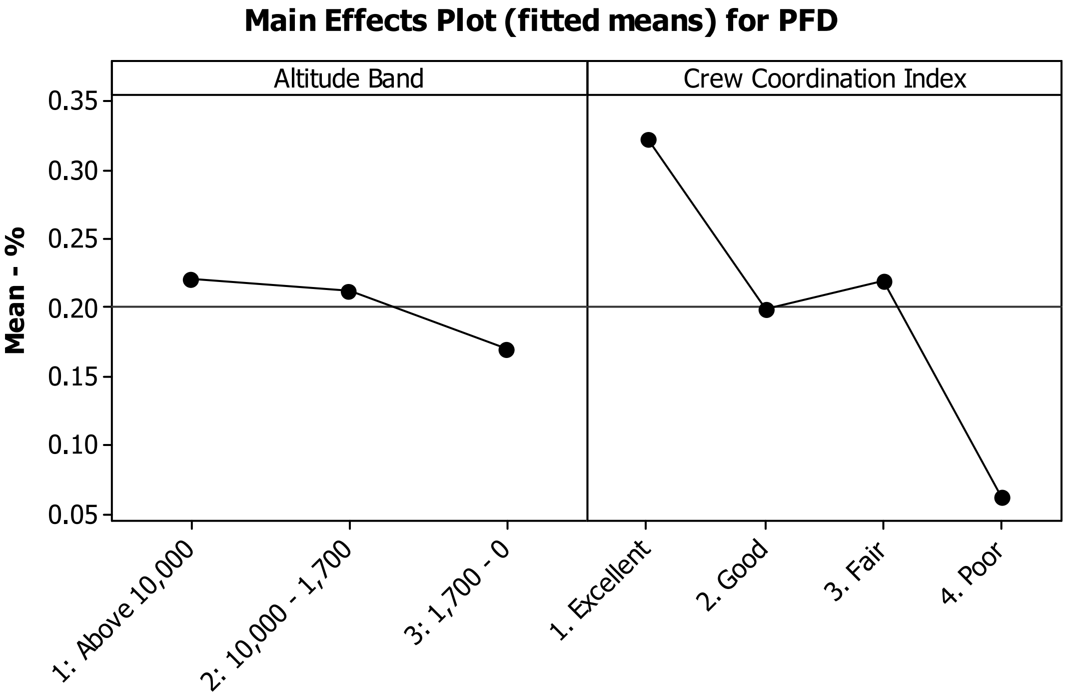


Figure 57. PFD PDT Difference Main Effects - (PF - PM)

www.manaraa.com

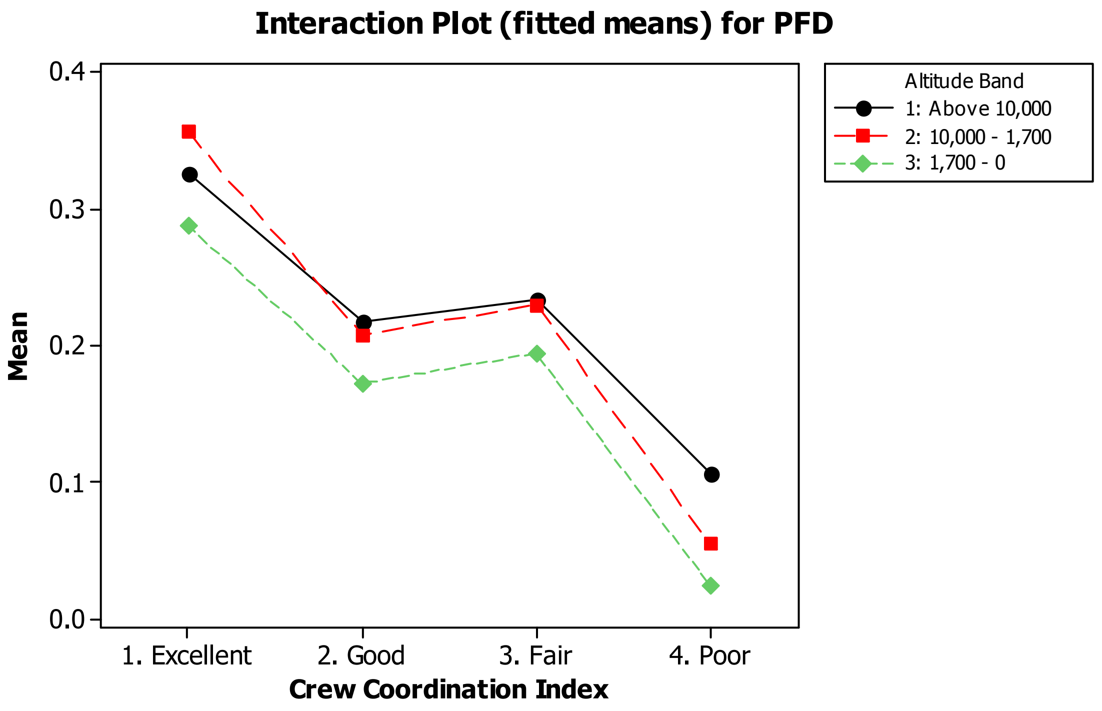


**Figure 58. PFD PDT Difference Interaction Plot - (PF - PM)**

### 5.3.3 Shared AOI, Dwell Time (5 second time window)

Statistical analyses for shared attention were performed across all AOIs for each pilot. The shared attention values range from zero to one, representing the percentage time both pilots dwelled on the same AOI within five seconds of each other. AOIs evaluated include the PFD, ND, CDU, AI, ALT, IP, and OTW. Shared attention analysis investigates the time pilots have shared visual attention within a time window of five seconds. The shared attention metric represents eye-scan behavior for each AOI that reveals how much time both the PF and PM have common awareness of specific sets of information. Observations for AOIs with statistically significant differences are included in this section and described with regard to flight operational impact. The comprehensive analysis of all shared attention AOI statistics is included in Appendix D.

www.manaraa.com

#### *5.3.3.1 Instrument Panel AOI, Shared Attention*

Analysis of crew shared attention on the IP AOI indicated statistically significant differences across altitude bands ($F(2,285)=13.88$, $p<0.000$), data is shown in Figure 59 and Figure 60. Analysis indicated statistically significant differences across the crew coordination index ratings, ($F(3,285)=4.36$, $p=0.005$). Pairwise comparisons across altitude bands show no statistically significant difference between the high and the middle altitude bands. Pairwise comparisons revealed a significant difference between the low altitude band and both the high altitude band (t=-3.608, $p=0.001$) and middle altitude band (t=-5.128, $p<0.001$)). Pairwise comparison findings were graphically shown in Figure 61. Pairwise comparisons across the crew coordination index ratings showed no significant differences between ratings of Excellent, Good, and Fair coordination. Significant differences appeared between a crew coordination index of Poor and all other index ratings: Excellent vs. Poor (t=-3.323, $p=0.005$), Good vs. Poor (t=-3.266, $p=0.006$), Fair vs. Poor (T=-2.951, $p<0.017$).

Findings revealed similar results to that of both the normative eye-scan behavior model difference and the PDT difference findings. Commonality across eye-scan metrics was that when there was a reduction in crew coordination, variation in eye-scan behavior was increasingly distinct. The significant differences across altitude band segments indicated variable magnitudes of shared attention, which was expected with task demands specific to each segment. A decrease in shared attention of the IP in combination with a reduction in crew coordination suggests that crew coordination was reliant on shared information between pilots. A threshold between crew coordination index ratings of Fair and Poor coordination with respect to a reduction in attention was observed with the shared attention metric. A threshold at the index rating of Fair was similar to the findings in the other eye-scan behavior metrics.

المنارة للاستشارات

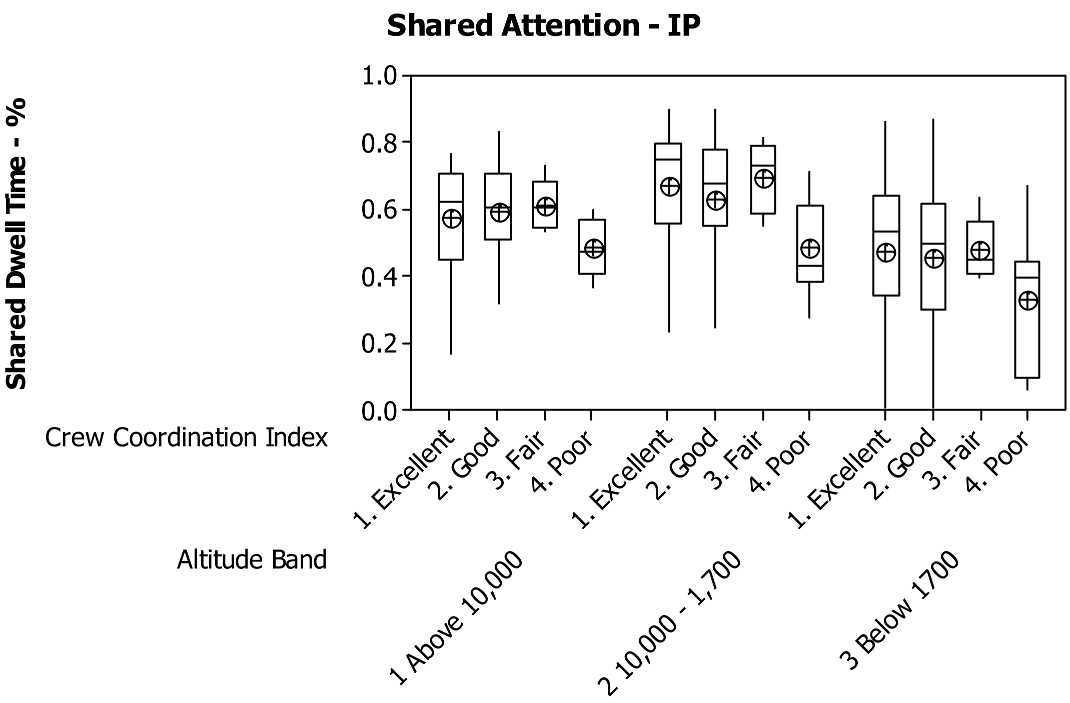


**Figure 59. Instrument Panel Shared Attention**

المنارة للاستشارات

www.manaraa.com

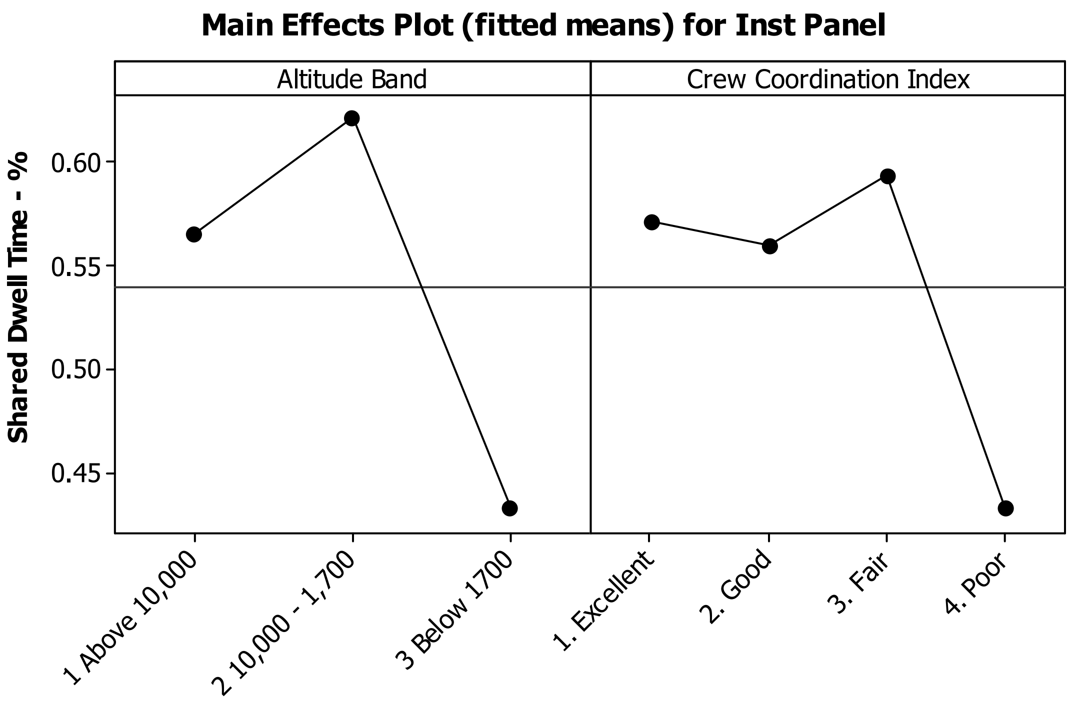


Figure 60. Instrument Panel Shared Attention Main Effects

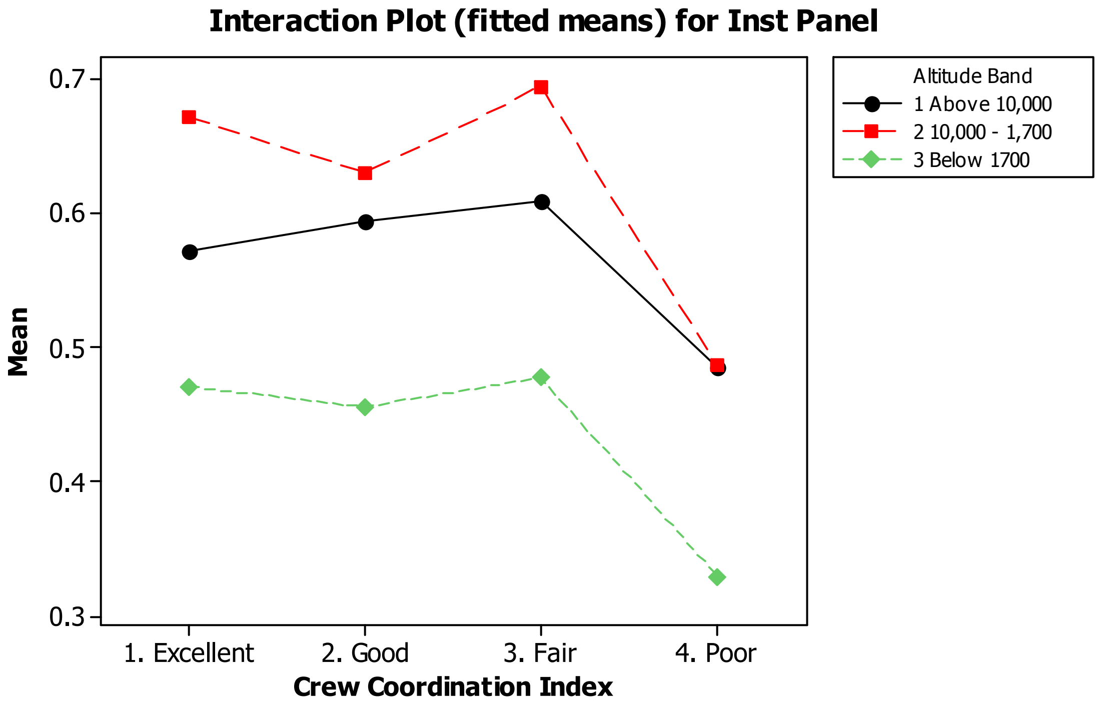


Figure 61. Instrument Panel Shared Attention Interaction Plot

المنارة للاستشارات

www.manaraa.com

### *5.3.3.2 Out the Window AOI, Shared Attention*

Analysis of crew shared attention on the OTW AOI indicated statistically significant differences across the altitude bands ($F(2,285)=127.66$, $p<0.000$), data shown in Figure 62 and Figure 63. No statistically significant differences were found across the crew coordination index ratings. Pairwise comparisons across the altitude bands showed no statistically significant difference between the high and the middle altitude bands. Pairwise comparison analyses revealed the significant difference in altitude bands exists between the low altitude band and the high altitude ($T=13.312$, $p<0.001$) and the middle altitude ($t=14.310$, $p<0.001$) bands. The pairwise comparisons are shown graphically in Figure 64.

Results quantitatively showed that pilot eye-scan behavior was almost never directed OTW until established past the final approach fix, captured in the low altitude band. There were no significant information references OTW until the aircraft was near the runway when the pilot must acquire the visual landing references. Lack of significant findings across the crew coordination index ratings suggest that the eye-scan behavior with regard to the OTW AOI was not as affected by a reduction in crew coordination. There appeared to be a common trend of reduced shared attention with reduced crew coordination index, and the reduction in shared attention may be operationally significant with a difference of approximately seven percent between coordination ratings of Excellent and Poor. While not statistically significant, the common trend of reduced shared awareness with reduced coordination observed in the other eye-scan metrics was also apparent in the shared attention data of the low altitude band.

المنارة للاستشارات

www.manaraa.com

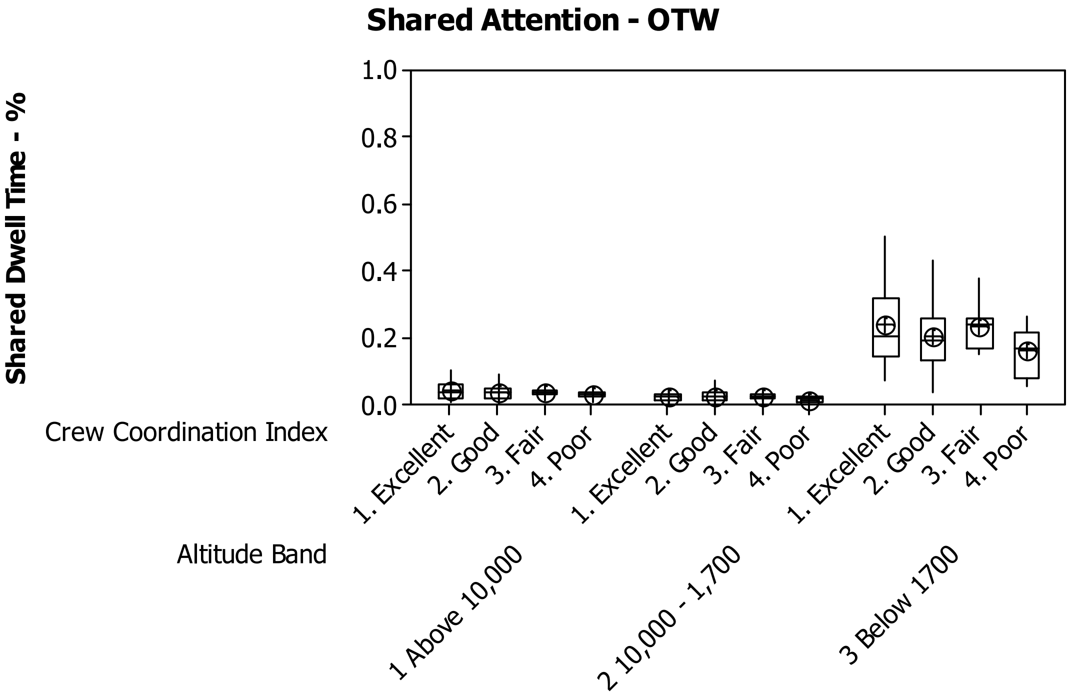


Figure 62. Out the Window Shared Attention

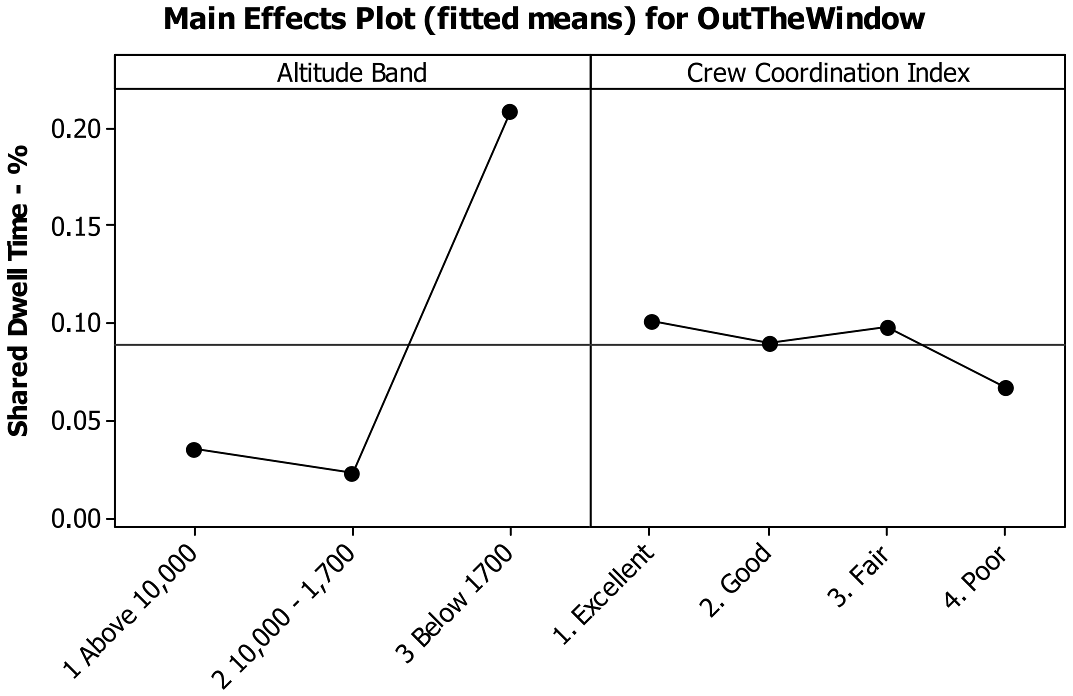


Figure 63. Out the Window Shared Attention Main Effects

المنارة للاستشارات

www.manaraa.com

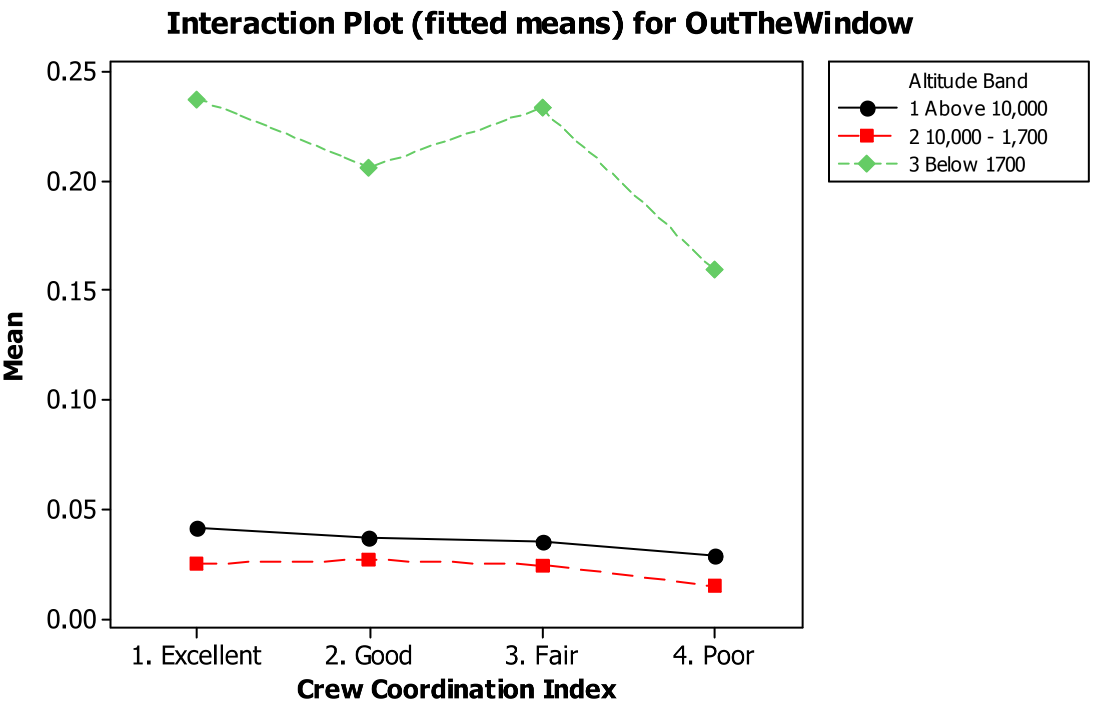


**Figure 64. Out the Window Shared Attention Interaction Plot**

*5.3.3.3 PFD Shared Attention*

Analysis of shared awareness of the PFD indicated statistically significant differences across the altitude band segments (*F*(2,285)=3.89, *p* =0.020), shown in Figure 65 and Figure 66. The analysis reported statistically significant differences across the crew coordination index ratings, (*F*(3,285)=10.89, *p*<0.000). Pairwise comparison across the altitude band segments revealed statistically significant differences that existed between the high and the middle altitude bands (t=2.637, *p*<0.023). Pairwise comparisons across the crew coordination index ratings showed no significant differences between ratings of Excellent and Fair coordination. The significant differences appeared between the crew coordination index ratings of Poor and all other ratings except for Fair, which was nearly significant -Excellent vs. Poor (t=-5.117, *p*<0.001), Good vs. Poor (t=-2.613, *p*=0.044), Fair vs. Poor (t=-2.544, *p*=0.053). There was also a significant difference

www.manaraa.com

between a crew coordination index of Good and a crew coordination index of Excellent (T=-4.401, *p*<0.001). Pairwise comparison findings are shown graphically in Figure 67.

There was an altitude dependent correlation between crew coordination index and the shared awareness of the PFD. Results of the shared attention analysis of the PFD were similar to the other statistically significant findings in the above analyses suggesting shared attention was a significant eye-scan behavior indicator and was capable of characterizing reduced coordination. The common trend of reduced shared attention of the PFD with reduced crew coordination across altitude bands suggests that regardless of the variable task sets, shared awareness of the PFD was indicative of good crew coordination. The PFD as a primary instrument across all phases of flight supports the findings of the shared attention analysis. Additionally, findings from the shared attention analysis also suggest that cross-checking of the PFD was conducted to a greater extent in well-coordinated crews, particularly below 10,000ft.

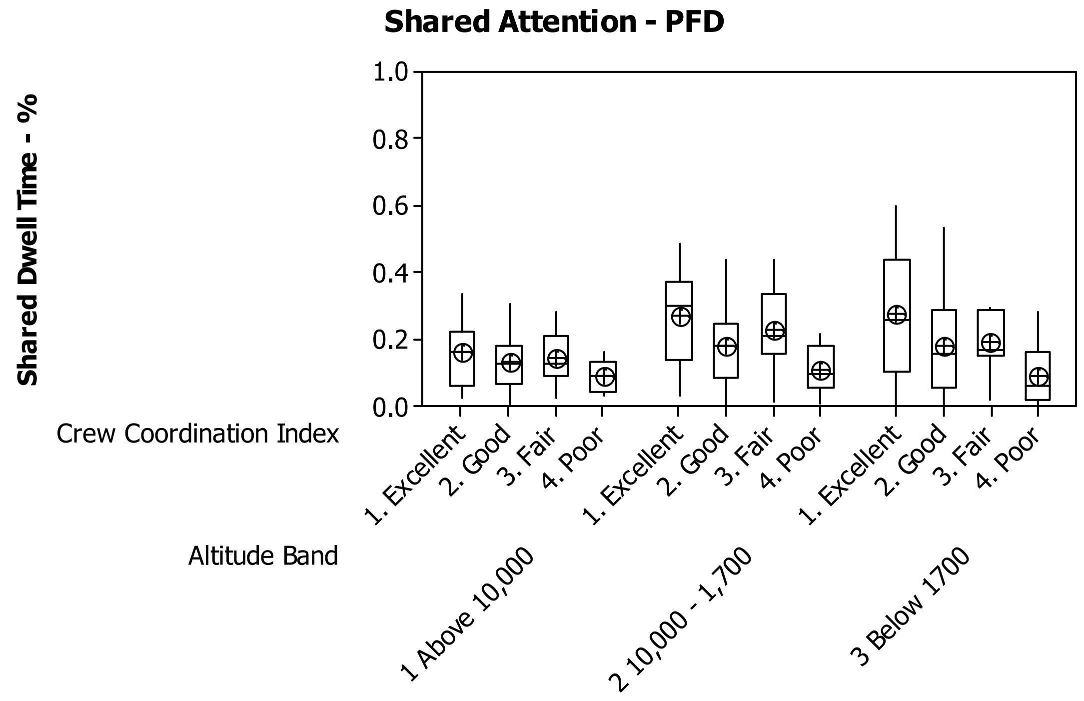


**Figure 65. PFD Shared Attention**

www.manaraa.com

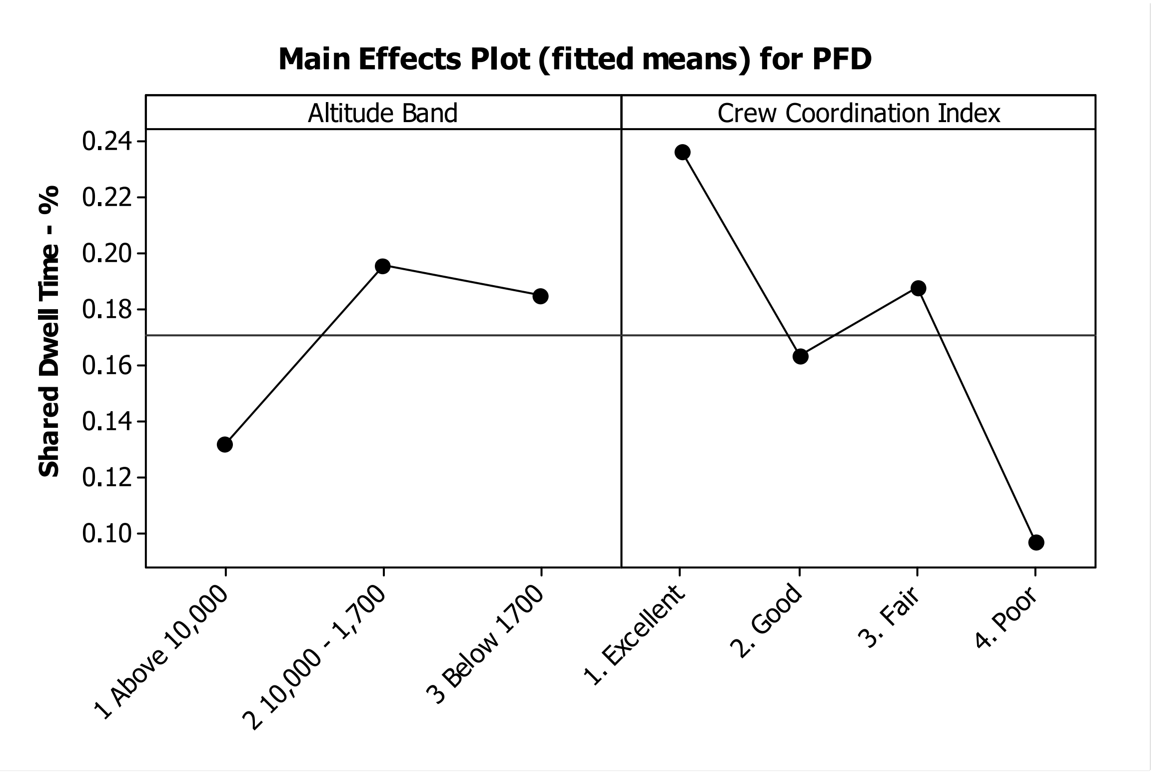


Figure 66. PFD Shared Attention Main Effects

www.manaraa.com

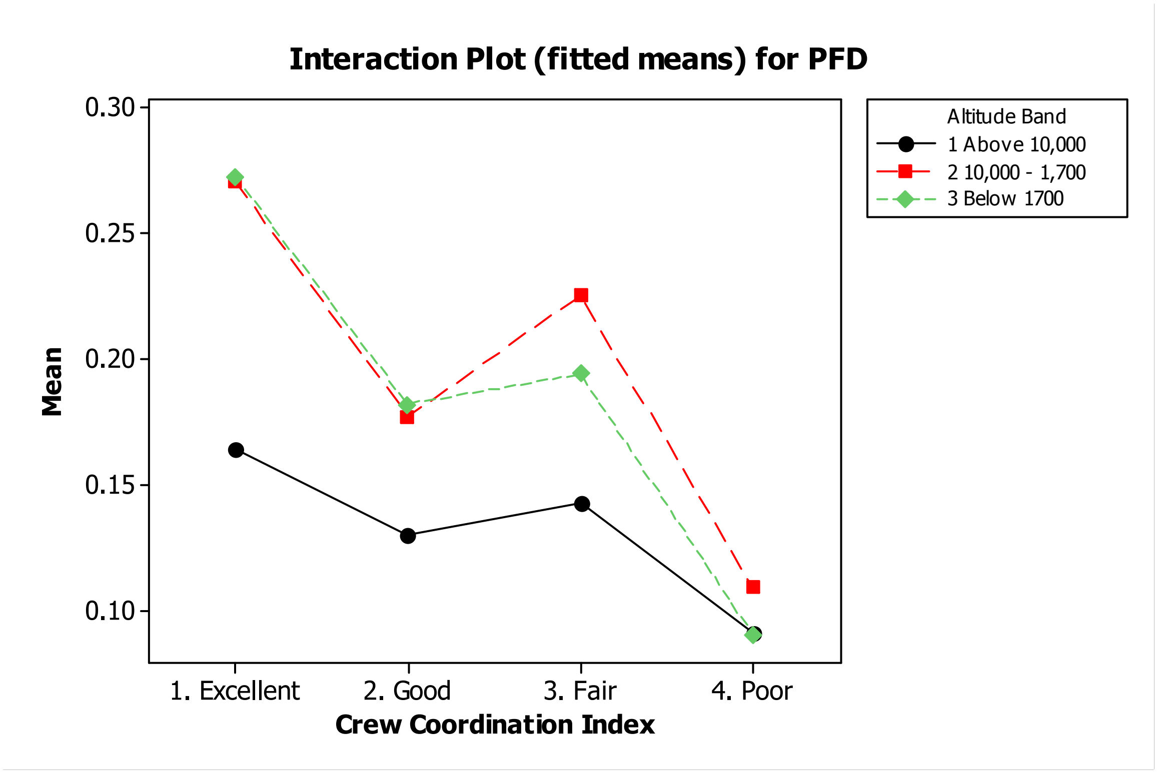


Figure 67. PFD Shared Attention Interaction Plot

www.manaraa.com

# CHAPTER 6.

# CLASSIFIER RESULTS

Based upon the findings of the statistical data analysis, data from the down selected AOI set for the normative eye-scan behavior model difference metric was processed using a machine learning classification technique. The WEKA classification software was used to transform the data using an unsupervised nominal to binary filter of the eye-tracking data with reference to the crew coordination index. The nominal to binary filter transformed the eye-tracking measures in the dataset to be categorically grouped with respect to crew coordination index. The filtered data was then processed through a classification method - the simple CART decision tree model.

Due to a class imbalance across crew coordination index ratings, the data was balanced by replicating observations for each crew coordination index. Index coordination values were replicated to be equal to the greatest number of observations for a single crew coordination index rating. Index rating sets were then doubled to ensure a minimum of 100 data points for each index rating was available to the classifier. The classification of the data was then compared using an index versus index approach, including crew coordination index rating of Excellent versus crew coordination index rating of Poor, Good versus Poor, Fair versus Poor. The index versus index approach was appropriate given the significant correlation between eye-scan behavior associated and crew coordination index, as summarized in the findings of Chapter 5.

**Results for Excellent versus Poor coordination rating classification are shown below in**
**Table 7,**

Table 8, and Table 9. Classification was made using the IP AOI for all three altitude bands. Further analysis of the data classification also indicates the PFD yields successful binary classification. Analysis of accuracy, precision, sensitivity, and specificity across all three altitude bands reported values of 100 percent. Results indicate the classifier was

www.manaraa.com

completely successful in classifying the dataset between index ratings of Excellent and Poor across all altitude bands.

**Table 7. Excellent vs. Poor Coordination Classifier Confusion Matrix, High Altitude Band**

Confusion Matrix

| *classified as:* | *Excellent Coordination* | *Poor Coordination* | |
|---|---|---|---|
| | 135 | 0 | *Excellent Coordination* |
| | 0 | 135 | *Poor Coordination* |
| | | | Truth |

**Table 8. Excellent vs. Poor Coordination Classifier Confusion Matrix, Middle Altitude Band**

Confusion Matrix

| *classified as:* | *Excellent Coordination* | *Poor Coordination* | |
|---|---|---|---|
| | 135 | 0 | *Excellent Coordination* |
| | 0 | 135 | *Poor Coordination* |
| | | | Truth |

**Table 9. Excellent vs. Poor Coordination Classifier Confusion Matrix, Low Altitude Band**

Confusion Matrix

| *classified as:* | *Excellent Coordination* | *Poor Coordination* | |
|---|---|---|---|
| | 135 | 0 | *Excellent Coordination* |
| | 0 | 135 | *Poor Coordination* |
| | | | Truth |

المنارة للاستشارات

www.manaraa.com

Results for the Good versus Poor crew coordination index rating classification are shown below in Table 10, Table 11, and Table 12. Classification was made utilizing the IP AOI for all three altitude bands. Accuracy analysis of the classification resulted in values of 93.75 percent, 95.54 percent, and 95.54 percent for the high, middle, and low altitude bands, respectively. Precision analysis of the classification resulted in values of 93.75 percent, 95.54 percent, and 95.54 percent for the high, middle, and low altitude bands, respectively. Sensitivity analysis of the classification resulted in a sensitivity of 100 percent across all three altitude bands. Specificity analysis of the classification resulted in values of 88.89 percent, 91.80 percent, and 91.80 percent for the high, middle, and low altitude bands, respectively. Results indicate the classifier was very successful in classifying the dataset between index ratings of Good and Poor across all altitude bands and increasingly accurate for the low and middle altitude bands.

**Table 10. Good vs. Poor Coordination Classifier Confusion Matrix, High Altitude Band**

Confusion Matrix

| *classified as:* | *Good Coordination* | *Poor Coordination* | |
|---|---|---|---|
| | 98 | 14 | Good Coordination |
| | 0 | 112 | Poor Coordination |
| | | | Truth |

www.manaraa.com

Table 11. Good vs. Poor Coordination Classifier Confusion Matrix, Middle Altitude Band

Confusion Matrix

| *classified as:* | *Good Coordination* | *Poor Coordination* | |
|---|---|---|---|
| | 102 | 10 | *Good Coordination* |
| | 0 | 112 | *Poor Coordination* |
| | | | Truth |

Table 12. Good vs. Poor Coordination Classifier Confusion Matrix, Low Altitude Band

Confusion Matrix

| *classified as:* | *Good Coordination* | *Poor Coordination* | |
|---|---|---|---|
| | 102 | 10 | *Good Coordination* |
| | 0 | 112 | *Poor Coordination* |
| | | | Truth |

Results for the Fair versus Poor crew coordination index rating classification are shown below in Table 13, Table 14, and Table 15. Classification was made utilizing the IP AOI for the high altitude band only. Accuracy analysis of the classification resulted in accuracies of 88.73 percent, 88.73 percent, and 85.59 percent for the high, middle, and low altitude bands, respectively. Precision analysis of the classification resulted in values of 100 percent, 100 percent, and 92.38 percent for the high, middle, and low altitude bands, respectively. Sensitivity analysis of the classification resulted in values of 100 percent, 100 percent, and 92.38 percent for the high, middle, and low altitude bands, respectively. Specificity analysis of the classification resulted in values of 88.89 percent, 91.80 percent, and 91.80 percent for the high, middle, and low altitude bands, respectively. The results appear promising; however, investigation of the classifier model show the classification is based solely on the ALT AOI values. The middle and low

المنارة للاستشارات
www.manaraa.com

altitude ranges were classified using a PDT range of approximately 1.75 percent, deeming the classification of crew coordination index ratings of Fair versus Poor operationally ineffective.

**Table 13. Fair vs. Poor Coordination Classifier Confusion Matrix, High Altitude Band**

Confusion Matrix

| *classified as:* | *Fair Coordination* | *Poor Coordination* | |
|---|---|---|---|
| | 105 | 0 | *Fair Coordination* |
| | 24 | 84 | *Poor Coordination* |
| | | | Truth |

**Table 14. Fair vs. Poor Coordination Classifier Confusion Matrix, Middle Altitude Band**

Confusion Matrix

| *classified as:* | *Fair Coordination* | *Poor Coordination* | |
|---|---|---|---|
| | 105 | 0 | *Fair Coordination* |
| | 24 | 84 | *Poor Coordination* |
| | | | Truth |

**Table 15. Fair vs. Poor Coordination Classifier Confusion Matrix, Low Altitude band**

Confusion Matrix

| *classified as:* | *Fair Coordination* | *Poor Coordination* | |
|---|---|---|---|
| | 97 | 8 | *Fair Coordination* |
| | 24 | 93 | *Poor Coordination* |
| | | | Truth |

المنارة للاستشارات

www.manaraa.com

Additionally, the normative eye-scan behavior model difference data was classified across all crew coordination index ratings. The classifier results indicate relatively high accuracy and precision of the classifier in its ability to distinguish between all index ratings in the same dataset. Results for the aggregate classification across all crew coordination index ratings for each altitude band are shown below in Table 16, Table 17, and Table 18. Accuracy analysis of the classification across the high, middle, and low altitude bands indicated values of 88.29 percent, 90.99 percent, and 90.54 percent, respectively. Precision analysis of the classification across the high, middle and low altitude bands indicated values of 89.40 percent, 93.20 percent, and 92.90 percent, respectively. Results of the aggregate classification proved most promising, providing a comparison across the range of crew coordination. Successful classification adds improved validity to the significance of the findings discussed in Chapter 5, suggesting there is a significant difference in eye-scan behavior metrics across the crew coordination indices, particularly when the crew coordination index rating of Poor.

**Table 16. All Crew Coordination Index Ratings Classifier Confusion Matrix, High Altitude Band**

Confusion Matrix

| *classified as:* | *Excellent Coordination* | *Good Coordination* | *Fair Coordination* | *Poor Coordination* | |
|---|---|---|---|---|---|
| | 108 | 0 | 0 | 0 | *Excellent Coordination* |
| | 2 | 94 | 4 | 12 | *Good Coordination* |
| | 0 | 0 | 89 | 23 | *Fair Coordination* |
| | 0 | 0 | 11 | 101 | *Poor Coordination* |
| | | | | | Truth |

www.manaraa.com

### Table 17. All Crew Coordination Index Ratings Classifier Confusion Matrix, Middle Altitude Band

Confusion Matrix

| *classified as:* | *Excellent Coordination* | *Good Coordination* | *Fair Coordination* | *Poor Coordination* | |
|---|---|---|---|---|---|
| | 108 | 0 | 0 | 0 | *Excellent Coordination* |
| | 2 | 96 | 14 | 0 | *Good Coordination* |
| | 0 | 0 | 112 | 0 | *Fair Coordination* |
| | 0 | 0 | 24 | 88 | *Poor Coordination* |
| | | | | | Truth |

### Table 18. All Crew Coordination Index Ratings Classifier Confusion Matrix, Low Altitude Band

Confusion Matrix

| *classified as:* | *Excellent Coordination* | *Good Coordination* | *Fair Coordination* | *Poor Coordination* | |
|---|---|---|---|---|---|
| | 108 | 0 | 0 | 0 | *Excellent Coordination* |
| | 2 | 94 | 16 | 0 | *Good Coordination* |
| | 0 | 0 | 112 | 0 | *Fair Coordination* |
| | 0 | 0 | 24 | 88 | *Poor Coordination* |
| | | | | | Truth |

Classifier results were most effective for the middle altitude band. Future research involving enhanced eye-tracking accuracy is critical to developing a stronger

www.manaraa.com

classification algorithm. Increasing the precision to determine attentional variations across finite regions of a display will also introduce the capability to incorporate more traditional global measures, such as fixation metrics, that have proven to correlate effectively with workload (Ellis & Schnell, 2009).

www.manaraa.com

# CHAPTER 7.

# DISCUSSION AND CONCLUSION

Crew coordination in the context of aviation is a specifically choreographed set of tasks performed by each pilot, defined for each phase of flight. Based on the constructs of CRM task load balancing and SOPs for each phase of flight, a shared understanding of crew workload and task responsibility is considered representative of well-coordinated crews. Nominal behavior is defined by SOPs and CRM theory, detectable through pilot eye-scan. This research effort investigates the relationship between the eye-scan exhibited by each pilot and the level of coordination between crewmembers.

This research presents three hypotheses that evaluate the relationship between crew coordination and the eye-scan of the PF and PM. The first hypothesis addresses crew coordination theory. Crew coordination, based on the constructs of aviation SOPs and CRM, is defined as adherence to trained responsibilities and effective balancing of task load between the PF and PM. Therefore, crew coordination was evaluated based on each pilot's understanding of the other crewmember's workload. By contrasting each pilot's workload-understanding, crew coordination was measured as the summed absolute difference of each pilot's understanding of the other crewmember's reported workload, resulting in a crew coordination index. The crew coordination index rates crew coordination on a scale ranging across Excellent, Good, Fair and Poor.

The second hypothesis focused on crew eye-scan behavior. The second hypothesis states it is possible to utilize eye-tracking of the PF and PM to identify eye-scan behaviors that are indicative of pilot tasking and well coordinated crews. Research has shown that pilot eye-scan behavior is successful in identifying common behaviors relative to tasking specific to each phase of flight (Ellis & Schnell, 2009). The eye-scan of each pilot exhibits a normative behavior when adhering to assigned tasks and high levels of coordination. Normative models were successfully developed using PF and PM

www.manaraa.com

eye-scan PDT data when crews reported Excellent coordination during the baseline scenarios.

The third hypothesis states that variations in attention allocation metrics correlate with a reduction in crew coordination. Several measures of crew eye-scan behavior were developed, including attention deviation from the normative eye-can behavior model, AOI PDT difference between the PF and PM, and shared attention of common AOIs. Differences in crew eye-scan behaviors were evaluated against crew coordination index ratings. A significant correlation was observed between all eye-scan behavior metrics and reduced crew coordination.

The eye-scan behavior of the PF was found to be the primary indicator of poor coordination between crewmembers. A decrease in the visual attention given to the IP and PFD was identified when the crew coordination index fell below a rating of Fair. Additionally, findings suggest a decrease in the visual attention given to the IP and PFD leads to increased cross-checking by the PM at the expense of primary tasking. Phase of flight was also found to be a significant factor affecting eye-scan behavior. Eye-scan behaviors of the PF and PM shifted in response to task load changes associated with specific phases of flight. Reduced crew coordination was identified across all phases of flight by decreased attention to primary AOIs and deviations from normative behavior..

To expand on the correlation between crew coordination and eye-scan behavior, eye-tracking data was processed through a classification algorithm in the attempt to characterize crew coordination index ratings using pilot eye-scan behavior. The classifier was successful in classifying the level of crew coordination using PF and PM eye-scan behavior metrics. Successful classification enables the use of eye-tracking data from each pilot to characterize crew coordination quantitatively in post-hoc analyses and in real-time. The classifier was able to characterize good versus bad crew coordination with a high level of success. Classification of all crew coordination index ratings was successful across all altitude bands, with an average classification accuracy of 89.94 percent.

المنارة للاستشارات

www.manaraa.com

In conclusion, eye-scan behavior metrics can reliably identify a reduction in crew coordination. Additionally, crew coordination was successfully characterized by eye-scan behavior data using machine learning classification methods. Identifying eye-scan behaviors on the flight deck indicative of reduced crew coordination can be used to inform training programs and design enhanced avionics that improve the overall coordination between the crewmembers and the flight deck interface. Ultimately, the ability to characterize crew coordination can be used to develop methods to increase shared situation awareness and crew coordination to reduce operational and flight technical errors. The ability to reduce operational and flight technical errors made by pilot crews improves the safety of aviation.

المنارة للاستشارات
www.manaraa.com

# CHAPTER 8.

# FUTURE RESEARCH

Research has yet to be conducted to evaluate the relationship between eye-scan behavioral indicators of reduced coordination and the impact on crew performance. In order to investigate the relationship between eye-scan behavior and crew performance, the crew coordination index used to evaluate eye-scan behavioral indicators must be validated. Validation of the crew coordination index is possible by evaluating the correlation between crew coordination index and crew performance determined by operational and flight technical errors. Development of scenarios that induce crew operational and flight technical error must be researched further to validate the crew coordination index.

The ability to characterize crew coordination and its impact on crew performance opens many research opportunities in the field of aviation safety. Optimization of pilot training, quantitative comparison analysis of flight deck configurations, and real-time feedback of pilot and crew state represent several applications that may benefit from characterization of crew coordination. Additionally, there is a need for research to close the feedback loop in the man-machine interface of the flight deck system. Piloting aircraft in today’s airspace requires pilots to manage an increased number of monitoring tasks, and humans have proven to be poor monitors over extended periods of time. A means to actively monitor crew state is important to provide information to the avionics to drive pilot attention to the correct areas when necessary. Monitoring both crew state and aircraft state information enables the feedback loop between the flight deck avionics and the pilot to be closed. Closing the feedback loop of the flight deck system enables research to potentially augment the avionics to react to deficiencies in pilot attention to prevent unsafe flight conditions.

المنارة للاستشارات

www.manaraa.com

# WORKS CITED

AirbusDriver.net. (2014). *Airbus Flight Control Laws*. Retrieved April 16, 2014, from AirbusDriver.net: http://www.airbusdriver.net/airbus_fltlaws.htm

Altonen, A., Hyrskykari, A., & Raiha, K. (1998). 101 Spots, or how do users read menus? *Proceedings of CHI 98 Human Factors in Computing Systems* (pp. 132-139). ACM Press.

AOPA. (2010). *2010 NALL Report, The Joseph T. Nall Report of Accident Trends and Factors*. Retrieved May 26, 2014, from AOPA: http://www.aopa.org/-/media/Files/AOPA/Home/News/All%20News/2012/April/VFR%20in%20to%20IMC%20Learn%20to%20escape%20the%20odds/10nall.pdf

Applied Science Laboratories. (2007). Eyenal (Eye-Analysis) Software Manual, Windows application Version 2.96, Manual Version 1.41. Applied Science Laboratories.

Australian Transport Safety Bureau. (2007). *An overview of spatial disorientation as a factor in aviation accidents and incidents.* Retrieved April 17, 2014, from Skybrary Aero: http://www.skybrary.aero/bookshelf/books/1124.pdf

Backs, R., & Walrath, L. (1992). Eye movement and pupillary response indices of mental workload during visual search of symbolic displays. *Applied Ergonomics* , 243-254.

Barker, J. M., Clothier, C. C., Woody, J. R., Mckinney, E. H., & Brown, J. L. (1996). Crew Resource Management: a simulator study comparing fixed versus formed aircrews. *Aviation Space and Environmental Medicine* , 67 (1):3-7.

Bonner, M. A., & Wilson, G. F. (2002). Heart Rate Measures of Flight Test and Evaluation. *The International Journal of Aviation Psychology* , 63-77.

Bower, G. H. (1981). Mood and Memory. *American Psychologist* , 129-148.

Boyce, C., Ross, A., Monaco, M., Hornak, L., & Xin, L. (2006). Multispectral Iris Analysis: A Preliminary Study. *Computer Vision and Pattern Recognition Workshop on Biometrics*, (p. 51). New York.

Brookings, J. B., Wilson, G. F., & Swain, C. R. (1996). *Psychophysiological responses to changes in workload during simulated air traffic control.* Dayton, OH: Biological Psychology.

Callan, D. (1998). Eye movement relationships to excessive performance error in aviation. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (pp. 1132-1136). Santa Moica, CA: Human Factors and Ergonomics Society.

www.manaraa.com

Card, S. (1984). Visual Search of computer command menus. In H. Bouma, & D. Bouwhuis, *Attention and Performance X, Control of Language Processes.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Crosby, M., & Peterson, W. (1991). Using eye movements to clasify search strategies. *Proceedings of the Human Factors Society 35th Annual Meeting* (pp. 1476-1480). Santa Monica: Human Factors and Ergonomics Society.

Di Nocera, F., Terenzi, M., & Camilli, M. (2006). Another look at scanpath: distance to nearest neighbor as a measure of mental workload. In D. de Waard, K. Brookhuis, & A. Toffetti, *Developments in Human Factors in Transportation, Design, and Evaluation* (pp. 1-9). Maastricht, the Netherlands: Shaker Publishing.

Egyptian Ministry of Civil Aviation. (2004, January 3). *Final Report of the Accident Investigation - Flash Airlines Flight 604.* Retrieved April 16, 2014, from Skybrary Aero: http://www.skybrary.aero/bookshelf/books/611.pdf

Ellis, K. K., & Schnell, T. (2009). *Eye Tracking Metrics for Workload Estimation in Flight Deck Operations.* Iowa City: University of Iowa.

Ellis, K. K., Kramer, L. J., Shelton, K. J., Arthur, J. T., Prinzel, L. J., & Norman, R. M. (2011). Transition of Attention in Terminal Area NextGen Operations Using Synthetic Vision Systems . *Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 55*, pp. 46-50. Las Vegas: Sage Publications.

Endsley, M. R., & Kaber, D. B. (1999). Level of Automation Effects on Performance, Situation Awareness, and Workload in a Dynamic Control Task. *Ergonomics* , 462-492.

Eye and Eyesight. (n.d.). *Anatomy of Human Eye*. Retrieved May 5, 2014, from Eyes and Eyesight: http://www.eyesandeyesight.com/2009/02/anatomy-of-the-eye/

Federal Aviation Administration. (2004). *AC120-51E, Crew Resource Management Training.* Department of Transportation, FAA. FAA.

Federal Aviation Administration. (2003). *AC120-71A.* Department of Transportation, Federal Aviation Administration. Federal Aviation Administration.

Federal Aviation Administration. (2012). *Instrument Flying Handbook (FAA--H--8083--15B).* Retrieved April 16, 2014, from FAA Regulations: http://www.faa.gov/regulations_policies/handbooks_manuals/aviation/media/FAA-H-8083-15B.pdf

Federal Aviation Administration. (2004, January 22). *www.faa.gov.* Retrieved from AC 120-51E - Crew Resource Management Training: http://www.faa.gov/regulations_policies/advisory_circulars/index.cfm/go/document.information/documentID/22879

المنارة للاستشارات

www.manaraa.com

Fischer, U., Orasanu, J., & Wich, M. (1995). Expert Pilots' Perceptions of Problem Situations. *Proceedings of the 8th International Symposium on Aviation Psychology* (pp. 777-872). Columbus, OH: Aviation Psychology.

Fitts, P. M., Jones, R. E., & Milton, J. L. (1950). Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review* , 24-29.

Frijda, N. H. (1987). *The Emotions: Studies in Emotion and Social Interaction.* New york: Cambridge University Press.

Gilland, J. (2008). *Driving, Eye Tracking and Visual Entropy: Exploration of Age and Task Effects.* University of South Dakota.

Goldberg, J., & Kotval, X. (1998). Eye movement-based evaluation of the computer interface. In S. Kumar, *Advances in Ocupational Ergonomics and Safety* (pp. 529-532). Amsterdam: ISO Press.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology* , 457-461.

Hancock, P. A., & Warm, J. S. (1989). A Dynamic Model of Stress and Sustained Attention. *Human Factors* , 519-537.

Hankins, T., & Wilson, G. (1998). A comparison of heart rate, eye activity, EEG and subjective measures of pilot mental workload during flight. *Aviation, space and Environmental Medicine* , 360-367.

Harris, R. L., Glover, B. J., & Spady, A. A. (1986). *Analytical Techniques of Pilot Scanning Behavior and Their Application.* Hampton, VA: NASA Technical Paper 2525.

Helmreich, R. L. (1984). Cockpit Management Attitudes. *Human Factors* , 583-589.

Helmreich, R. L. (1982). Pilot Selection and Training. *American Psychological Association.* Washington DC.

Hendrickson, J. (1989). Performance, preference, and visual scan patterns on a menu-based system: implications for interface design. *Proceedings of the ACM CHI '89 Human Factors in Computing Systems Conference* (pp. 217-222). ACM Press.

Hilburn, B. (2004). Measuring head-down time via area-of interest analysis: operational and experimental data. In D. De Waard, K. Brookhui, R. Van Egmond, & T. Boersma, *Human Factors in Design, Safety, and Management* (pp. 427-436). Maastricht, the Netherlands: Shaker Publishing.

Hilburn, B., & Jorna, P. G. (2001). Worklaod and Air Traffic Control. In P. A. Hancock, & P. A. Desmond, *Stress, Workload, and Fatigue.* Mahwah, NJ: Erlbaum.

المنارة للاستشارات
www.manaraa.com

Hogarth, R. M., & Einhorn, H. J. (1992). Order Effects in Belief Updating: The Belief-adjustment Model. *Cognitive Psychology* , 1-55.

Interstate Aviation Committee. (2009). *Final Report B 737-505 VP-BKO.* Retrieved April 16, 2014, from AAIB UK: http://www.aaib.gov.uk/cms_resources.cfm?file=/VP-BKO_Report_en.pdf

Itoh, Y., Hayashi, Y., Tsukui, I., & Saito, S. (1990). The ergonomic evaluation of eye movements and mental workload in aircraft pilots. *Ergonomics* , 719-733.

Jacob, R. J., & Karn, K. S. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. *The Mind's eye: Cognitive The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research* , 573 - 603.

Just, M. A., & Carpenter, P. A. (1993). The Intensity Dimension of Thought: Pupillometric Indices of Sentence Processing. *Canadian journal of Experimental Psychology* , 310-339.

Katz, L. C., Kambe, G., Kline, K. F., & Grubb, G. N. (2006). *Nonverbal Communication and Aircrew Coordination in Army Aviation: Annotated Bibliography.* Arlington, VA: United States Army Research Institute for the Behavioral and Social Sciences.

Kotval, X., & Goldberg, J. (1998). Eye movements and interface components grouping: an evaluation method. *Proceedings of the 42nd Annual Meeting of the Human Factors and Ergonomics Society* (pp. 486-490). Santa Monica: Human Factors and Ergonomics Society.

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2004). Focused attention on one contextual attribute does not reduce source memory for a different attribute. *Memory* , 183-192.

McGrath, J. E. (1976). Stress and Behavior in Organizations. In M. D. Dunnette, *Handbook of industrial and organizational psychology* (pp. 1951-1395). Chicago: Rand McNally.

Neiss, R. (1988). Reconceptualizing Arousal: Psychological states in motor performance. *Psychological Bulletin* , 345-366.

Neuberg, S. L., & Newsome, J. T. (1993). Personal Need For Structure: Individual Differences in the Desire for Simple Structure. *Journal of Personality and Social Psychology* , 113-131.

Norman, R. M. (2010, January). Normative Visual Behavior In DataComm Operations. (K. K. Ellis, Interviewer)

Norman, R. M., Baxley, B. T., Adams, C. A., Ellis, K. K., Latorella, K. A., & Comstock, J. R. (2013). *Flight Crew Workload, Acceptability, and Performance When Using Data*

www.manaraa.com

*Comm in a High-Density Terminal Area Simulation.* NASA, Crew Systems and Aviation Operations. Hampton, VA: NASA.

Norman, R. M., Baxley, B. T., Ellis, K. K., Adams, C. A., Latorella, K. A., & Comstock, J. R. (2010). *NASA/FAA Data Comm Airside Human-In-The-Loop Simulation.* Washington DC: Federal Aviation Administration.

Noton, D., & Stark, L. (1971). Eye movements and visual perception (Eye movements and visual perception, describing scan path for memory traces). *Scientific American* , 35-43.

Panksepp, J. (1996). Affective Neuroscience: A Paradigm to Study the Animate Circuits for Human Emotions. In R. D. Kavanaugh, B. Zimmerberg, & S. Fein, *Emotions: An Interdisciplinary Approach* (pp. 22-60). Mahwah, NJ: Erlbaum.

Pope, A. T., & Bogart, E. H. (1992). Identification of Hazardous Awareness States in Monitoring Environments. *SAE, International Conference on Environmental Systems.* Seattle: NTRS.

Rao, R. P., Zelinsky, G. J., Hayhoe, M. M., & Ballard, D. H. (1997). *Eye Movements in Visual Cognition A Computational Study.* National Resource Laboratory for the Study of Brain Behavior, University of rochester.

Rayner, K., & Bertera, J. H. (1979). Reading without a Fovea. In *Science* (pp. 468-469). American Association for the Advancement of Society.

Razmjou, S. (1996). Mental Workload in Heat: Toward a Framework for Analyses of Stress States. *Aviation, Space, & Environmental Medicine* , 530-538.

Roscoe, A. H. (1987). *In-Flight Assessment of Workload Using Pilot Ratings and Heart Rate.* England: Britannia Airways Ltd Luton.

Schutte, P. C., & Trujillo, A. C. (1996). Flight Crew Task Management in Non-Normal Situations. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting.* Human Factors and Ergonomics Society.

Simon, P., Rousseau, F., & Angue, J.-C. (1993). *Quantitative Analysis of Mental Workload Influence on Eye Scanning Movements.* Le Touquet, France: Systems, Man and Cybernetics.

Smart Eye Inc. (2011). *Smart Eye Market Segments for Eye Tracking*. Retrieved February 6, 2012, from http://smarteye.se/solutions/

Staal, M. A. (2004). *Stress, Cognition, and Human Performance: A Literature Review and Conceptual Framework.* Moffett Field, CA: NASA.

المنارة للاستشارات

www.manaraa.com

Staal, M. A. (2004). *Stress, Cognition, and Human Performance: A Literature Review and Conceptual Framework.* Moffett Field, CA: NASA Ames Research Center.

Stout, R. J., Salas, E., & Carson, R. (1994). Individual Task Proficiency and Team Process Behavior: What's Important for Team Functioning. *Military Psychology* , 177-192.

Taylor, R. M. (1990). *Situational Awareness Rating Technique (SART): The Development of a Tool for Aircrew Systems Design.* Neuilly Sur Seine, France: AGARD.

Thornton, R. C., Kaempf, G. L., Zeller, J. L., & McAnulty, M. (1992). *An Evaluation of Crew Coordination and Performance During a Simulated UH-60 Helicopter Mission.* Fort Rucker, AL: Army Research Institute.

Van Orden, K. F., Limbert, W., Makeig, S., & Jung, T. -P. (2001). Eye Activity Correlates of Workload during a Visuospatial Memory Task. *Human Factors* , 111-121.

Wang, H., Lin, S., Liu, X., & Kang, S. B. (2005). Separating Reflections in Human Iris Images for Illumination Estimation. *Tenth IEEE International Conference on Computer Vision*, (pp. 1691-1698). Beijing, China.

Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman, & D. R. DAvies, *Varieties of Attention* (pp. 63-101). New York: Academic Press.

Wickens, C. D. (1991). Processing resources in attention. In D. L. Damos, *Multiple-task performance* (pp. 3-34). Washington DC: Taylor & Francis.

Wickens, C. D., Goh, J., Helleberg, J., Horrey, W., & Talleur, D. A. (2003). Attentional Models of Multitask Pilot Performance Using Advanced Display Technology. *Human Factors: The Journal of the Human Factors and Ergonomics Society* , 360-380.

Wickens, C. D., Helleberg, J., Goh, J., Xu, X., & Horrey, W. J. (2001). *Pilot task management: Testing an attentional expected value model of visual scanning.* Savoy, IL: University of Illinois. Aviation Research Lab.

Wickens, C. D., McCarley, J. S., & Thomas, L. (2003). Attention-situation Awareness (A-SA) model. *Human Performance Modeling Workshop Proceedings.* NASA Ames Research Center.

Wilson, G. F., Purvis, B., Skelly, J., Fullenkamp, P., & Davis, I. (1987). Physiological data used to measure pilot workload in actual flight and simulator conditions. *Proceedings from the 31st Annual Meeting of the Human Factors Socitey* (pp. 779-783). New York: Human Factors Society.

Wooding, D. S. (2002). *Fixation Maps: Quantifying Eye-Movement Traces.* New Orleans: ETRA.

المنارة للاستشارات

www.manaraa.com

Yamamoto, S., & Kuto, Y. (1992). A method of evaluating VDT Scren layout by eye movement analysis. *Ergonomics* , 591-606.

Yarbus, A. (1967). *Eye Movements and Vision.* Plenum Press.

المنارة للاستشارات
www.manaraa.com

# APPENDIX A: BIOGRAPHICAL QUESTIONNAIRE

Appendix A is an exact copy of the Biographical Questionnaire completed by the subject pilots.

This questionnaire requests the most up to date information about the Subject Pilot. This data may be used during data analysis; however, no personal information will be connected to any of the data recorded in this simulation.

Age ____________

Gender (please circle) MALE FEMALE

Commercial aircraft type / hours ____________ ____________

Military aircraft type / hours ____________ ____________

Total flight hours / total simulator hours ____________ ____________

Date of last flight (airline transport) ____________

Will you wear glasses during this experiment? YES NO

Have you had eye surgery? (Please describe your surgery below) YES NO

Do you have any known eye or eyelid abnormalities (astigmatism, etc)? (Describe) YES NO

Are your eyes corrected to different distances? (Describe) YES NO

Do you have experience using DataComm equipment and procedures? (Describe) YES NO

How often have you flown into and out of Boston Logan airport in the past five years? ______

المنارة للاستشارات
www.manaraa.com

# APPENDIX B: POST SCENARIO QUESTIONNAIRE

Appendix B contains all the questions in the Post Scenario Questionnaire completed by the subject pilots on a Tablet PC (personal computer) after the last training run, and after every data collection run.

**Appendix B Table of Contents**

## B.1 Workload during scenario by phase of flight

Using the chart below, read the descriptions that define a particular workload level during a particular phase of flight or during ground operations. Move vertically up the scale until you find a description that accurately portrays the level of workload based on the scenario you have just flown. Move to the right and read the choices. Below the chart, record the appropriate ratings associated with receiving messages on the specified phase of flight from 1 to 10, 1 being lowest and 10 being the highest workload. If the scenario is a departure there will only be one question to rate. (NOTE: the entire scale was visible to the subject pilot while answering the workload rating questions.)

www.manaraa.com

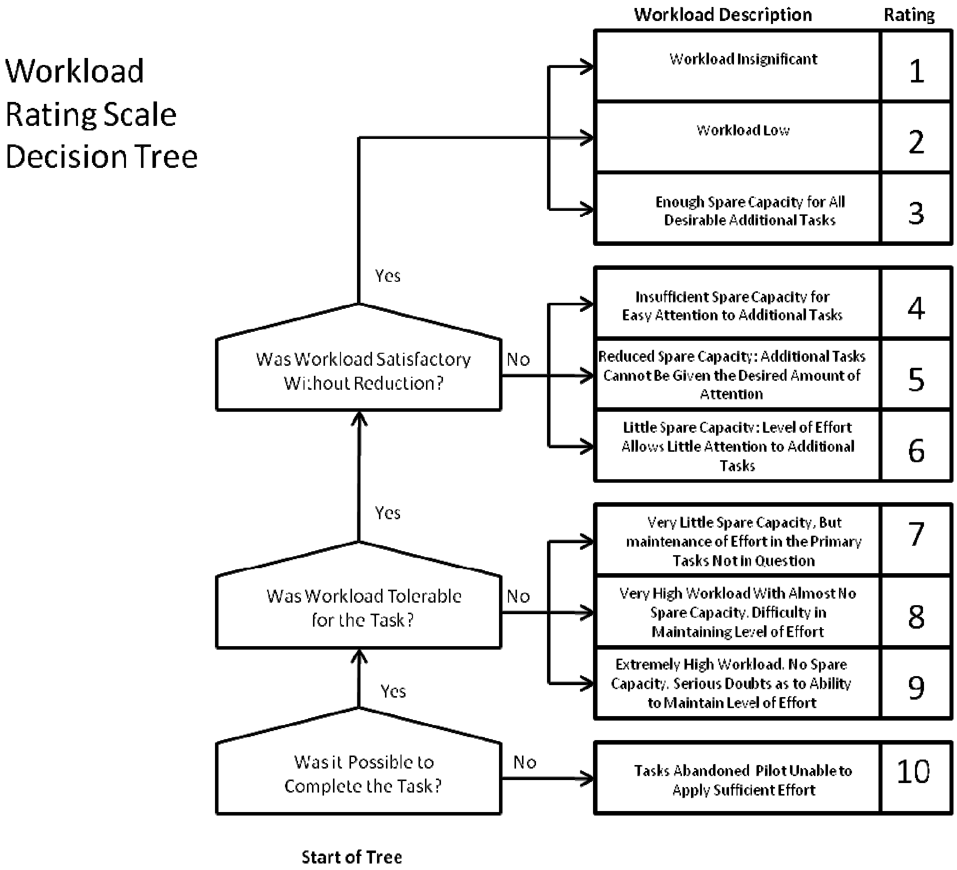


Figure B 1. Bedford Workload Scale

Workload by Phase of Flight

1) Your workload in flight ___________ (1-10)

2) Your workload during surface / taxi operations ___________ (1-10)

3) Your crewmember’s workload in flight ___________ (1-10)

4) Your crewmember’s workload during surface / taxi operations ___________ (1-10)

## B.2 Situation Awareness by phase of flight

Please answer the questions below with respect to the impact of Voice or Data

www.manaraa.com

Communications between the controller and pilot during the scenario. Select the rating that reflects your understanding of the dimensions described at the left for the appropriate phase of flight (all phases for the arrival scenarios, and surface operations only for the departure scenarios).

| | |
|---|---|
| DEMAND ON ATTENTIONAL RESOURCES:<br>Rate your overall impression of the scenario in terms of how much attention and effort was required to successfully perform the tasks. Items to consider include: the likelihood of the situation changing suddenly, the degree of complexity associated with this scenario; and the number of variables changing during the scenario. | (1) High Low (7)<br>2A) 1 2 3 4 5 6 7 during flight<br>2B) 1 2 3 4 5 6 7 surface ops |
| SUPPLY OF ATTENTIONAL RESOURCES:<br>Rate the degree of spare attention that you had available to perform tasks other than your primary task of piloting the aircraft was performed. Items to consider include: how much focus and concentration was necessary and how you divided your attention between the flying task and other tasks. High = plenty of spare capacity; Low = little spare capacity.. | (1) High Low (7)<br>2C) 1 2 3 4 5 6 7 during flight<br>2D) 1 2 3 4 5 6 7 surface ops |
| UNDERSTANDING OF THE SITUATION:<br>Rate your overall understanding of what was happening with the aircraft during this scenario. Items to consider include: the quantity of information received and understood; the quality of the information; and the familiarity you may have had with what was taking place during the scenario. | (1) High Low (7)<br>2E) 1 2 3 4 5 6 7 during flight<br>2F) 1 2 3 4 5 6 7 surface ops |

www.manaraa.com

## B.3 Sources of information

Please rate the following with “1” as Very Important, and “7” as Not Important, areas that contributed to your Situation Awareness given all available resources in the flight test scenario. Place an “X” by those areas that did not contribute to your SA.

1. Visual information on the Primary Flight Display
2. Visual perception on the NAV Display
3. Visual information on the charts
4. Visual information available out the window
5. Visual information on the CDU pages
6. Visual information that your crew member directed your attention to
7. Auditory information conveyed by ATC
8. Auditory information conveyed by your crew member
9. Your perception of your crew member’s actions

## B.4 Crew interaction

<table>
<tr><td>1) Your performance was proficient in this scenario.</td><td>1) Strongly Agree, (7) Strongly Disagree<br>|____|____|____|____|____|____|<br>1 7</td></tr>
<tr><td>2) My crewmember’s performance was proficient in this scenario.</td><td>1) Strongly Agree, (7) Strongly Disagree<br>|____|____|____|____|____|____|<br>1 7</td></tr>
<tr><td>3) You aware of operational plans, decisions, and had appropriate SA throughout the flight.</td><td>(1) Strongly Agree, (7) Strongly Disagree<br>|____|____|____|____|____|____|<br>1 7</td></tr>
<tr><td>4) The other pilot was aware of operational plans, decisions, and had appropriate SA throughout the flight.</td><td>(1) Strongly Agree, (7) Strongly Disagree<br>|____|____|____|____|____|____|<br>1 7</td></tr>
</table>

المنارة للاستشارات

www.manaraa.com

5) There was adequate communication.

(1) Strongly Agree, (7) Strongly Disagree
|____|____|____|____|____|____|
1 7

6) The Captain and FO maintained their roles throughout the scenario.

(1) Strongly Agree, (7) Strongly Disagree
|____|____|____|____|____|____|
1 7

المنارة للاستشارات
www.manaraa.com

## B.5 Acceptability of “Expected Taxi” and “Taxi” Clearances

1) Did the display of the OWNSHIP POSITION on the navigation display make the taxi clearance easier to understand and to carry out? [NA for runs without ownship displayed]

| | | | | | | |

| | | |
|---|---|---|
| instructions were easier to understand | sometimes easier to understand | did not make easier to understand |

2) Did the display of the ROUTE on the navigation display make the taxi clearance easier to understand and to carry out? [NA for runs without route displayed]

| | | | | | | |

| | | |
|---|---|---|
| instructions were easier to understand | sometimes easier to understand | did not make easier to understand |

3) Did you have confidence that the taxi route was accurately depicted based on the ATC instruction?

| | | | | | | |

| | | |
|---|---|---|
| confident the taxi route was accurate & followed the route | confident route accurate but verified the route | not confident taxi route displayed accurately |

4) Did you have a sufficient amount of time to respond to the Voice or transmitted messages?

| | | | | | | |

| | | |
|---|---|---|
| I had more than enough time to respond | just about the right amount of time | I did not have enough time to respond |

5) Was the amount of Head Down time required to receive and respond to just the “Expected Taxi” messages acceptable in this scenario?

| | | | | | | |

| | | |
|---|---|---|
| Minimal increase in Head Down time | Acceptable amount of Head Down time | Too much head down time |

6) Was the amount of heads-down time required to receive and respond to other non-time critical messages acceptable in this scenario? (e.g., frequency changes, new altimeter setting, etc)

| | | | | | | |

| | | |
|---|---|---|
| Minimal increase in Head Down time | Acceptable amount of Head Down time | Too much heads down time |

7) Overall, was the communication mode (Voice or ) for receiving Expected Taxi and Taxi clearances acceptable during this scenario? (Include consideration of message intrusiveness, amount of heads-down time required, effect of party line information, expected response and timing of the response, ease of use, etc.)

www.manaraa.com

| | | | | | |

Completely acceptable Neither unacceptable nor acceptable Completely unacceptable

8) How much operational risk was introduced by the communication mode (Voice or ) used during this scenario?

| | | | | | |

extremely low risk neither high or low risk extremely high risk

9) Was there a point at which you did not feel that the transmitted taxi instructions were accurate?

| | | | | | |

the message was accurate some aspects were inaccurate or in question I did not feel the message was accurate

المنارة للاستشارات

www.manaraa.com

# APPENDIX C:  EYE TRACKING APPARATUS

A ten-camera oculometer system was installed in the IFD to support unobtrusive collection of eye tracking and head position data for both flight crew subjects. To collect eye-tracking information from both crewmembers (PF and PM) a state-of-the-art eye tracking system was tested and integrated into the IFD. The Smart-Eye Inc. eye tracker used in this experiment is a remote eye tracking system that used facial recognition to calculate the position of defined points on a subjects head relative to the calibrated position of two or more cameras. The camera's use the facial features to locate the corners of each of the subject's eyes and digitally zoom to enhance the image of the eye.

To calculate eye gaze vectors from the head origin, infrared light emitting diodes project infrared light to illuminate the pilots face and to create two ocular reflections; a static corneal reflection and a pupil reflection that moves in conjunction with eye movements. Triangulating the angular difference between the corneal reflection and pupil reflection, the Smart-Eye eye tracking system creates a vector between the two points, which creates an eye gaze vector originating from the corneal reflection at the center of the pilot's eyes (Smart Eye Inc., 2011).

Ten cameras in total are utilized, with one eye tracking system for the PF and one for the PM, each with five cameras to capture the gaze vectors of both pilots simultaneously. To synchronize the systems, Smart-Eye Inc. created a modified eye tracking system network, tethering two systems together using a primary-secondary relationship. Each system is time stamped synchronously with global positioning system time so eye gaze vector data from both pilots can be accurately compared to scenario events.

In order to achieve robust eye tracking data over the span of coverage required for normal cockpit operations, the system had to be capable of covering +/-45 degrees of center, and +10 degrees from horizon and to the base of the CDU for each pilot. This

المنارة للاستشارات

www.manaraa.com

requirement had to be met while still maintaining a high level of simulator fidelity by making the cameras as inconspicuous as possible on the flight deck. Camera placement was optimized for coverage within constraints imposed by limited available real estate.

To test which available locations for installation on the flight deck provided the greatest coverage capability, a mockup of the IFD was created using 80/20™ aluminum. Test results concluded with five locations per side being chosen (mirrored locations between left and right seat) that yielded sufficient coverage to perform flight testing while remaining minimally obtrusive in the flight deck. System spatial accuracy was tested to be no greater than 2 degrees gaze angle for any calibration point on the display panels.

The oculometer provided the following raw data in real-time:

- Gaze vectors for each eye of both crew members (raw)
- Head and eye position (each eye) for each crew
- Eyelid closure distance for each eye for each crew
- Pupil size for each eye for each crew

المنارة للاستشارات

www.manaraa.com

# APPENDIX D: STATISTICAL ANALYSIS OF EYE TRACKING

All boxplots and graphs depict PDT for the PF and PM across all available AOIs for each visual behavior metric. CCwl(In Flight)RATING is representative of Crew Coordination Index, with ratings 1, 2, 3, and 4 representing Excellent, Good, Fair, and Poor coordination, respectively.

## D1: Altitude Band Analysis Boxplots – All AOIs

### D1.1 Normative Model Difference

*D1.1.1 Pilot Flying*

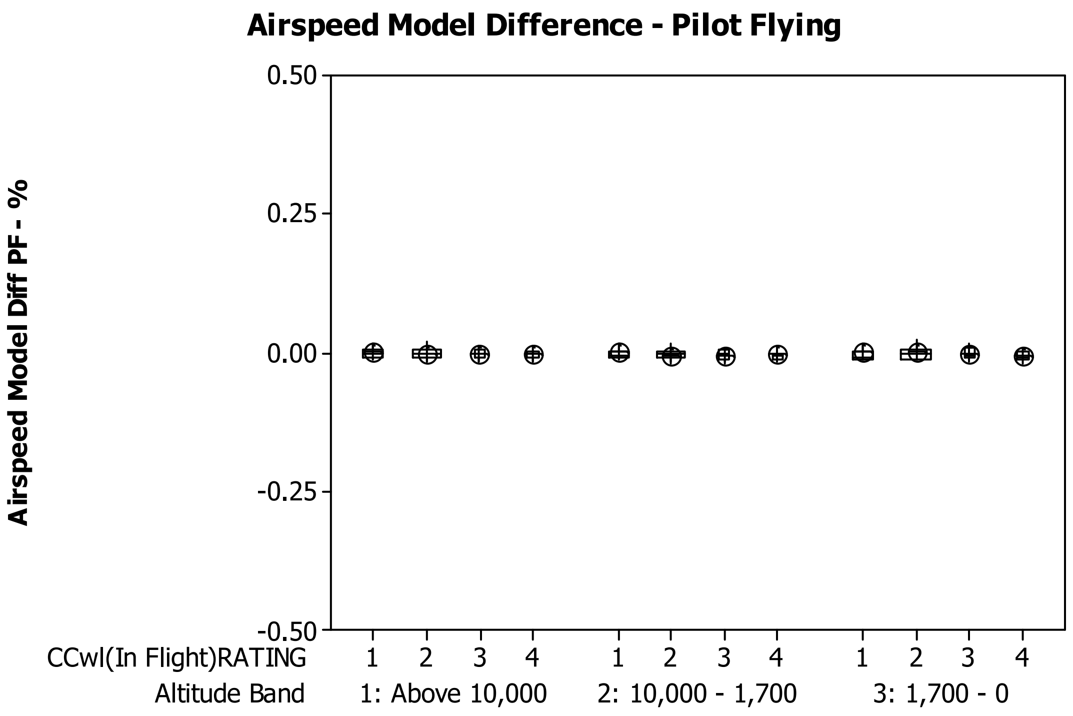


**Figure D 1. Airspeed Model Difference - PF**

www.manaraa.com

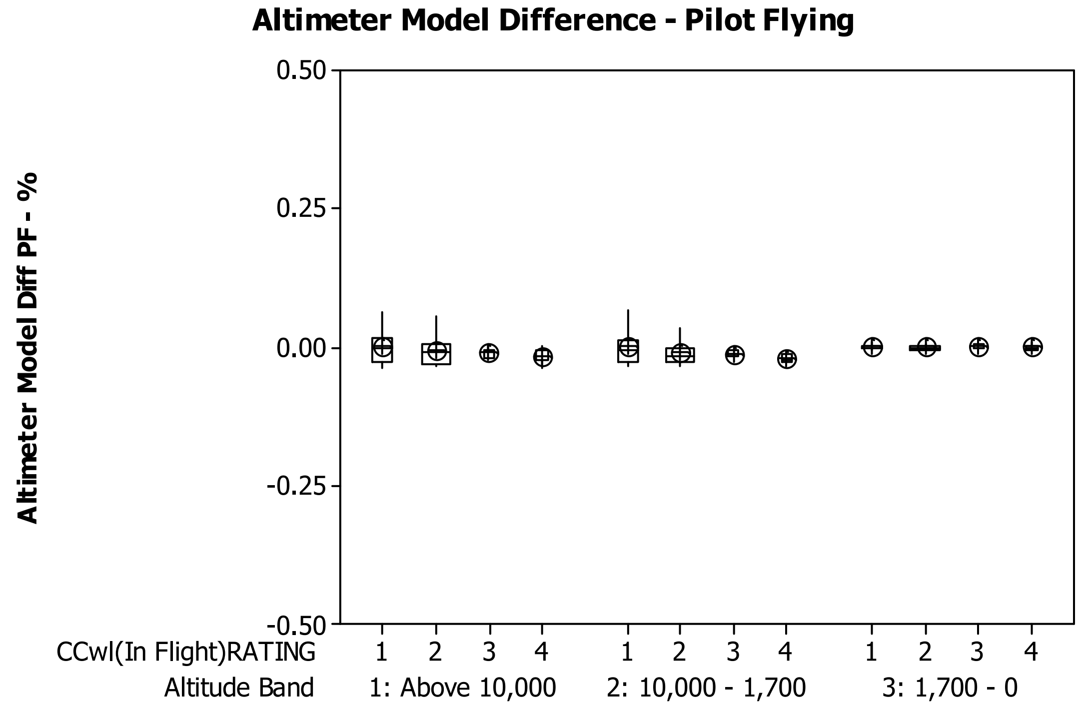


Figure D 2. Altimeter Model Difference - PF

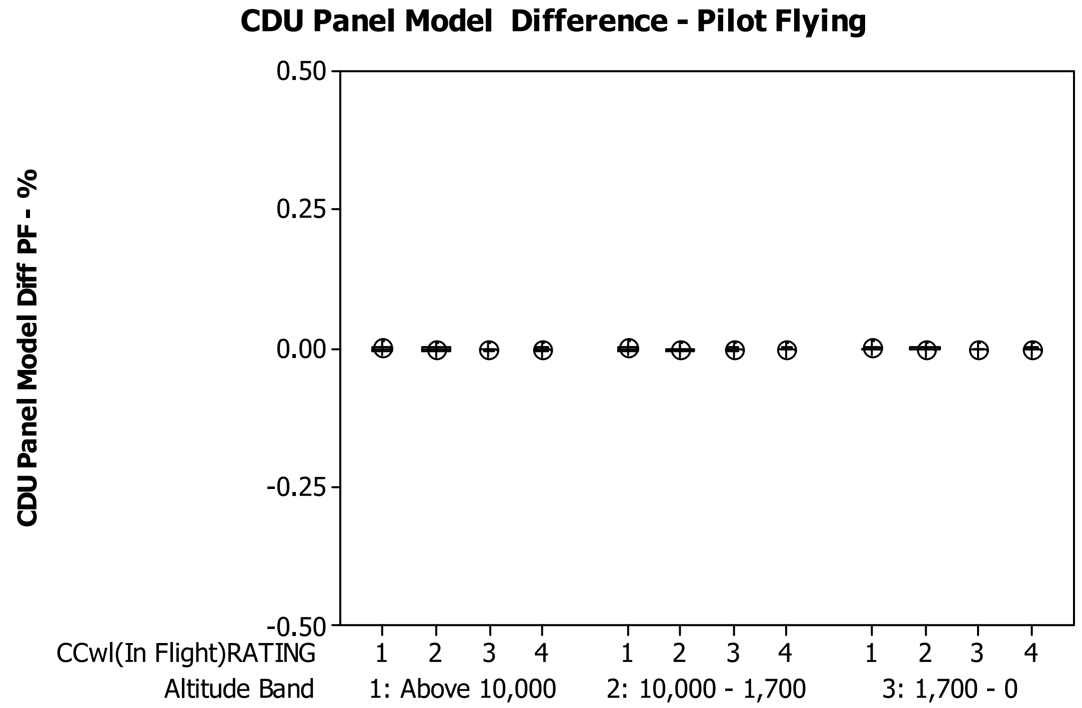


Figure D 3. CDU Panel Model Difference - PF

المنارة للاستشارات

www.manaraa.com

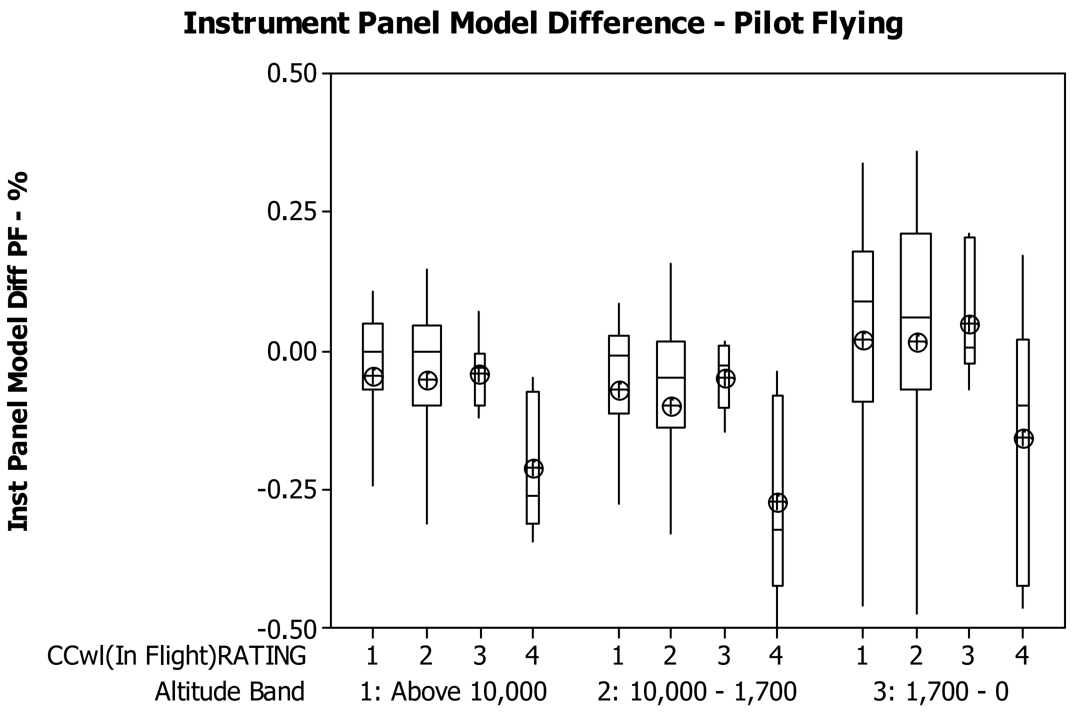


Figure D 4. Instrument Panel Model Difference - PF

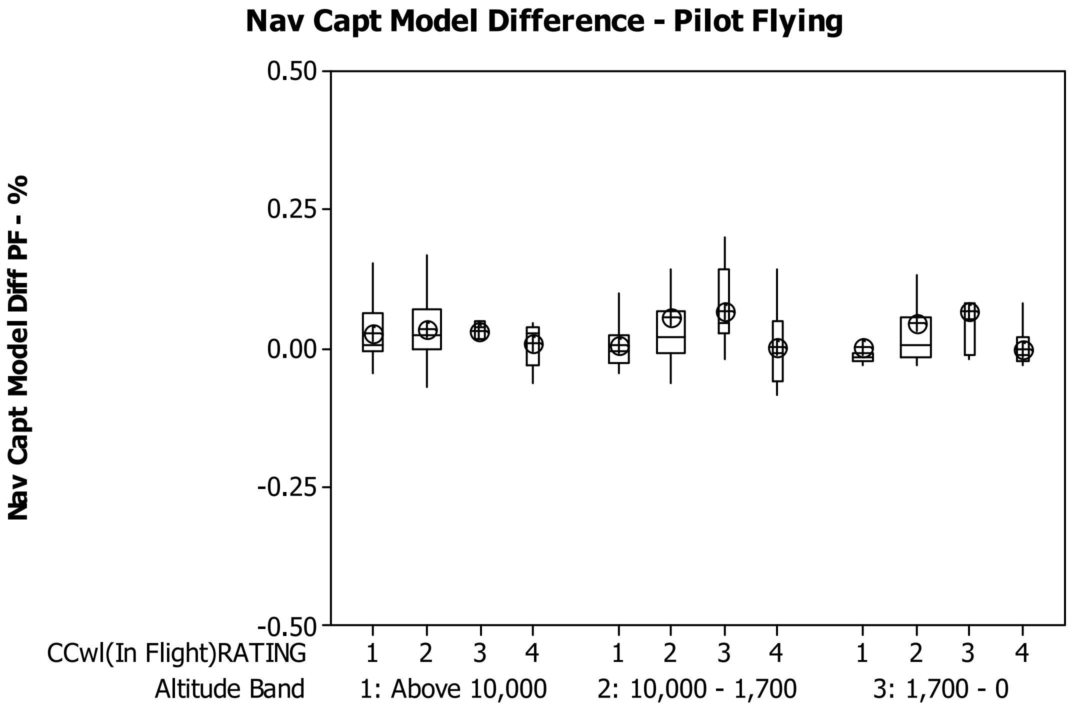


Figure D 5. Navigation Display Model Difference - PF

المنارة للاستشارات

www.manaraa.com

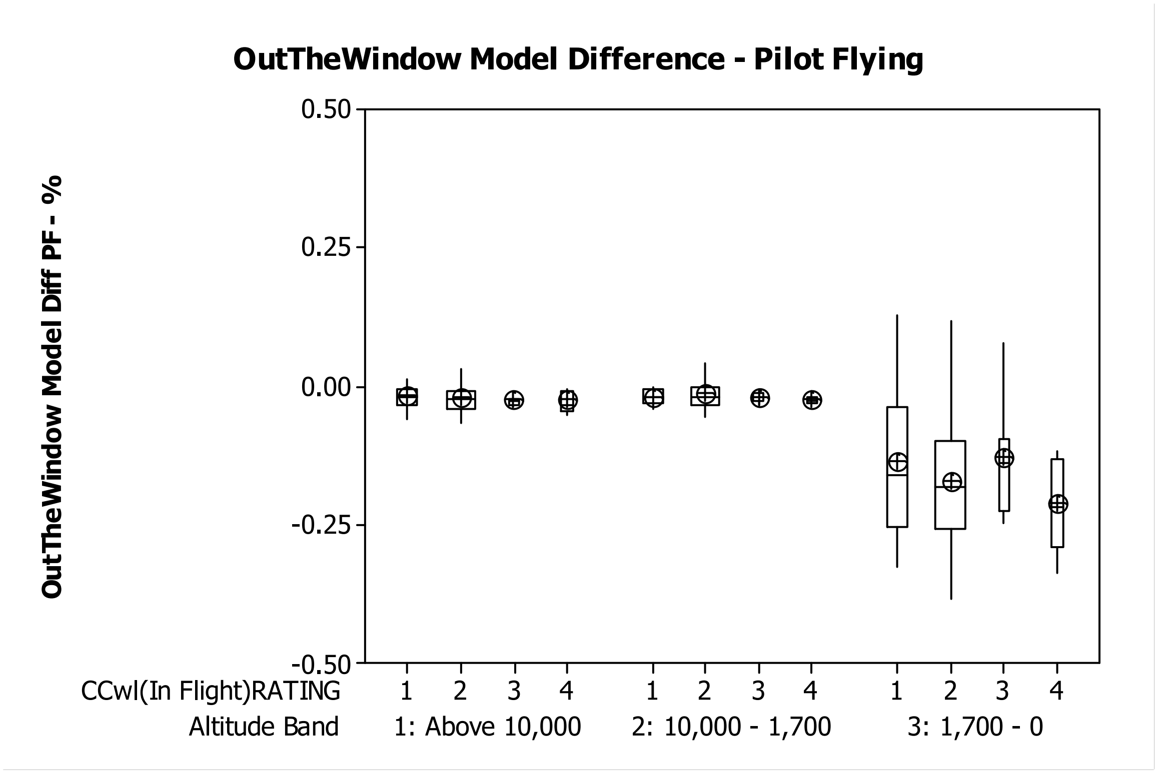


Figure D 6. Out the Window Model Difference - PF

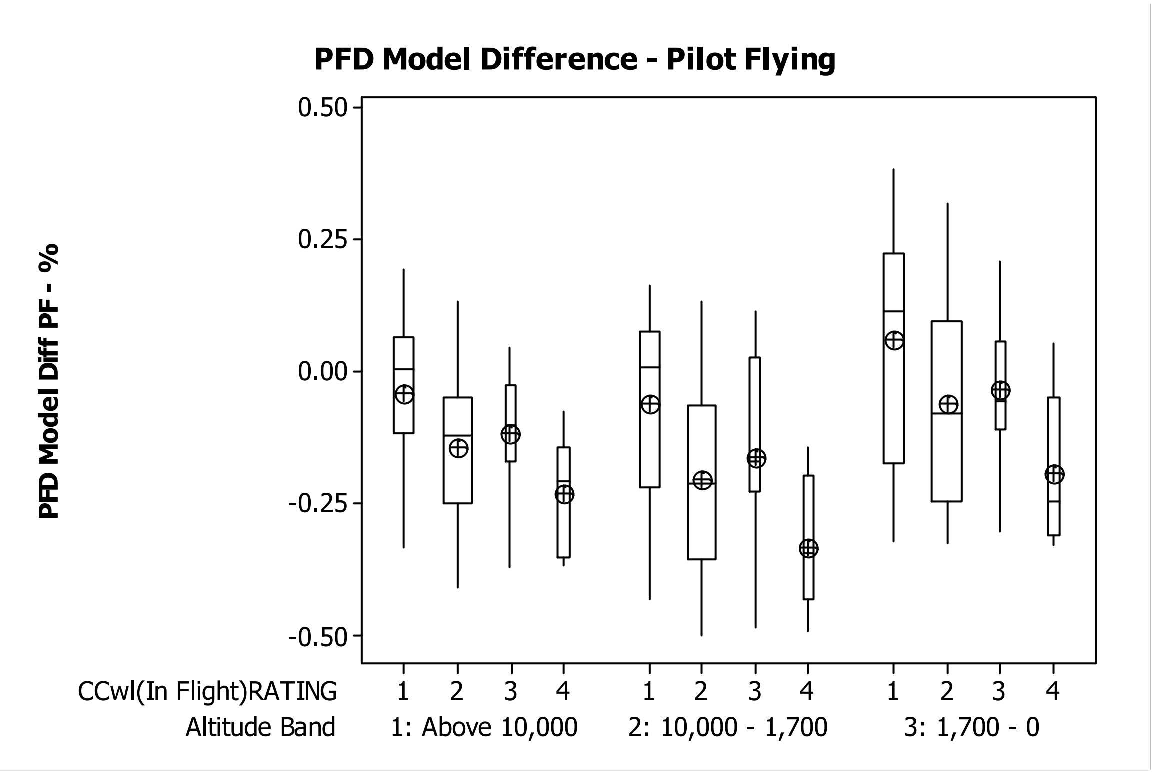


Figure D 7. PFD Model Difference - PF

www.manaraa.com

### *D1.1.2 Pilot Monitoring*

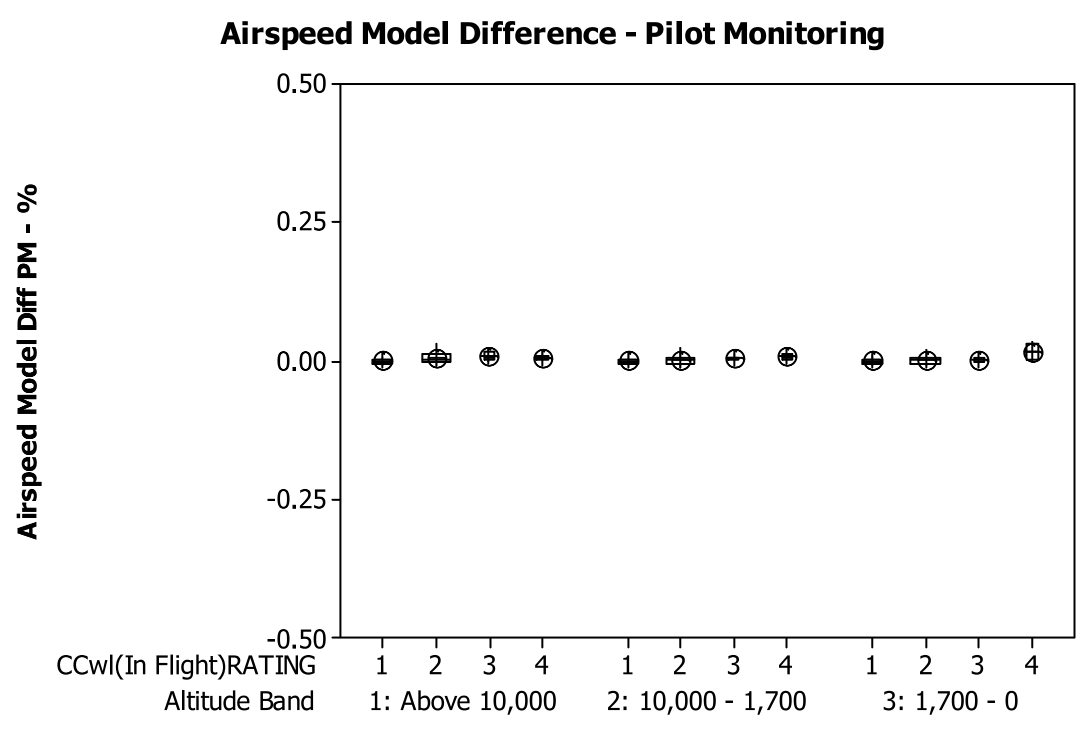


**Figure D 8. Airspeed Indicator Model Difference - PM**

www.manaraa.com

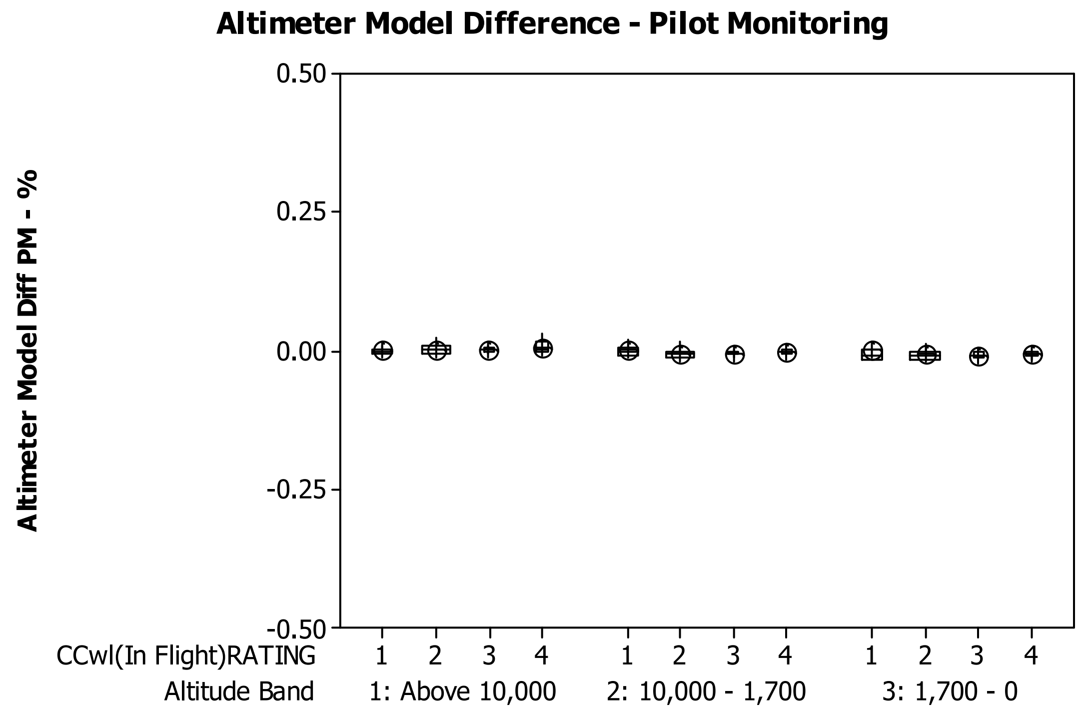


Figure D 9. Altimeter Model Difference - PM

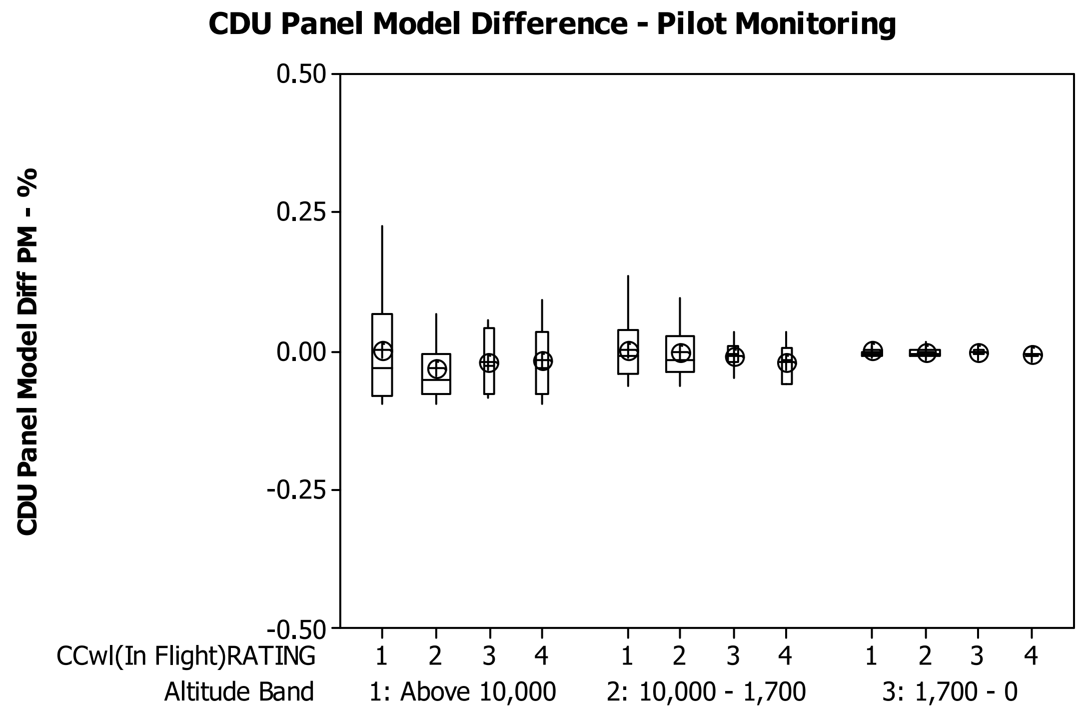


Figure D 10. CDU Panel Model Difference - PM

www.manaraa.com

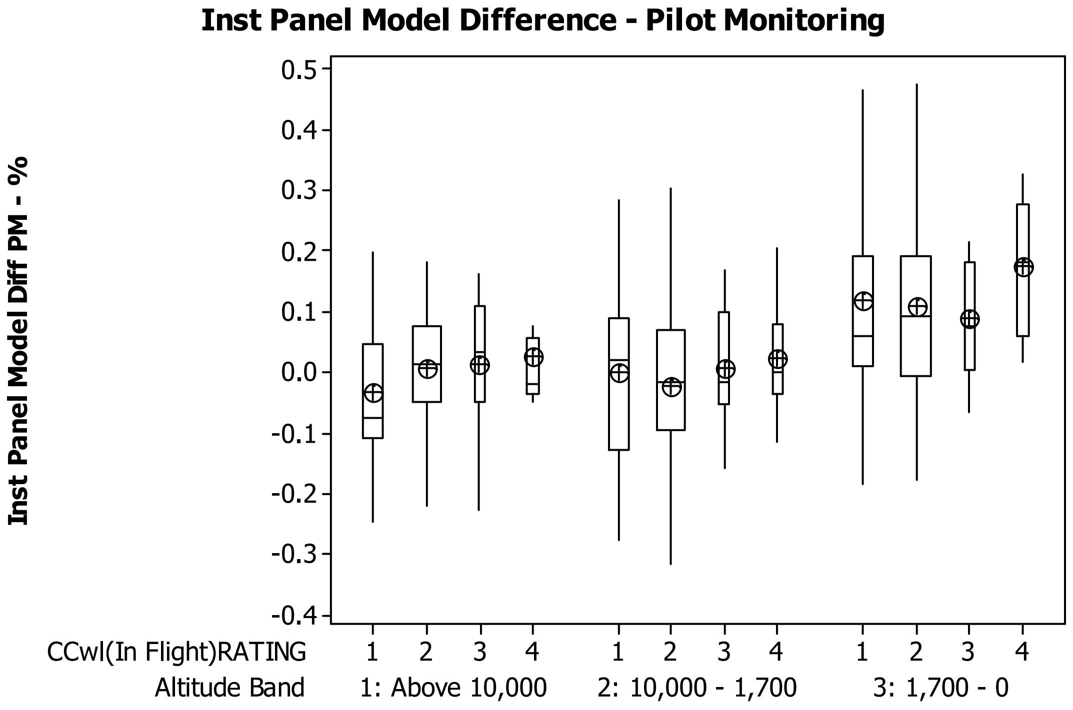


Figure D 11. Instrument Panel Model Difference - PM

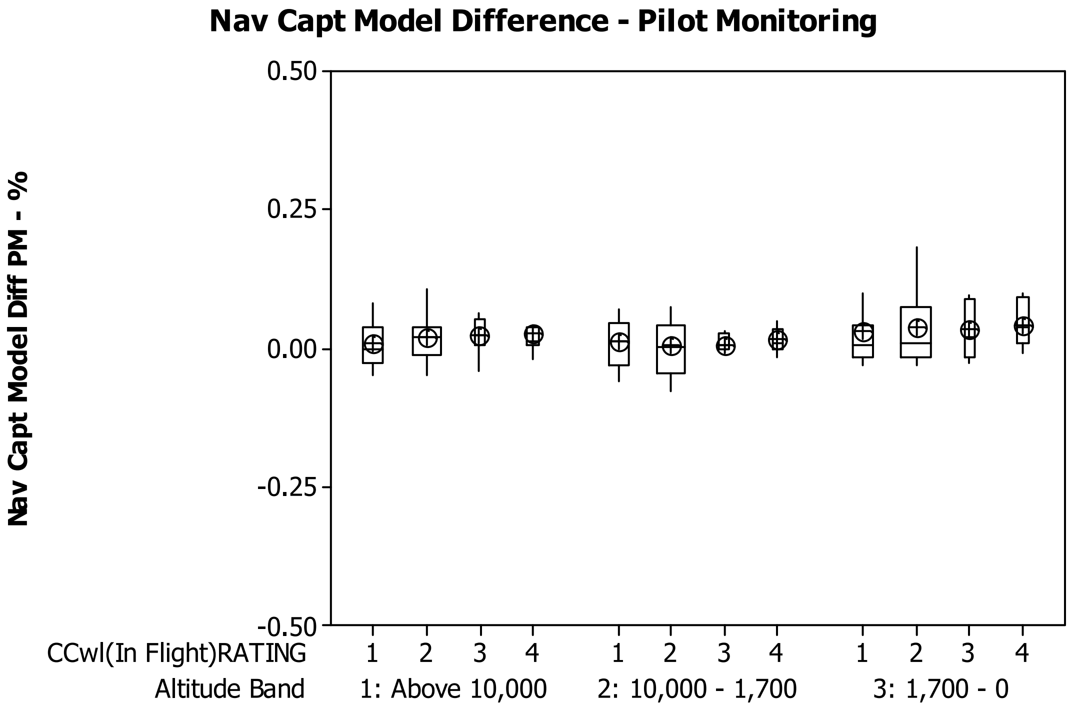


Figure D 12. Navigation Display Model Difference - PM

www.manaraa.com

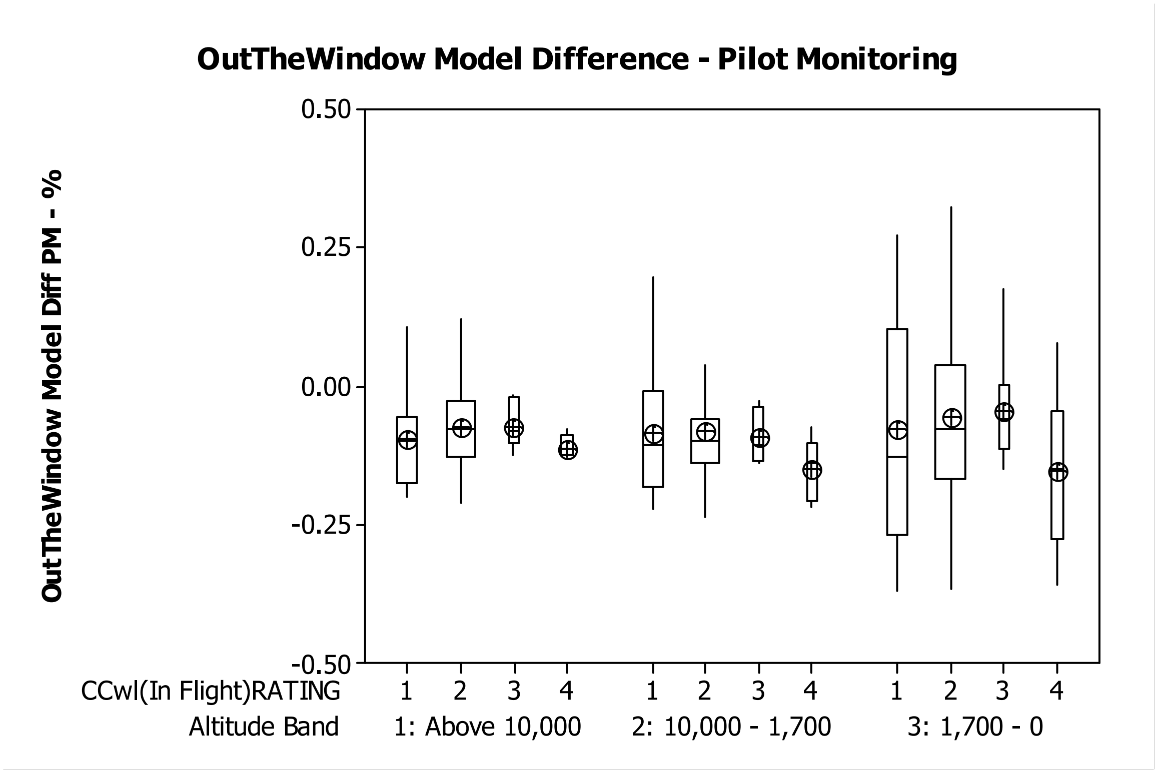


Figure D 13. Out the Window Model Difference - PM

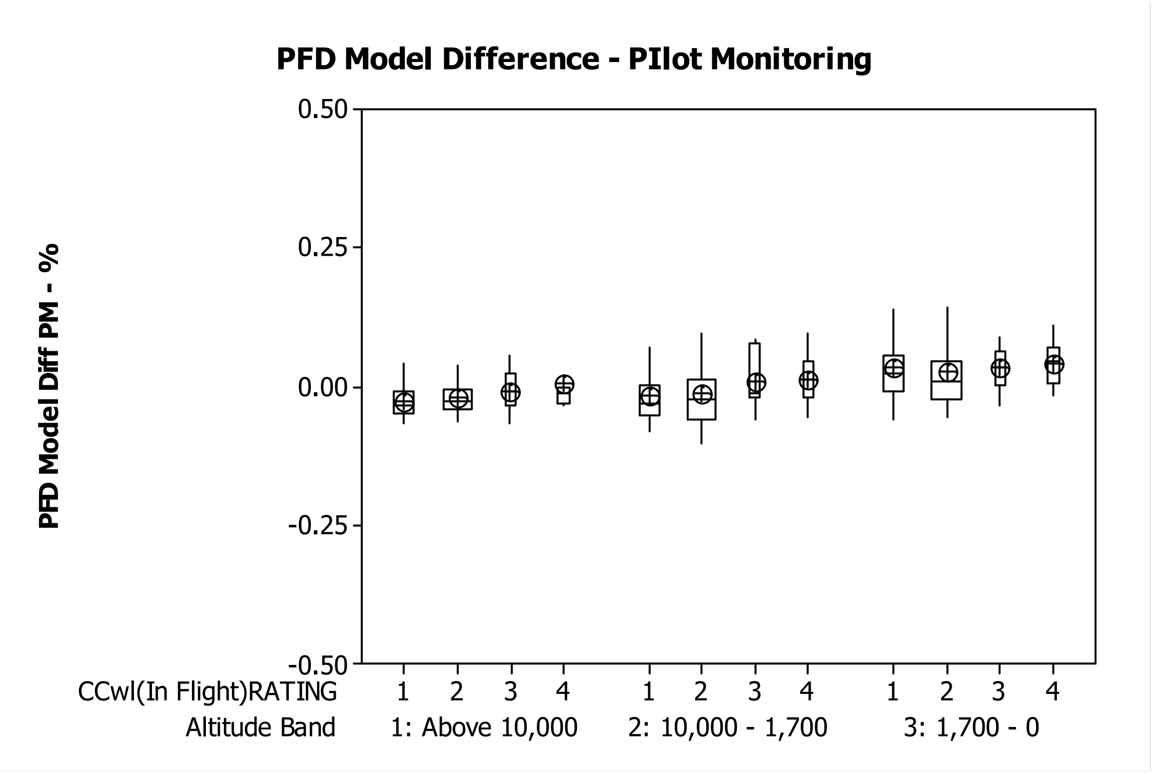


Figure D 14. PFD Model Difference - Pilot Monitoring

المنارة للاستشارات

www.manaraa.com

## D1.2 Pilot Flying PDT Difference from Pilot Monitoring PDT

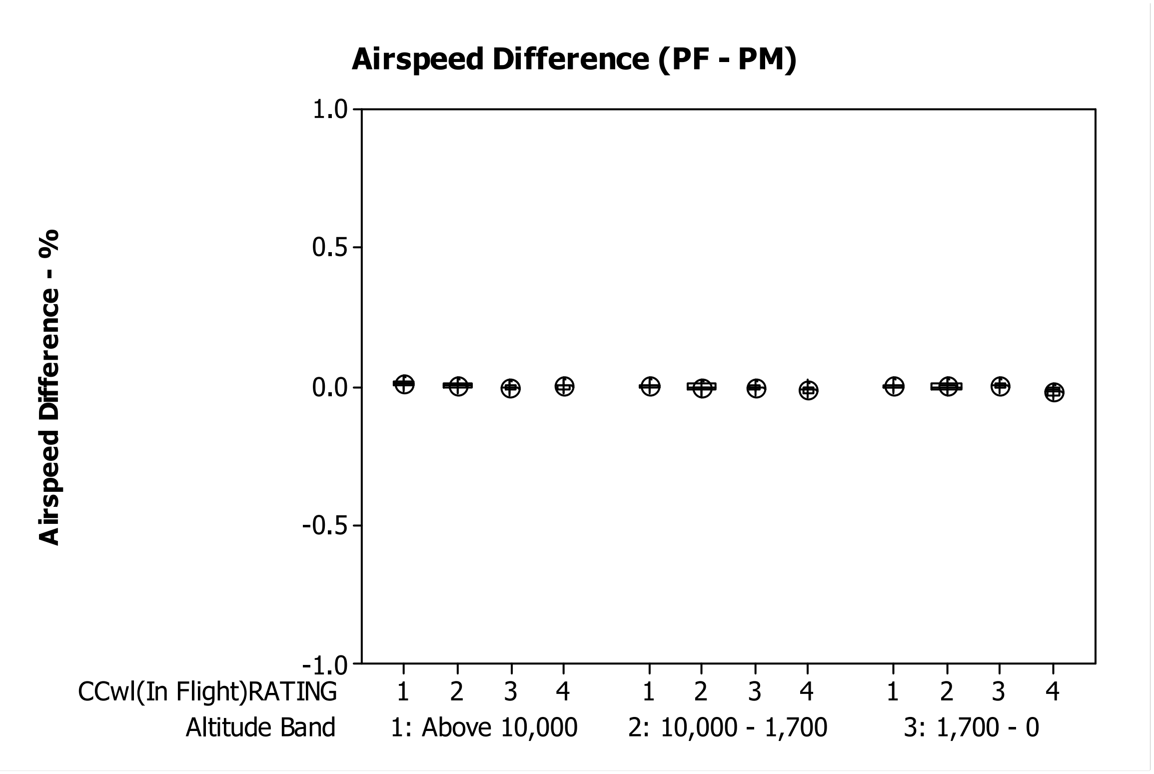


**Figure D 15. Airspeed Indicator PDT Difference (PF - PM)**

www.manaraa.com

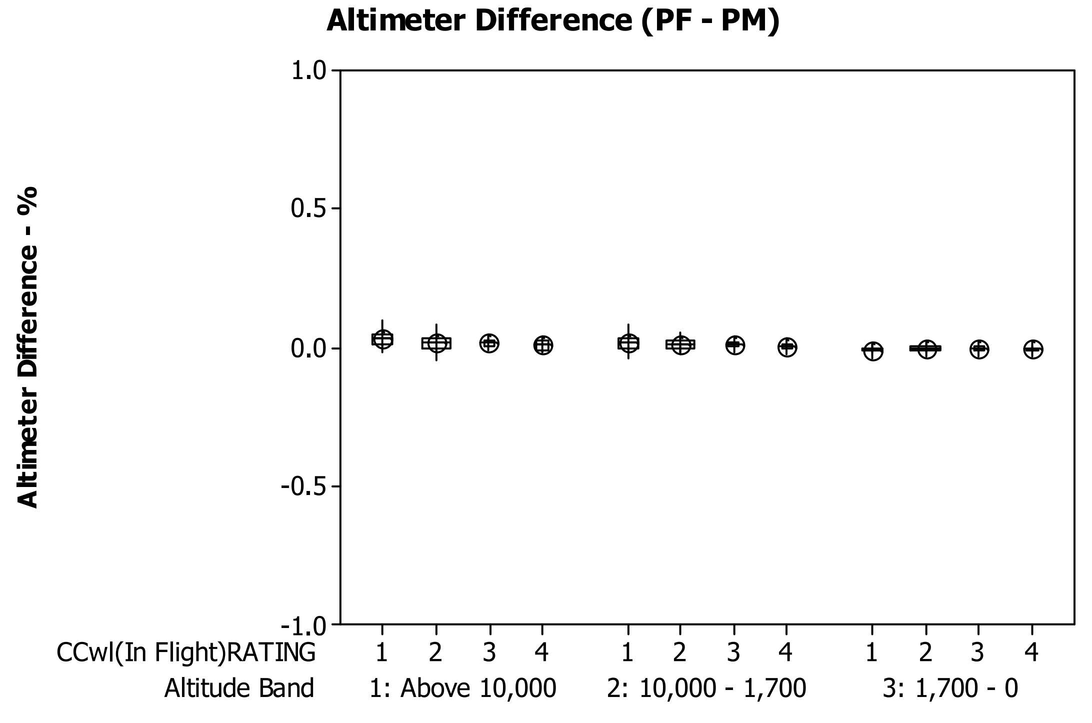


**Figure D 16. Altimeter PDT Difference – (PF – PM)**

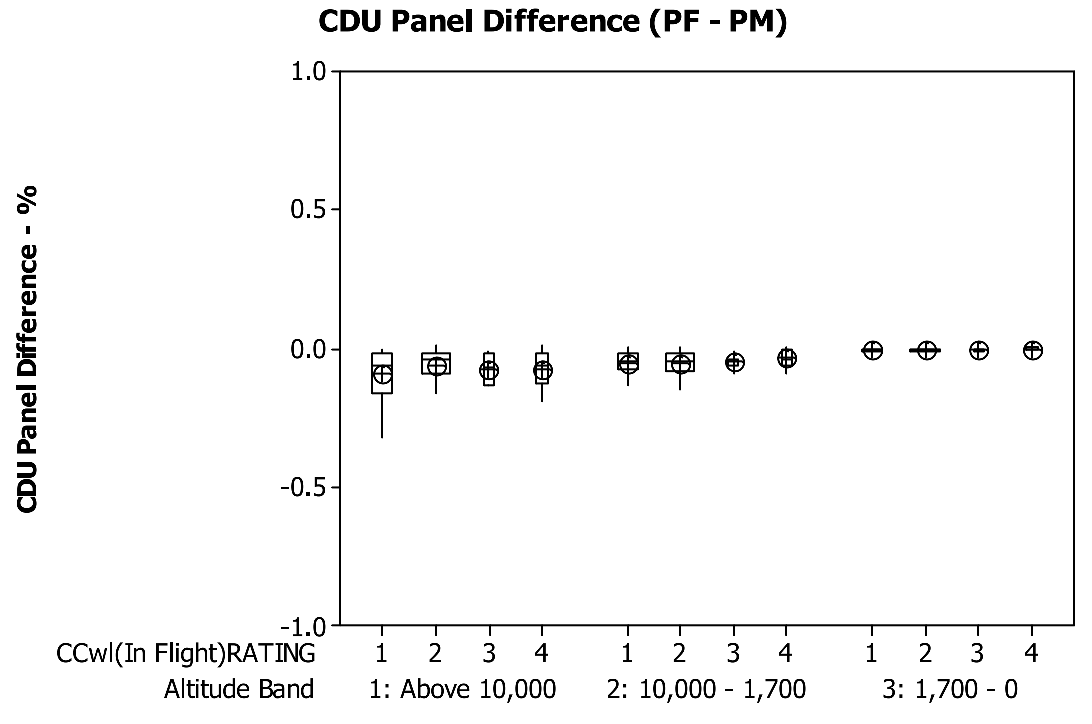


**Figure D 17. CDU Panel PDT Difference - (PF - PM)**

المنارة للاستشارات

www.manaraa.com

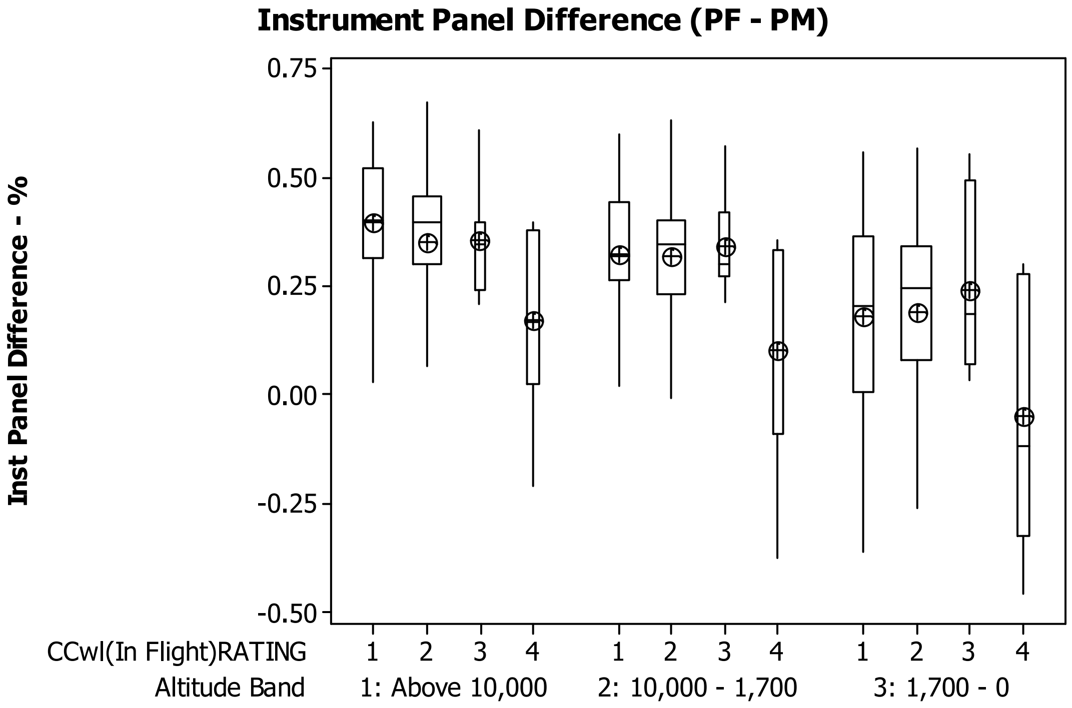


Figure D 18. Instrument Panel PDT Difference - (PF - PM)

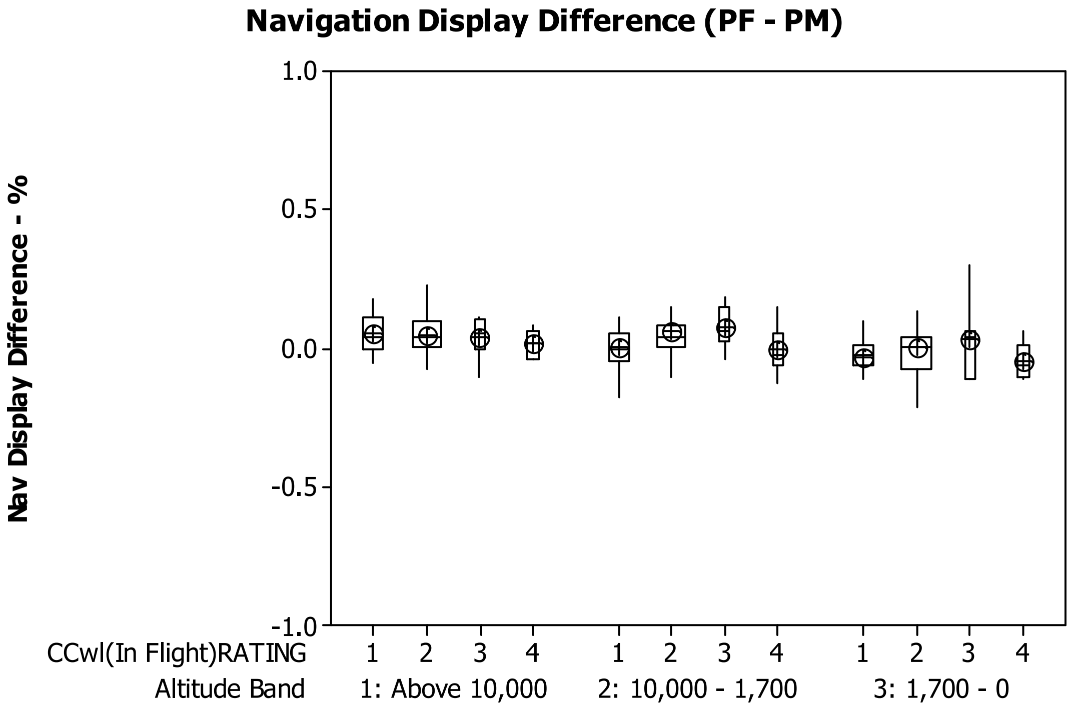


Figure D 19. Navigation Display PDT Difference - (PF - PM)

www.manaraa.com

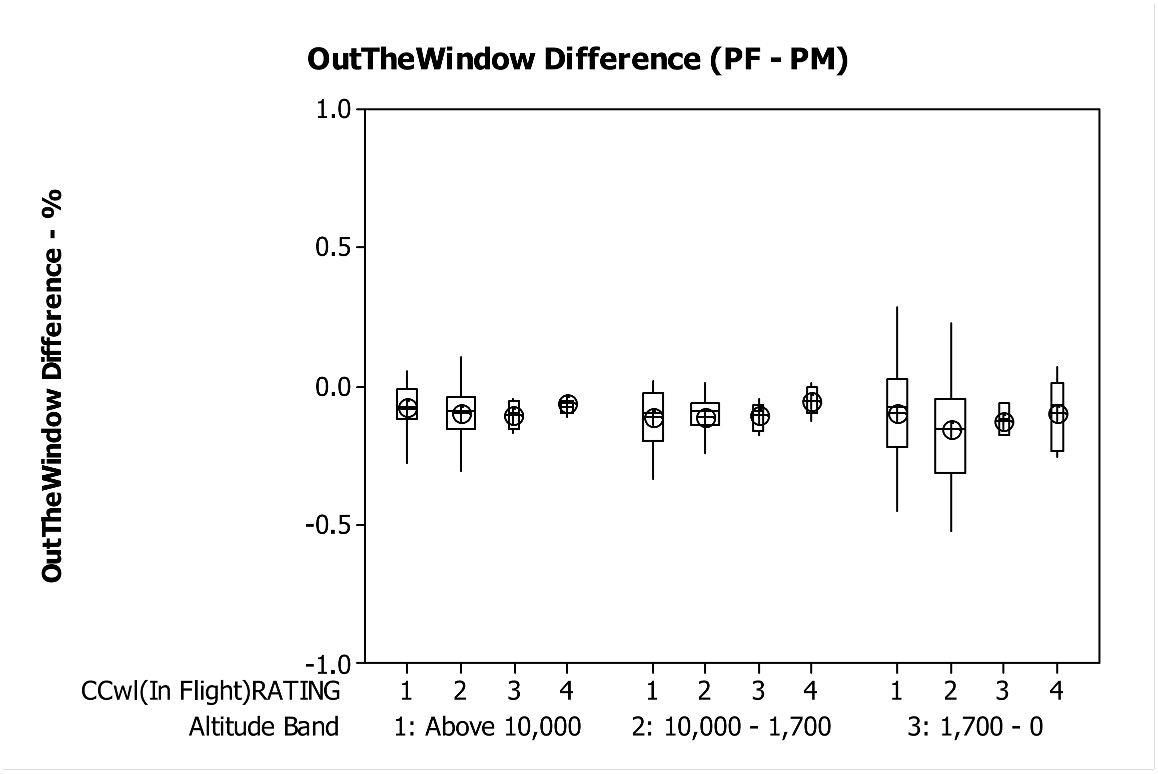


Figure D 20. Out the Window PDT Difference - (PF - PM)

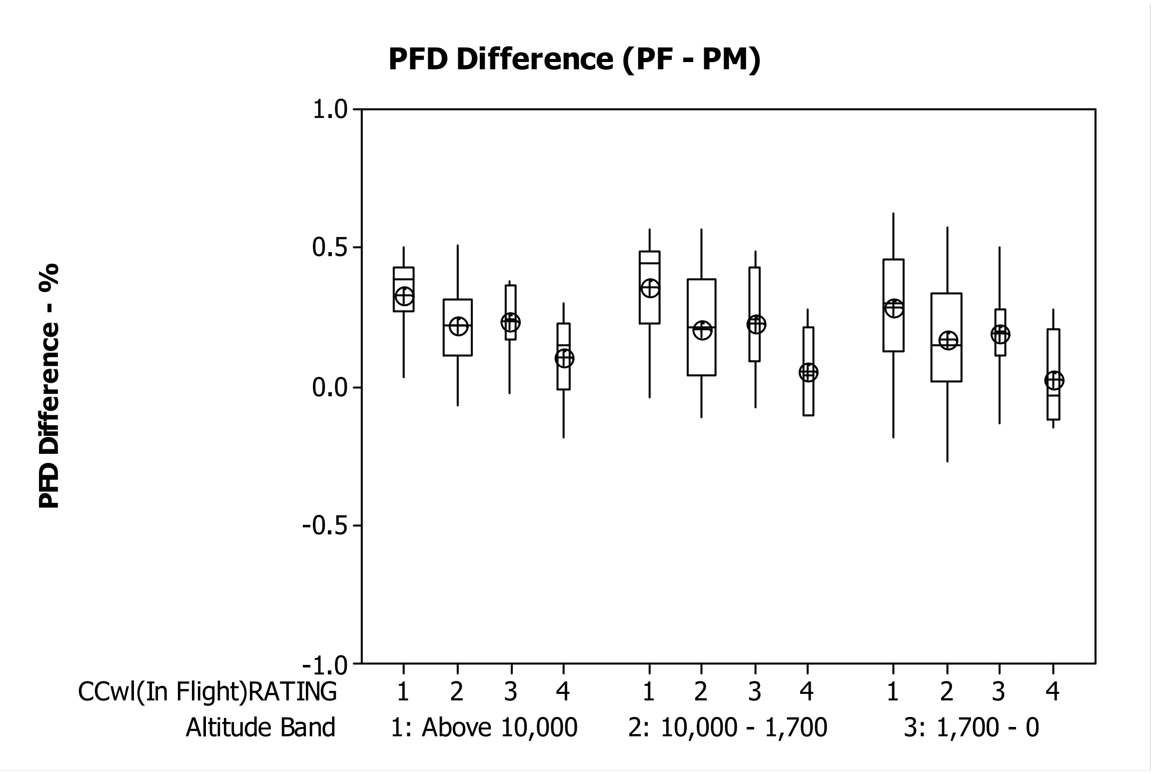


Figure D 21. PFD PDT Difference - (PF - PM)

المنارة للاستشارات

www.manaraa.com

### D1.3 Shared Awareness (5 seconds)

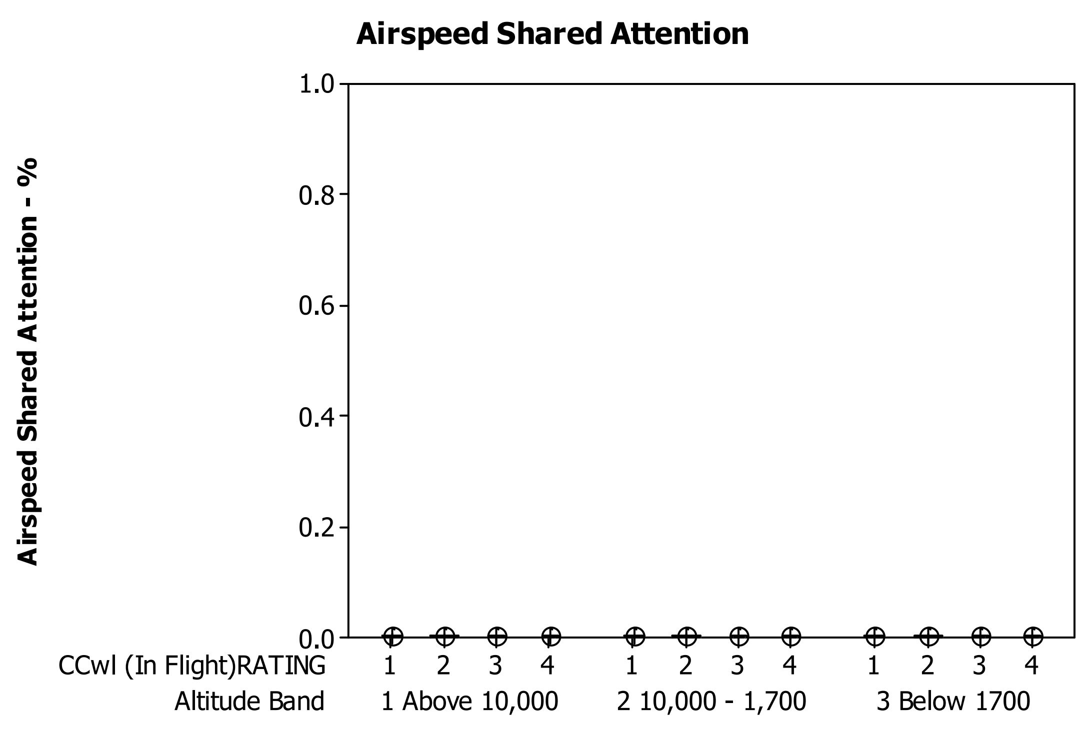


Figure D 22. Airspeed Indicator Shared Attention

www.manaraa.com

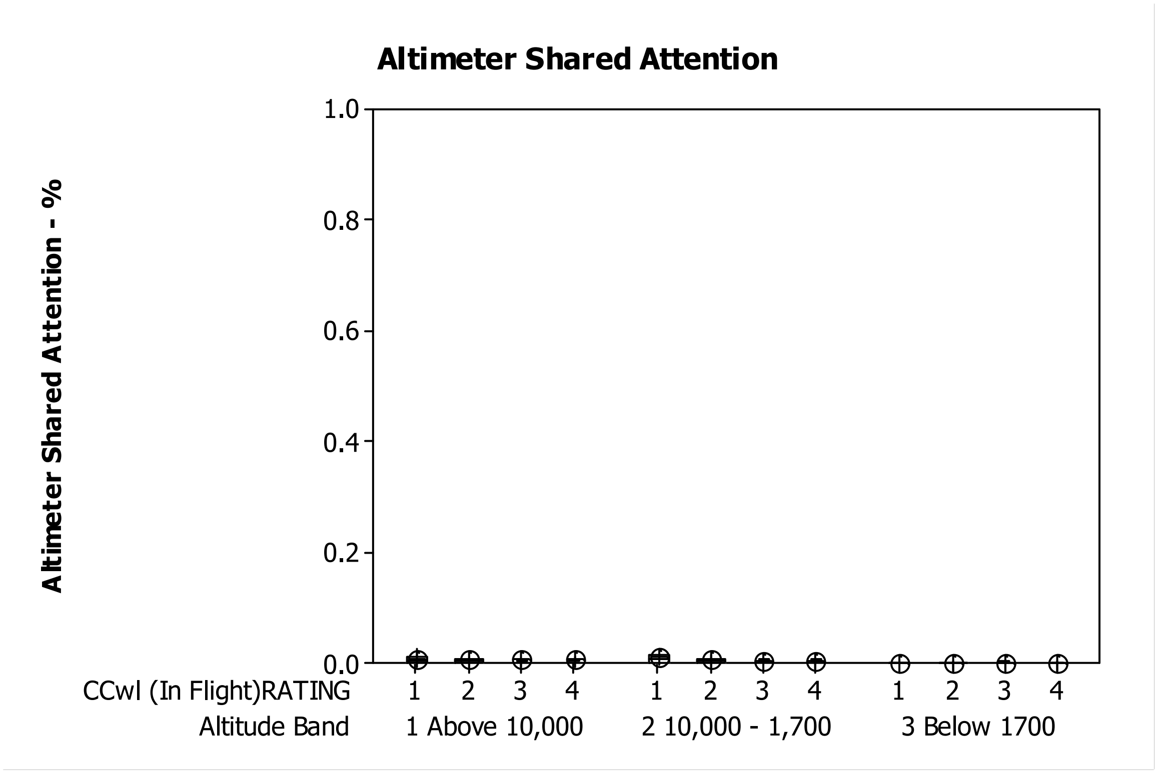


Figure D 23. Altimeter Shared Attention

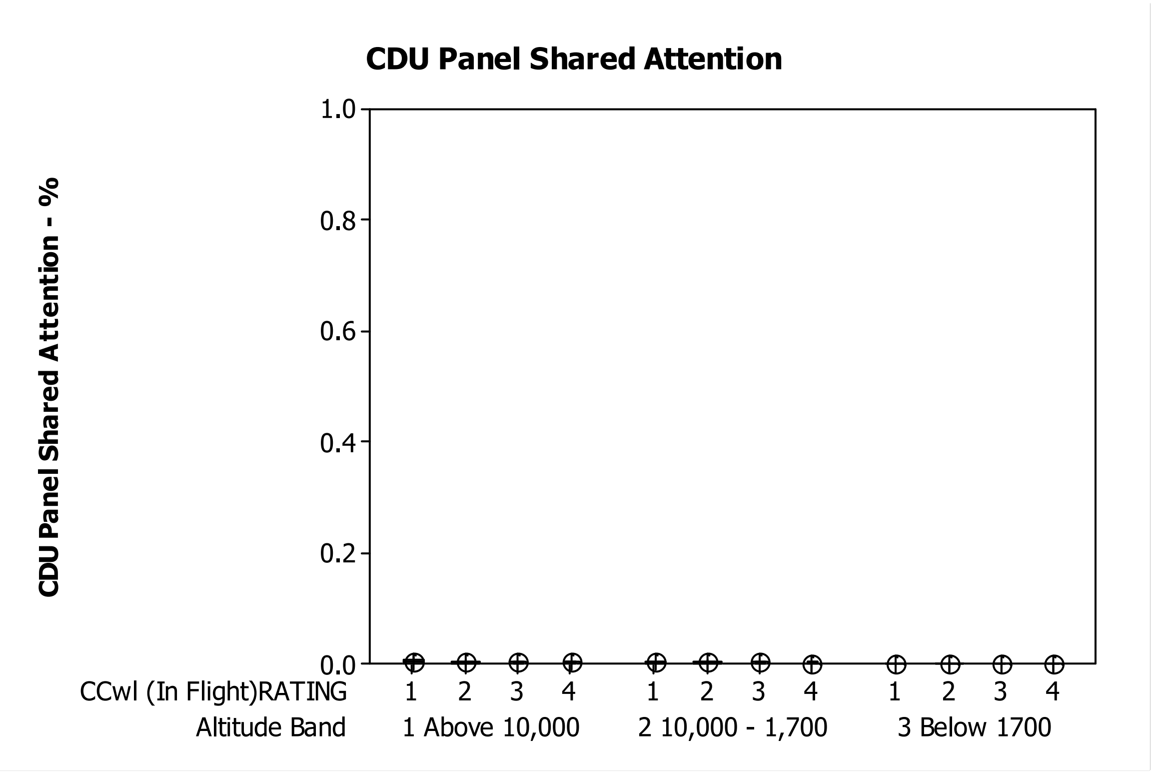


Figure D 24. CDU Panel Shared Attention

www.manaraa.com

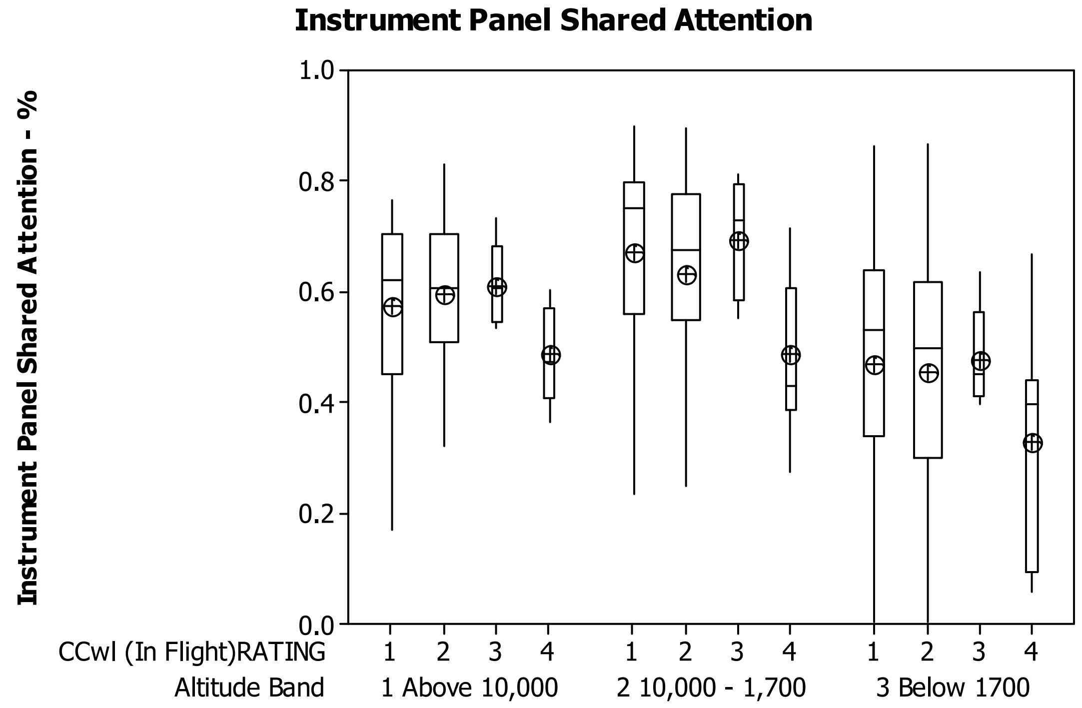


Figure D 25. Instrument Panel Shared Attention

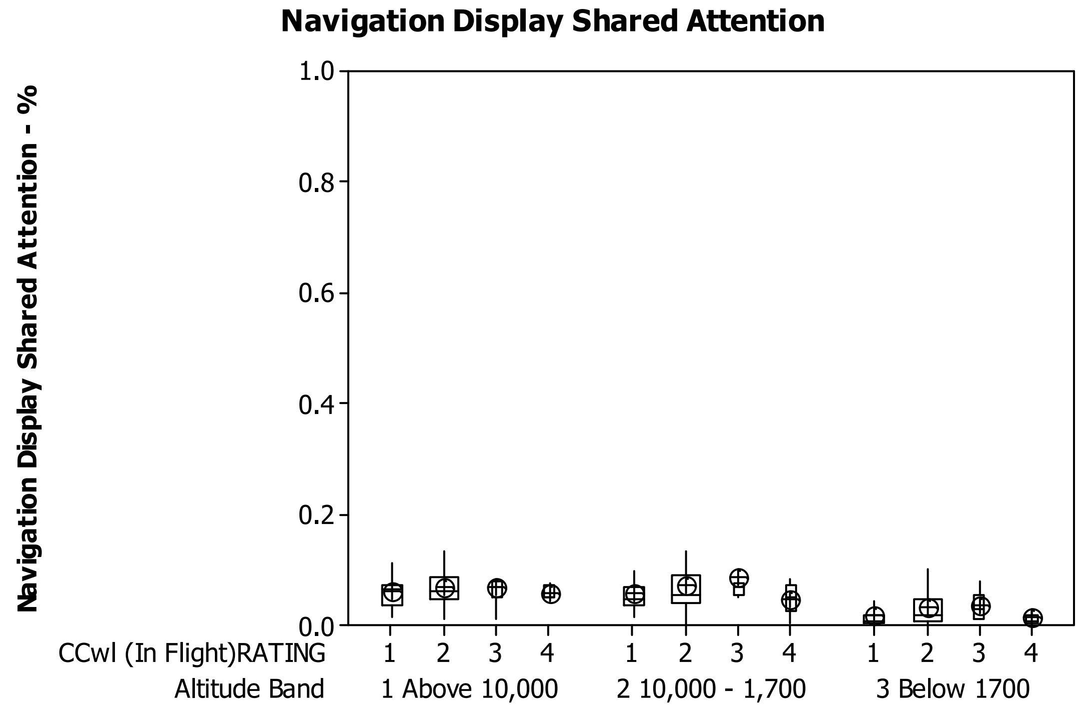


Figure D 26. Navigation Display Shared Attention

المنارة للاستشارات

www.manaraa.com

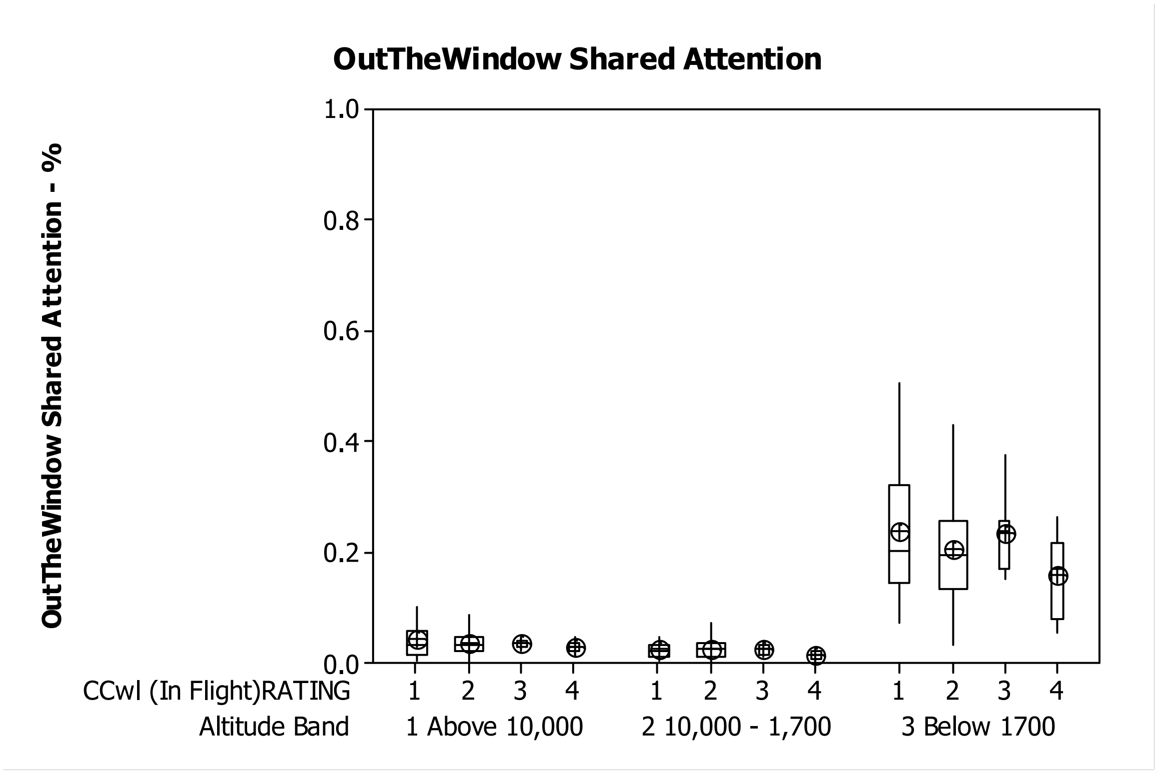


Figure D 27. Out the Window Shared Attention

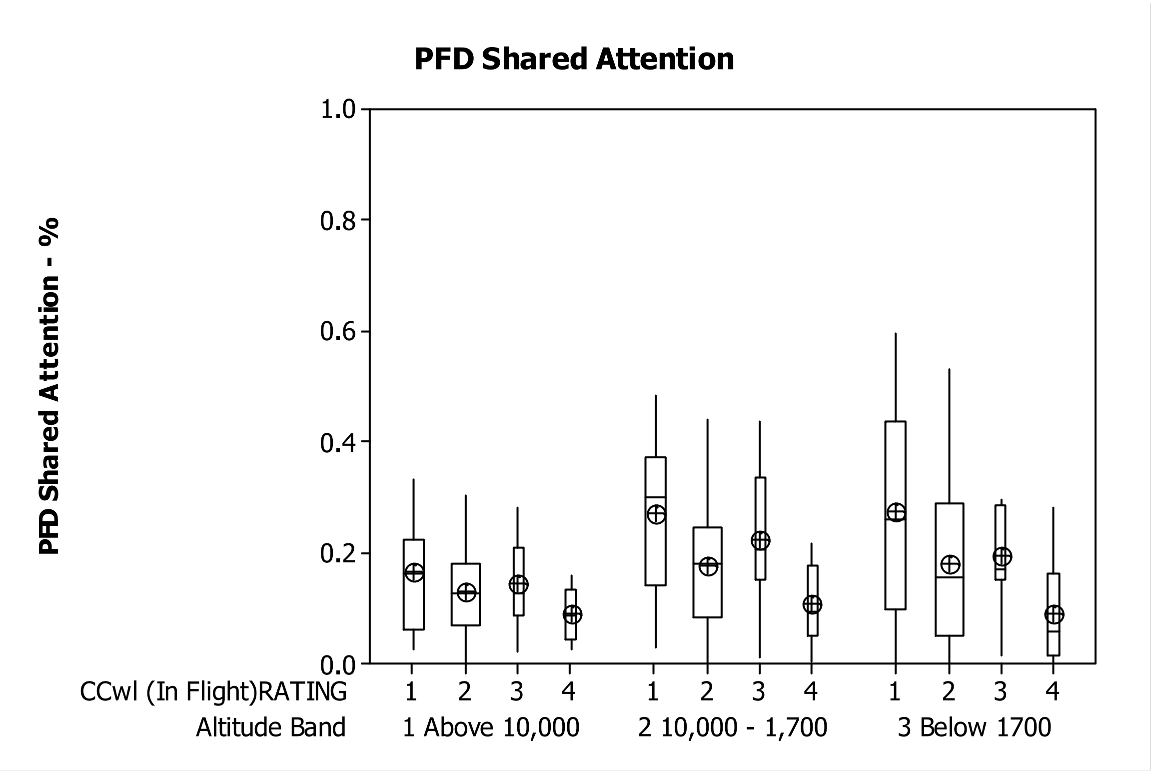


Figure D 28. PFD Shared Attention

المنارة للاستشارات

www.manaraa.com

## D2. ANOVA Results of Selected AOIs

### D2.1 Normative Model Difference

#### D2.1.1 PF

### General Linear Model: Inst Panel M versus Altitude Ban, CCwl(In Flig

```
Factor         Type  Levels Values
Altitude Band     fixed    3 1: Above 10,000, 2: 10,000 - 1,700, 3:
                   1,700 - 0
CCwl(In Flight)RATING fixed    4 1, 2, 3, 4


Analysis of Variance for Inst Panel Model Diff PF, using Adjusted SS for Tests

Source                DF  Seq SS  Adj SS  Adj MS   F
Altitude Band             2  0.58088 0.29791 0.14896 4.46
CCwl(In Flight)RATING         3  0.74439 0.73941 0.24647 7.38
Altitude Band*CCwl(In Flight)RATING  6  0.01129 0.01129 0.00188 0.06
Error                282  9.42042 9.42042 0.03341
Total                293 10.75698

Source                  P
Altitude Band            0.012
CCwl(In Flight)RATING        0.000
Altitude Band*CCwl(In Flight)RATING 0.999
Error
Total


S = 0.182772  R-Sq = 12.43%  R-Sq(adj) = 9.01%
```

**Tukey Simultaneous Tests**
**Response Variable Inst Panel Model Diff PF**
**All Pairwise Comparisons among Levels of Altitude Band**

```
Altitude Band = 1: Above 10,000 subtracted from:

          Difference    SE of      Adjusted
Altitude Band    of Means Difference T-Value  P-Value
2: 10,000 - 1,700  -0.03595   0.03631 -0.9898  0.5833
3: 1,700 - 0     0.06964   0.03633  1.9168  0.1339


Altitude Band = 2: 10,000 - 1,700 subtracted from:

        Difference    SE of      Adjusted
Altitude Band  of Means Difference T-Value  P-Value
3: 1,700 - 0    0.1056   0.03593  2.939  0.0092
```

```
Tukey Simultaneous Tests
Response Variable Inst Panel Model Diff PF
All Pairwise Comparisons among Levels of CCwl(In Flight)RATING
CCwl(In Flight)RATING = 1 subtracted from:

CCwl(In     Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
2          -0.0139   0.02477  -0.559  0.9440
3          0.0181   0.04476  0.405  0.9775
4          -0.1810   0.04125  -4.389  0.0001


CCwl(In Flight)RATING = 2 subtracted from:

CCwl(In     Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
3          0.0320   0.04233  0.756  0.8741
4          -0.1672   0.03860  -4.330  0.0001


CCwl(In Flight)RATING = 3 subtracted from:

CCwl(In     Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
4          -0.1992   0.05366  -3.711  0.0012
```

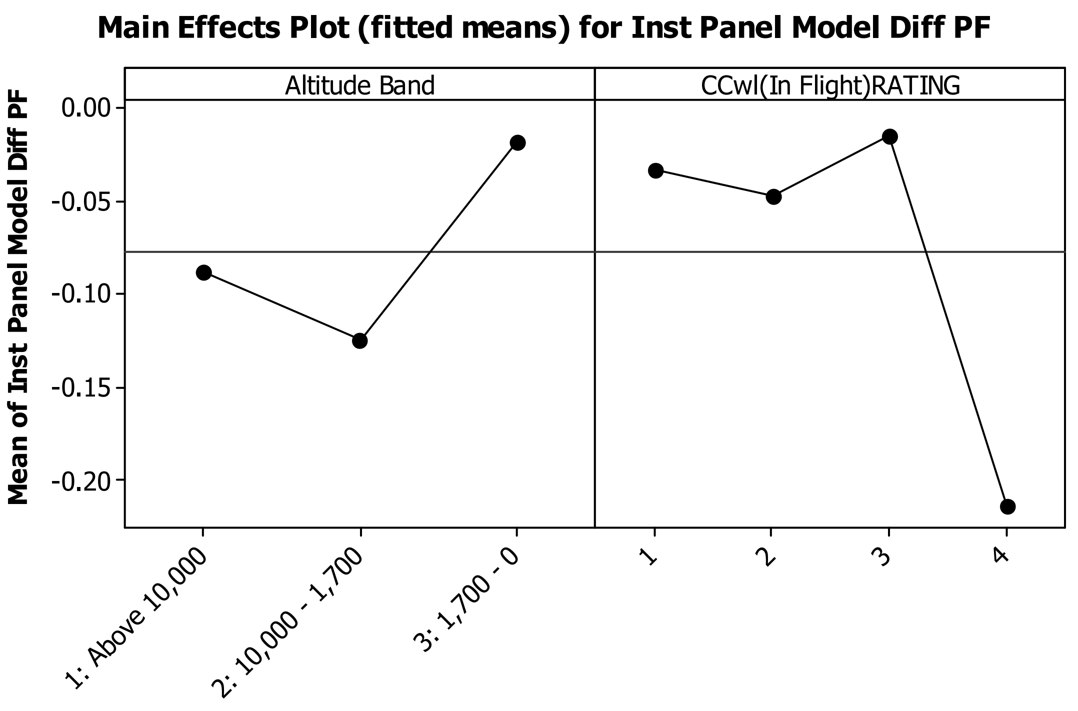


Figure D 29. Instrument Panel Model Difference Main Effects - PF

www.manaraa.com

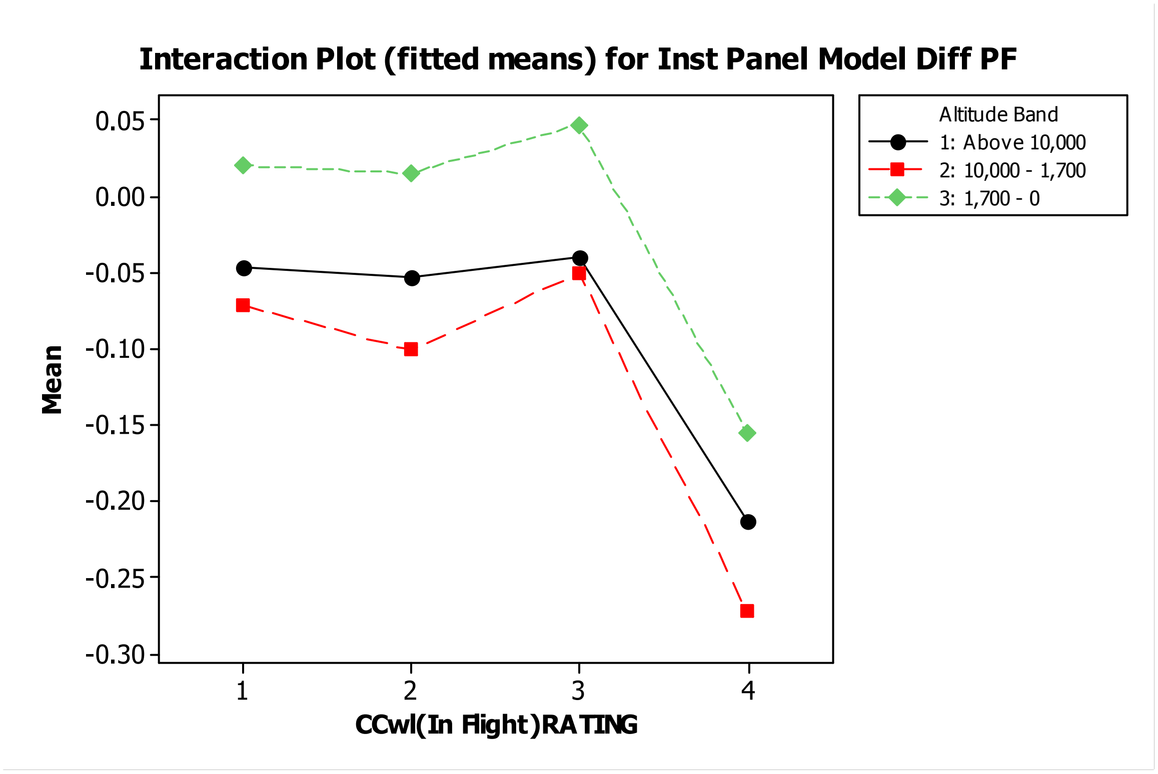


Figure D 30. Instrument Panel Model Difference Interaction Plot - PF

## General Linear Model: OutTheWindow versus Altitude Ban, CCwl(In Flig

```
Factor          Type  Levels Values
Altitude Band     fixed     3 1: Above 10,000, 2: 10,000 - 1,700, 3:
                   1,700 - 0
CCwl(In Flight)RATING fixed    4 1, 2, 3, 4


Analysis of Variance for OutTheWindow Model Diff PF, using Adjusted SS for
   Tests

Source                DF  Seq SS  Adj SS  Adj MS   F
Altitude Band              2 1.33912 0.69023 0.34511 61.47
CCwl(In Flight)RATING          3 0.02168 0.02095 0.00698  1.24
Altitude Band*CCwl(In Flight)RATING  6 0.03579 0.03579 0.00597  1.06
Error                    282 1.58330 1.58330 0.00561
Total                    293 2.97989

Source                  P
Altitude Band             0.000
CCwl(In Flight)RATING         0.294
Altitude Band*CCwl(In Flight)RATING 0.385
Error
Total


S = 0.0749301  R-Sq = 46.87%  R-Sq(adj) = 44.79%
```

```
Tukey Simultaneous Tests
Response Variable OutTheWindow Model Diff PF
All Pairwise Comparisons among Levels of Altitude Band
Altitude Band = 1: Above 10,000 subtracted from:

          Difference    SE of      Adjusted
Altitude Band    of Means Difference T-Value  P-Value
2: 10,000 - 1,700   0.0026   0.01489  0.174  0.9835
3: 1,700 - 0     -0.1407   0.01490  -9.443  0.0000


Altitude Band = 2: 10,000 - 1,700 subtracted from:

        Difference    SE of      Adjusted
Altitude Band  of Means Difference T-Value  P-Value
3: 1,700 - 0   -0.1432   0.01473  -9.724  0.0000



Tukey Simultaneous Tests
Response Variable OutTheWindow Model Diff PF
All Pairwise Comparisons among Levels of CCwl(In Flight)RATING
CCwl(In Flight)RATING = 1 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
2        -0.01184   0.01016  -1.165  0.6488
3        -0.00123   0.01835  -0.067  0.9999
4        -0.03038   0.01691  -1.797  0.2748


CCwl(In Flight)RATING = 2 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
3         0.01060   0.01735  0.611  0.9286
4        -0.01855   0.01583  -1.172  0.6446


CCwl(In Flight)RATING = 3 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
4        -0.02915   0.02200  -1.325  0.5469
```

www.manaraa.com

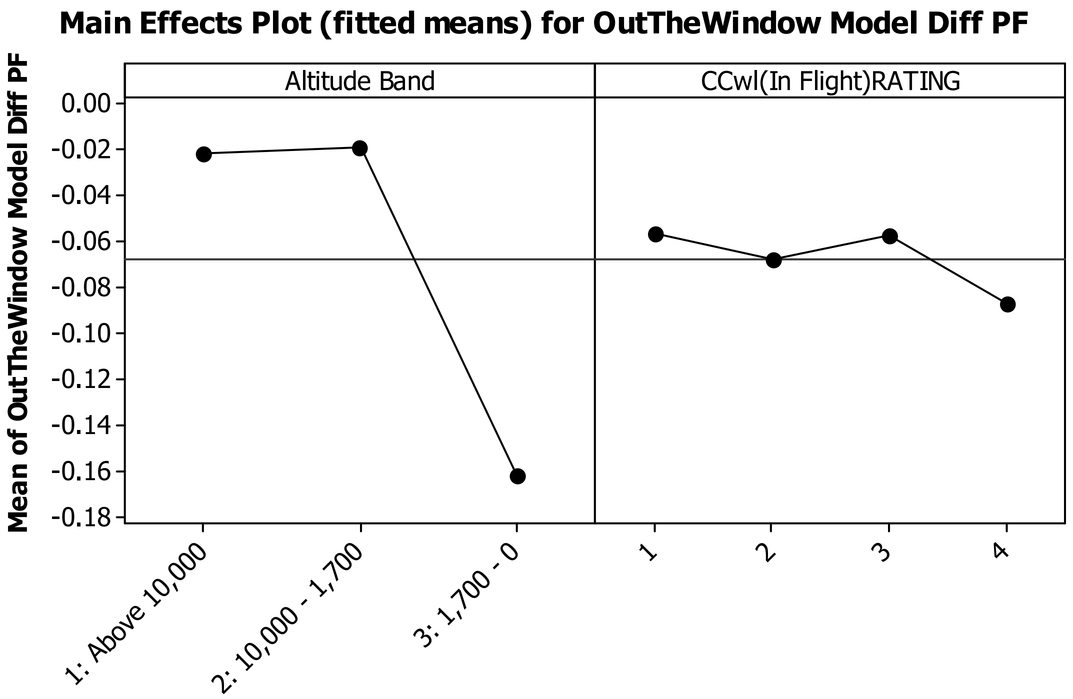


Figure D 31. Out the Window Model Difference Main Effects - PF

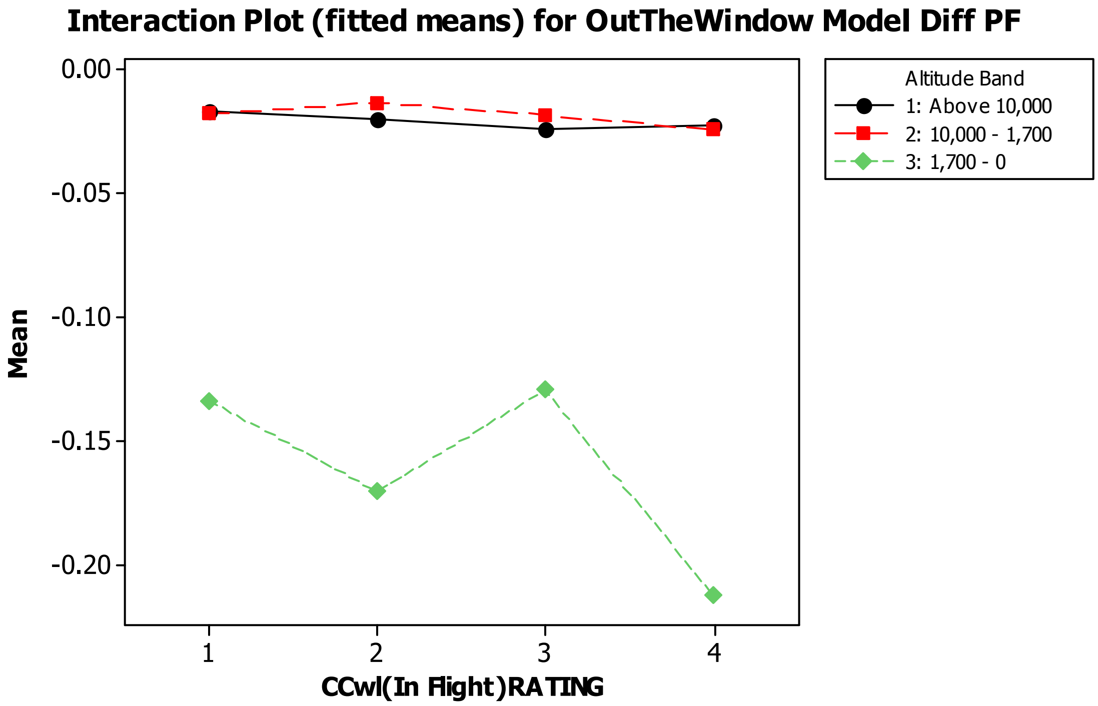


Figure D 32. Out the Window Model Difference Interaction Plot - PF

www.manaraa.com

## General Linear Model: PFD Model Di versus Altitude Ban, CCwl(In Flig

```
Factor          Type  Levels Values
Altitude Band     fixed    3 1: Above 10,000, 2: 10,000 - 1,700, 3:
                   1,700 - 0
CCwl(In Flight)RATING fixed    4 1, 2, 3, 4


Analysis of Variance for PFD Model Diff PF, using Adjusted SS for Tests

Source                DF  Seq SS  Adj SS  Adj MS   F
Altitude Band             2  0.93716 0.46411 0.23205  7.77
CCwl(In Flight)RATING         3  1.40777 1.39673 0.46558 15.60
Altitude Band*CCwl(In Flight)RATING  6  0.03166 0.03166 0.00528  0.18
Error                282  8.41889 8.41889 0.02985
Total                293 10.79548

Source                  P
Altitude Band            0.001
CCwl(In Flight)RATING        0.000
Altitude Band*CCwl(In Flight)RATING 0.983
Error
Total


S = 0.172784  R-Sq = 22.01%  R-Sq(adj) = 18.97%
```

```
Tukey Simultaneous Tests
Response Variable PFD Model Diff PF
All Pairwise Comparisons among Levels of Altitude Band
Altitude Band = 1: Above 10,000 subtracted from:

          Difference    SE of      Adjusted
Altitude Band    of Means Difference T-Value  P-Value
2: 10,000 - 1,700  -0.05708   0.03433  -1.663  0.2196
3: 1,700 - 0     0.07640   0.03435  2.224  0.0671


Altitude Band = 2: 10,000 - 1,700 subtracted from:

        Difference    SE of      Adjusted
Altitude Band  of Means Difference T-Value  P-Value
3: 1,700 - 0    0.1335   0.03397  3.930  0.0003
```

```
Tukey Simultaneous Tests
Response Variable PFD Model Diff PF
All Pairwise Comparisons among Levels of CCwl(In Flight)RATING
CCwl(In Flight)RATING = 1 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
2         -0.1232   0.02342  -5.262  0.0000
3         -0.0899   0.04231  -2.125  0.1452
4         -0.2386   0.03899  -6.118  0.0000
```

www.manaraa.com

```
CCwl(In Flight)RATING = 2 subtracted from:

CCwl(In     Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
3          0.0333   0.04002  0.833  0.8389
4          -0.1153   0.03649  -3.160  0.0086


CCwl(In Flight)RATING = 3 subtracted from:

CCwl(In     Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
4          -0.1487   0.05073  -2.930  0.0178
```

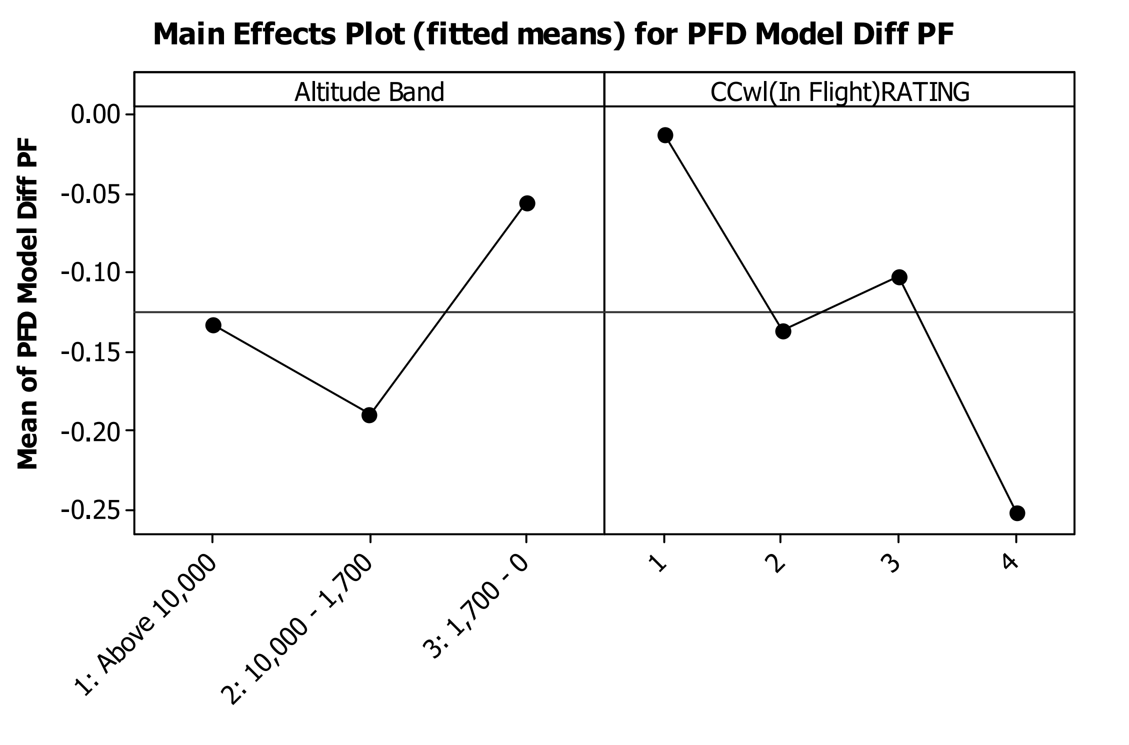


Figure D 33. PFD Model Difference Main Effects - PF

المنارة للاستشارات
www.manaraa.com

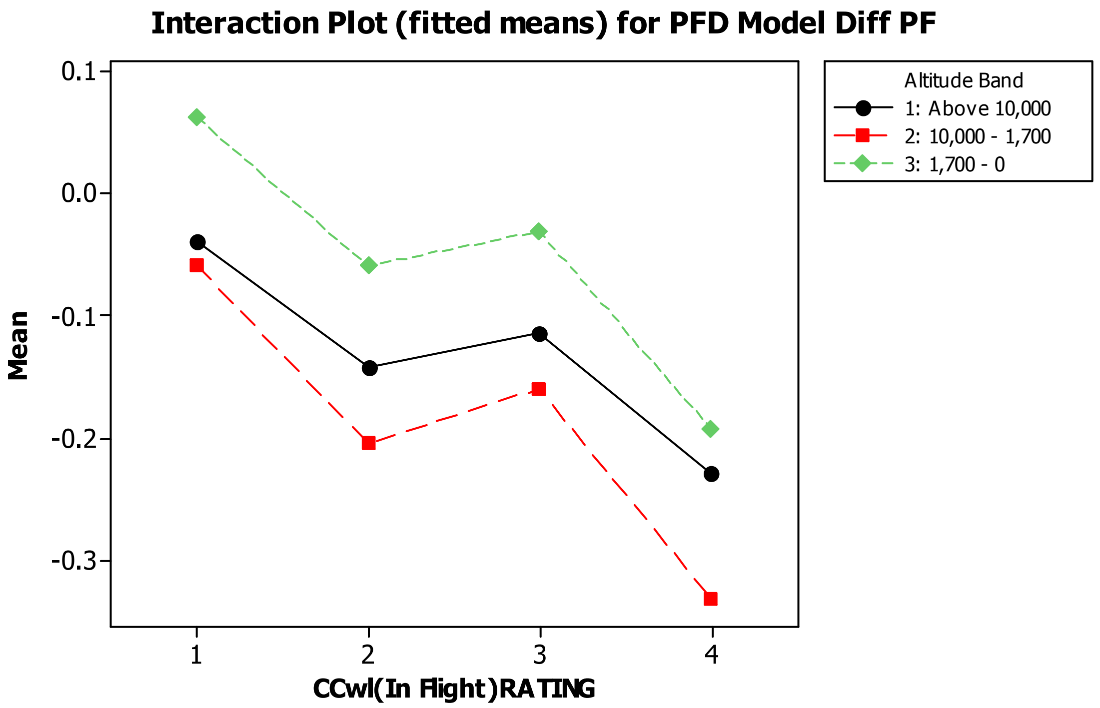


**Figure D 34. PFD Model Difference Interaction Plot - PF**

### D2.1.2 PM

## General Linear Model: Inst Panel M versus Altitude Ban, CCwl(In Flig

```
Factor          Type  Levels Values
Altitude Band      fixed    3 1: Above 10,000, 2: 10,000 - 1,700, 3:
                   1,700 - 0
CCwl(In Flight)RATING fixed    4 1, 2, 3, 4


Analysis of Variance for Inst Panel Model Diff PM, using Adjusted SS for Tests

Source                DF   Seq SS   Adj SS   Adj MS    F
Altitude Band             2 0.99146 0.50374 0.25187 12.29
CCwl(In Flight)RATING         3 0.04912 0.04808 0.01603  0.78
Altitude Band*CCwl(In Flight)RATING  6 0.05353 0.05353 0.00892  0.44
Error                282 5.77907 5.77907 0.02049
Total                293 6.87319

Source                  P
Altitude Band            0.000
CCwl(In Flight)RATING        0.505
Altitude Band*CCwl(In Flight)RATING 0.855
Error
Total
```

www.manaraa.com

```
S = 0.143154  R-Sq = 15.92%  R-Sq(adj) = 12.64%
```

**Tukey Simultaneous Tests**
**Response Variable Inst Panel Model Diff PM**
**All Pairwise Comparisons among Levels of Altitude Band**

```
Altitude Band = 1: Above 10,000 subtracted from:

          Difference    SE of      Adjusted
Altitude Band    of Means Difference  T-Value  P-Value
2: 10,000 - 1,700  -0.002219   0.02844 -0.07801  0.9967
3: 1,700 - 0     0.120153   0.02846  4.22217  0.0001


Altitude Band = 2: 10,000 - 1,700 subtracted from:

        Difference    SE of      Adjusted
Altitude Band  of Means Difference T-Value  P-Value
3: 1,700 - 0    0.1224   0.02814  4.348  0.0000
```

**Tukey Simultaneous Tests**
**Response Variable Inst Panel Model Diff PM**
**All Pairwise Comparisons among Levels of CCwl(In Flight)RATING**

```
CCwl(In Flight)RATING = 1 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
2        0.002417   0.01940  0.1246  0.9993
3        0.007525   0.03506  0.2147  0.9965
4        0.046723   0.03231  1.4463  0.4704


CCwl(In Flight)RATING = 2 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
3        0.005108   0.03316  0.1541  0.9987
4        0.044306   0.03023  1.4654  0.4586


CCwl(In Flight)RATING = 3 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
4         0.03920   0.04203  0.9326  0.7873
```

www.manaraa.com

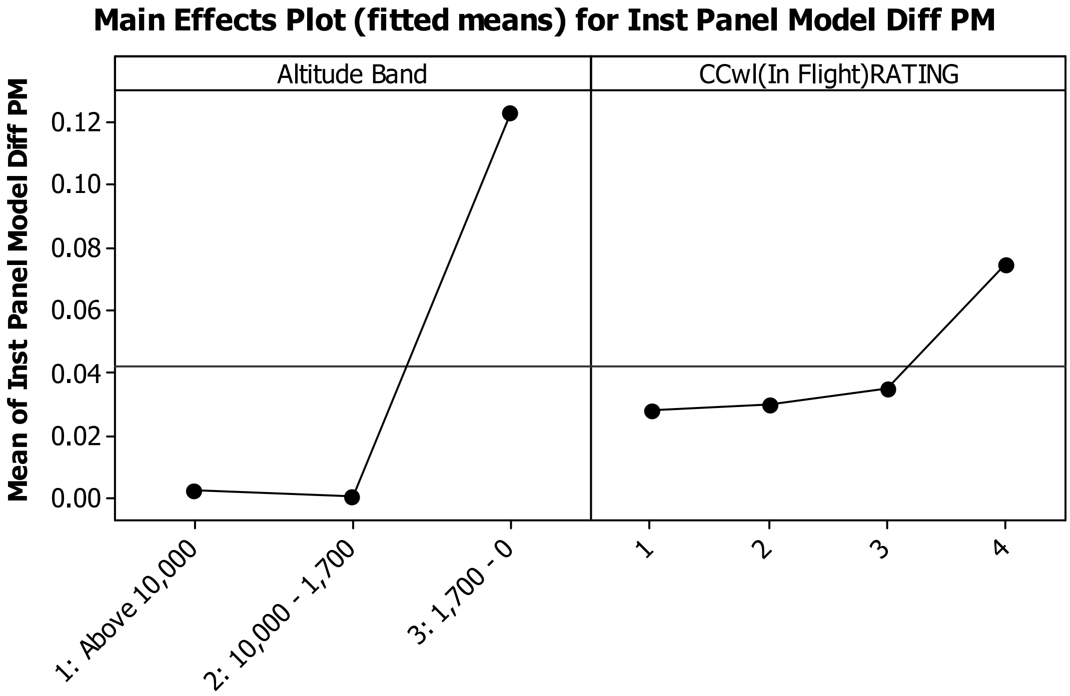


Figure D 35. Instrument Panel Model Difference Main Effects - PM

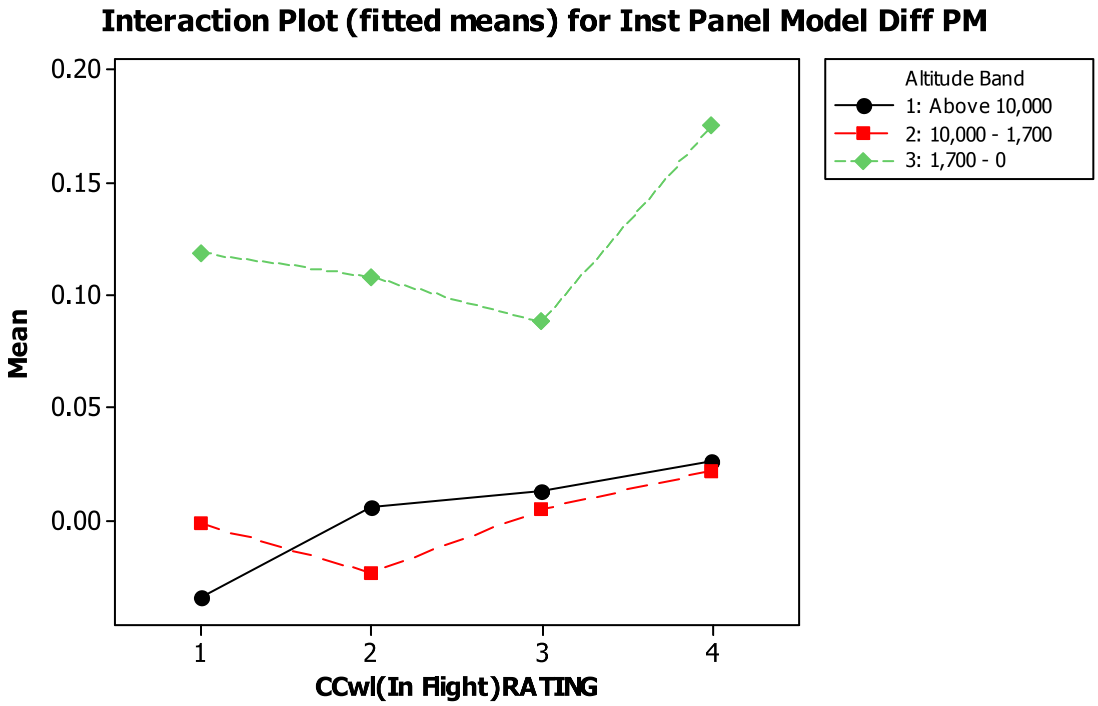


Figure D 36. Instrument Panel Model Difference Interaction Plot - PM

المنارة للاستشارات
www.manaraa.com

## General Linear Model: OutTheWindow versus Altitude Ban, CCwl(In Flig

```
Factor         Type  Levels Values
Altitude Band     fixed    3 1: Above 10,000, 2: 10,000 - 1,700, 3:
                   1,700 - 0
CCwl(In Flight)RATING fixed    4 1, 2, 3, 4


Analysis of Variance for OutTheWindow Model Diff PM, using Adjusted SS for
   Tests

Source                DF  Seq SS  Adj SS  Adj MS   F
Altitude Band             2 0.01737 0.00975 0.00488 0.31
CCwl(In Flight)RATING         3 0.11708 0.11374 0.03791 2.44
Altitude Band*CCwl(In Flight)RATING  6 0.01900 0.01900 0.00317 0.20
Error                282 4.38515 4.38515 0.01555
Total                293 4.53860

Source                  P
Altitude Band            0.731
CCwl(In Flight)RATING        0.065
Altitude Band*CCwl(In Flight)RATING 0.975
Error
Total


S = 0.124700  R-Sq = 3.38%  R-Sq(adj) = 0.00%
```

**Tukey Simultaneous Tests**
**Response Variable OutTheWindow Model Diff PM**
**All Pairwise Comparisons among Levels of Altitude Band**

```
Altitude Band = 1: Above 10,000 subtracted from:

          Difference    SE of      Adjusted
Altitude Band    of Means Difference T-Value  P-Value
2: 10,000 - 1,700  -0.01344    0.02478 -0.5425  0.8503
3: 1,700 - 0     0.00543   0.02479  0.2189  0.9739


Altitude Band = 2: 10,000 - 1,700 subtracted from:

        Difference    SE of      Adjusted
Altitude Band  of Means Difference T-Value  P-Value
3: 1,700 - 0   0.01887   0.02452  0.7696  0.7217
```

**Tukey Simultaneous Tests**
**Response Variable OutTheWindow Model Diff PM**
**All Pairwise Comparisons among Levels of CCwl(In Flight)RATING**

```
CCwl(In Flight)RATING = 1 subtracted from:

CCwl(In    Difference    SE of      Adjusted
```

المنارة للاستشارات
www.manaraa.com

```
Flight)RATING  of Means Difference T-Value  P-Value
2          0.01518   0.01690  0.898  0.8058
3          0.01558   0.03054  0.510  0.9567
4         -0.05463   0.02814  -1.941  0.2107


CCwl(In Flight)RATING = 2 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
3          0.00040   0.02888  0.014  1.0000
4         -0.06981   0.02634  -2.651  0.0401


CCwl(In Flight)RATING = 3 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
4         -0.07021   0.03661  -1.918  0.2205
```

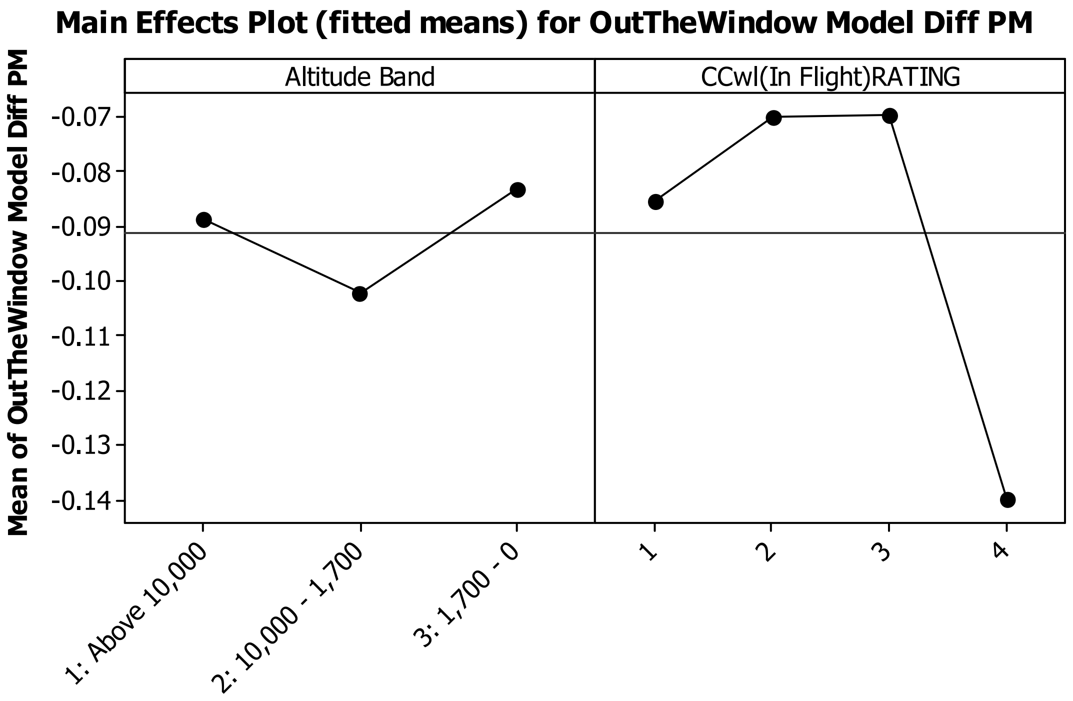


**Figure D 37. Out the Window Model Difference Main Effects - PM**

www.manaraa.com

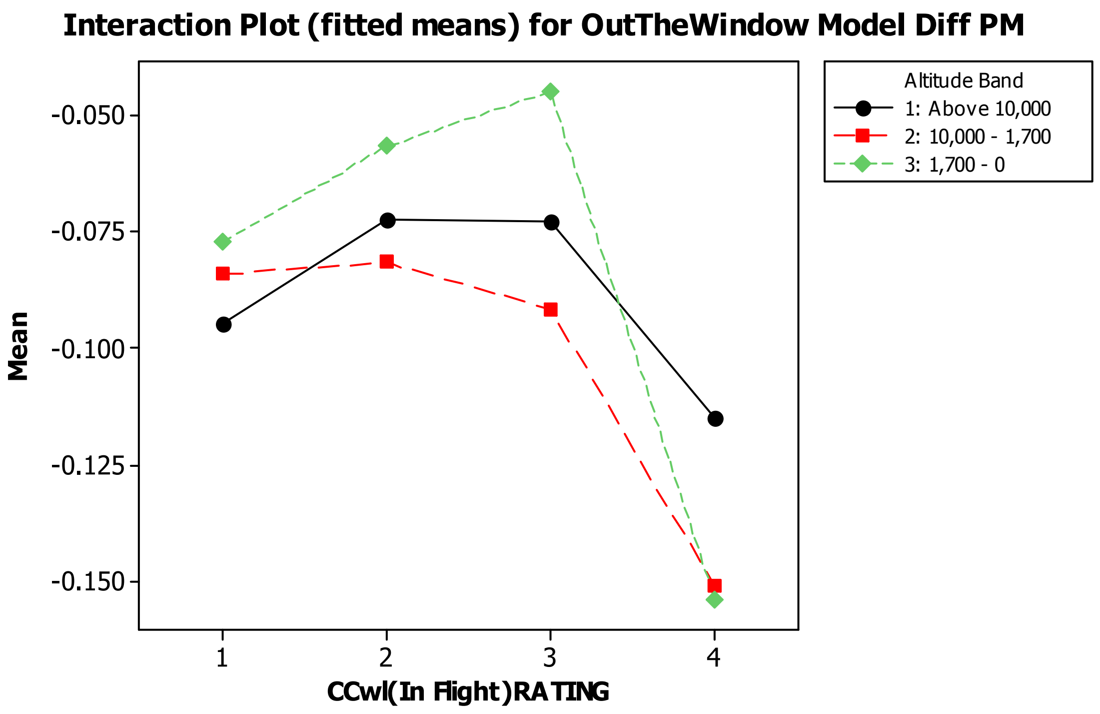


**Figure D 38. Out the Window Model Difference Interaction Plot - PM**

## General Linear Model: PFD Model Di versus Altitude Ban, CCwl(In Flig

```
Factor         Type  Levels Values
Altitude Band     fixed    3 1: Above 10,000, 2: 10,000 - 1,700, 3:
                   1,700 - 0
CCwl(In Flight)RATING fixed    4 1, 2, 3, 4


Analysis of Variance for PFD Model Diff PM, using Adjusted SS for Tests

Source                DF   Seq SS   Adj SS   Adj MS    F
Altitude Band             2 0.134721 0.059813 0.029907 10.12
CCwl(In Flight)RATING         3 0.014249 0.014388 0.004796  1.62
Altitude Band*CCwl(In Flight)RATING  6 0.003405 0.003405 0.000567  0.19
Error                282 0.833614 0.833614 0.002956
Total                293 0.985989

Source                  P
Altitude Band            0.000
CCwl(In Flight)RATING        0.184
Altitude Band*CCwl(In Flight)RATING 0.979
Error
Total


S = 0.0543698  R-Sq = 15.45%  R-Sq(adj) = 12.16%
```

www.manaraa.com

```
Tukey Simultaneous Tests
Response Variable PFD Model Diff PM
All Pairwise Comparisons among Levels of Altitude Band
Altitude Band = 1: Above 10,000 subtracted from:

          Difference    SE of      Adjusted
Altitude Band    of Means Difference T-Value  P-Value
2: 10,000 - 1,700  0.009721   0.01080  0.8999  0.6404
3: 1,700 - 0     0.045921   0.01081  4.2487  0.0001


Altitude Band = 2: 10,000 - 1,700 subtracted from:

        Difference    SE of      Adjusted
Altitude Band  of Means Difference T-Value  P-Value
3: 1,700 - 0   0.03620   0.01069  3.387  0.0020
```

```
Tukey Simultaneous Tests
Response Variable PFD Model Diff PM
All Pairwise Comparisons among Levels of CCwl(In Flight)RATING
CCwl(In Flight)RATING = 1 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
2        0.000984  0.007369  0.1336  0.9992
3        0.014028  0.013314  1.0536  0.7177
4        0.023162  0.012270  1.8878  0.2333


CCwl(In Flight)RATING = 2 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
3         0.01304   0.01259  1.036  0.7284
4         0.02218   0.01148  1.931  0.2148


CCwl(In Flight)RATING = 3 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
4        0.009134   0.01596  0.5722  0.9404
```

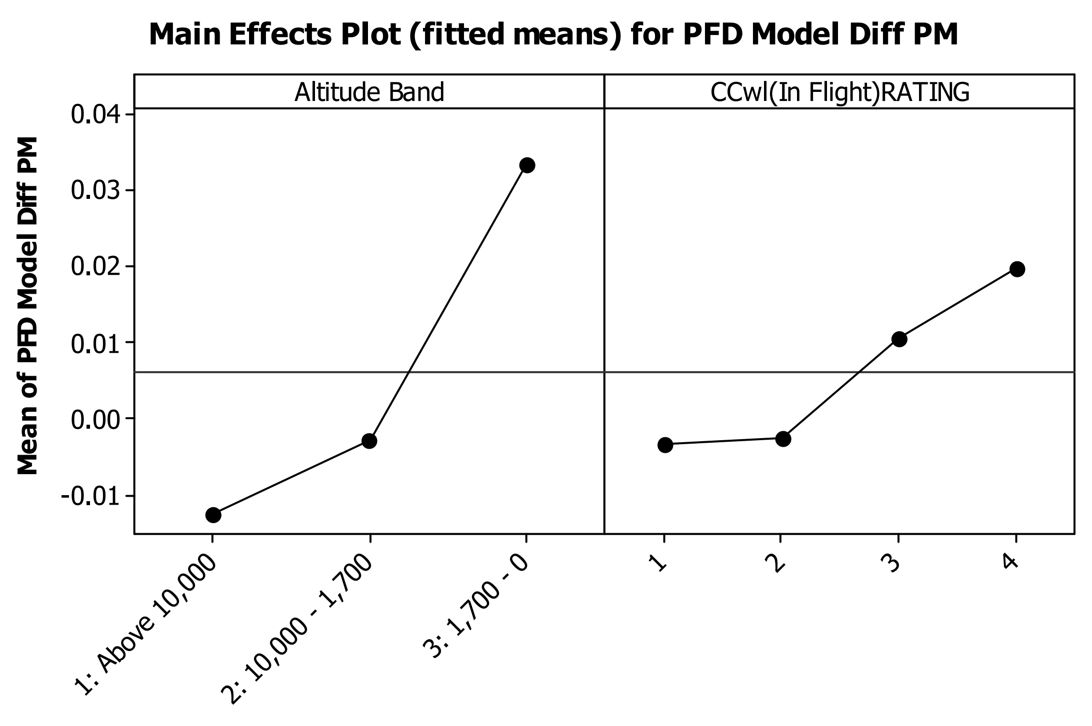


Figure D 39. PFD Model Difference Main Effects - PM

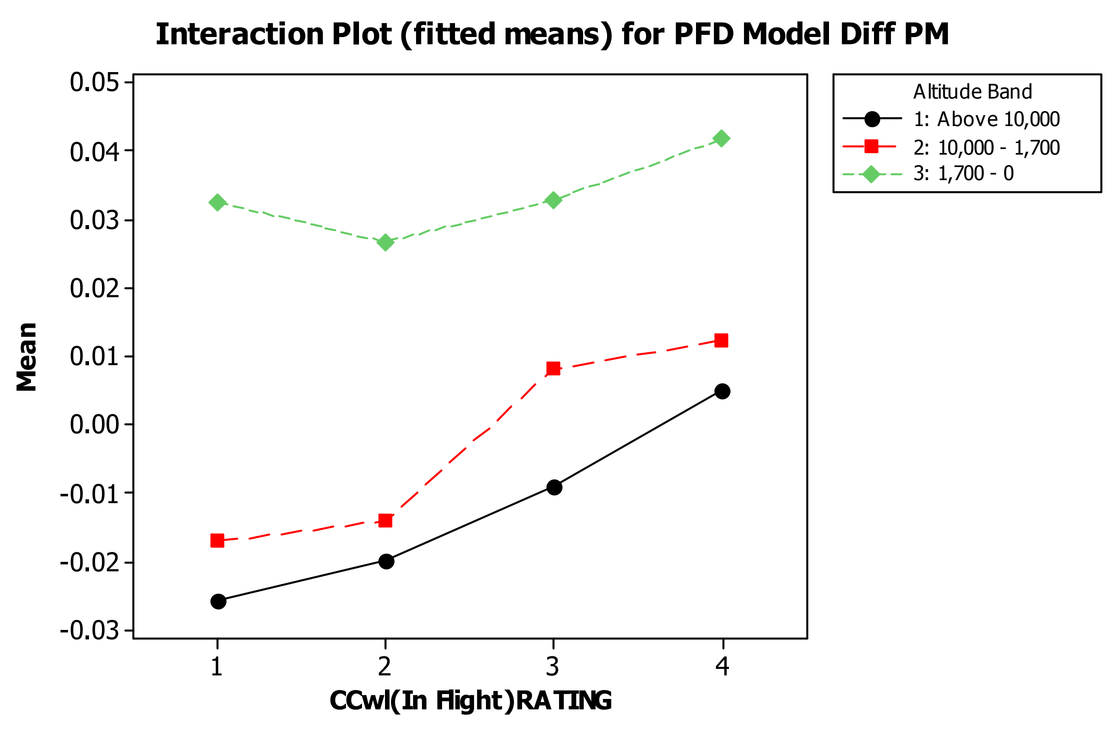


Figure D 40. PFD Model Difference Interaction Plot - PM

www.manaraa.com

## D2.2 Difference (PF – PM) PDT

### General Linear Model: Inst Panel versus Altitude Band, CCwl(In Flight)R

```
Factor         Type  Levels Values
Altitude Band     fixed    3 1: Above 10,000, 2: 10,000 - 1,700, 3:
                   1,700 - 0
CCwl(In Flight)RATING fixed    4 1, 2, 3, 4


Analysis of Variance for Inst Panel, using Adjusted SS for Tests

Source                DF  Seq SS  Adj SS  Adj MS   F
Altitude Band             2  1.69847  0.86428 0.43214 9.56
CCwl(In Flight)RATING         3  1.16311  1.15167 0.38389 8.49
Altitude Band*CCwl(In Flight)RATING  6  0.04901  0.04901 0.00817 0.18
Error                282 12.74905 12.74905 0.04521
Total                293 15.65963

Source                  P
Altitude Band            0.000
CCwl(In Flight)RATING        0.000
Altitude Band*CCwl(In Flight)RATING 0.982
Error
Total


S = 0.212625  R-Sq = 18.59%  R-Sq(adj) = 15.41%
```

**Tukey 95.0% Simultaneous Confidence Intervals**
**Response Variable Inst Panel**
**All Pairwise Comparisons among Levels of Altitude Band**

```
Altitude Band = 1: Above 10,000 subtracted from:

Altitude Band    Lower  Center   Upper
2: 10,000 - 1,700 -0.1460 -0.0471  0.05175
3: 1,700 - 0    -0.2766 -0.1777 -0.07878

Altitude Band   --------+---------+---------+--------
2: 10,000 - 1,700        (---------*---------)
3: 1,700 - 0    (---------*---------)
          --------+---------+---------+--------
            -0.20   -0.10   -0.00


Altitude Band = 2: 10,000 - 1,700 subtracted from:

Altitude Band  Lower  Center   Upper
3: 1,700 - 0  -0.2284 -0.1306 -0.03275

Altitude Band --------+---------+---------+--------
```

www.manaraa.com

```
3: 1,700 - 0    (---------*---------)
        --------+---------+---------+--------
          -0.20   -0.10   -0.00


Tukey Simultaneous Tests
Response Variable Inst Panel
All Pairwise Comparisons among Levels of Altitude Band
Altitude Band = 1: Above 10,000 subtracted from:

          Difference    SE of      Adjusted
Altitude Band    of Means Difference T-Value  P-Value
2: 10,000 - 1,700   -0.0471   0.04225  -1.116  0.5044
3: 1,700 - 0     -0.1777   0.04227  -4.204  0.0001


Altitude Band = 2: 10,000 - 1,700 subtracted from:

        Difference    SE of      Adjusted
Altitude Band  of Means Difference T-Value  P-Value
3: 1,700 - 0   -0.1306   0.04180  -3.124  0.0051
```

**Tukey Simultaneous Tests**
**Response Variable Inst Panel**
**All Pairwise Comparisons among Levels of CCwl(In Flight)RATING**

```
CCwl(In Flight)RATING = 1 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
2         -0.0163   0.02882  -0.565  0.9425
3         0.0106   0.05207  0.204  0.9970
4         -0.2277   0.04798  -4.746  0.0000


CCwl(In Flight)RATING = 2 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
3         0.0269   0.04925  0.546  0.9476
4         -0.2115   0.04491  -4.709  0.0000


CCwl(In Flight)RATING = 3 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
4         -0.2384   0.06243  -3.818  0.0008
```

www.manaraa.com

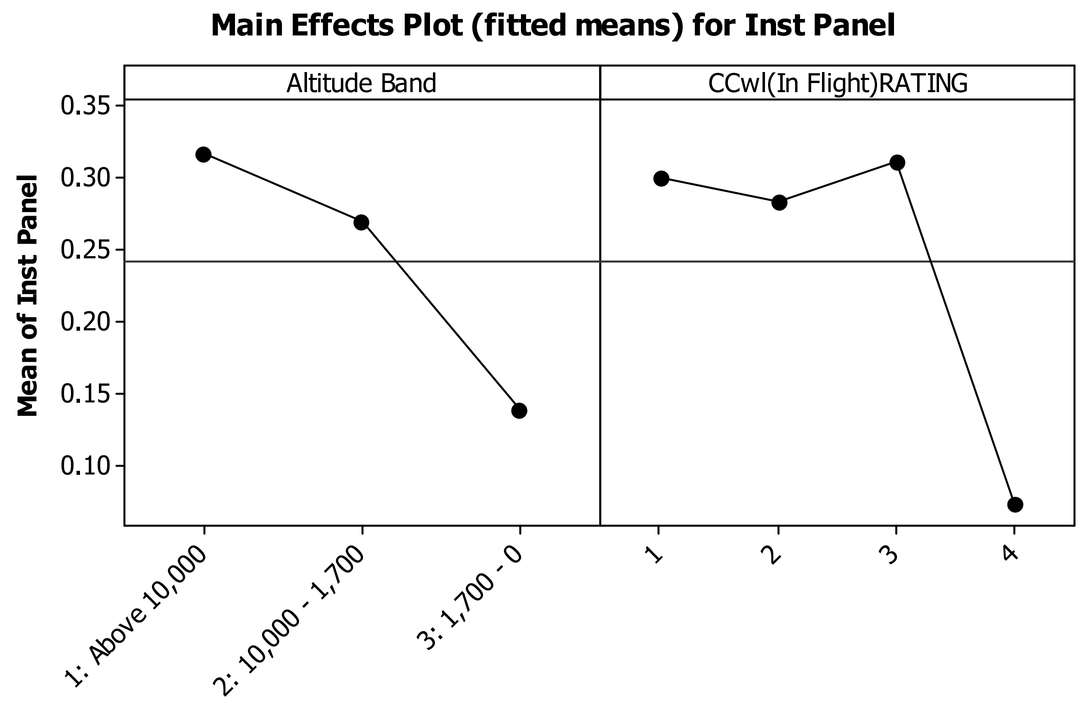


Figure D 41. Instrument Panel PDT Difference Main Effects - (PF - PM)

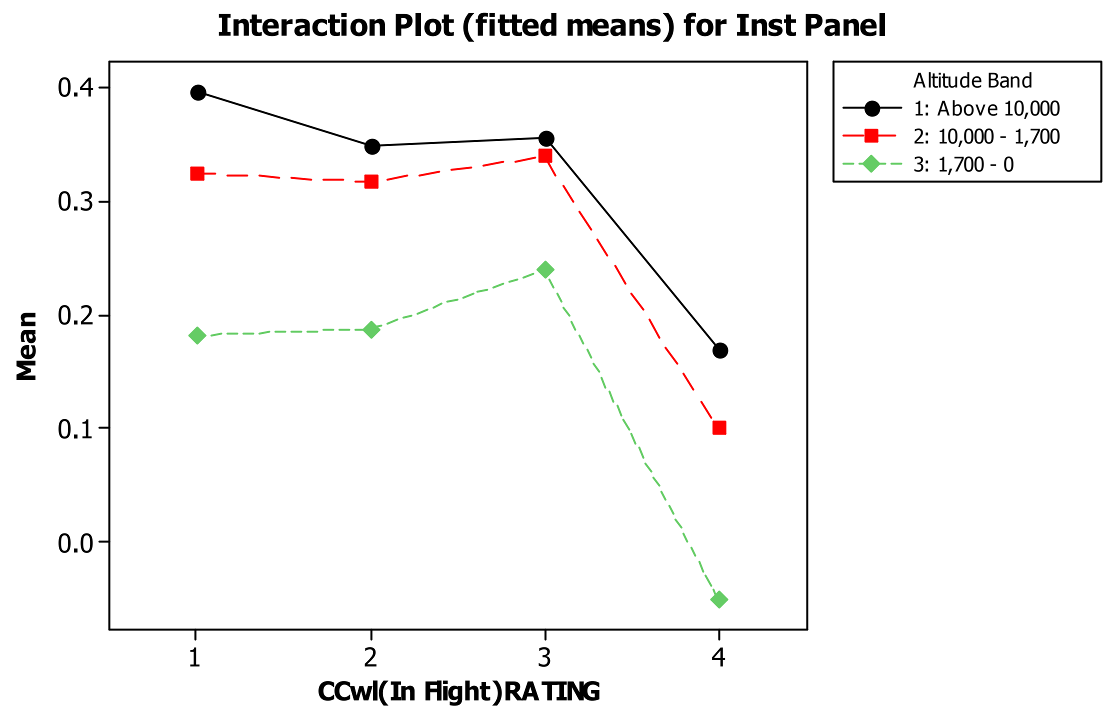


Figure D 42. Instrument Panel PDT Difference - (PF - PM)

www.manaraa.com

## General Linear Model: PFD versus Altitude Band, CCwl(In Flight)RATING

```
Factor         Type  Levels Values
Altitude Band     fixed    3 1: Above 10,000, 2: 10,000 - 1,700, 3:
                   1,700 - 0
CCwl(In Flight)RATING fixed    4 1, 2, 3, 4


Analysis of Variance for PFD, using Adjusted SS for Tests

Source                DF  Seq SS  Adj SS  Adj MS   F
Altitude Band             2  0.13472 0.07702 0.03851  1.21
CCwl(In Flight)RATING         3  1.59684 1.58537 0.52846 16.55
Altitude Band*CCwl(In Flight)RATING  6  0.02917 0.02917 0.00486  0.15
Error                282  9.00543 9.00543 0.03193
Total                293 10.76615

Source                  P
Altitude Band            0.301
CCwl(In Flight)RATING        0.000
Altitude Band*CCwl(In Flight)RATING 0.989
Error
Total


S = 0.178701  R-Sq = 16.35%  R-Sq(adj) = 13.09%
```

```
Tukey Simultaneous Tests
Response Variable PFD
All Pairwise Comparisons among Levels of Altitude Band
Altitude Band = 1: Above 10,000 subtracted from:

          Difference    SE of      Adjusted
Altitude Band    of Means Difference T-Value  P-Value
2: 10,000 - 1,700  -0.00881   0.03551  -0.248  0.9667
3: 1,700 - 0     -0.05132   0.03552  -1.445  0.3180


Altitude Band = 2: 10,000 - 1,700 subtracted from:

        Difference    SE of      Adjusted
Altitude Band  of Means Difference T-Value  P-Value
3: 1,700 - 0   -0.04252   0.03513  -1.210  0.4471
```

```
Tukey Simultaneous Tests
Response Variable PFD
All Pairwise Comparisons among Levels of CCwl(In Flight)RATING
CCwl(In Flight)RATING = 1 subtracted from:

CCwl(In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
2         -0.1242   0.02422  -5.128  0.0000
3         -0.1039   0.04376  -2.375  0.0819
4         -0.2617   0.04033  -6.490  0.0000
```

www.manaraa.com

```
CCwl(In Flight)RATING = 2 subtracted from:

CCwl(In     Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
3          0.0203   0.04139  0.490  0.9614
4          -0.1375   0.03774  -3.643  0.0015


CCwl(In Flight)RATING = 3 subtracted from:

CCwl(In     Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
4          -0.1578   0.05247  -3.007  0.0140
```

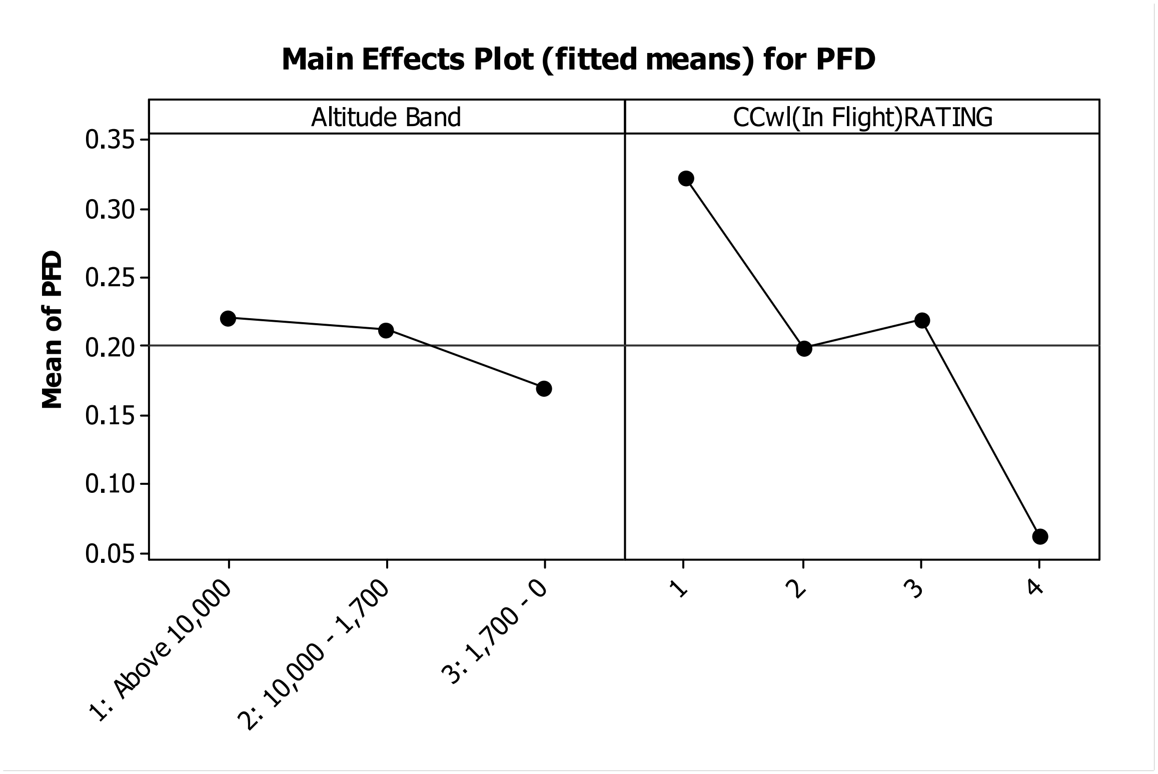


**Figure D 43. PFD PDT Difference Main Effects - (PF - PM)**

www.manaraa.com

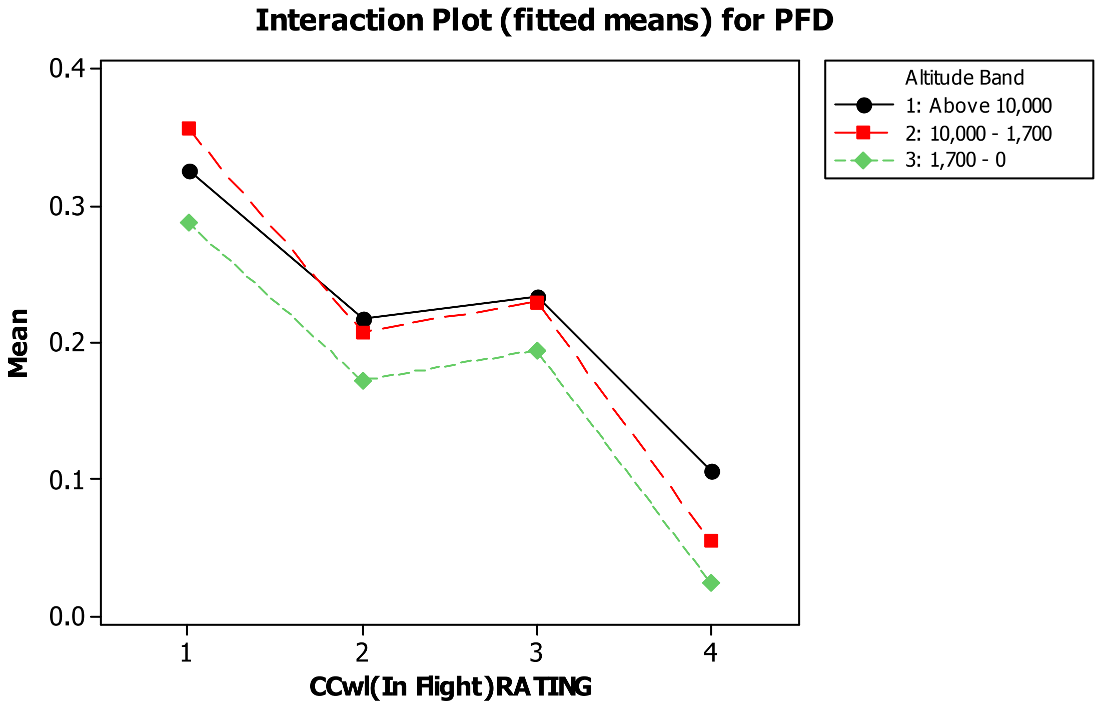


**Figure D 44. PFD PDT Difference Interaction Plot - (PF - PM)**

### D2.3 Shared AOI (5 Second Frame)

## General Linear Model: Inst Panel versus Altitude Band, CCwl (In Flight)

```
Factor           Type   Levels Values
Altitude Band       fixed    3 1 Above 10,000, 2 10,000 - 1,700, 3
                        Below 1700
CCwl (In Flight)RATING fixed    4 1, 2, 3, 4


Analysis of Variance for Inst Panel, using Adjusted SS for Tests

Source                 DF   Seq SS   Adj SS   Adj MS    F
Altitude Band              2  1.76010 0.96324 0.48162 13.88
CCwl (In Flight)RATING         3  0.45375 0.45375 0.15125   4.36
Altitude Band*CCwl (In Flight)RATING  6  0.05366 0.05366 0.00894  0.26
Error                 285  9.89243 9.89243 0.03471
Total                 296 12.15994

Source                  P
Altitude Band             0.000
CCwl (In Flight)RATING        0.005
Altitude Band*CCwl (In Flight)RATING 0.956
```

www.manaraa.com

```
Error
Total


S = 0.186307  R-Sq = 18.65%  R-Sq(adj) = 15.51%


Tukey Simultaneous Tests
Response Variable Inst Panel
All Pairwise Comparisons among Levels of Altitude Band
Altitude Band = 1 Above 10,000 subtracted from:

         Difference    SE of      Adjusted
Altitude Band    of Means Difference T-Value  P-Value
2 10,000 - 1,700   0.0556   0.03661  1.520  0.2817
3 Below 1700     -0.1321   0.03661  -3.608  0.0009


Altitude Band = 2 10,000 - 1,700 subtracted from:

        Difference    SE of      Adjusted
Altitude Band  of Means Difference T-Value  P-Value
3 Below 1700   -0.1877   0.03661  -5.128  0.0000



Tukey Simultaneous Tests
Response Variable Inst Panel
All Pairwise Comparisons among Levels of CCwl (In Flight)RATING
CCwl (In Flight)RATING = 1 subtracted from:

CCwl (In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
2         -0.0114   0.02520  -0.452  0.9693
3         0.0224   0.04562  0.491  0.9612
4         -0.1376   0.04140  -3.323  0.0049


CCwl (In Flight)RATING = 2 subtracted from:

CCwl (In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
3         0.0338   0.04312  0.783  0.8620
4         -0.1262   0.03863  -3.266  0.0060


CCwl (In Flight)RATING = 3 subtracted from:

CCwl (In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
4         -0.1600   0.05421  -2.951  0.0167
```

www.manaraa.com

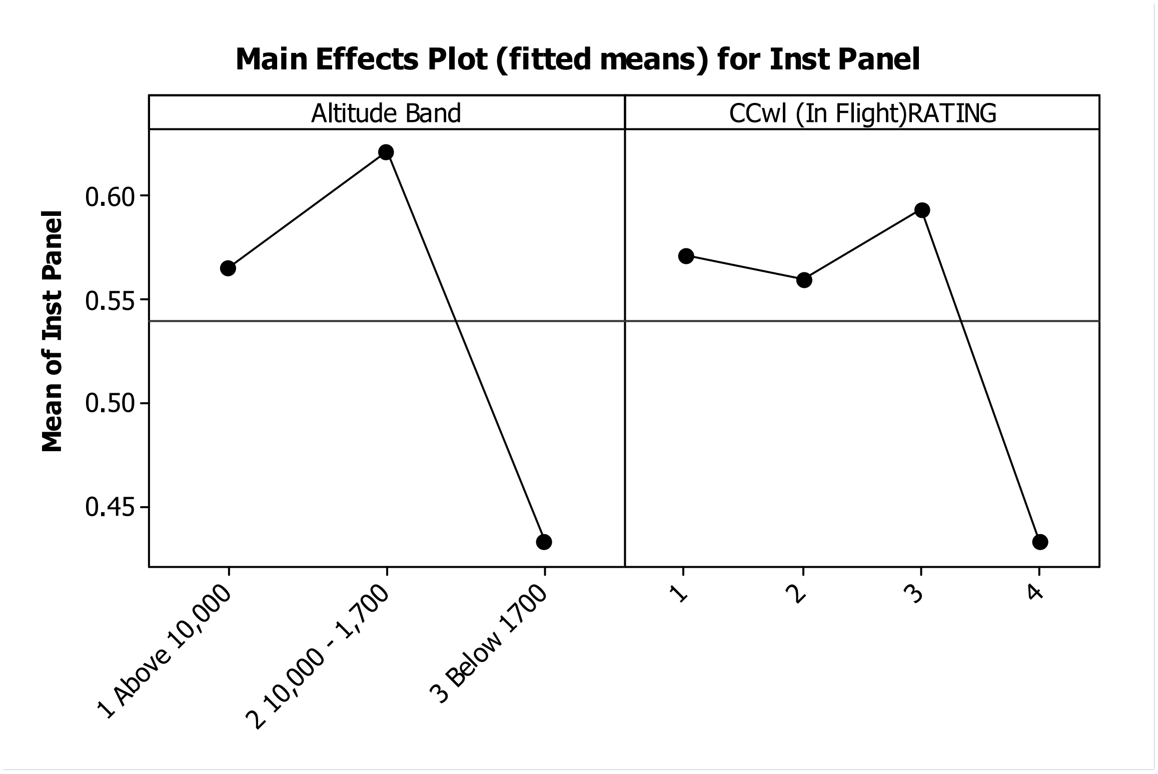


Figure D 45. Instrument Panel PDT Difference Main Effects - (PF - PM)

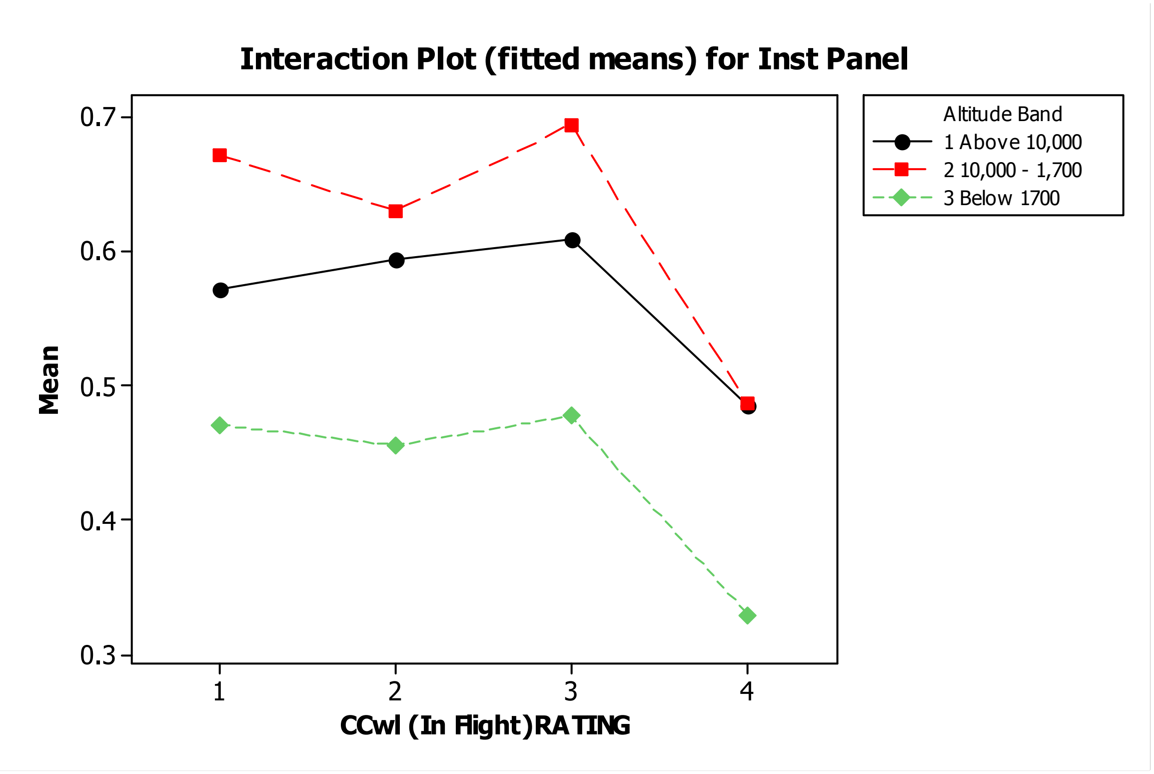


Figure D 46. Instrument Panel PDT Difference Interaction Plot - (PF - PM)

المنارة للاستشارات
www.manaraa.com

## General Linear Model: OutTheWindow versus Altitude Ban, CCwl (In Fli

```
Factor         Type   Levels Values
Altitude Band      fixed    3 1 Above 10,000, 2 10,000 - 1,700, 3
                    Below 1700
CCwl (In Flight)RATING fixed    4 1, 2, 3, 4


Analysis of Variance for OutTheWindow, using Adjusted SS for Tests

Source                 DF  Seq SS  Adj SS  Adj MS    F
Altitude Band              2 2.17586 1.12291 0.56146 127.66
CCwl (In Flight)RATING         3 0.02470 0.02470 0.00823  1.87
Altitude Band*CCwl (In Flight)RATING  6 0.02551 0.02551 0.00425  0.97
Error                 285 1.25348 1.25348 0.00440
Total                 296 3.47956

Source                  P
Altitude Band             0.000
CCwl (In Flight)RATING        0.134
Altitude Band*CCwl (In Flight)RATING 0.448
Error
Total


S = 0.0663189  R-Sq = 63.98%  R-Sq(adj) = 62.59%
```

**Tukey Simultaneous Tests**
**Response Variable OutTheWindow**
**All Pairwise Comparisons among Levels of Altitude Band**

```
Altitude Band = 1 Above 10,000 subtracted from:

         Difference    SE of      Adjusted
Altitude Band    of Means Difference T-Value  P-Value
2 10,000 - 1,700  -0.01301   0.01303 -0.9984  0.5779
3 Below 1700     0.17346   0.01303 13.3116  0.0000


Altitude Band = 2 10,000 - 1,700 subtracted from:

        Difference    SE of      Adjusted
Altitude Band  of Means Difference T-Value  P-Value
3 Below 1700    0.1865   0.01303  14.31  0.0000
```

**Tukey Simultaneous Tests**
**Response Variable OutTheWindow**
**All Pairwise Comparisons among Levels of CCwl (In Flight)RATING**

```
CCwl (In Flight)RATING = 1 subtracted from:

CCwl (In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
2        -0.01155  0.008971  -1.287  0.5709
3        -0.00397  0.016240  -0.244  0.9949
4        -0.03391  0.014738  -2.301  0.0978


CCwl (In Flight)RATING = 2 subtracted from:
```

www.manaraa.com

```
CCwl (In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
3          0.00758   0.01535  0.494  0.9605
4         -0.02236   0.01375  -1.626  0.3639


CCwl (In Flight)RATING = 3 subtracted from:

CCwl (In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
4         -0.02994   0.01930  -1.552  0.4066
```

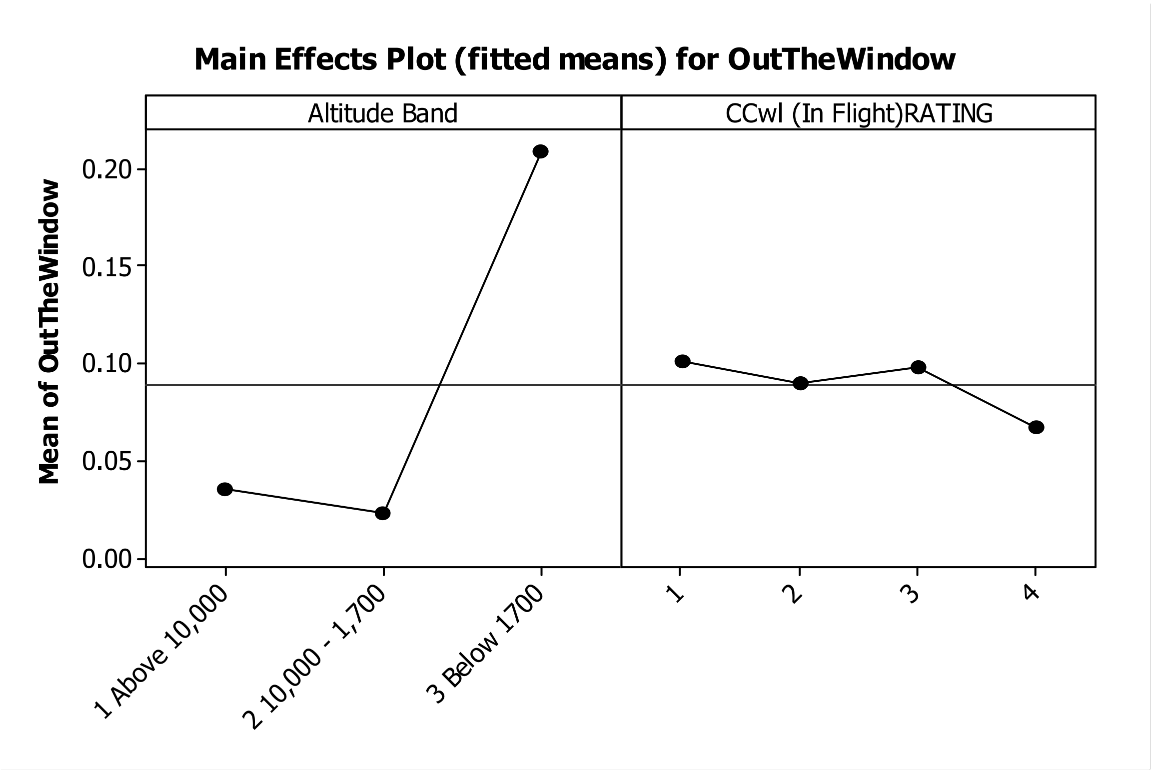


Figure D 47. Out the Window PDT Difference Main Effects - (PF - PM)

المنارة للاستشارات

www.manaraa.com

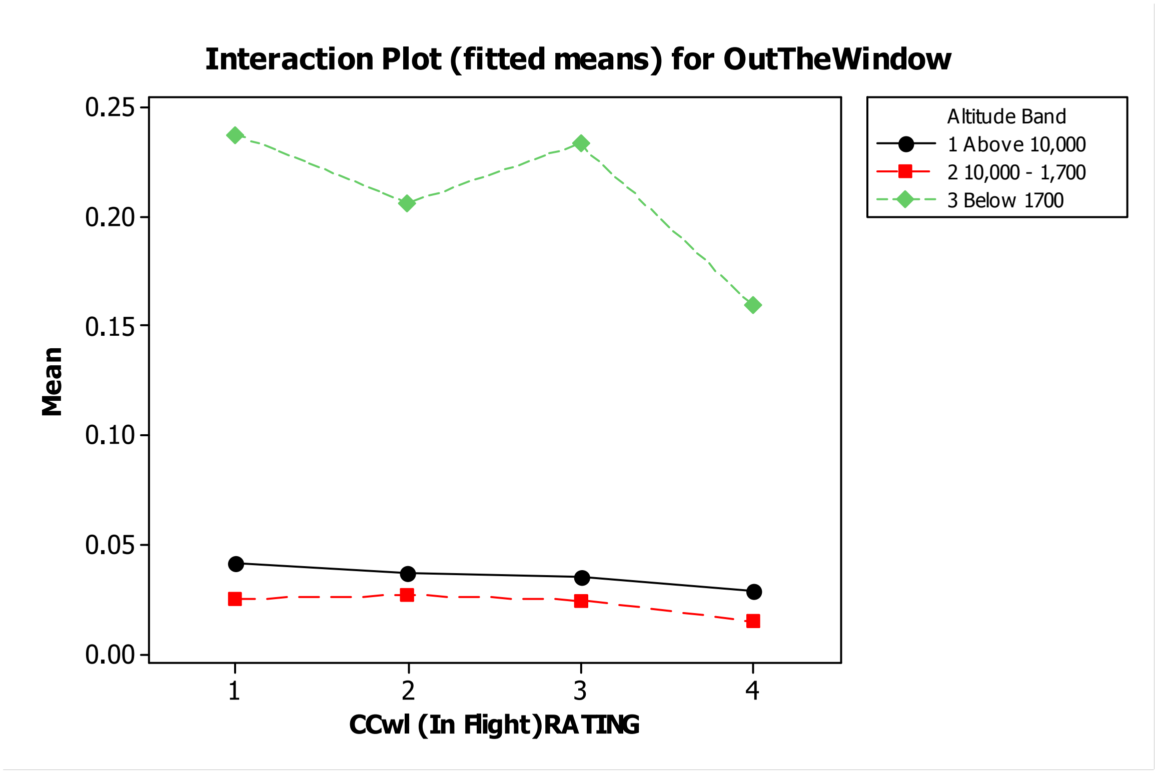


Figure D 48. Out the Window PFD Difference - (PF - PM)

## General Linear Model: PFD versus Altitude Band, CCwl (In Flight)RATING

```
Factor              Type  Levels Values
Altitude Band          fixed    3 1 Above 10,000, 2 10,000 - 1,700, 3
                        Below 1700
CCwl (In Flight)RATING fixed    4 1, 2, 3, 4


Analysis of Variance for PFD, using Adjusted SS for Tests

Source                     DF  Seq SS  Adj SS  Adj MS    F
Altitude Band                  2 0.26015 0.11940 0.05970  3.98
CCwl (In Flight)RATING             3 0.49032 0.49032 0.16344 10.89
Altitude Band*CCwl (In Flight)RATING  6 0.06640 0.06640 0.01107  0.74
Error                     285 4.27863 4.27863 0.01501
Total                     296 5.09550

Source                      P
Altitude Band                 0.020
CCwl (In Flight)RATING            0.000
Altitude Band*CCwl (In Flight)RATING 0.620
Error
Total


S = 0.122526  R-Sq = 16.03%  R-Sq(adj) = 12.79%
```

www.manaraa.com

**Tukey Simultaneous Tests**
**Response Variable PFD**
**All Pairwise Comparisons among Levels of Altitude Band**

```
Altitude Band = 1 Above 10,000 subtracted from:

         Difference    SE of      Adjusted
Altitude Band    of Means Difference T-Value  P-Value
2 10,000 - 1,700   0.06348   0.02408  2.637  0.0228
3 Below 1700     0.05260   0.02408  2.185  0.0738


Altitude Band = 2 10,000 - 1,700 subtracted from:

        Difference    SE of      Adjusted
Altitude Band  of Means Difference T-Value  P-Value
3 Below 1700   -0.01088   0.02408 -0.4518  0.8936
```

**Tukey Simultaneous Tests**
**Response Variable PFD**
**All Pairwise Comparisons among Levels of CCwl (In Flight)RATING**

```
CCwl (In Flight)RATING = 1 subtracted from:

CCwl (In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
2         -0.0729   0.01657  -4.401  0.0001
3         -0.0486   0.03000  -1.621  0.3666
4         -0.1393   0.02723  -5.117  0.0000


CCwl (In Flight)RATING = 2 subtracted from:

CCwl (In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
3         0.02430   0.02836  0.857  0.8270
4        -0.06639   0.02540  -2.613  0.0444


CCwl (In Flight)RATING = 3 subtracted from:

CCwl (In    Difference    SE of      Adjusted
Flight)RATING  of Means Difference T-Value  P-Value
4        -0.09069   0.03565  -2.544  0.0534
```

www.manaraa.com

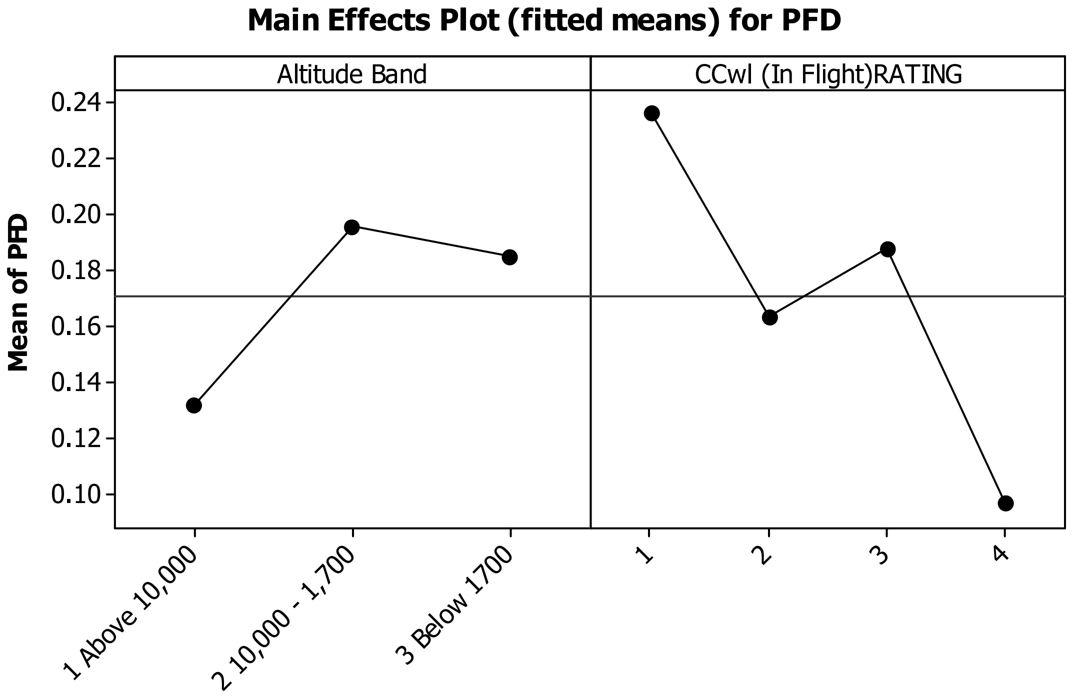


Figure D 49. PFD PDT Difference Main Effects - (PF - PM)

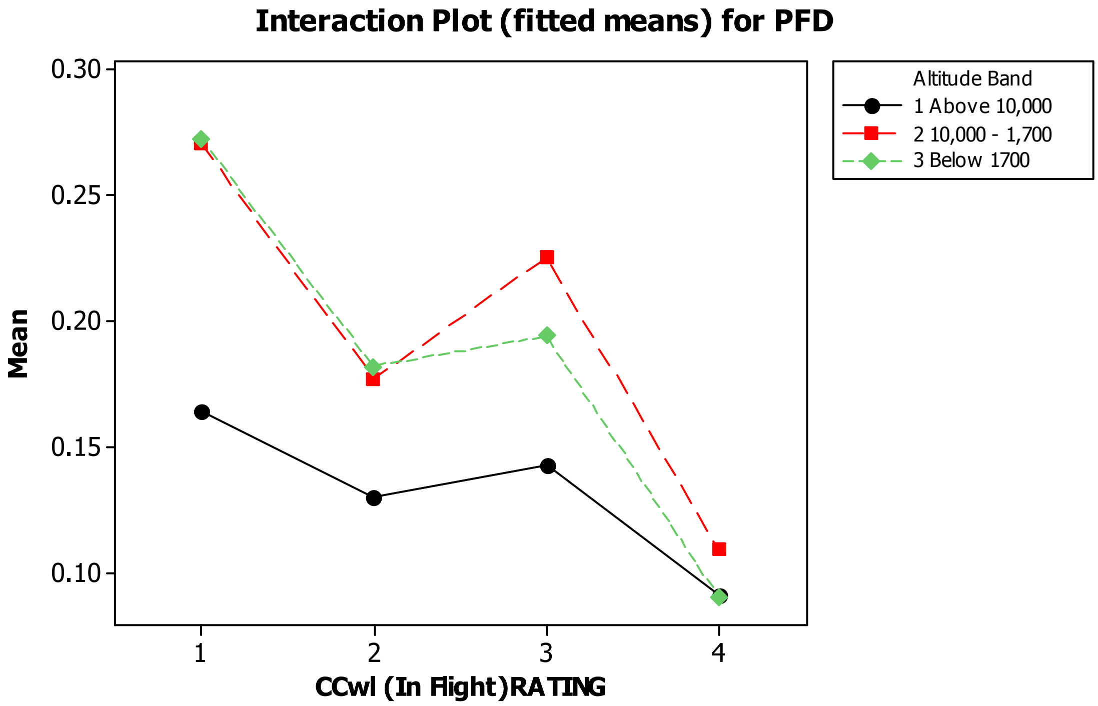


Figure D 50. PFD PDT Difference Interaction Plot - (PF - PM)

المنارة للاستشارات
www.manaraa.com