12-2008

A Flight Test Study to Assess the Utility of an Aircraft Referenced 3D Audio Display to Improve Pilot Performance under High Workload Conditions

Mehendi Hassan Naqvi

University of Tennessee - Knoxville

Recommended Citation

Naqvi, Mehendi Hassan, "A Flight Test Study to Assess the Utility of an Aircraft Referenced 3D Audio Display to Improve Pilot Performance under High Workload Conditions. " Master's Thesis, University of Tennessee, 2008.
https://trace.tennessee.edu/utk_gradthes/481

To the Graduate Council:

I am submitting herewith a thesis written by Mehendi Hassan Naqvi entitled "A Flight Test Study to Assess the Utility of an Aircraft Referenced 3D Audio Display to Improve Pilot Performance under High Workload Conditions." I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Aviation Systems.

Richard Ranaudo, Major Professor

We have read this thesis and recommend its acceptance:

Stephan Corda, Peter Solies

Accepted for the Council:
Carolyn R. Hodges

Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

To the Graduate Council:

I am submitting herewith a thesis written by Mehndi Hassan Naqvi entitled “A Flight Test Study to Assess the Utility of an Aircraft Referenced 3D Audio Display to Improve Pilot Performance under High Workload Conditions.” I have examined the final electronic copy of this thesis for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Master of Science, with a major in Aviation Systems.

Richard Ranaudo
Major Professor

We have read this thesis
and recommend its acceptance:

Dr. Stephan Corda

Dr. Peter Solies

Accepted for the Council:

Carolyn R. Hodges
Vice Provost and Dean of the Graduate School

(Original signatures are on file with official student records.)

A Flight Test Study to Assess the Utility of an Aircraft Referenced 3D Audio Display to Improve Pilot Performance under High Workload Conditions

A Thesis
Presented for the
Master of Science
Degree
The University of Tennessee, Knoxville

Mehndi Hassan Naqvi
December 2008

DEDICATION

This thesis is dedicated to my wife for all her unfaltering love, support, and encouragement and to my son for all his joyous distractions.

ACKNOWLEDGEMENTS

I must acknowledge Professor Richard Ranaudo for his hard work and direct contribution to this study. Thank you for sharing your knowledge and experiences and providing your guidance.

I would also like to acknowledge the administrative and technical staffs at the University of Tennessee Space Institute, the Middle Tennessee State University and the Canadian Defense Academy for their work to ensure I had all the resources I needed in a timely manner.

And finally, I'd like to acknowledge the evaluation pilots from Middle Tennessee State University Professional Pilot program for volunteering their time to this study.

ABSTRACT

A study to assess the utility of an aircraft referenced 3D audio display was undertaken to determine if there could be any improvements to pilot performance when operating under high workload conditions. Test subjects flew a general aviation light twin-engine aircraft under simulated single-pilot instrument flight rule conditions. Workload was elevated by ensuring each test subject had to execute an unexpected missed approach procedure and simultaneously handle a simulated engine failure. Subjective data was gathered using the NASA Task Load Index and a post-flight questionnaire on perceived performance, workload and situational awareness. Objective data on pilot performance was gathered using the research aircraft's onboard instrumentation system. Within the limitations of having a low number (5) of test subjects available, subjective data results showed a perceived increase in situational awareness, performance, and a statistically significant reduction in workload. Although not statistically significant, the only objective impact to performance was a slight increase in heading control and course intercept. There was no corresponding performance increase in airspeed control, angle of bank control, or improvements to aircraft track. Overall, the results indicate that a 3D audio display would have utility and pilot acceptance as a supplemental navigational display, but would not result in any substantial improvements to pilot performance.

TABLE OF CONTENTS

Chapter Page

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

α	Significance Level
n	Sample Size
AFRL	Air Force Research Laboratory
AGL	Above Ground Level
AHRS	Attitude and Horizontal Reference System
ANR	Active Noise Reduction
ATC	Air Traffic Control
ATF	Anatomical Transfer Function
CDI	Course Deviation Indicator
EP	Evaluation Pilot
FAR	Federal Aviation Regulation
FAF	Final Approach Fix
FTE	Flight Test Engineer
GA	General Aviation
GPS	Global Positioning System
HRTF	Head-Related Transfer Function
HSI	Horizontal Situation Indicator
IAF	Initial Approach Fix
ICS	Intercommunications System
IFR	Instrument Flight Rules
IMC	Instrument Metrological Conditions
IPSS	Internet Protocol Sound Lab Server
KIAS	Nautical miles per hour Indicated Airspeed
KTHA	Tullahoma Regional Airport, Tullahoma, Tennessee
MAWP	Missed Approach Way Point
MDA	Minimum Descent Altitude
MSL	Mean Sea Level

MTSU	Middle Tennessee State University
NM	Nautical Miles
PAR	Precision Approach Radar
PCMCIA	Personal Computer Memory Card International Association
POH	Pilot Operating Handbook
RHCB	Battlespace Acoustics Branch
RNAV	Area Navigation
RWY	Runway
SLAB	Sound Lab
SM	Statue Miles
SP	Safety Pilot
SPIFR	Single Pilot Instrument Flight Rules
SWAT	Subjective Workload Assessment Technique
TLX	Task Load Index
UTSI	University of Tennessee Space Institute
VFR	Visual Flight Rules
VNAV	Vertical Navigation

1. INTRODUCTION

In General Aviation (GA) one of the most demanding and intense flying situations occurs when operations are conducted as a Single-Pilot operating under Instrument Flight Rules (SPIFR)[1]. While flying SPIFR the GA pilot must independently fly the aircraft, handle all communications, manage all navigational systems and monitor all aircraft equipment[2]. Without a co-pilot available to assist during spikes in workload, to help in the maintenance of situational awareness or simply offer a second sober opinion, SPIFR requires near perfect performance right from flight planning to engine shutdown[1,2]. No other type of GA flying requires as much concentration, or skill, as SPIFR[1].

The situation is exacerbated when the SPIFR pilot must also contend with a critical emergency, such as an engine failure, during a critical phase of flight, such as during the missed approach procedure. Even during normal aircraft operations the execution of missed approaches can represent an elevated level of workload. In such a situation, the pilot needs to contend with power, aircraft configuration and airspeed changes while ensuring any altitude clearance limits and airframe restrictions are not exceeded. In some cases, the pilot may also need to ensure that noise abatement procedures have been considered and can be complied with. Additionally the pilot needs to re-configure navigational aids, develop the proper positional awareness relative to a new navigational fix, and make appropriate radio transmissions. These factors may also be affected by the elevated emotional stress associated with having poorly executed an approach, deteriorating weather, and the potential of having to divert to an alternate airport.

Throughout this intense experience the SPIFR pilot is gathering nearly every piece of information needed through the visual channel. Piloting an aircraft has always been, and will most likely always remain, a visually intensive activity[3,4,5]. This is true whether operating in Visual Meteorological Conditions (VMC) or Instrument Meteorological Conditions (IMC); pilots have always primarily used the visual channel to obtain the information they need to safely fly the aircraft. This is, of course, because visual displays have traditionally been the means by which information is provided in aviation. Thus, the modern GA cockpit presents a significant amount of information to the pilot through the visual modality. There are some secondary displays that use the aural modality, such as system alerts, alarms and warnings. However, in GA there is no commercially available audio display that takes advantage of the aural attentiveness and localization capabilities of the human auditory system.

Over the past several decades, the Air Force Research Laboratory's Battlespace Acoustics Branch (AFRL/RHCB) at Wright-Patterson AFB, Ohio, has made significant advancements in the field of 3D audio displays[6,7]. Experiments conducted by AFRL/RHCB have resulted in the development of audio displays that provide spatial localized cues to a pilot's headset, as opposed to conventional stereo or monaural systems[3,4]. Some of the potential advantages of 3D audio displays which have been demonstrated in the laboratory, or in laboratory conditions, include decreased reaction time, reduced work load and increased performance for visual scanning tasks[8] and supplemental navigation[4] or aircraft control[9].

In early 2007 the University of Tennessee Space Institute (UTSI) established a technical information exchange partnership with the AFRL/RHCB to design and install an aircraft referenced 3D audio display system in its Piper PA-31 Navajo research aircraft. The 3D audio display system was designed to provide spatialized aural cues that appear to be outside the pilots head and coincident with selected navigational waypoint.

A preliminary research effort by UTSI was conducted in late 2007 with this 3D audio display system to evaluate up and away navigation, workload, pilot performance, and situational awareness[10]. Results from that experiment indicated that the aircraft referenced 3D audio display was not a hindrance to any cockpit tasks and may be helpful when workload is very high. However, it was concluded that the benefits of the 3D audio display on pilot performance and situational awareness were not realized due to the relatively low workload of the flight task[10].

1.1. Purpose, Objectives, and Scope

The purpose of this experiment was to explore the utility of an aircraft referenced 3D audio display in real-world scenarios and address the recommendation from previous 3D audio display research conducted at UTSI by Wigdhal[10] by increasing test subject workload.

The objectives of this study were to assess the higher workload presented to each test subject (also referred to as evaluation pilots) in the following manner:

1. Determine if a difference in pilot subjective workload exists between the 3D audio display off condition and the 3D audio display on condition by using the NASA Task Load Index (TLX)[11];

2. Assess pilot opinion on 3D audio display system implementation, workload from using the 3D audio display, impact on flight and navigational performance, and impact of the 3D audio display on situational awareness;

3. Quantitatively evaluate the difference in pilot technical performance between the non-3D audio display condition and the 3D audio display condition by measuring tracking task performance of angle of bank, indicated airspeed and intercept heading; and

4. Qualitatively evaluate the difference in pilot technical performance between the non-3D audio display condition and the 3D audio display condition by plotting the flight path of the aircraft when turning to the initial fix in the revised missed approach procedure.

The scope of this experiment required that:

1. The evaluation pilots fly actual IFR approaches in a SPIFR scenario under simulated instrument meteorological conditions (IMC);

2. The evaluation pilots were exposed to a high workload condition, which consisted of the tasks associated with the execution of a missed approach procedure while handling a simulated engine failure;

3. An unannounced power cut to zero-thrust be used to simulate the engine failure at a critical and high workload point during each test run;

2. BACKGROUND

2.1. Audio Displays in Aviation

Although the pilot's vision is the primary information channel in aviation, audio displays have become a normal part of nearly all GA, airline and military aircraft cockpits. The main reason for this is that audio displays: allow for the quicker assimilation of information regardless of the pilot's gaze, they are usually less expensive than the equivalent visual display, and they require little to no instrument panel area[12].

Based on the literature reviewed for this experiment, there are currently two different types of audio displays. These are non-speech audio displays and speech audio displays[12]. Non-speech audio displays typically consist of horns, buzzers, tones, and other various electronically generated sounds at specific frequencies that grab the attention[12] when played over the headset or the cockpit loud speaker. The speech audio displays are computer synthesized human voice, recorded human voice, or even real-time human voice played over the headset.

We use our natural auditory system to acquire information via sound waves and to assist in our gaze. Similarly, audio displays are typically used in two main ways in aviation. The first is when the audio display is used strictly to provide information about the state of the aircraft or a system on the aircraft. This is a system state audio display used to draw the attention, but not necessarily the visual attention, of the pilot. Typically these are non-speech displays, but more often, modern aircraft tend to employ speech displays. An example of non-speech audio display is a stall horn used on most GA aircraft. This alert sounds when the wing angle of attack reaches some pre-determined value close to the stall angle of attack. An example of a speech display is the "ENGINE FIRE RIGHT, ENGINE FIRE RIGHT" alert in the CF18 Hornet fighter aircraft. In both cases, the audio display is compelling enough to draw the pilot's attention to the system state it is presenting (angle of attack or engine fire), but does not spatially localize the system in terms of where the problem exists. That is, this type of audio display only provides information to the pilot aurally and does not direct his or her gaze.

The second way audio displays have been used in aviation is to provide critical information with the intent of directing the gaze of the pilot. These “aid-to-visual scan” audio displays are designed to compel the pilot to look directly at a system of interest. An example is the “C” chord, which sounds when approaching or deviating from an altitude selected by the pilot, thus forcing the visual scan to the altitude display. Another is the Traffic Advisory and Resolution Advisory audio alerts in Traffic Alert and Collision Avoidance Systems (TCAS), which compels the pilot to look at the visual display for more information about the conflicting traffic’s relative bearing, range and altitude. In this case, the audio display is compelling enough to get the pilot’s visual attention onto the system of interest.

The real-time human voice, although not commonly mentioned in the literature that was reviewed, is very often used as an aid-to-visual scan audio display. Some examples are when the pilot is flying on vectors provided by Air Traffic Control (ATC), or when conducting a precision-approach-radar (PAR) approach, or when receiving traffic information from ATC (e.g. “…traffic at your 2 o’clock for 3 miles…”). In this case, the real-time human voice is compelling enough to get the pilot’s visual attention on the system of interest, be it a specific aircraft heading (ATC vectors), a specific course and glideslope (PAR approach), or another conflicting aircraft.

Current audio displays used by GA are limited by the fact that they only utilize our ability to hear, but not our ability to localize sound. Sound localization is our ability to determine the relative spatial direction from which the sound source emanates. Sound localization has the obvious advantage of cueing our gaze. When someone calls our name, we naturally turn and look in the direction of the sound source. The impulse to look at the source of unexpected sounds is typically unavoidable. Sound localization also gives us a sense of spatial awareness relative to a known sound source. It is this sound localization and spatial relationship that a 3D audio display can provide. The utility of such a system in the context of a high pilot workload scenario was explored in this study.

2.2. 3D Audio Displays

Although aviation has regularly used auditory displays, they have traditionally not contained any spatial cuing or been designed to take advantage of our natural sound localization abilities. If the sound were to emanate from the system of interest in an aid-to-visual scan audio display the rate of information transfer, along with task performance, might go up. With this premise in mind 3D audio displays (also referred to as spatial audio[3,7,13] or virtual audio[4]) and their application to aviation were studied by various institutions over the past two decades.

At the time of this writing, 3D audio displays in the public domain have only been discussed in the context of research and experimentation; there is no known 3D audio display or application actually employed in routine operations. Most of the known experiments with 3D audio displays have been conducted in the laboratory or in simulators for possible aviation applications[7]. However, other research into 3D audio has been done for augmented reality[14], computer gaming[15], and human machine interfaces for the visually impaired[16].

Within the context of aviation, experiments into 3D audio displays have typically fallen into the two main audio display classes that were described earlier in this paper; information on "system state" and "aid-to-visual scan" displays.

The majority of the studies and experiments with 3D audio displays appear to be with the aid-to-visual scan class of displays. This class of 3D audio display compels the pilot to put full visual attention onto the system of interest. This typically involves some kind of a target search, as discussed in references 5, 8, 17, and 18, where results of experiments showed improved reaction time performance.

Other experiments have focused on using this technology towards the "system state" situation, where the audio display compels the pilot to place attention onto the system of interest, but not necessarily any visual attention. Work in this area has studied the use of a 3D audio display to control of aircraft roll, pitch, yaw, airspeed and vertical velocity. The ultimate goal of these studies was to determine if spatialized audio could supplement or by-pass the visual channel[3,9,19] for aircraft control tasks Flight simulation results showed that one can learn to control an aircraft with spatialized audio, but in all cases the visual channel was still found to be superior.

2.3. Sound Localization

Under natural surroundings those of us with normal hearing can localize sounds easily and rapidly. Sound localization in humans is widely accepted to be the result of eight factors, which are: interaural time difference, interaural intensity difference, pinnae response, shoulder echo, head motion, early echo response, reverberation and vision[13,20,21,22]. The goal of a 3D audio display is to emulate and synthesize all the natural qualities of a sound source that enable humans to localize. Although the process of sound localization has been studied extensively for over 100 years[23], it is only recently that computer technology has been able to produce sounds approaching the natural qualities that allow localization over headphones[7,13, 21].

For localization to be conducted over headphones there is a requirement for the development of a Head-Related Transfer Function (HRTF)[13,21,24]. The HRTF, also known as the Anatomical Transfer Function (ATF)[24,25], is a linear function that is developed by placing microphones in the auditory canals of a test subject, or of a manikin, and placing that test subject in an anechoic chamber. A pre-recorded sound with a known frequency spectrum from a known relative location in space is played and subsequently recorded by the left and right auditory canal microphones. Afterwards, the recordings from each ear are compared to the original sound source to compute the HRTF. This HRTF is only valid for the relative location used in the original recording. To emulate sounds from other directions, finite impulse response filters need to be developed for each location of interest. Therefore, if a complete three-dimensional localization is required in a virtual environment, recordings must be made from every point on a sphere around the test subject.

Individualized HRTF's typically yield the best localization results, with a laboratory minimum audible angle on the order of 5° in azimuth and 30° in elevation[26]. However, individualized HRTF's tend to be very costly and time consuming to obtain. Therefore, most experiments use a generic HRTF collected from a manikin. Typically, for gross sound localization (within ± 10°) where front/back confusion along the sagittal plane is not a significant factor, a generic HRTF is sufficient. This is usually the case when most sound sources are to emulate mainly from one side or the other on the horizontal plane containing the ears. For more precise azimuth localization, and for vertical localization, an individualized HRTF is necessary. With respect to this experiment, only localization in azimuth within the forward hemisphere was investigated. Therefore, non-individualized HRTF's were considered acceptable.

One important factor to consider is that sound localization is, apparently, a learned behavior. As initially studied by Hofman, Van Riswick, and Van Opstal[27] and further studied by Zahorik et al[28], it appears the original spatial maps that people develop over time can be relearned, and remarkably quickly too. This is most likely via aural-localization and visual feedback, as explored by Zahorik et al[28] with non-individualized HRTF's. The ability to relearn sound localization when using non-individualized HRTF's will most likely play a significant factor in determining if 3D audio displays can be made economically viable for commercial use.

3. METHODOLOGY

3.1. General

Five data collection test flights were conducted for this experiment, each averaging approximately 2.4 hours. All flight tests were conducted during the day under Visual Flight Rules (VFR). The test area was confined to airspace within 12 Nautical Miles (NM) surrounding the Tullahoma Regional Airport (KTHA) and between 1900 feet Above Ground Level (AGL) to 4900 feet AGL. The weather condition during all flights required, as a minimum, a ceiling of 5000 ft AGL and visibility of 5 Statue Miles (SM). The flight crew consisted of an evaluation pilot (EP), safety pilot, and flight test engineer (FTE).

A series of area navigation / global positioning system (RNAV/GPS) approaches and associated approach and missed approach waypoints were used to develop the test sequence for each data collection run. The GPS approaches that were used during the test flights were the RNAV (GPS) RWY 06, 18, 24 and 36 approaches for the KTHA airport. A sample of each approach is provided in Figure 1[*], Figure 2, Figure 3 and Figure 4.

Each flight consisted of six approaches and subsequent missed approach procedures with the 3D audio display off and six with the 3D audio display on, in an alternating fashion. During the final stages of each approach a revised missed approach procedure was given to the evaluation pilot. The evaluation pilot would acknowledge the revised missed approach procedure, and reprogram the Garmin GNS 530W GPS. The evaluation pilots would then advance the throttles to execute the revised missed approach. Almost immediately afterward, the evaluation pilot would be presented with a simulated engine failure. The evaluation pilot would continue flying the missed approach, simulate all immediate emergency procedure actions, and follow up with the required abnormal procedures checklist.

---

[*] All figures and tables are located in the Appendix.

Therefore, in this experiment, each evaluation pilot was exposed to the workload associated with a revised missed approach procedure, reprogramming the primary navigational aid, executing a simulated engine emergency and associated checklist, and flying a laterally out of trim aircraft with reduced climb performance. With all these tasks combined, the evaluation pilots were operating in a workload environment that was higher than that previously presented during on-aircraft 3D audio research[10].

All flights were conducted under simulated single-pilot, instrument flight rule (IFR) operations by the use of a view limiting device worn by the evaluation pilot. The view limiting device used was similar to a safety goggle, but had an opaque frosting on the top and sides to limit the view outside the cockpit. This type of view limiting device is commonly used for IFR training flights by many flight training institutions and provides the required simulation of instrument metrological conditions (IMC) to the evaluation pilot. This type of view limiting device was also chosen since it could be removed easily in the case of a real aircraft emergency and did not impair the safety pilot’s field of view out the left side of the aircraft.

3.2. Description of Test Aircraft

3.2.1. Basic Aircraft

The basic aircraft type from which the research aircraft was derived was a Piper PA-31 Navajo, shown in Figure 5. The basic PA-31 Navajo was a conventional, twin engine, multi-purpose GA aircraft with retractable landing gear that was available in non-pressurized and pressurized models. It was originally designed, built and marketed by Piper Aircraft Company and entered the GA aircraft market in the 1970’s. It was certified under Federal Aviation Regulation (FAR) Part 23 in the normal category and was approved to operate in day and night VFR conditions. Certain variants of the PA-31, which had appropriate optional equipment installed were operated in instrument and known icing conditions. The cockpit area of the PA-31 was designed to accommodate two pilots, however most often GA aircraft in its class have been used for air taxi and air charter operations with only one pilot. The cabin area of the PA-31 aircraft was typically configured to carry a total of 8 persons (including the pilots) or equivalent weight in cargo. The basic PA-31 typically had a maximum gross takeoff weight of 6500 lbs, a cruise speed of 170 Nautical Miles per hour (knots) true airspeed, and was equipped with two Lycoming TIO-540 engines and Hartzel constant-speed, variable-pitch propellers.

3.2.2. Test Aircraft Modifications / Instrumentation

Figure 6 and Figure 7 show the Piper PA-31 Navajo, registration number N11UT, which was used for these series of 3D audio display flight tests. The test aircraft was equipped with air data, body rate, attitude, control position, pilot force, and inertial sensors for use in stability and control testing. Specifically for the purposes of this evaluation the test aircraft was equipped with a Garmin GNS 530W GPS. This particular GPS system is also capable of providing vertical navigational (VNAV) guidance.

3.3. 3D Audio Display System

Integration of the 3D audio display system into the test aircraft was performed by UTSI. A schematic of this system is shown in Figure 8. System operation was facilitated by two laptop computers. Laptop 2, shown in Figure 8, functioned as a data acquisition computer. Inputs from the GNS 530W, the Attitude and Horizontal Reference System (AHRS), and the static and dynamic pressure transducers were feed into Laptop 2 through a four port PCMCIA adaptor card.

Laptop 1, also shown in Figure 8, was configured with the Internet Protocol Sound Lab Server (IPSS) software. The IPSS accessed the location of the active waypoint in spherical coordinates from the GNS 530W, provided by Laptop 2, and converted it into the appropriate calls to NASA’s Sound Lab (SLAB) software version 2.0.2 audio render. SLAB was the software that generated the spatial audio cue sent to the test aircraft’s stereo Intercommunications System (ICS) and, thus, into the headsets worn by the evaluation pilot.

The LabVIEW data acquisition program was run on Laptop 2 and was used to combine all the input data into a composite test file that was exported to the Microsoft EXCEL spreadsheet program for post flight data analysis. Data was recorded at a 1 Hz rate. Laptop 2 was also programmed to take information on the aircraft's location and the location of the active waypoint selected in the GNS 530W, convert this information into the relative spherical coordinates of the selected navigational waypoint's location, and feed this information to Laptop 1. Laptop 2 could also be used to change the characteristics of the audio display cues by requesting any of the stored sound sources from Laptop 1 and mute the display for the 3D audio off test points.

The other supporting pieces of equipment that made up the 3D Audio Display System were the three Bose active noise reduction (ANR) headsets that were used to provide the 3D audio display to the evaluation pilot, safety pilot, and flight test engineer. Also, a video camera recording system was installed in the cockpit to record evaluation pilot actions during the test flight. The video camera system was connected into the aircraft's ICS, thus recording all audio signals presented to the evaluation pilot, including the 3D audio display.

The entire 3D audio display system was operated by a FTE from a control station located in the cabin area of the aircraft. The FTE ensured the 3D audio display system was functioning properly and ensured the 3D audio cue was either off (muted) or on, as appropriate, at the start of each test run. Additionally, the FTE also performed real-time track file data integrity monitoring on Laptop 2 and recorded any events of note for post-flight analysis.

The 3D audio display system as installed in UTSI's Navajo was an aircraft referenced system, versus a head referenced system. Unlike head referenced systems, in an aircraft referenced system the pilot must be looking forward and essentially parallel to the longitudinal axis of the aircraft to correctly localize the 3D audio cues. Since GA pilots look forward most of the time when performing navigation and control tasks the aircraft referenced implementation may be a viable alternative to head referenced systems. Aircraft referenced systems also have the advantage of not requiring a head tracking device; making it lighter, less complex, less costly and more suitable for GA aircraft types. In addition to the potential benefits mentioned above, localized cues may be useful for improving pilot performance and situational awareness by distributing total workload across both visual and auditory modalities.

For a GA application, an aircraft referenced audio display system was considered suitable for the terminal phase of flight based on the findings of reference 3 and 4. Specifically, reference 4 found that for GA aircraft an aircraft referenced 3D audio display actually had better pilot performance than that of a head referenced 3D audio system. This result was also supported by the study performed by Wigdahl[10]. For the tasks explored in this study the pilot will typically spend most of the time scanning instruments on the forward panel, with occasional head movement to perform other mission or operational tasks.

3.3.1. 3D Audio Display Cue

The 3D audio cue that was used in this research effort was developed by the AFRL/RHCB and can best be described as a pulsetrain approximately one-half-second in total duration, consisting of three pulses of broadband white noise that is repeated every five seconds. The cue is a Microsoft wav file which resides on Laptop 1 and was called up by SLAB. It was chosen during the testing conducted at reference 10 as being the most salient 3D audio cue available at that time for the planned flight experiments.

3.3.2. Head Related Transfer Functions

The HRTF's used in this experiment were developed by AFRL/RHCB in their Auditory Localization Facility. This facility is a 4.3 meter-diameter geodesic sphere with 277 loudspeakers mounted on its inner surface. The loudspeakers are located approximately 15 degrees apart when viewed from the geometric center; however, only those speakers above -45 degrees elevation are used. The test model or subject for the HRTF's measurements stands, or sits, in the center of the sphere.

Due to time, cost, and shear practicality constraints of this experiment, individualized HRTF's for the evaluation pilots were not developed. Therefore, each evaluation pilot in this experiment used non-individualized HRTF's, which had been previously developed by AFRL/RHCB from testing with a live person. These HRTF's were processed using the "snowman" model described by Algazi, Duda & Thompson[29], and were contained in a binary file called "SnowmanDSB2.slh", which was used by the NASA's SLAB software that run on Laptop 1 (see Figure 8).

It should be noted that the same non-individualized HRTF's used by Wigdahl[10] were used in this experiment. As in the case of Wigdahl[10], this experiment also found satisfactory accuracy (± 10°) in the lateral localization of the 3D audio display cue during the familiarization training given to each evaluation pilot.

Although using individualized HRTF's would have been ideal, using non-individualized HRTF's was not expected to have a significant impact on the results for three main reasons. Firstly, all the sound localization planned for this experiment was to be conducted in azimuth along the horizontal plane containing the ears and no localization in elevation was planned. Secondly, all the sound cues were presented at angles no less than 25° from the sagittal plane, on either side. Thirdly, localization accuracy of ± 10° was considered acceptable for the purposes of this experiment.

3.4. Evaluation Pilots

All five evaluation pilots were selected from the Middle Tennessee State University (MTSU) Professional Pilot program. Each evaluation pilot had flying experience as detailed in Table 1, a current commercial pilot medical, and a valid commercial pilot license with current multi-engine and IFR ratings. Although the evaluation pilots had various total and multi-engine flight times, they were all considered low-time pilots trained to the same standard at the same institution. For the purposes of this experiment, the evaluation pilot group was considered homogenous.

3.5. Familiarization Training

Before they flew the test flight all five evaluation pilots were given approximately three hours of classroom time for a background and introduction to the test program. This included a briefing on the standard operating procedures for flying in the KTHA area, Piper Navajo cockpit layout, overview of the Garmin GNS 530W, normal and emergency operations for UTSI’s Piper PA-31 Navajo, and an overview of the test sequence and familiarization on the 3D audio display.

Each evaluation pilot was also given an in-cockpit ground familiarization with UTSI’s Piper PA-31 Navajo, which included reviewing cockpit layout, normal and emergency checklists and an opportunity to experience the 3D audio display while conducting basic localization tasks.

After the ground familiarization, each evaluation pilot was given approximately one hour of in-flight training. A summary of the training time provide to each evaluation pilot is provided in Table 2. The familiarization flight covered the basic aircraft normal and simulated emergency handling, practice in executing the flight tasks that will be performed during the actual test flight, and use of the 3D audio display while conducting basic waypoint localization and tracking tasks. The familiarization flight was also used to give each evaluation pilot exposure to completing the NASA TLX prior to using it during the actual test flights[30].

3.6. Test Flight Execution

Throughout the entire flight the safety pilot acted as the controlling agency providing all the necessary IFR clearances and instructions to the evaluation pilot at the appropriate time. By having the safety pilot simulate the communication inputs of ATC there was a controlled and consistent level of verbal communication (workload) between all evaluation pilots. The evaluation pilot made all appropriate radio calls to the safety pilot.

The detailed sequence of the 12 runs conducted on each flight is described in the flight test matrix, provided in Table 3. A summary of the date and evaluation time for each evaluation pilot is provided in Table 2. The details of how each run was conducted are described in the sample test card, provided in Figure 9. The items in bold on the sample test card shown in Figure 9 were changed for each run as per the parameters described in Table 3.

The first run was initiated with the aircraft positioned by the safety pilot at approximately 3 NM laterally offset from a point halfway between the Initial Approach Fix (IAF) and the Final Approach Fix (FAF). For safety purposes, the altitude flown on each run was 2000 feet above all altitudes published on the approach procedures contained in Figure 1 to Figure 4. The aircraft was trimmed for straight and level flight at 110 KIAS approach speed with the landing gear extended. At this point the safety pilot gave control of the aircraft to the evaluation pilot. Once the evaluation pilot had control of the aircraft the safety pilot read out the simulated ATC clearance for the approach, as provided on the appropriate test card. The evaluation pilot read back the clearance, loaded the correct approach in the GNS 530W and activated it. The evaluation pilot then conducted the pre-landing check.

At the FAF the safety pilot called for the FTE to turn on the data recording. The FTE ensured the recording system was on and the 3D audio display was set to the condition indicated on the test card. The evaluation pilot flew the approach as per the published approach procedure and established a 500 feet per minute rate of descent after passing the FAF.

Upon reaching the Minimum Descent Altitude (MDA), the safety pilot made a simulated ATC transmission stating that the weather at the airport was below approach minimums and provided missed approach instructions that were different than those published for the approach. The evaluation pilot read back the clearance, loaded the new approach and initial fix in the GNS 530W and activated it. Once the new approach was loaded in the GNS 530W and course information was provided the evaluation was to determine a target heading that would result in 30° course intercept. The evaluation pilot announced the target heading to the safety pilot and FTE, and then set climb power and commenced the turn towards the revised Missed Approach

Waypoint (MAWP) provided in the clearance. The importance of pre-determining a target heading that would result in a 30° course intercept was that it ensured that front-back confusion errors associated with non-individualized HRTF's were avoided (see section 2.3). As soon as the evaluation pilot commenced the turn the safety pilot simulated an engine failure by retarding the throttle on the appropriate engine as depicted on the test card for that run. The safety pilot set approximately 10 inches of manifold pressure on the simulated "failed" engine in order to simulate a zero thrust, feathered propeller condition. The evaluation pilot responded by carrying out all the necessary emergency items and flew the revised missed approach clearance as provided.

As soon as each evaluation pilot was able they turned to intercept the new course and navigated to the revised MAWP while completing the engine failure and engine restart checklists. Once the evaluation pilot had completed all checklist items and had stabilized on course to the revised MAWP (within 2 dots on the CDI display) for two minutes the safety pilot terminated the run. The safety pilot then took control of the aircraft from the evaluation pilot and called for the FTE to turn off data recording. The FTE ensured the recording system was turned off and the 3D audio cue was set to the next condition as indicated on the test card. The safety pilot then gave the evaluation pilot the NASA TLX workload survey form to fill out while repositioning the aircraft for the next run.

In order to maintain the same workload and randomness between runs, the evaluation pilots were not predisposed to which engine was to be simulated as failed nor did they have prior knowledge of what the new approach procedure was in the revised missed approach clearance. The evaluation pilot was also not presented with the same revised missed approach instructions consecutively. This was to avoid learning bias and potentially cause a pre-programmed response instead of a situation assessment. Also, the simulated failed engine was pre-selected such that a turn direct to the new initial fix presented a worst case turning performance problem for the evaluation pilot. That is, the revised missed approach instructions and the simulated engine failure were designed such that the aircraft had a natural turning tendency that was away from the shortest direction of turn toward the new initial fix.

As shown in the test matrix contained in the appendix, the 3D audio cue was alternated off and on between each subsequent run and specifically chosen so as to avoid a potential learning bias. The only exception was between runs 6 and 7, where the 3D audio display was on during both runs in order to ensure a proper flow in the sequence of test points. In the test planning process, this small, but necessary deviation was not anticipated to introduce any significant bias.

3.7. Test Flight Risk Assessment

All ground and flight risks associated with executing this test plan were mitigated to an acceptable level through UTSI's flight safety review process. In general, this process required that the Aviation Safety Committee initially assess previous safety findings for 3D audio display flight experiments and revise them as necessary to incorporate the identified hazards and risk assessments for this particular test. To that end, a new hazard associated with potential loss of control when simulating an engine failure during the missed approach procedure had to be mitigated by additional flight limitations and procedures. Mitigation was largely accomplished by: raising the altitudes for all the approach and missed approach procedures by 2000 ft, specifying safety pilot "take control procedures", providing narrow airspeed and bank limits, and a requirement that evaluation pilots demonstrate an adequate and safe level of aircraft handling during practice flights. These experiment controls are on record and specified in UTSI's airworthiness process. The complete safety and risk assessment is contained in a separate document on file in the Aviation Systems Program office. These documents were reviewed and accepted by the MTSU Internal Review Board before flight testing began.

3.8. Test Envelope

In addition to the flight envelope depicted in the Piper PA-31 Navajo Pilot Operating Handbook (POH) the following flight limitations were also adhered to:

1. The minimum altitude during any stage of the approach was no lower than that equal to the sum of the published altitude for that phase of the approach plus 2000 ft;

2. The safety pilot ensured that the airspeed did not inadvertently go below Blueline (94 KIAS);

3. The targeted airspeed during the entire missed approach procedure was Blueline + 5 KIAS; and

4. The maximum targeted bank angle during the simulated single engine missed approach procedure was 15°.

4. DATA REDUCTION, FINDINGS, AND DISCUSSION

4.1. Correction to Airspeed Data

During data reduction and analysis it became obvious that two separate instrumentation problems had occurred, both of which impacted the recorded indicated airspeed data. The first problem was that the test aircraft's instrumentation pitot system had a leak. This was first suspected during initial data reduction and was confirmed by a ground check of the research instrumentation pitot system and the ship's pitot system. The solution to this problem was to fly a re-calibration flight that consisted of flying an airspeed sweep from 80 knots to 120 KIAS at an altitude between 3500 and 4500 feet MSL. From this re-calibration flight, a correction between the aircraft's indicated airspeed and the recorded instrumentation airspeed was derived. The derived correction was applied to all the indicated airspeed data for all runs of each evaluation pilot.

The second problem had to do with the re-initialization of the instrumentation recording system following shutdown. The instrumentation system on the test aircraft was powered by the aircraft electrical system; however, the data acquisition laptop (laptop 2 as depicted in Figure 8) was powered by its internal batteries. Unknown to the test crew, when power is removed from the instrumentation system and then powered on again without a reboot of the data acquisition system's laptop; the airspeed channel data became corrupted. This is exactly what happened to run's 10, 11 and 12 during evaluation pilot number 1's flight, since the test crew had to stop to refuel after run 9. Indicated airspeed data for all of evaluation pilot number 5 was also lost in a similar fashion. This instrumentation problem has now been documented and resolved and should not impact further testing, however, for the purposes of this experiment there was no way to recover the data. Therefore, indicated airspeed data from runs 10, 11, and 12 for evaluation pilot 1 and all of the indicated airspeed data for evaluation pilot number 5 were not available for data analysis.

4.2. NASA TLX Workload Scores

NASA TLX is a six-dimensional subjective workload instrument. It is generally used to predict performance, and is based on the assumption that subjective workload represents the cost to achieve a certain level of performance[11,30]. It was chosen for this experiment over other subjective workload instruments, such as the Subjective Workload Assessment Technique (SWAT) or Workload Profile, because it was considered to have the advantages of: ease of use for subjects[30,31], applicability to a broad range of tasks[30,31], and was found applicable to real-world tasks such as piloting aircraft and simulators[31]. NASA TLX was also considered to have: good sensitivity, low cost to administer, and low to nil interference with the test subject's task[30]. Another important characteristic that made NASA TLX attractive was that, by design, it was supposed to reduce between-subject variability[30].

For this experiment, a 20, five point step scale from 0-100 was used for each dimension. A sample of the NASA TLX questionnaire test card used is provided in Figure 10. Each evaluation pilot was asked to complete the test card shown in Figure 10, by making a rating for each dimension and making a pair wise comparison of all the dimensions. Ratings were made by all five evaluation pilots immediately following each run, resulting in a final data set of 60 NASA TLX cards completed; 30 for the 3D audio display off condition and 30 for the 3D audio display on condition. Ratings for each dimension were analyzed using the sample calculations described in references 11 and 30 to obtain the weighted NASA TLX workload score.

NASA TLX data reduction typically involves a calculation of a weighted workload score based upon the pair-wise comparison weighting assigned by each evaluation pilot to each dimension after each run. The main purpose of including the weighting calculation is to reduce the between-subject variability[11,30]. However, literature exists, which concludes that the pair-wise comparison of the dimensions has little to nil impact on variability and can be dropped[32,33]. For this experiment, non-weighted and weighted NASA TLX workload scores were calculated to determine if there would be any influence on the results.

For the purposes of testing for statistical significance the null hypothesis was “whether the 3D audio display is off or on, there will be no difference in the dependent variable”. The dependent variable was the subjective rating assigned to each dimension of the NASA TLX and the non-weighted and weighted NASA TLX subjective workload scores. The independent variable was the 3D audio display either off or on.

Based on the experimental design and the fact the dependent variable was the subjective ratings of the evaluation pilots, the requirements for a typical independent, parametric statistical test were not satisfied. Firstly, the experiment design had one group of five subjects tested twice against the same dependent variable. Therefore, a dependent test for the statistical difference between 3D audio display off and on was considered more appropriate[34,35]. Second, the subjective ratings are ranked data, and cannot be assumed to be random samples from a normal distribution. Also, subjective ratings data were considered to be ordinal[36]; in which case the median was considered a more appropriate indicator of central tendency. Although independent, parametric statistical tests are considered robust to violations in assumptions about their distribution, it was deemed more appropriate to use a dependent, non-parametric statistic test[34,36], like the Wilcoxon signed-rank test.

A summary of the NASA TLX ratings and workload scores are presented in Table 4 and in Figure 11 to Figure 14. From these data some important characteristics and trends were observed. Firstly, each evaluation pilot had a significant variation of between-subject baseline workload, which was readily apparent in Figure 11 and Figure 13. Although a variation in baseline was expected, it is interesting to note the high degree of variation, which was approximately 300% between the lowest baseline (EP 4) and the highest baseline (EP 5). It is also interesting to note that the non-weighted NASA TLX workload scores and the weighted NASA TLX workload scores had nearly the same degree of variation in workload. The weighted NASA TLX method was suppose to reduce between-subject variability[11,30], but did not appear to be the case. The minimal to nil impact on between-subject variability with weighted NASA TLX had also been noted by other researchers[32,33].

Secondly, as also highlighted in Figure 11 and Figure 13, there was a drastic reduction in non-weighted and weighted NASA TLX workload scores for evaluation pilots 1, 2, and 4 as the number of runs increased. Evaluation pilot 5 data showed a slight reduction trend and evaluation pilot 3 showed no visible reduction trend. This reduction in perceived workload was most likely due to a learning effect that was occurring for 4 (80%) of the evaluation pilots over the course of the experiment. Although familiarization training was provided on the aircraft, on the 3D audio display and on the NASA TLX questionnaire, the trend highlighted in Figure 11 and Figure 13 was most likely the result of not enough training[30] provided to the evaluation pilots on the aircraft, the task, or on the NASA TLX form, prior to the start of the experiment. Correlation between the learning effect observed in Figure 11 and Figure 13 and evaluation pilot demographics (Table 1) or training time provided (Table 2) was not conducted. However, it should be noted that the most experienced evaluation pilot (EP 4, 1300 hours total) exhibited nearly the same learning effect as the least experienced evaluation pilot (EP 2, 325 hours total).

Finally, from Table 4, there is a high degree of intra-subject variance, as indicated by the relatively high values of standard deviation and range, and differences between the mean and the median. Note that the evaluation pilots with the highest variance also exhibited higher degrees of the learning effect, suggesting that the learning effect is probably the source of the high intra-subject variance in the data. The high variance in the intra-subject data also gives more support for using the median as a more appropriate indicator of central tendency than the mean[34]. Median data on the subjective ratings for each dimension of NASA TLX contained in Table 4 was plotted in Figure 15 through Figure 20.

To test for statistical significance, the Wilcoxon signed-rank test was used. It is a common non-parametric, dependent statistical test for small sample sizes[34,35]. The significance level (α) selected was 0.05 and the sample size (n) was 5. Since the hypothesis statement was unidirectional, one-tailed testing tables were used. Also, as discussed earlier, the median was considered to be a more appropriate indicator of the central tendency of the dimension ratings and overall workload scores for each evaluation pilot. Therefore, the median of the six paired ratings and workload scores for each pilot was used in the Wilcoxon signed-rank test.

A summary of the statistical test results are presented in Table 5. In summary, there was a statistically significant (α=0.05, n=5, one-tail) reduction in workload for the non-weighted and weighted NASA TLX scores between the 3D audio display off and on conditions. However, out of the six dimensions of NASA TLX, the Mental Demand dimension was found to not have statistical significance and for all other dimensions the Wilcoxon signed-rank test yield at least one signed-rank difference of zero. Signed-rank differences of zero must be dropped from the Wilcoxon signed-rank test calculations[34,35]. Since the lower limit of the sample size in the Wilcoxon signed-rank test calculation was five, calculations for statistical significance for all the other dimensions, aside from Mental Demand, could not be conducted.

However, for each of the dimensions of NASA TLX, a comparison analysis was conducted to determine general trends in the data. Using the data presented in Figure 15 to Figure 20, a count of how many of evaluation pilots' median data indicated a decrease, or increase, between the 3D audio display off and on condition was conducted. From Table 6 the results of the comparison analysis shows a general trend that most of the evaluation pilots gave lower ratings for each NASA TLX dimension when the 3D audio display was presented. The only dimension that had a flat trend was Temporal Demand; for which 2 (40%) evaluation pilots indicated a reduction, 2 indicated an increase, and 1 (20%) evaluation pilot was neutral. Conversely, Table 6 shows a majority (4, 80%) of the evaluation pilots felt the 3D audio display reduced their Mental Demand. Although most of the evaluation pilots felt a reduction in Mental Demand, it was found earlier not be a statistically significant reduction.

In summary, the NASA TLX workload data indicates:

1. significant between-subject variability, even though NASA TLX was considered to reduce this variance[30];

2. significant intra-subject variability, most likely the result of learning due to the repetitive nature of the tasks in the experiment or not enough time spent training the subjects;

3. no appreciable difference in results between the non-weight NASA TLX workload scores and the weighted workload scores, which was similar to results found by other researchers[32,33]; and

4. there appeared to be a statistically significant (α=0.05, n=5, one-tail) reduction in the overall NASA TLX scores when the 3D audio display was presented; and

5. a majority (4, 80%) of evaluation pilots indicated that the Mental Demand dimension was reduced by the presence of the 3D audio display. Temporal Demand had the least frequency, with only 2 (40%) evaluation pilots indicating lower ratings. All the other dimensions had a high number (3, 60%) of the evaluation pilots frequently indicating lower ratings when the 3D audio display was present.

4.3. Post Flight Questionnaires

Each evaluation pilot completed a post-flight questionnaire immediately following the flight. Questions were designed to gather evaluation pilot opinion in each of four categories: implementation of the 3D audio display cue, improvements to flight or navigation performance, reduction in workload, and improvements to situational awareness. Data is presented on a per evaluation pilot basis in Figure 21, Figure 22, Figure 23, and Figure 24.

With respect to implementation of the 3D audio display, a few (2, 40%) of the evaluation pilots agreed the sound (cue) used in the display was attention getting, with a few (2, 40%) neutral about the sound’s ability to grab attention and one (20%) disagreeing. However, a majority (4, 80%) agreed, and one (20%) strongly agreed, that the 3D audio display did not disturb their ability to concentrate on other cockpit tasks, interfere with hearing other communication, or find the aircraft reference confusing when turning the head. A high number (3, 60%) of evaluation pilots agreed, and one (20%) strongly agreed, that the display was intuitive and easy to adapt to. However, one (20%) evaluation pilot also disagreed with that statement. All evaluation pilots (5, 100%) agreed the 3D audio display functioned similarly during the test runs as it did during training on the ground. All of these results are expressed in Figure 21 and point to an implementation that was generally well accepted. In fact, the implementation questions received a score of 118 out of a possible 150. Any improvements to the implementation of the system should be focused on the actual sound used for localization.

Questionnaire responses regarding improvements to flight and navigational performance were varied. While several (3, 60%) of the evaluation pilots agreed that bank angle control improved when the 3D audio display was available, one (20%) was neutral on this matter and one (20%) disagreed. With respect to improvements in airspeed control when the 3D audio display was available, one (20%) agreed, and few (2, 40%) were either neutral or disagreed. However, a majority (4, 80%) agreed, and one (20%) was neutral, that the 3D audio display improved accuracy of turns to the target heading and improved course intercept (questions 9 and 10). Although a majority (4, 80%) of the evaluation pilots either agreed that the 3D audio display improved heading control and course intercept, only one (20%) of them agreed that it could be used solely to navigate to the fix without a visual display, such a the horizontal situation indicator (HSI) or GPS. A few (2, 40%) of the respondents where neutral, and a few (2, 40%) disagreed, that navigation to the fix could be conducted with just the localization of the sound cue. Overall, majority (4, 80%) of the evaluation pilots agreed that the 3D audio display had improved flight and navigational performance, with one (20%) respondent neutral. Figure 22 summarizes these results. The improvements to flight and navigational performance questions scored a 102 out of a possible 150. In general, pilot opinion indicates that this 3D audio display did improve their heading control and course intercept performance.

Workload, in the context of this experiment and the questions in the post flight questionnaire, deals with time and how rushed the evaluation pilots were in getting their tasks accomplished. With respect to workload, when the 3D audio display was available, a majority (4, 80%) of the evaluation pilots agreed, and one (20%) was neutral, that they had more time to scan non-navigational instruments. However, several (3, 60%) of the evaluation pilots agreed and a few (2, 20%) were neutral, that they could efficiently scan the navigational display, had more time to scan the attitude indicator, or did not feel rushed. Figure 23 summarizes these results and indicates that, most (3, 60%) the evaluation pilots felt there was workload reduction when the 3D audio display was available to them. Overall, the decrease in workload questions scored a 73 out of a possible 100.

Situational awareness, in the context of aviation navigation, can be grossly defined as the pilot's awareness of where the aircraft is, where it needs to go, and when it will get there. Situational Awareness also encompasses other dimensions, which include system and environmental states. Arguably, it is the single most important contributor to aviation safety and accident prevention[37]. Furthermore, nearly all displays to date that contribute to situational awareness utilize the visual channel. Therefore, any activity that causes visual fixation or overload will negatively impact situational awareness. Subjective data on situational awareness shows that a few (2, 40%) strongly agree, a few (2, 40%) agree, and one (20%) was neutral with regards to having better awareness of which direction to turn towards the revised fix when the 3D audio display was available. A few (2, 40%) agreed, a few (2, 40%) were neutral, and one (20%) disagreed with having better awareness of unintentional deviations in course. However, all (5, 100%) of the evaluation pilots agreed that they were better able to attend to the simulated engine failure and still maintain a continuous awareness of aircraft position relative to the revised fix when the 3D audio display was presented. Furthermore, all (5, 100%) of the evaluation pilots agreed that the presence of the 3D audio cue enabled them to create a better mental picture of their flight position while flying the missed approach procedure. In addition, a few (2, 40%) strongly agreed and the remaining (3, 60%) agreed that the 3D audio display was useful in planning the turn towards the fix when they had to attend to the simulated engine failure. All of these results are presented in Figure 24 and indicate that, subjectively, the 3D audio display had utility increasing pilot situational awareness. Overall, the increase in situational awareness questions scored 99 out of a possible 125.

In summary, the post flight questionnaire data indicates that the evaluation pilots:

1. preferred having the 3D audio display available and generally found the implementation of the display agreeable;

2. felt the 3D audio display improved their heading control and course intercept performance;

3. felt the 3D audio display did reduce their workload; and

4. felt they had more situation awareness when the 3D audio display was available.

4.4. Pilot Performance

As a means to quantify the utility of the 3D audio display, three parameters that were collected by the aircraft's instrumentation system were compared for deviations from the ideal value. During the period of interest for each run, which was from the MDA until course intercept, the evaluation pilot flew the aircraft in a single-engine, full power turning climb. During this time the primary flight parameters the evaluation pilots were tasked to track were: the angle of bank (ideally 15° left or right), the indicated airspeed (ideally 99 KIAS), and the difference in actual and target heading (ideally 0°). Any deviations from the ideal were considered decreases in performance. Therefore, changes in the evaluation pilot's performance were quantified as the difference in the amount of deviation from the ideal for the two treatment conditions of 3D audio display off or on.

The angle of bank tracking task was initiated from the onset of the first turn towards the intercept heading until the evaluation pilot made the final roll to wings level on the intercept heading. Any overshoots in heading were included in the calculations. Upon study of the angle of bank data presented in Table 7 and Table 8 it is clear that mean angle of bank in the turn toward the intercept heading is nearly the same for all pilots. Similarly, the variability in maintaining the desired bank angle, the root-mean-square, is the same between the 3D audio display on and off condition for all pilots. Therefore, within the scope of this study none of the evaluation pilots had an increase, or decrease, in the angle of bank tracking task performance with the addition of the 3D audio display. This result is somewhat subjectively contradicted by post flight questionnaire data, which indicated several (3, 60%) of the evaluation pilots agreed that bank angle control improved when the 3D audio display was available, one (20%) was neutral, and one (20%) disagreed (see Figure 22). This difference in subjective and objective data could be an indication that although the 3D audio display did not actually improve pilot bank angle tracking performance, it led the evaluation pilots to think they were performing the bank angle tracking task better.

The indicated airspeed tracking task initiated as soon the evaluation pilot commenced the missed approach procedure ended when the course to the revised fix was intercepted to within 2-dots of center (less than half of full scale error) on the course deviation indictor (CDI) display on the HSI. As described in section 4.1 above data from runs 10, 11, and 12 for evaluation pilot 1 and all of evaluation pilot 5 data was corrupted and could not be used. The missing data is indicated with a dash in Table 9. The indicated airspeed control data presented in Table 9 and Table 10 shows that there was also little to no increase, or decrease, in the indicated airspeed task performance when the 3D audio display was presented. This result is subjectively supported by the post flight questionnaire data, which indicated only one (20%) evaluation pilot agreed that the airspeed tracking task was improved when the 3D audio display was presented. Two (40%) others were neutral and two (40%) others disagreed.

The heading tracking task was conducted while the evaluation pilot was established on the intercept heading to the inbound course towards the revised missed approach waypoint. The intercept heading was determined by the evaluation pilot prior to executing the missed approach procedure. The intercept heading was the heading that the evaluation pilot must turn towards and track in order for the aircraft to be on the GPS track from the missed approach point to the revised missed approach holding fix. Each pilot was pre-briefed to make all intercepts of the track to the revised missed approach waypoint at 30°, declare and set this heading on the HSI, and to maintain this heading until the CDI bar on the HSI was within 2-dots of center. The significance of using an intercept heading was that it placed the 3D audio cue away from the front/back ambiguity that typically exists in 3D audio displays.

The data for difference in actual heading and the required intercept heading is provided in Table 11 and Table 12. Based on the data in Table 12 it appears the evaluation pilots had an improvement in their heading control task performance when the 3D audio display was available. However, a two-tailed t-test for difference in means does not indicate a statistical difference. Although not statistically significant, the result does correlate with the post flight questionnaire results discussed in section 4.3, where a majority (4, 80%) of the evaluation pilots agreed, and one (20%) strongly agreed, that the 3D audio display improved accuracy of turns to the target heading and improved course intercept (Figure 22). In this case the subjective data and objective data agree and support the conclusion that the evaluation pilots in this study found some utility from the 3D audio display.

4.5. Aircraft Track Plots

Aircraft tracks for the two 3D audio display conditions were plotted for each pair-wise run combination and are presented in Figure 25 to Figure 54. During the familiarization training, each evaluation pilot was instructed that they were to execute the turn to the revised missed approach fix in an expeditious manner. Therefore, the aircraft track plots for the 3D audio display on/off pair-wise runs were compared to each other to determine under which display condition the evaluation pilots were able to execute their turn toward the revised missed approach waypoint in the most expeditious manner. For the purposes of this experiment, the most expeditious turn was determined by the aircraft track which was furthest from the missed approach holding waypoint for the approach just conducted. For the RNAV (GPS) RWY 18 approach (see Figure 2) the missed approach holding waypoint for this approach is KOJAK; for the RNAV (GPS) RWY 36 the missed approach holding waypoint is LOYSI, and so on for the other approaches. As an example, from Figure 25, the aircraft track from run 7 is furthest from

KOJAK and, therefore, the aircraft was turned towards the revised missed approach fix in a more expeditious manner than that from run 1. In other words, the most expeditious turn would result in the shortest aircraft track to the revised MAWP.

Applying the most expeditious turn criterion for all the pair-wise track plots for each evaluation pilot produced the results contained in Table 13. From the results in Table 13 it is apparent that a greater total number of expeditious turns (57%) were completed when the 3D audio display was off. Conversely, this result shows that when the 3D audio display was on, it resulted in an increased aircraft tracks to the revised MAWP. It should be noted that evaluation pilot 5 had a significantly greater number of expeditious turns compared to all the other evaluation pilots. If the data from evaluation pilot 5 is ignored, then the results in Table 13 are nearly the same between the two conditions of display on and off, but still slightly favoring the display off condition. Unfortunately, with the data collected it was not possible to determine why evaluation pilot 5 data was so different from the rest of the evaluation pilots.

Regardless of whether or not evaluation pilot 5 data is ignored, the results in Table 13 were contrary to what was expected, which was shorter aircraft tracks to the revised MAWP when the 3D audio display was on. With the data at hand, it was not possible to determine why there were a greater percentage of expeditious turns when the 3D audio display was off.

5. CONCLUSIONS AND RECOMMENDATIONS

Within the limitations of having a low number (5) of test subjects available, the main result of the subjective data shows a perceived increase in situational awareness (post flight questionnaires), but also a perceived decrease in workload (NASA TLX scores). This result supports the conclusion that with 3D audio displays, pilots can receive additional information without additional workload, which was also found by McKinley et al[Error! Reference source not found.]. The major difference, however, is that with McKinley et al[Error! Reference source not found.], workload variation with and without a 3D audio display was studied in the context of a visual tracking task. In this experiment, workload and situational awareness, together, were explored in the context of real-world navigational tasks. The results from this experiment indicate there is potential utility of a 3D audio display system as a supplemental navigational aid in general aviation.

From the data gathered in this experiment it was not possible to determine why there was a perceived increase in situational awareness with a corresponding decrease in workload. Possibly, the 3D audio display provided the evaluation pilots with redundant information on aircraft position and relative heading without taxing the already heavily burdened visual channel. As discussed by Broadbent[38] and Woods[39], and briefly discussed more recently by Seagull, Wickens and Loeb[40], the increase in situational awareness could have resulted from pre-attentive referencing. That is, the constant positional and relative heading information made available with the 3D audio display provided a constant background navigational cue that did not require an effort on the part of the evaluation pilot to obtain.

Although not statistically significant, the only objective impact to performance was a slight increase in heading control and course intercept. There was no corresponding performance increase in airspeed control, angle of bank control, or improvements in aircraft track. There was not enough data collected to determine exactly why the subjective data indicates a perceived significant utility from the 3D audio display, but only a slight, non-significant improvement of one objective performance metric.

Based on the results of the NASA TLX scores, it appears that more practice in conducting approaches in the experimental aircraft should have been given to the evaluation pilots before data collection was performed. This would likely have reduced the intra-subject variation in NASA TLS scores for the same task, run to run. Based on the learning effect present in the NASA TLX data, an additional ten to 15 approaches, probably over 2 flights, should have been given to each evaluation pilot.

All the evaluation pilots in this study, and those used by Wigdhal[10], were relatively new pilots and all had recently completed training to the commercial pilot level. Future studies should use pilots with more advanced experience to determine if there is any correlation between flying experience and perceived situational awareness or subjective workload with using a 3D audio display as a supplemental navigational aid.

Although all the evaluation pilots were given ground training and a familiarization flight with the 3D audio display, future studies should consider conducting visual feed-back training in order to improve accommodation to the non-individualized HRTF's[28].

Due to the low number (5) of subjects available, the experimental design employed had a single group tested twice against the same variable. This situation led to the use of a less desirable non-parametric, dependent, statistical technique. It is recommended that future experiments into 3D audio displays at UTSI utilize two groups with as large a number of subjects as possible. One group should fly with the 3D audio display off and the other with the 3D audio display on. Even if a large number of test subjects in not available, the experimental design should be conducted with two independent groups, instead of the same group tested twice.

LIST OF REFERENCES

LIST OF REFERENCES

[1] Lawhon, G., "Single-Pilot IFR," *Safety Advisor*, Aircraft Owners and Pilots Association (AOPA) Air Safety Foundation Operations & Proficiency No. 1, 2006.

[2] "Advisory Circular, AC 91-11, Single Pilot IFR", Civil Aviation Authority of New Zealand, Revision 00, 28 February 2006.

[3] Simpson, B., Brungart, D., Dallman, R., Joffrion, J., Presnar, M., & Gilkey, R., "Spatial Audio as a Navigation Aid and Attitude Indicator", *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting, Human Factors and Ergonomics Society*, 2005, pp. 1602-1606.

[4] Brungart, D., Simpson, B., Dallman, R., Romigh, G., Yasky, R.., & Raquet, J., "A Comparison of Head-Tracked and Vehicle-Tracked Virtual Audio Cues in an Aircraft Navigation Task," *Proceedings of the 13th International Conference on Auditory Display*, International Community for Auditory Display, 2007, pp. 32-37.

[5] Veltman, J., Oving, A., & Bronkhorst, A., "3-D Audio in the Fighter Cockpit Improves Task Performance," *The International Journal of Aviation Psycholog*y, Vol. 14, No. 3, 2004, pp. 239-256.

[6] Furness, T., "The Super Cockpit and Its Human Factors Challenges," *Proceedings of the Human Factors Society 30th Annual Meeting*, Dayton, OH, 1986, pp. 48-52.

[7] Bolia, R., "Special Issue: Spatial Audio Displays for Military Aviation," *The International Journal of Aviation Psychology*, Vol. 14, No. 3, 2003, pp. 233-238.

[8] Simpson, B. et al, "3D Audio Cueing for Target Identification in a Simulated Flight Task," *Human Factors and Ergonomics Society 48th Annual Meeting Proceedings*, New Orleans, LA, 2004, pp. 1836-1840.

[9] Lyon, T., Gillingham, K., Teas, D., Ercoline, W., & Oakley, C., "The Effects of Acoustic Orientation Cues on Instrument Flight Performance in a Flight Simulator," *Aviation, Space, and Environmental Medicine,* Vol. 61, 1990, pp. 699-706.

[10] Wigdahl, A., "A Limited Evaluation of Pilot Technical Performance, Situational Awareness, and Workload when Flying with an Airframe-Referenced 3D Audio Display," M.Sc. Thesis, Aviation Systems Dept., University of Tennessee Space Institute, Tullahoma, TN, 2008.

[11] Hart, S., & Lowell, S., "Development of NASA-TLX: Results of Empirical and Theoretical Research," in Hancock, P., and Meshkati, N., (Eds.), *Human Mental Workload*, North Holland Press, Amsterdam, 1988, pp. 239-250.

[12] Stokes, A., & Wickens, C., "Aviation Displays," *Human Factors in Aviation*, Edited by E. Wiener and D. Nagel, Academic Press, San Diego, 1988, pp. 387-431.

[13] Tonnesen, C., & Steinmetz, J., "3D Sound Synthesis," *Human Interface Technology Laboratory*, URL: http://www.hitl.washington.edu/scivw/EVE/I.B.1.3DSoundSynthesis.html [cited 24 July 2008].

[14] Sundareswaran, V. et al, "3D Audio Augmented Reality: Implementation and Experiments," *Proceedings of the Second IEEE and ACM International Symposium on Mixed and Augmented Reality*, IEEE, 2003.

[15] Wang, K., Sundareswaran, V., Tam, C., Bangayan, P., & Zahorik, P., "Efficient and Effective use of Low-Cost 3D Audio Systems," *Proceedings of the 2002 International Conference on Auditory Display*, International Community for Auditory Display, Kyoto, Japan, 2002, pp. 1-5.

[16] Cooper, M., & Petrie, H., "Three Dimensional Auditory Display: Issues in Applications for Visually Impaired Students," *Proceedings of ICAD 04-Tenth Meeting of the International*

Conference on Auditory Display, International Community for Auditory Display, Sydney, Australia, 2004.

[17] Bolia, R., D'Angelo, W., & Mckinley, R., "Aurally Aided Visual Search in Three-Dimensional Space," *Human Factors*, Vol. 41, No. 4, 1999, pp. 664-669.

[18] Bronkhorst, A., Veltman, J., & van Breda, L., "Application of a Three-Dimensional Auditory Display in a Flight Task," *Human Factors*, Vol. 38, No. 1, 1996, pp. 23-33.

[19] McAnally, K. & Martin, R., "Localization of Sound Presented Via a Spatial Audio Display during Visually Induced Vection in Pitch, Roll, and Yaw," *Aviation, Space, and Environmental Medicine*, Vol. 79, No. 6, 2008, pp. 611-615.

[20] "Sound Localization," Wikipedia, URL: http://en.wikipedia.org/wiki/Sound_localization, [cited 13 January 2008].

[21] Dell, W., "The Use of 3D Audio to Improve Auditory Cues in Aircraft," Class CS4H, Department of Computer Science, University of Glasgow, Glasgow, Scotland, 2000.

[22] Wallach, H., "On Sound Localization," *Journal of the Acoustical Society of America*, Vol. 10, 1939, pp. 270-274.

[23] Wightman, F., & Kistler, D., "Hearing in Three Dimensions: Sound Localization," NASA-CR-187706, 1990.

[24] "Head-related transfer function," Wikipedia, URL: http://en.wikipedia.org/wiki/Head-related_transfer_function, [cited 13 January 2008].

[25] Hartmann, W., "How We Localize Sound," *Physics Today on the web*, URL: http://www.aip.org/pt/nov99/locsound.html [cited 27 July 2008].

[26] McKinley, R., Erickson, M., & D'Angelo, W., "3-Dimensional Auditory Displays: Development, Applications, and Performance," *Aviation, Space, and Environmental Medicine*, 1994, Vol. 65, No. 5 (Suppl.), pp. A31-38.

[27] Hofman, P., van Riswick, J., & van Opstal, A., "Relearning sound localization with new ears," *Nature Neuroscience*, Vol. 1, No. 5, 1998, pp. 417-421.

[28] Zahorik, P., Tam, C., Bangayan, P., & Sundareswaran, V., "Localization Accuracy in 3-D Sound Displays: The Role of Visual-Feedback Training," *Proceedings of the Advanced Displays and Interactive Displays ARL Federated Laboratory 5th Annual Symposium*, College Park, MD, 2001, pp. 17-22.

[29] Algazi, V., Duda, R., & Thompson, D., "The use of Head-and-Torso Models for Improved Spatial Sound Synthesis," *Audio Engineering Society 113th Convention*, Los Angeles, CA, 2002.

[30] Geddie, J., et al, "NATO Guidelines on Human Engineering Testing and Evaluation," *NATO Research and Technology Organization*, RTO-TR-021, 2001.

[31] Rubio, S., Diaz, E., Martin, J., & Puente, J., "Evaluation of Subjective Mental Workload: A Comparison of SWAT, NASA-TLX, and Workload Profile Methods," *Applied Psychology: An International Review*, Vol. 53, No. 1, 2004, pp. 61-86.

[32] Nygren, T., "Psychometric Properties of Subjective Workload Measurement Techniques: Implications for their use in the Assessment of Perceived Mental Workload," *Human Factors*, Vol. 33, No. 1, 1991, pp. 17-33.

[33] Hart, S., "NASA-Task Load Index (NASA-TLX); 20 Years Later," *Proceedings of the 50th Annual Meeting of the Human Factors and Ergonomics Society*, San Francisco, CA, 2006.

[34] Thomas, J., & Nelson, J., *Research Methods in Physical Activity*, 3rd ed., Human Kinetics, Champaign, IL, 1996, Chaps. 6, 10.

[35] Scheaffer, R., & McClave, J., *Probability and Statistics for Engineers*, 3rd ed., PWS-Kent Publishing, Boston, MA, 1990, Chaps. 8, 12.

[36] Cain, B., "A Review of Mental Workload Literature," NATO, Rept. RTO-TR-HFM-121-Part-II, Defence Research and Development Canada, Human Systems Integration, Toronto, ON, Canada, July 2007.

[37] Endsley, M., & Garland, D., "Pilot Situation Awareness Training in General Aviation," *Proceedings of 14th Triennial Congress of the International Ergonomics Association and the 44th Annual Meeting of the Human Factors and Ergonomics Society*, Santa Monic, CA, 2000.

[38] Broadbent, D., "The hidden preattentive process," *American Psychologist*, Vol. 32, No. 2, 1977, pp. 109-118.

[39] Woods, D., "The Alarm problem and directed attention in dynamic fault management," *Ergonomics*, Vol. 38, No. 11, 1995, pp. 2371-2394.

[40] Seagull, F., Wickens, C., and Loeb, R., "When is Less More? Attention and Workload in Auditory, Visual, and Redundant Patient-Monitoring Conditions," *Proceedings of the Human Factors and Ergonomics Society 45th Annual Meeting*, Minneapolis, MN, 2001.

[41] "RNAV (GPS) RWY 6, Tullahoma RGNL Airport/WM Northern Field (THA)," National Aeronautical Charting Office of the Federal Aviation Administration, 10 April 2008, URL: http://204.108.4.16/index.asp?xml=naco/online/d_tpp [cited 26 April 2008].

[42] "RNAV (GPS) RWY 18, Tullahoma RGNL Airport/WM Northern Field (THA)," National Aeronautical Charting Office of the Federal Aviation Administration, 10 April 2008, URL: http://204.108.4.16/index.asp?xml=naco/online/d_tpp [cited 26 April 2008].

[43] "RNAV (GPS) RWY 24, Tullahoma RGNL Airport/WM Northern Field (THA)," National Aeronautical Charting Office of the Federal Aviation Administration, 10 April 2008, URL: http://204.108.4.16/index.asp?xml=naco/online/d_tpp [cited 26 April 2008].

[44] "RNAV (GPS) RWY 36, Tullahoma RGNL Airport/WM Northern Field (THA)," National Aeronautical Charting Office of the Federal Aviation Administration, 10 April 2008, URL: http://204.108.4.16/index.asp?xml=naco/online/d_tpp [cited 26 April 2008].

APPENDIX

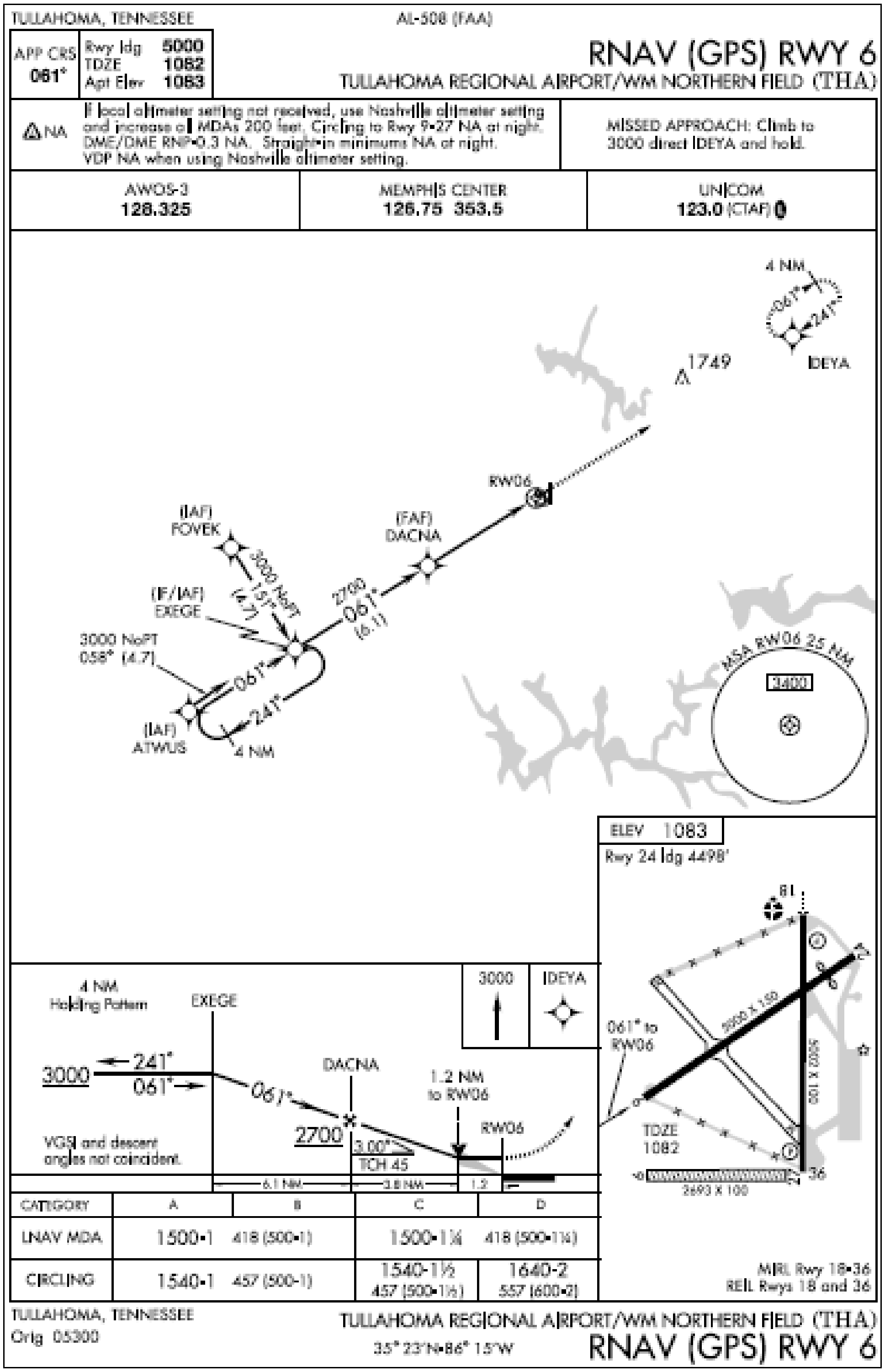


Figure 1. RNAV (GPS) RWY 6 Approach at KTHA.[41]

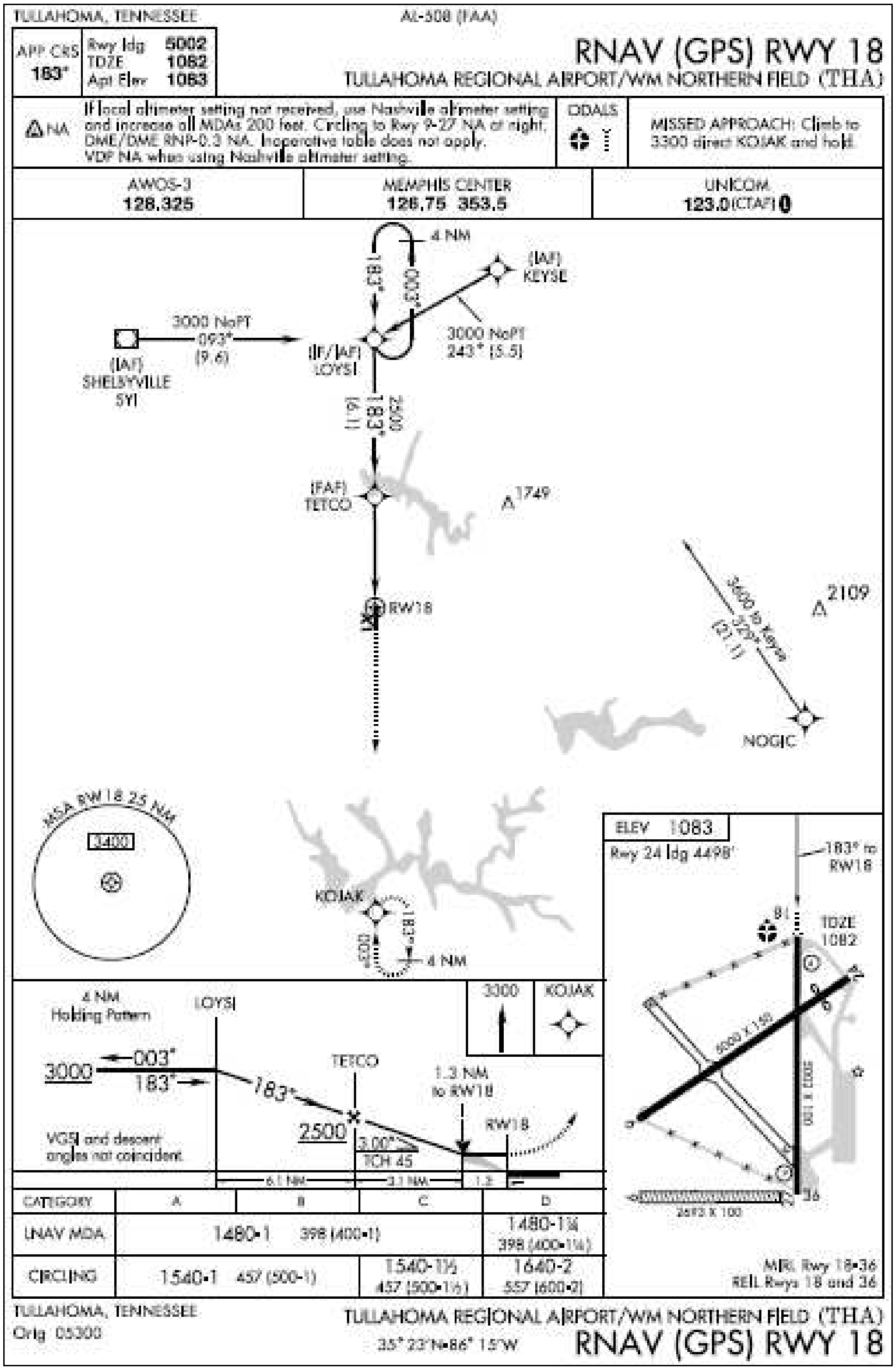
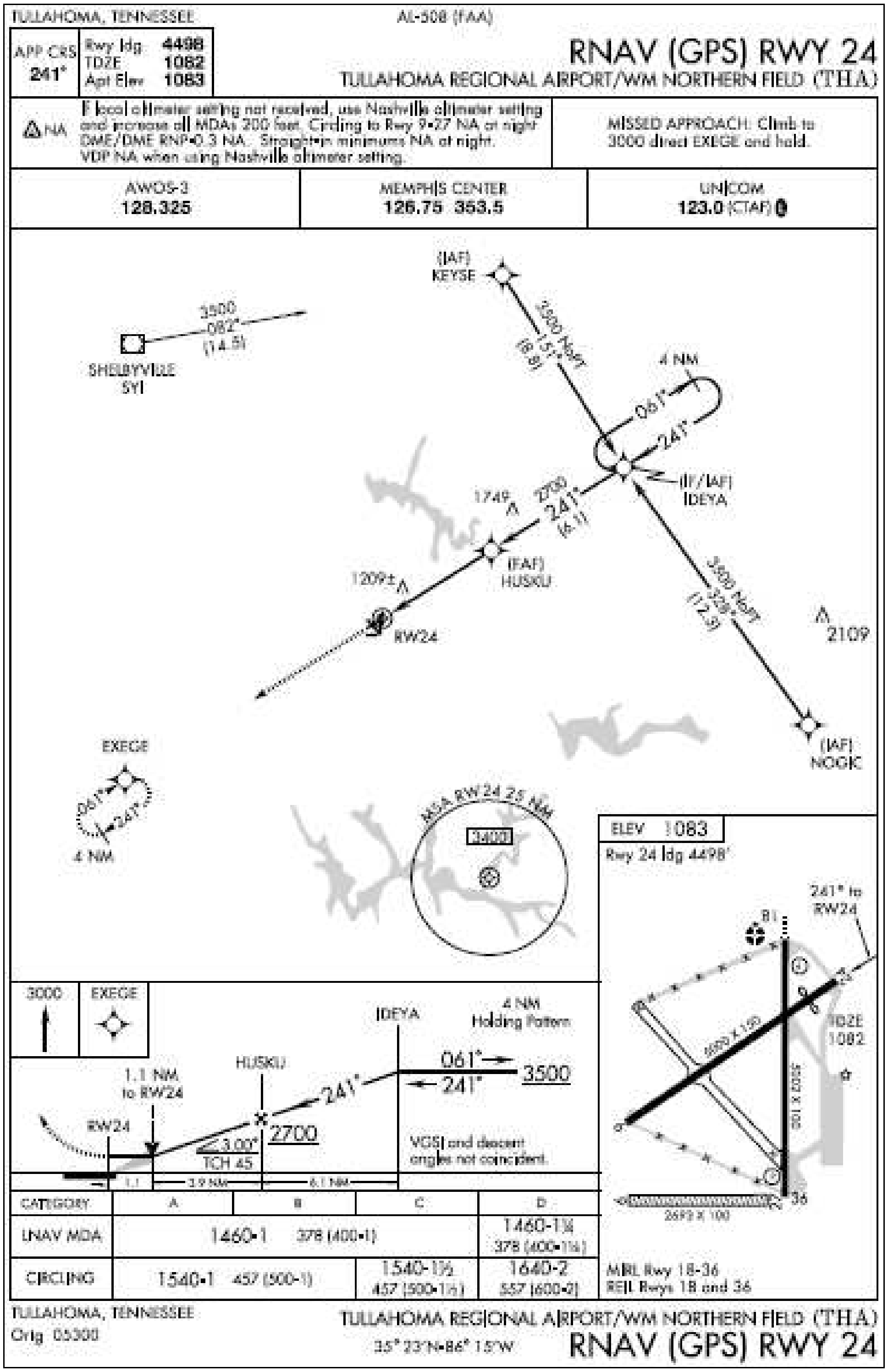


Figure 2. RNAV (GPS) RWY 18 Approach at KTHA.[42]



NA If local altimeter setting not received, use Nashville altimeter setting and increase all MDAs 200 feet. Circling to Rwy 9-27 NA at night. DME/DME RNP-0.3 NA. Straight-in minimums NA at night. VDP NA when using Nashville altimeter setting.

MISSED APPROACH: Climb to 3000 direct EXEGE and hold.

CATEGORY	A	B	C	D
LNAV MDA	1460-1 378 (400-1)			1460-1¼ 378 (400-1¼)
CIRCLING	1540-1 457 (500-1)		1540-1½ 457 (500-1½)	1640-2 557 (600-2)

Figure 3. RNAV (GPS) RWY 24 Approach at KTHA.[43]

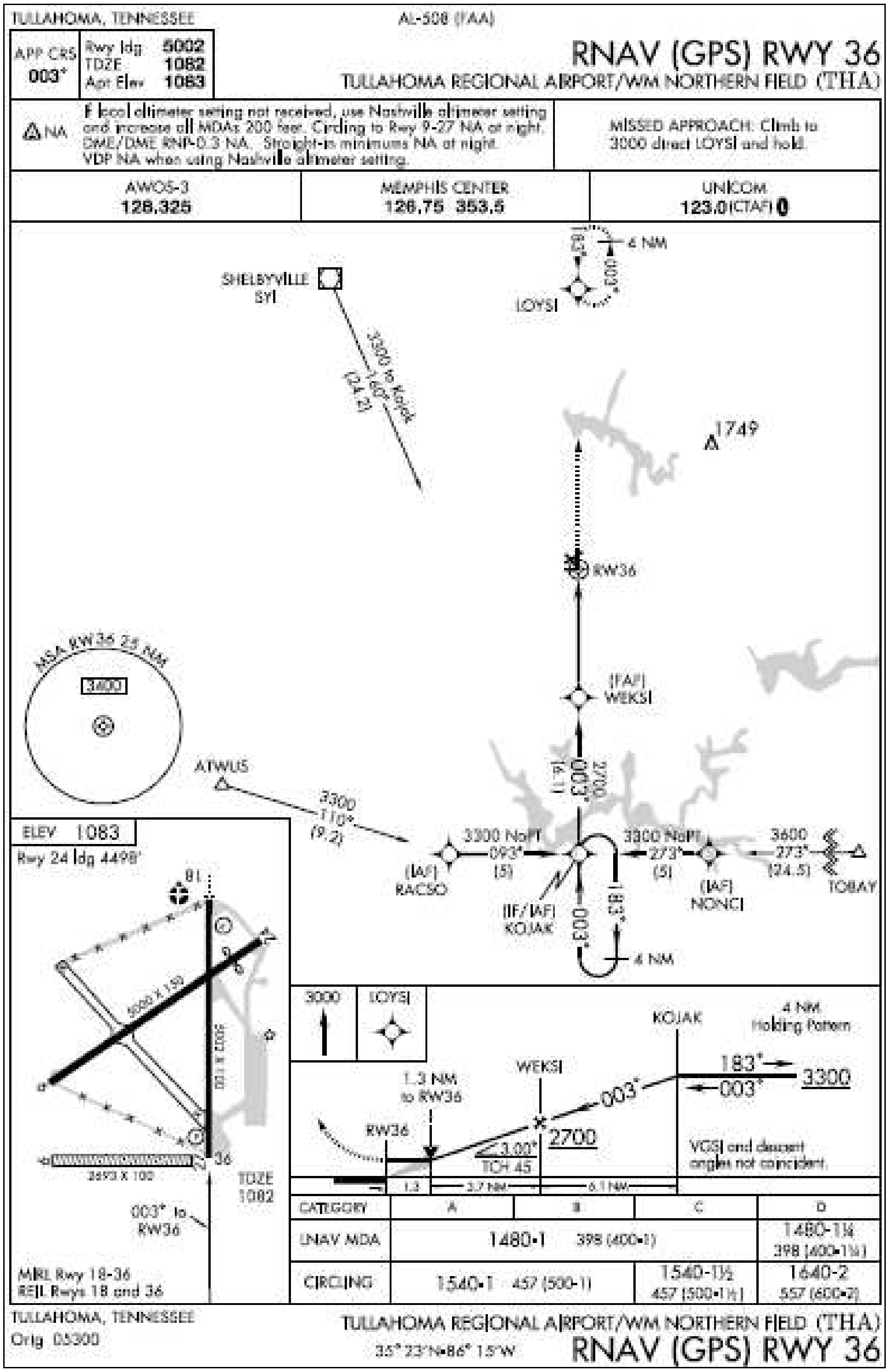
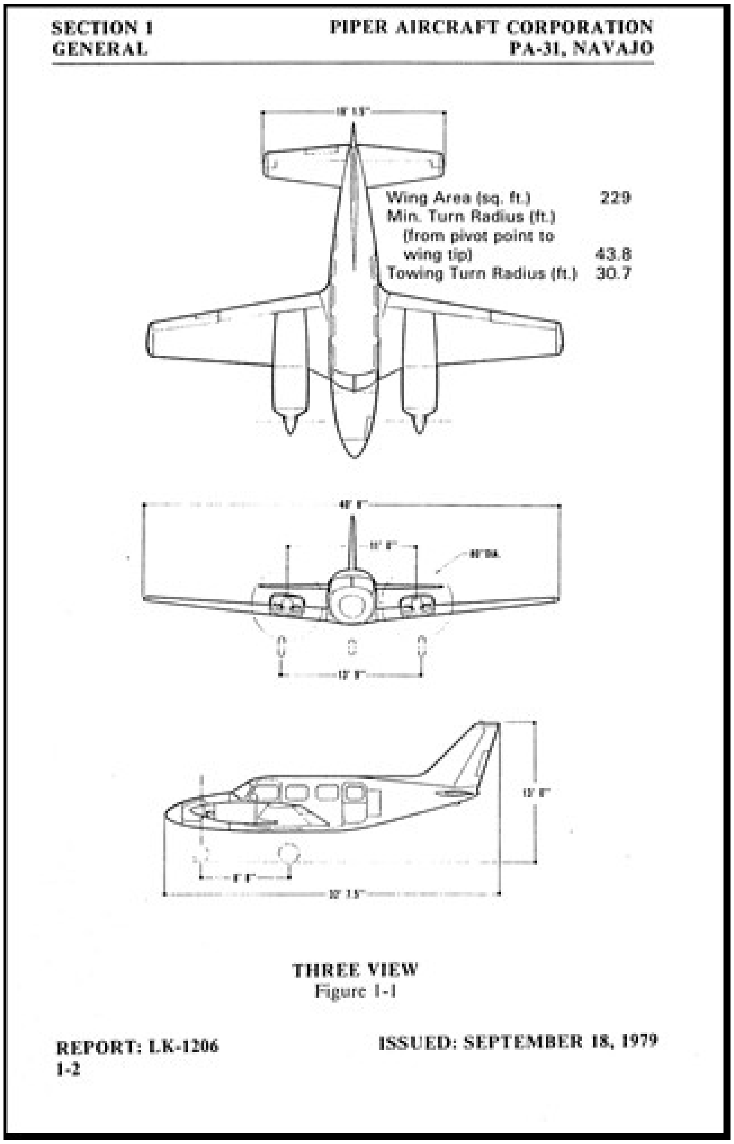


Figure 4. RNAV (GPS) RWY 36 Approach at KTHA.[44]



SECTION 1
GENERAL

PIPER AIRCRAFT CORPORATION
PA-31, NAVAJO

Wing Area (sq. ft.)	229
Min. Turn Radius (ft.) (from pivot point to wing tip)	43.8
Towing Turn Radius (ft.)	30.7

THREE VIEW
Figure 1-1

REPORT: LK-1206
1-2

ISSUED: SEPTEMBER 18, 1979

Figure 5. Three-view of Piper PA-31 Navajo aircraft.

Figure 6. UTSI 3D Audio Display Research Aircraft - Piper PA-31 Navajo.

Figure 7. Cockpit of UTSI's Piper PA-31 Navajo.

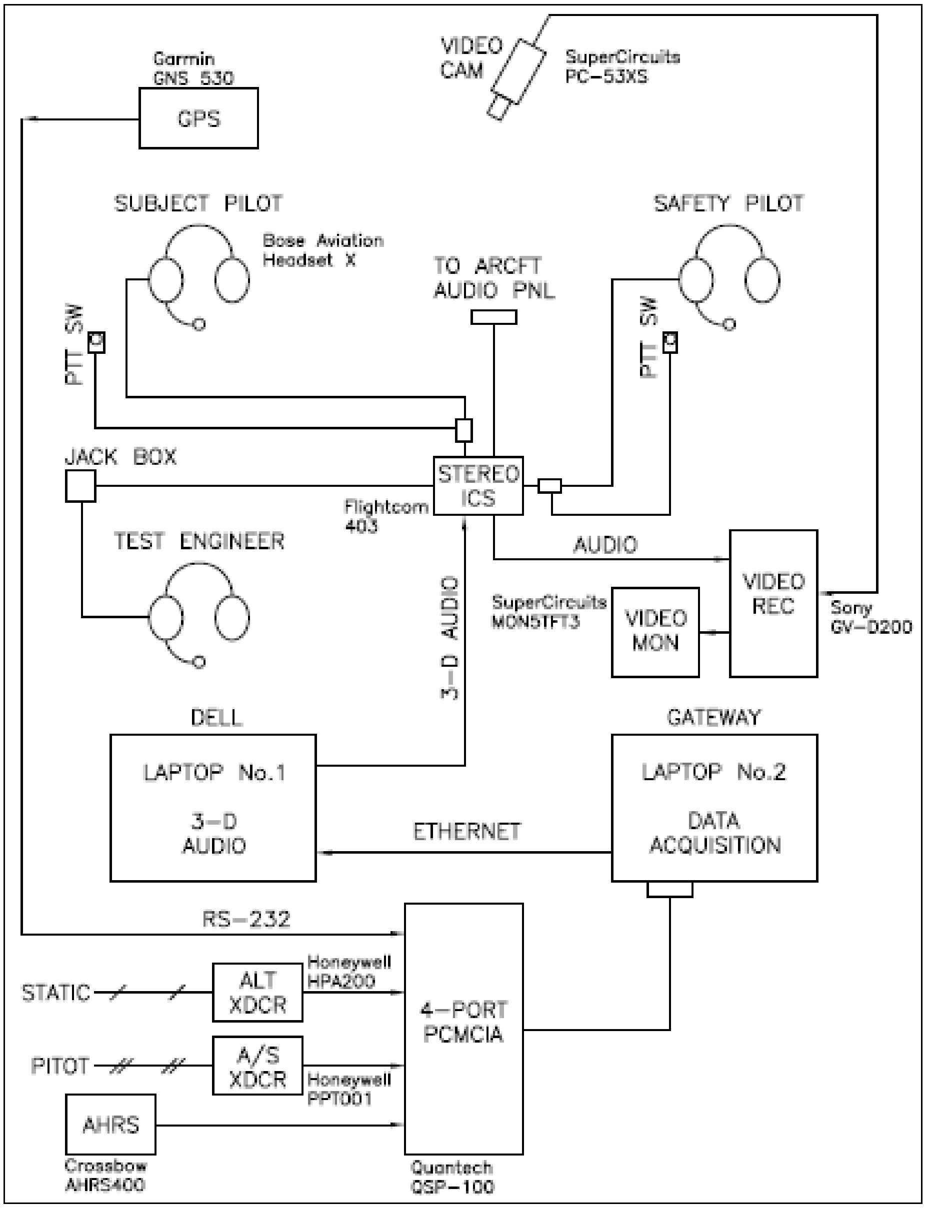


Figure 8. Schematic of the 3D Audio Display Installation.

FLT # -___	**EVALUATION PILOT # - ___**	**RUN # - 1**

Initial Test Conditions:	Test Limitations:
Trim for 120 KIAS	ALT - Not below 3500 ft MSL
Fly Pub Alt + 2500 ft	AOB - 15° max
3D Audio Display – **OFF**	KIAS - > Blue line + 5 (99 KIAS)

INITIAL APPROACH CLEARANCE

1. SP – "*N11UT, cleared **GPS RWY 18** approach via radar vectors. In the event of a missed approach, execute the published missed approach instructions.*"
2. EP – Clearance read-back. Loads the approach and reviews.
3. SP – Positions aircraft at 5500 ft MSL and gives control to EP – TIME:____________

AT **TETCO**

4. SP – "*3D Audio **OFF***" – TIME:_____________
5. FTE – Confirms 3D Audio **OFF** - "*3D Audio **OFF***"
6. EP – Is at the controls, on conditions establishes 700 fpm ROD.
7. SP – Calls for "*DATA ON*".
8. FTE – "*DATA ON*" – TIME:_____________

APPROACHING MINIMUM DESCENT ALTITUDE ~ 4000 ft MSL

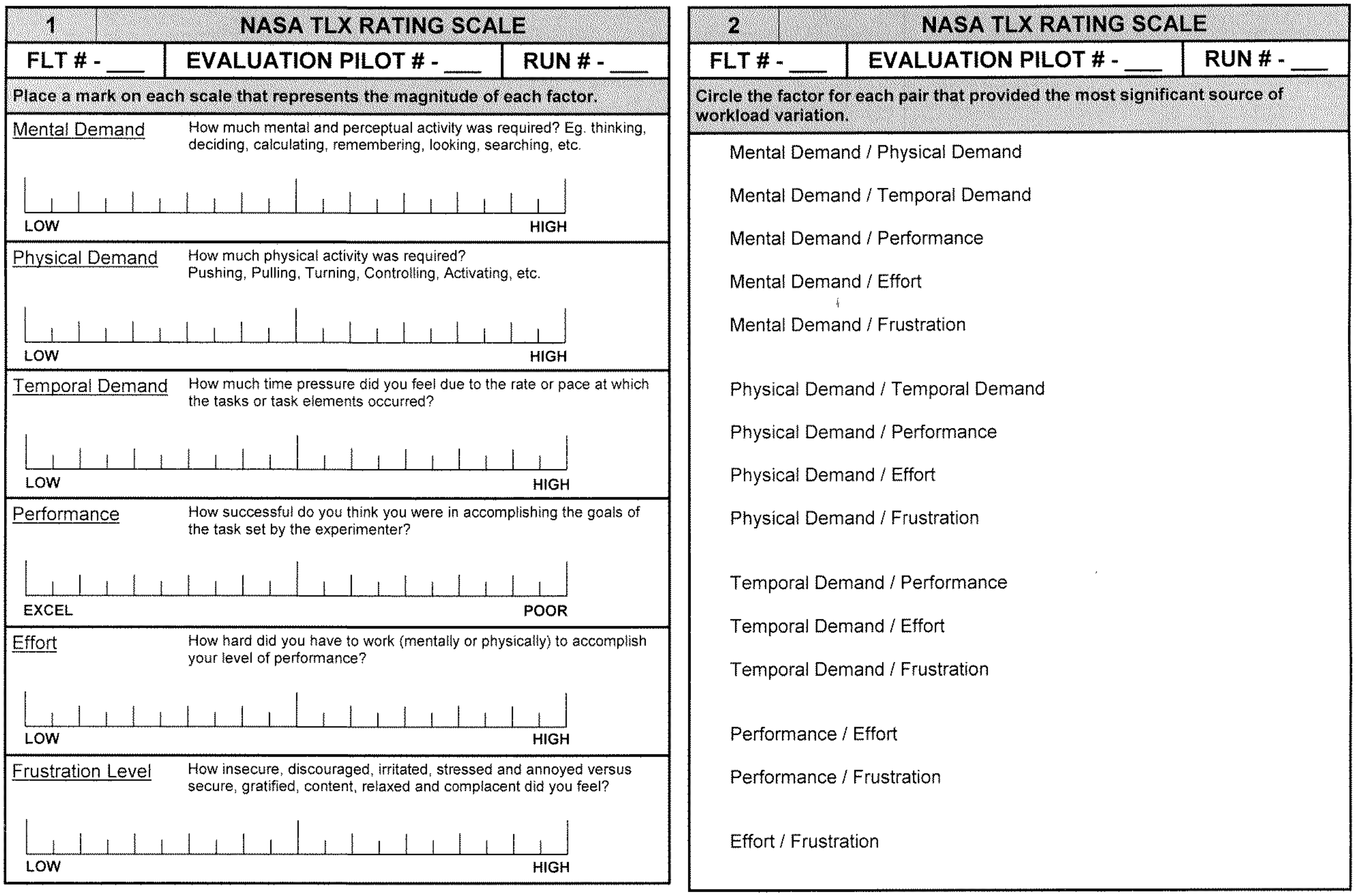
9. SP – "*N11UT the field is below minimums, maintain **3500** ft and rwy heading. When able proceed direct **EXEGE** for the **GPS RWY 06** approach. Climb and maintain **5500** ft.*"
10. EP – Clearance read-back. Loads **GPS RWY 06** via **EXEGE**, no holding, activates approach.
11. EP – Sets Power, retracts the gear, min speed "BLUELINE"; 30 deg intercept heading _____° on course **EXEGE**
12. SP – Retards **LEFT** throttle immediately (10 in.) and adjusts **RIGHT** throttle as required (37 in.) – TIME:____________

WHEN EP INSIDE OF 2 DOTS FOR 1 MINS OR AT **EXEGE**

13. SP – Calls "*2 DOTS*" – TIME:_____________
14. SP – When 2 dots for 1 min, Calls for "*DATA OFF*"
15. FTE - "*DATA OFF, CUE OFF*" – TIME:_____________
16. SP – Takes Control. Positions aircraft for next test point.
17. EP – Completes TLX Data Card.

NEXT: RUN #2

Figure 9. Sample Test Card



1 NASA TLX RATING SCALE

FLT # - ___ EVALUATION PILOT # - ___ RUN # - ___

Place a mark on each scale that represents the magnitude of each factor.

Mental Demand	How much mental and perceptual activity was required? Eg. thinking, deciding, calculating, remembering, looking, searching, etc.	LOW – HIGH
Physical Demand	How much physical activity was required? Pushing, Pulling, Turning, Controlling, Activating, etc.	LOW – HIGH
Temporal Demand	How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred?	LOW – HIGH
Performance	How successful do you think you were in accomplishing the goals of the task set by the experimenter?	EXCEL – POOR
Effort	How hard did you have to work (mentally or physically) to accomplish your level of performance?	LOW – HIGH
Frustration Level	How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel?	LOW – HIGH

2 NASA TLX RATING SCALE

FLT # - ___ EVALUATION PILOT # - ___ RUN # - ___

Circle the factor for each pair that provided the most significant source of workload variation.

Mental Demand / Physical Demand
Mental Demand / Temporal Demand
Mental Demand / Performance
Mental Demand / Effort
Mental Demand / Frustration

Physical Demand / Temporal Demand
Physical Demand / Performance
Physical Demand / Effort
Physical Demand / Frustration

Temporal Demand / Performance
Temporal Demand / Effort
Temporal Demand / Frustration

Performance / Effort
Performance / Frustration

Effort / Frustration

Figure 10. Sample of NASA TLX Questionnaire Test Card.

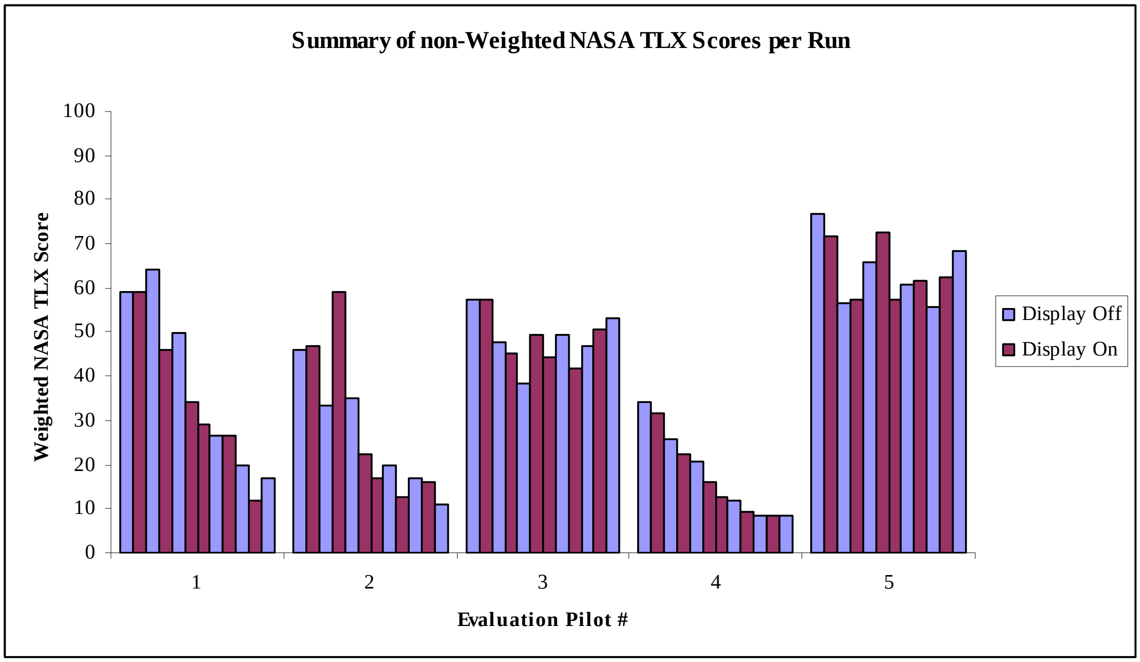


Figure 11. Summary of non-Weight NASA TLX Workload Scores.

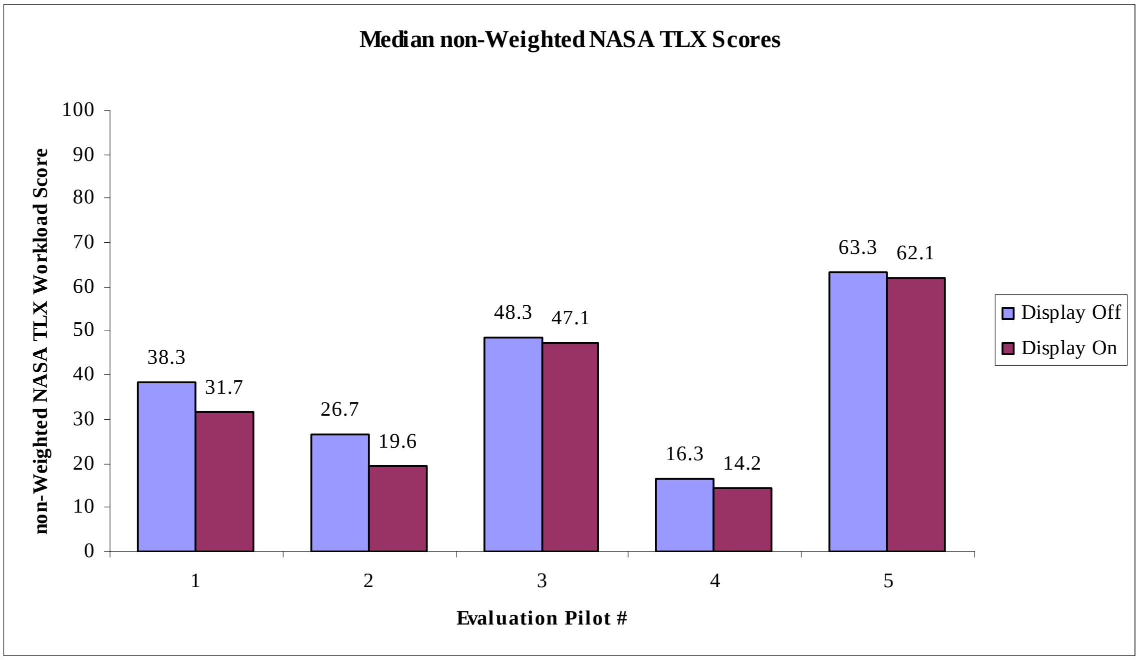


Figure 12. Median of non-Weighted NASA TLX Workload Scores.

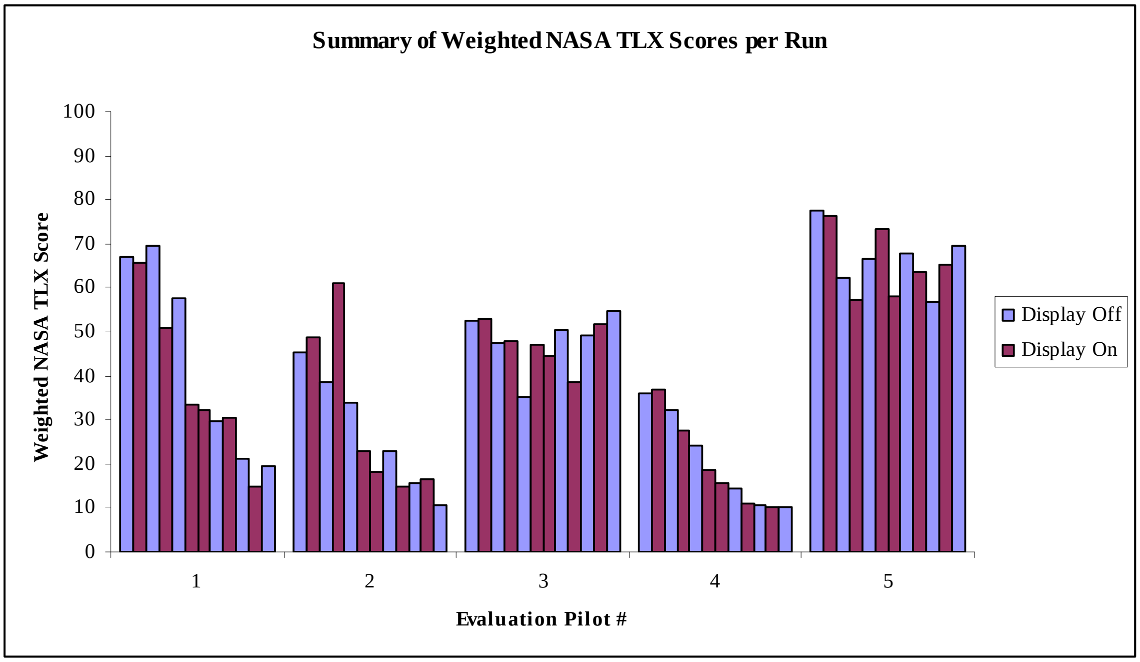


Figure 13. Summary of Weighted NASA TLX Workload Scores.

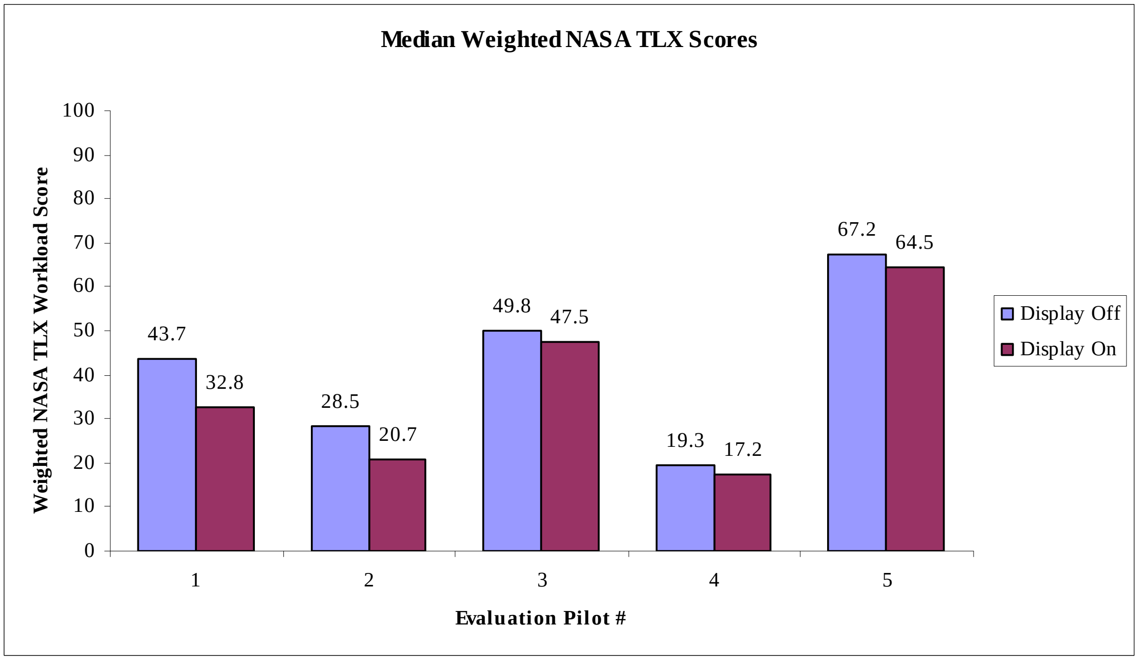


Figure 14. Median of Weighted NASA TLX Workload Scores.

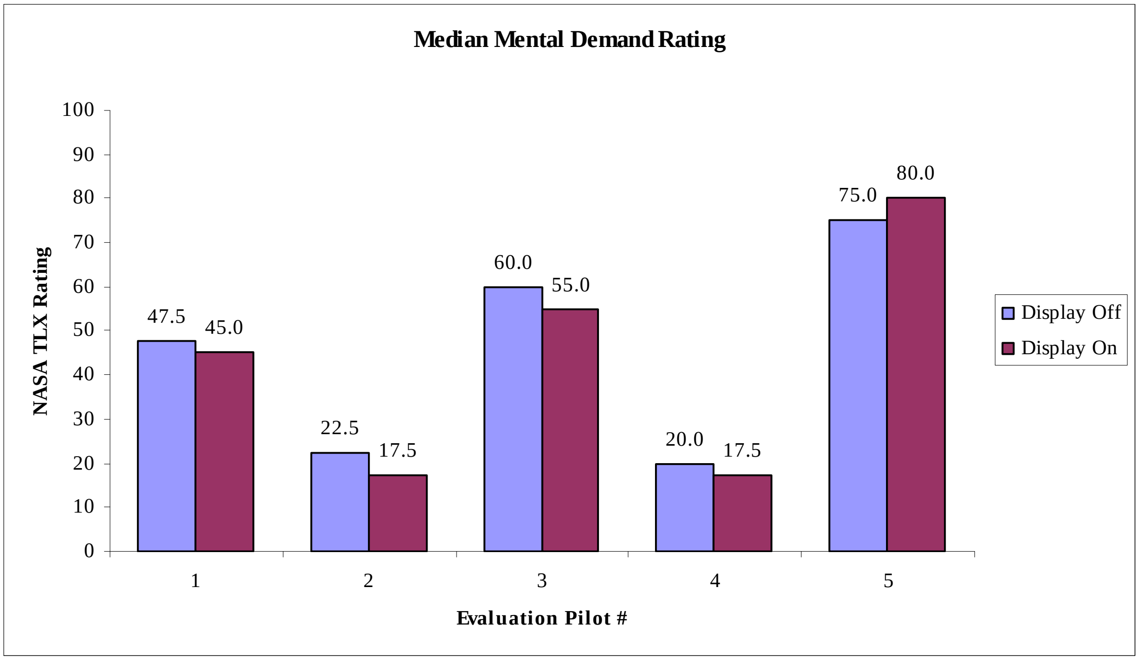


Figure 15. Median of Mental Demand Ratings.

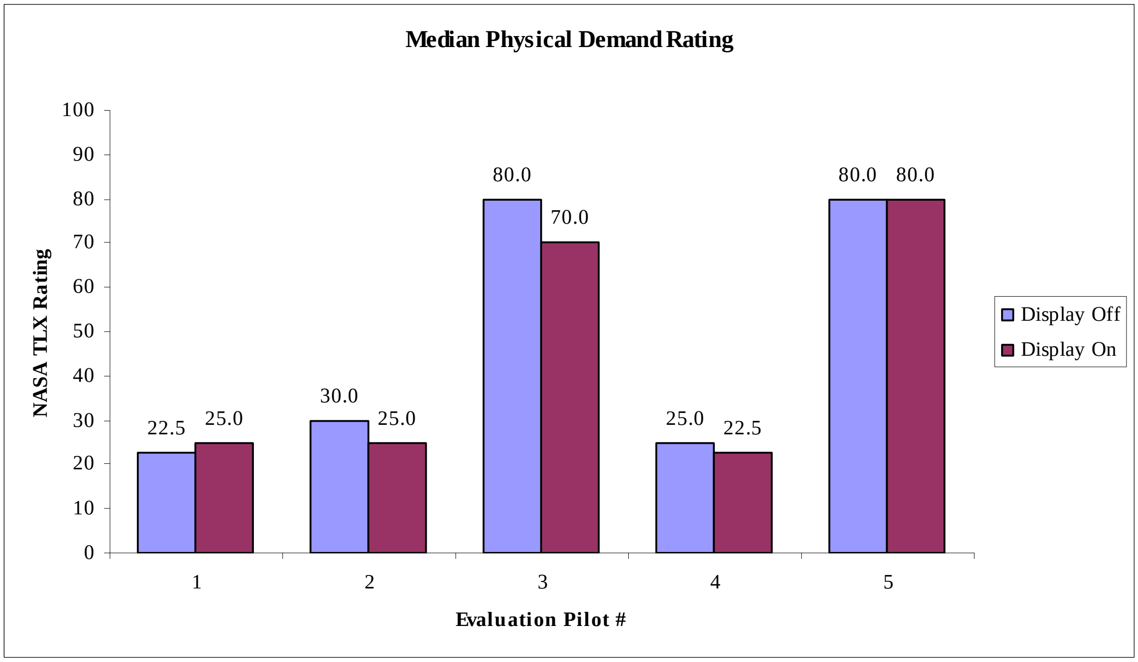


Figure 16. Median of Physical Demand Ratings.

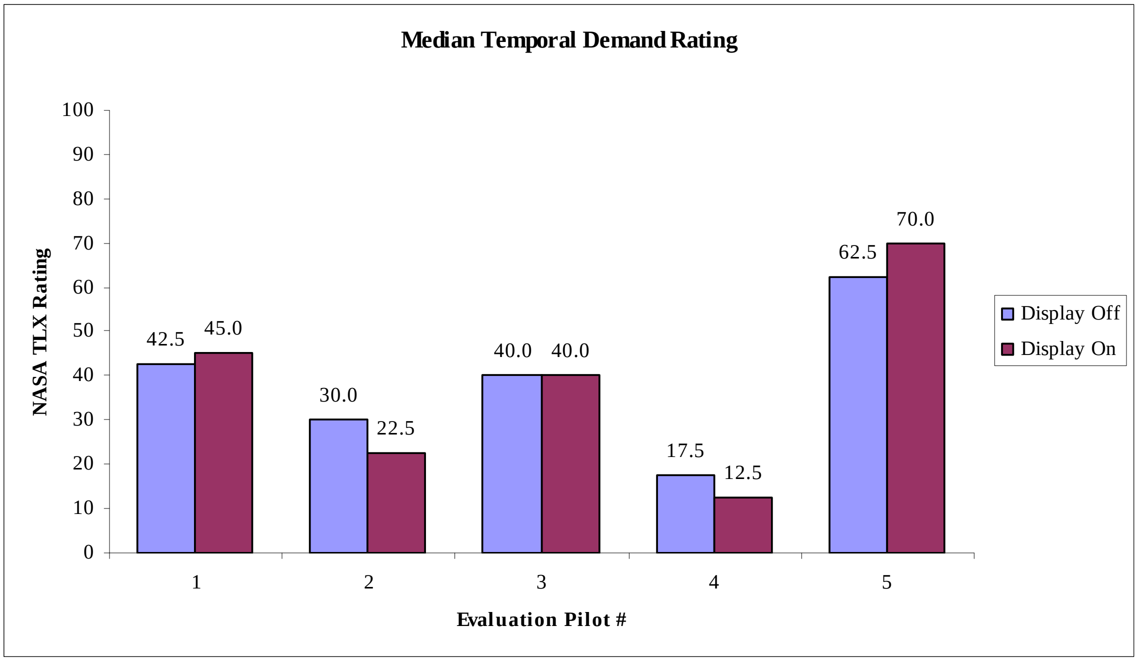


Figure 17. Median of Temporal Demand Ratings.

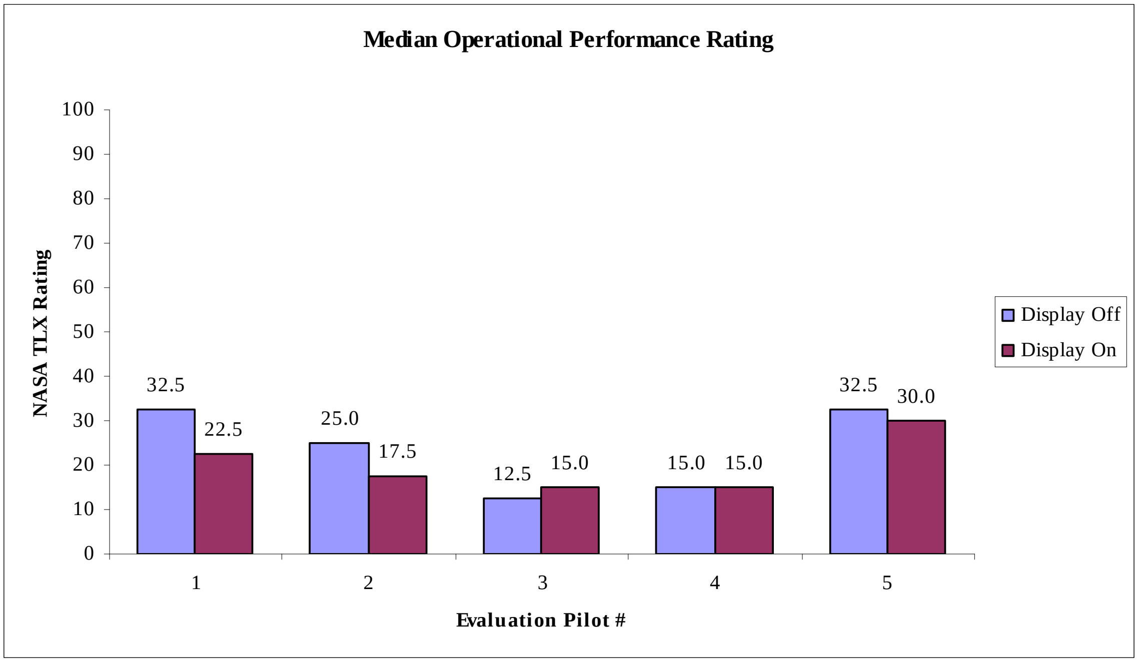


Figure 18. Median of Operational Performance Ratings.

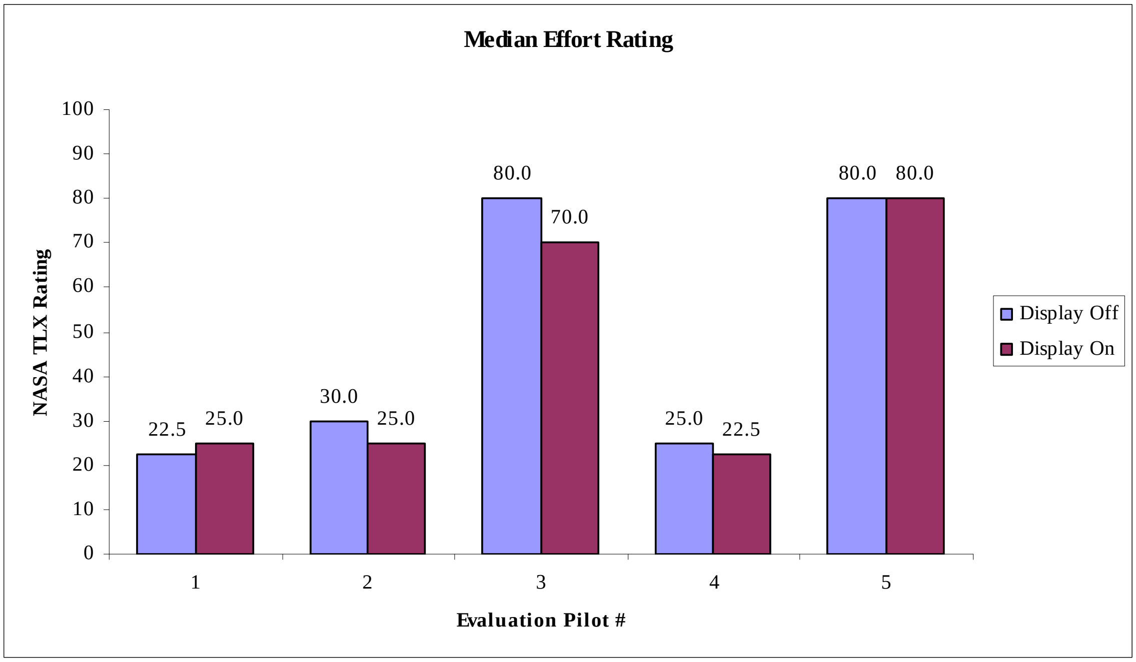


Figure 19. Median of Effort Ratings.

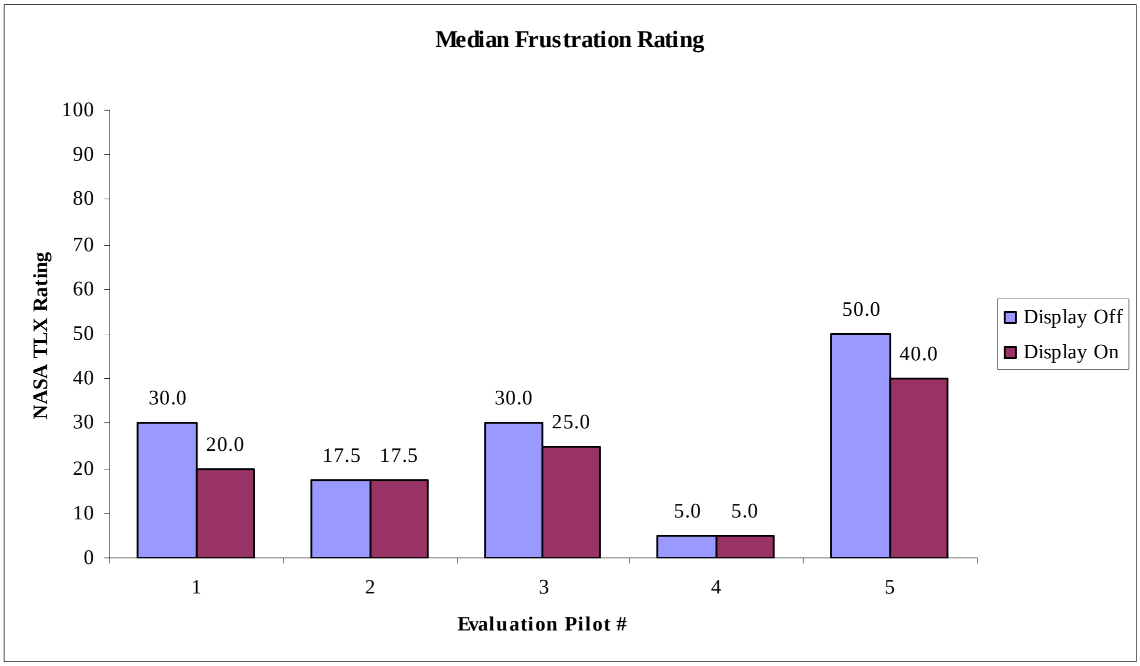


Figure 20. Median of Frustration Ratings.

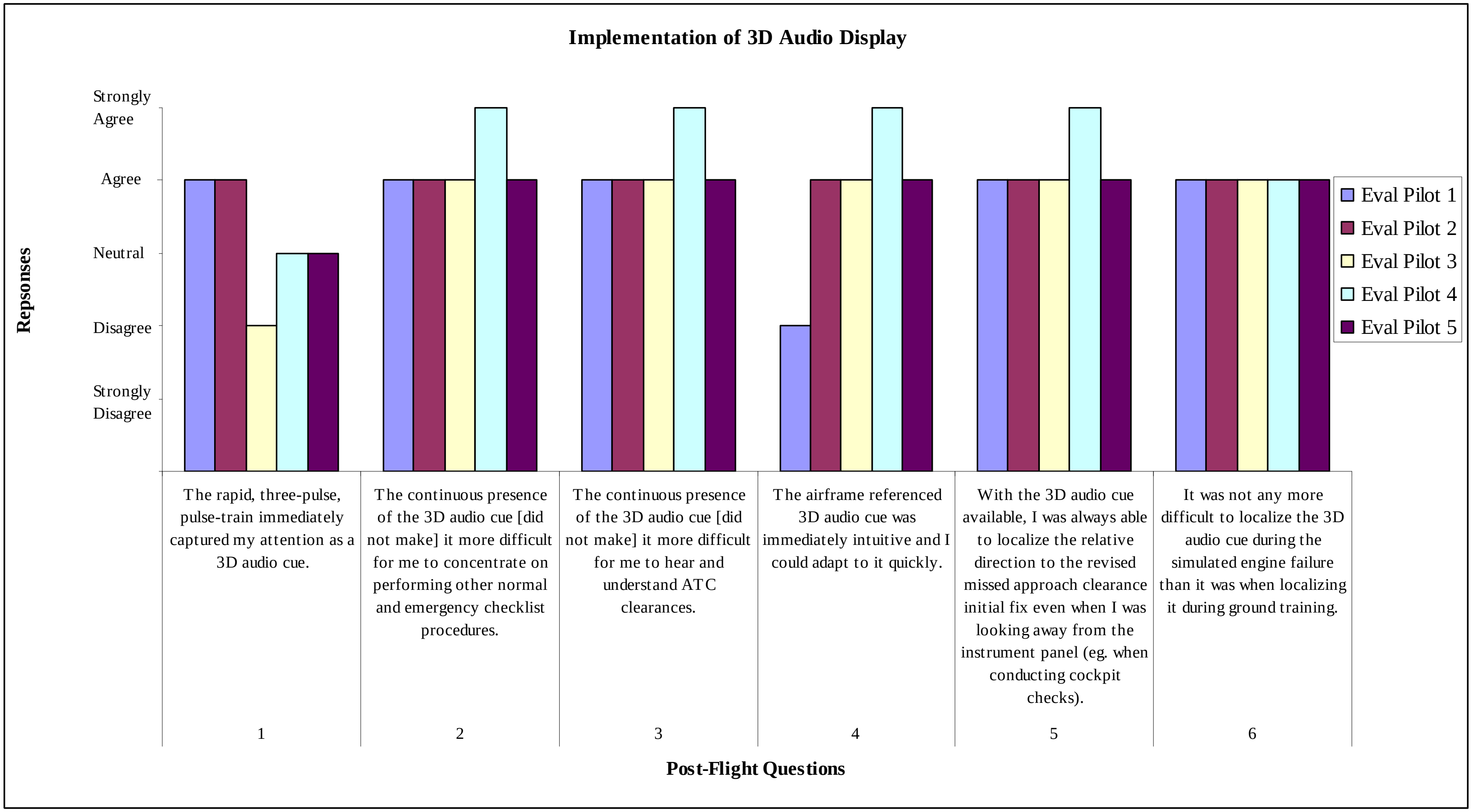


Figure 21. Pilot Responses to Questions on Implementation.

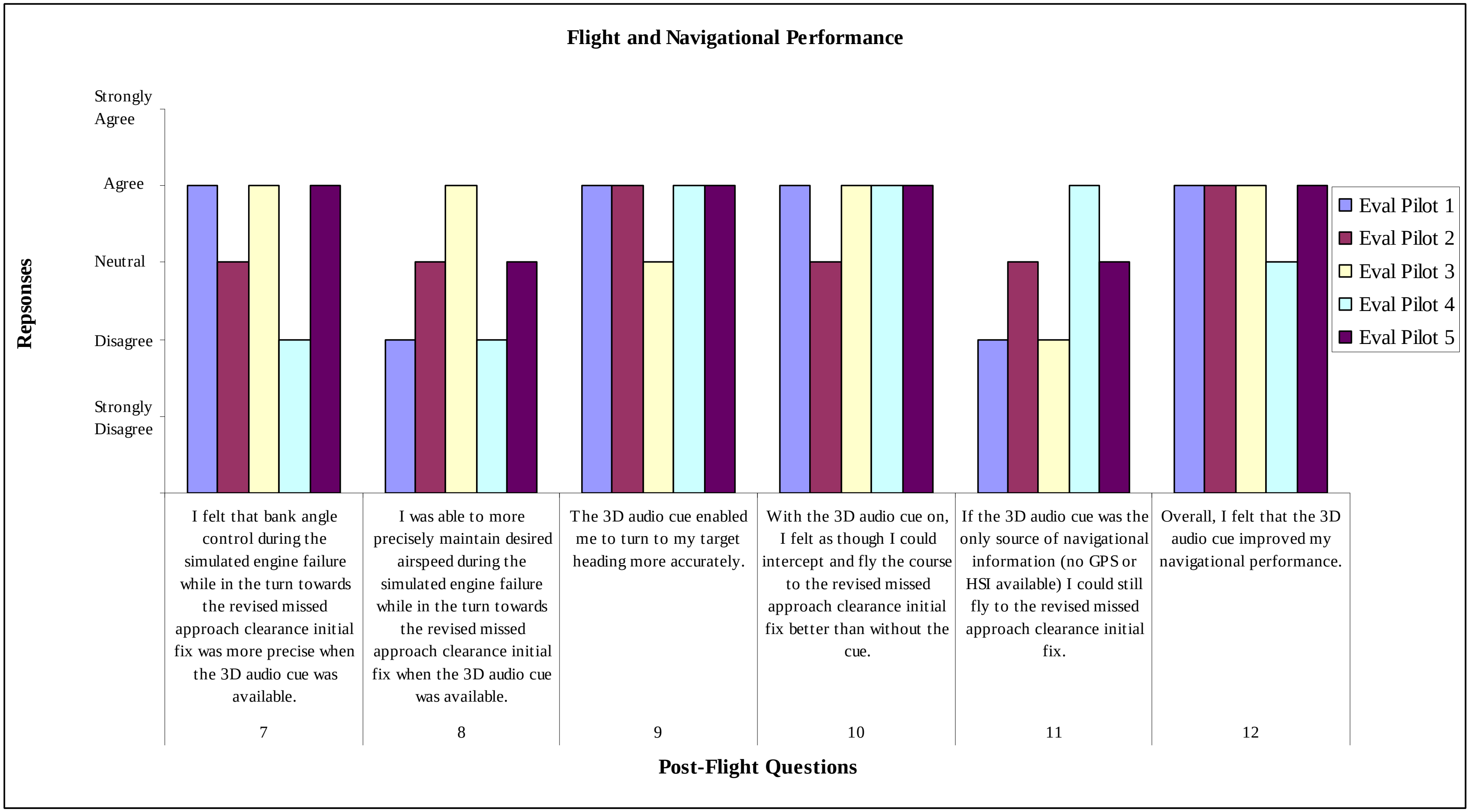


Figure 22. Pilot Responses to Questions on Flight and Navigational Performance.

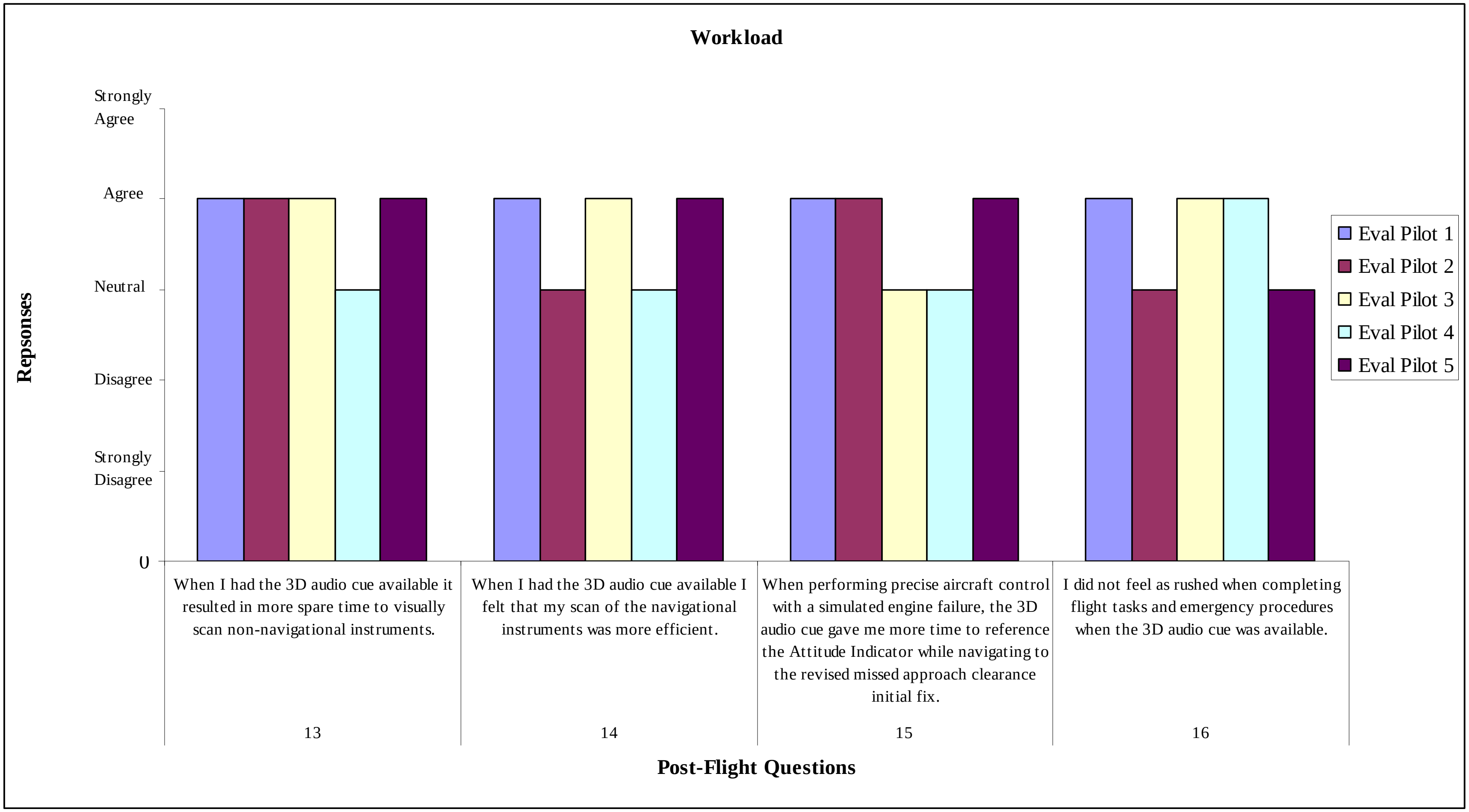


Figure 23. Pilot Responses to Questions on Workload.

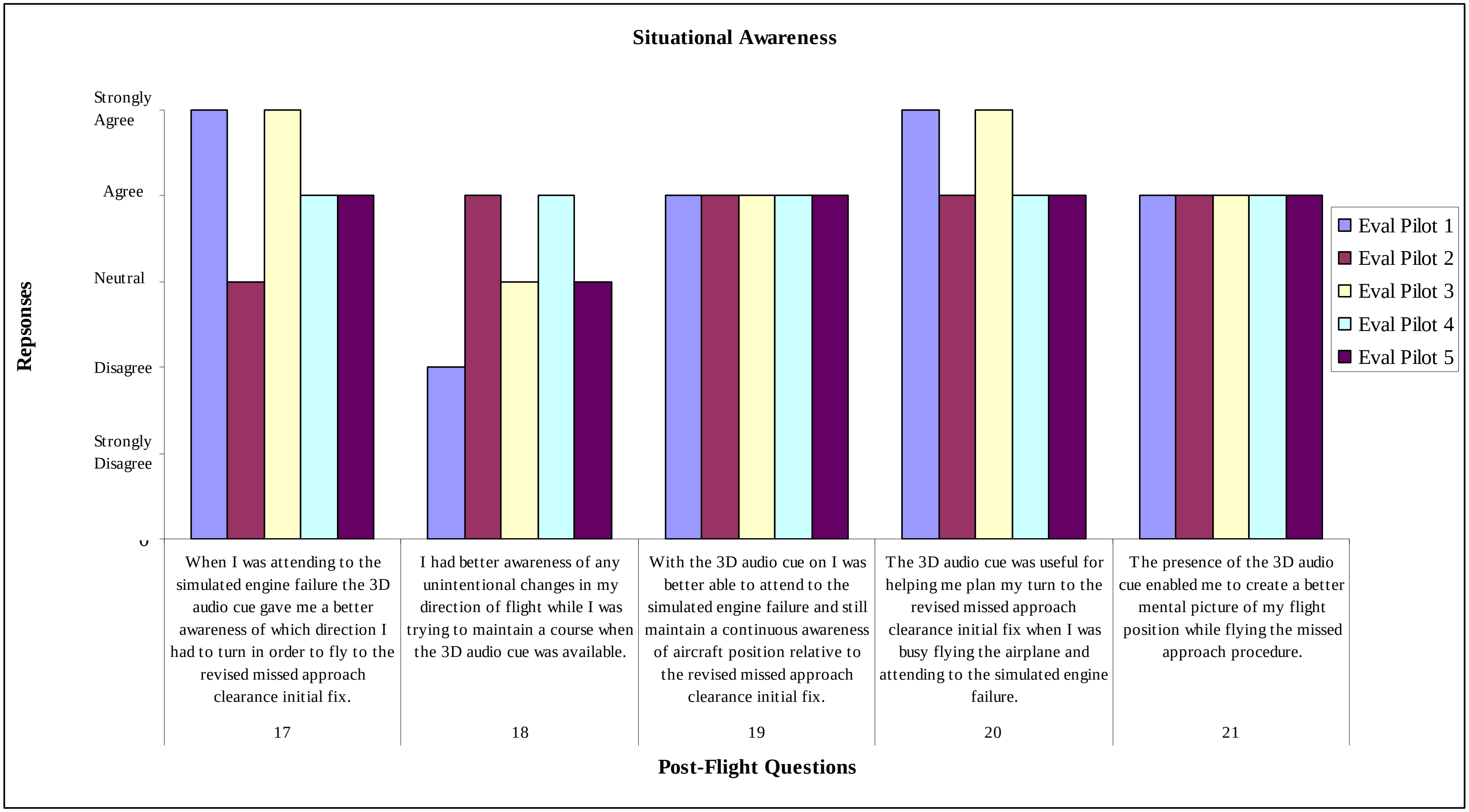


Figure 24. Pilot Responses to Questions on Situational Awareness.

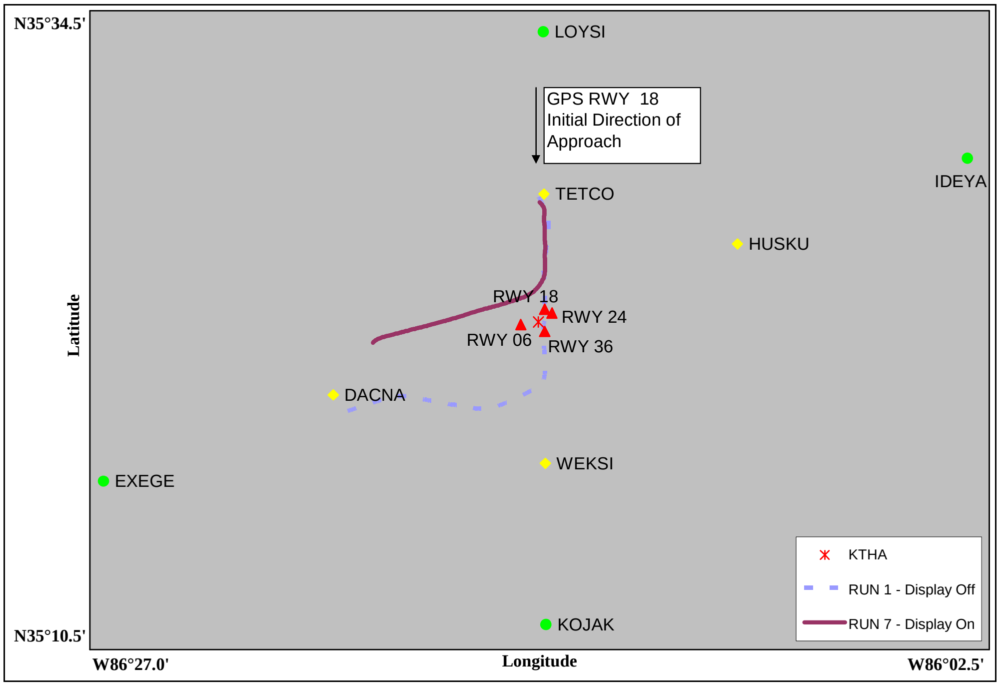


Figure 25. Aircraft Track for Evaluation Pilot 1, Run 1 & 7.

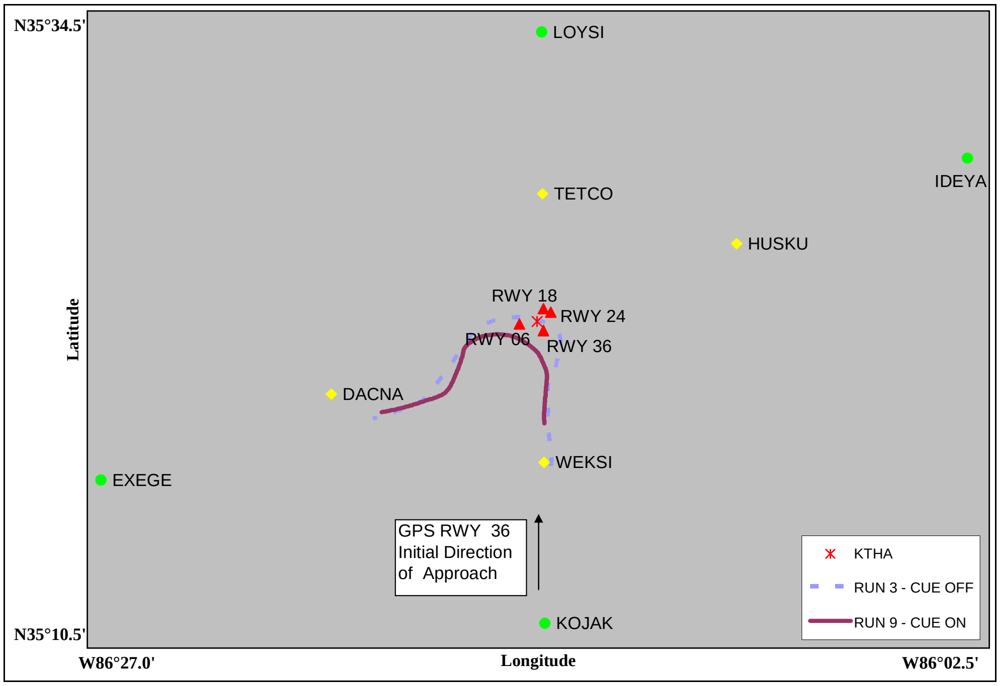


Figure 26. Aircraft Track for Evaluation Pilot 1, Run 3 & 9.

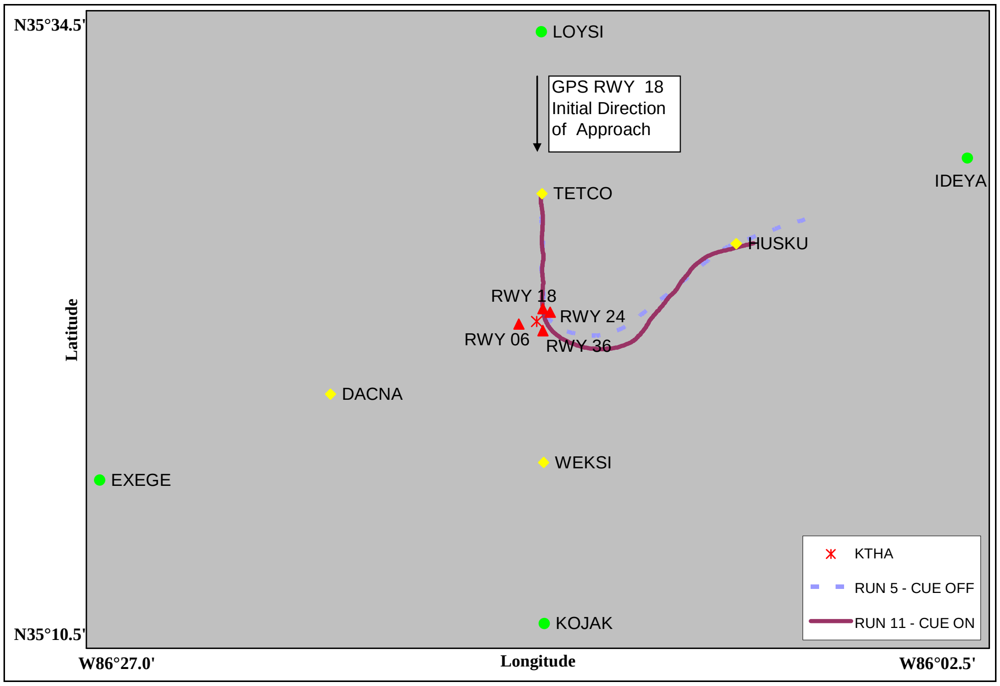


Figure 27. Aircraft Track for Evaluation Pilot 1, Run 5 & 11.

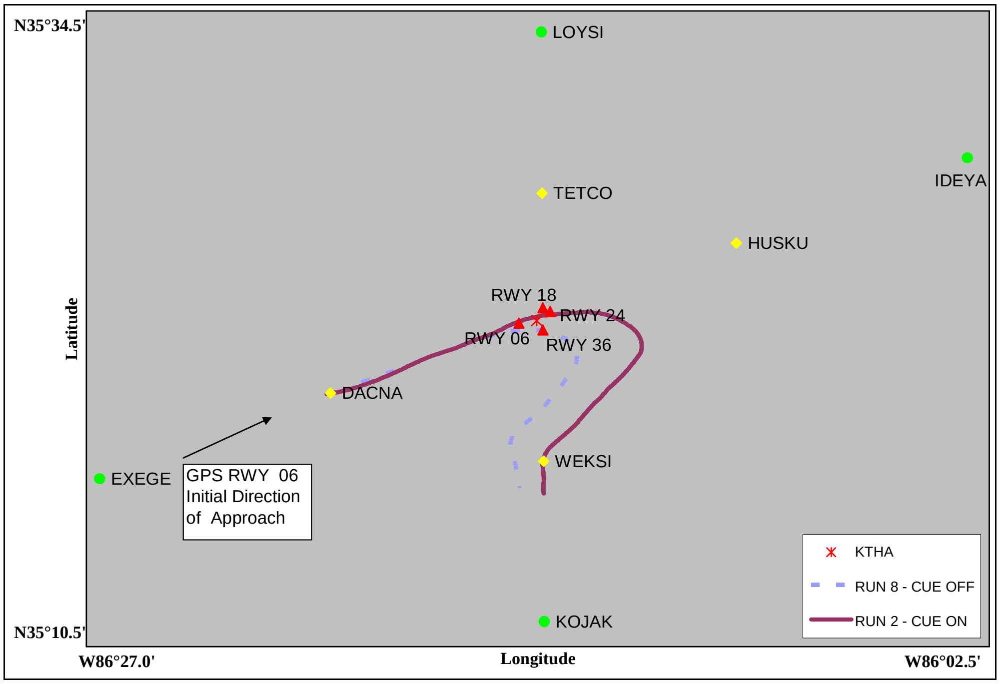


Figure 28. Aircraft Track for Evaluation Pilot 1, Run 8 & 2.

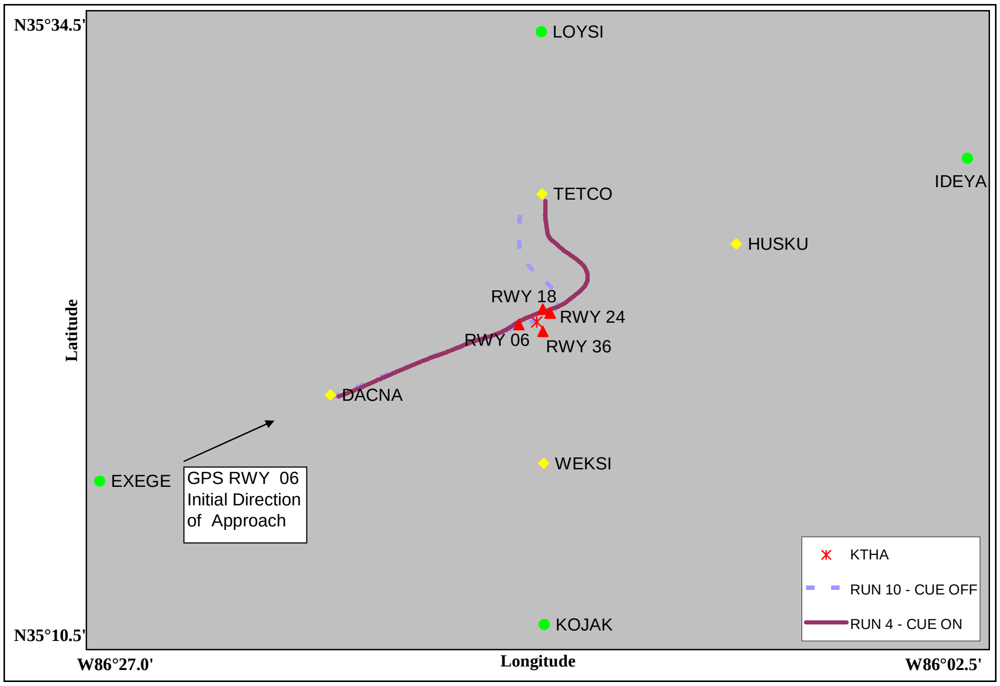


Figure 29. Aircraft Track for Evaluation Pilot 1, Run 10 & 4.

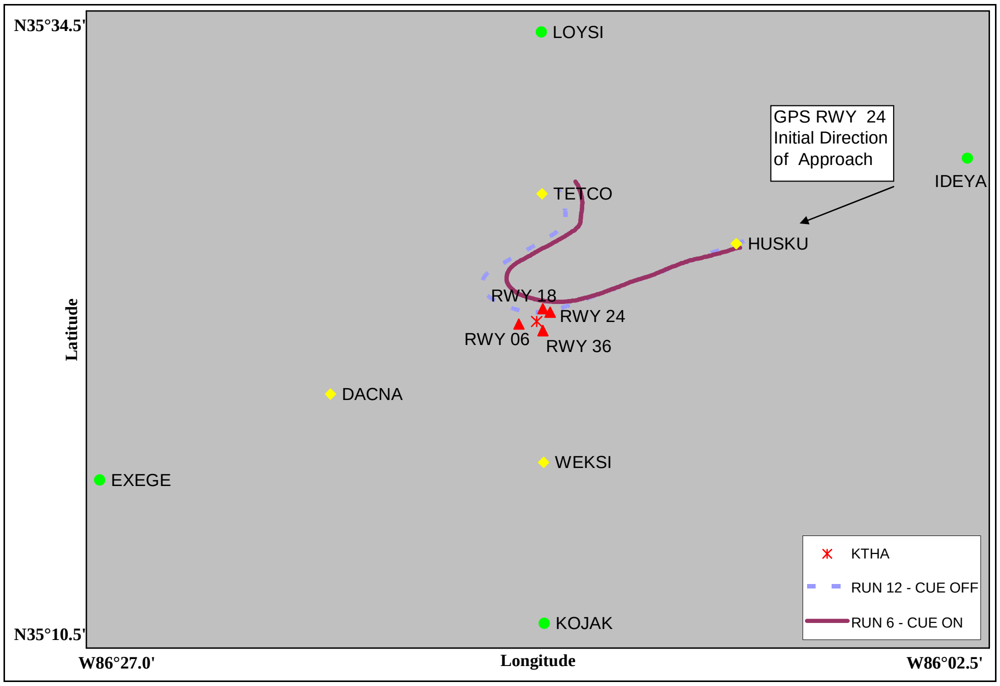


Figure 30. Aircraft Track for Evaluation Pilot 1, Run 12 & 6.

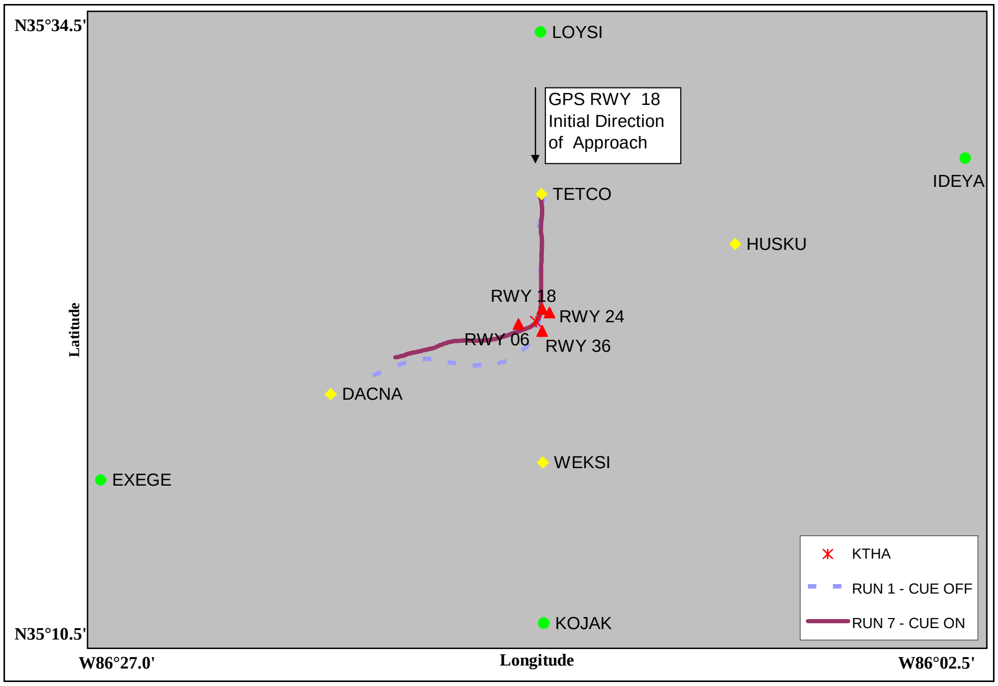


Figure 31. Aircraft Track for Evaluation Pilot 2, Run 1 & 7.

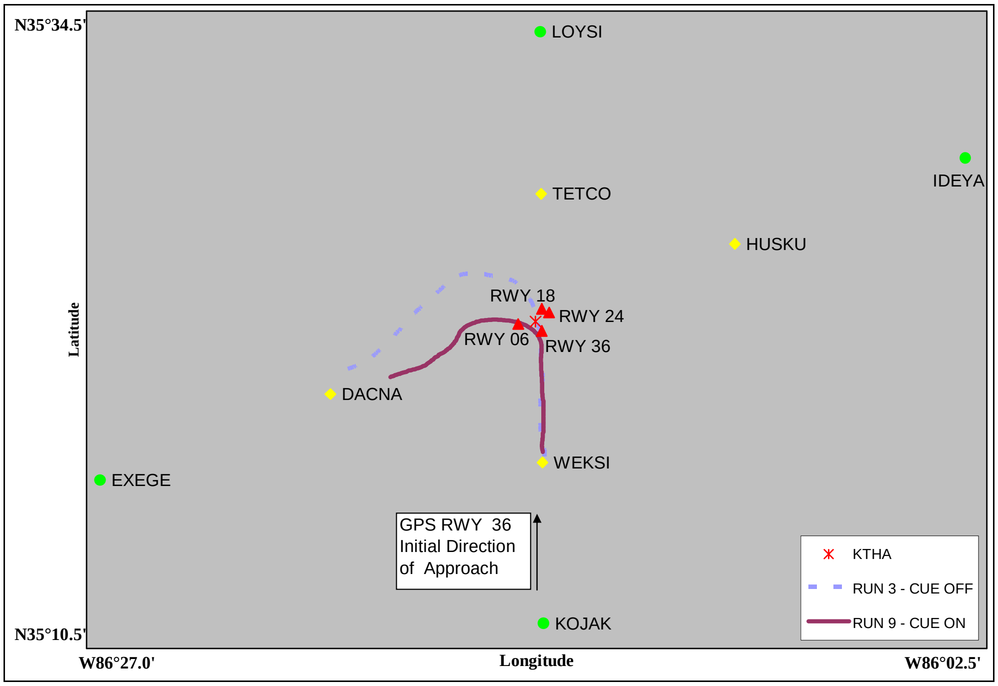


Figure 32. Aircraft Track for Evaluation Pilot 2, Run 3 & 9.

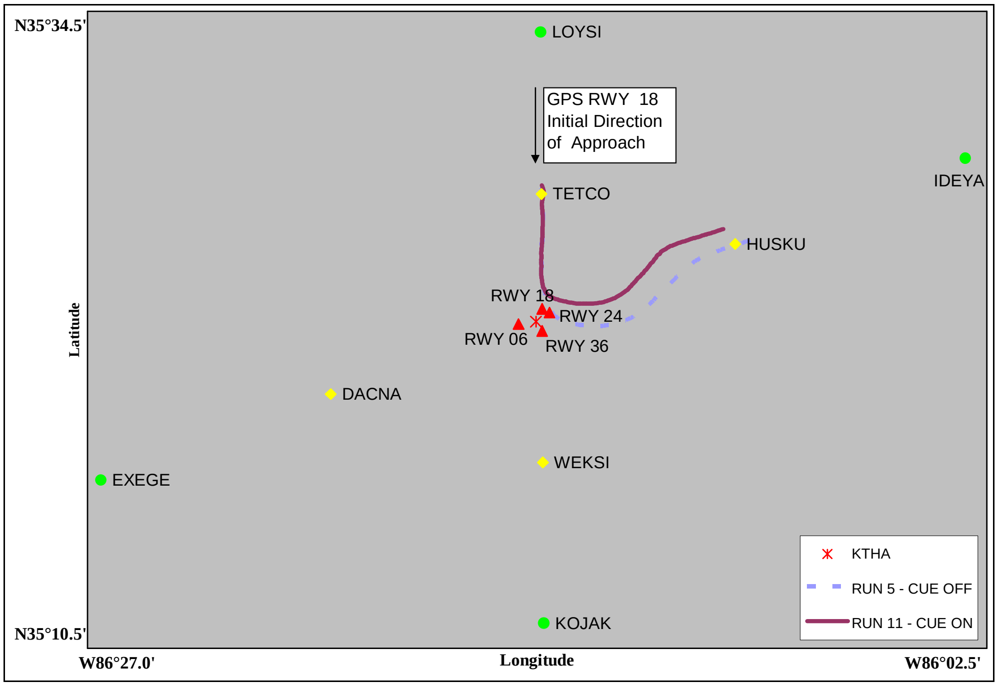


Figure 33. Aircraft Track for Evaluation Pilot 2, Run 5 & 11.

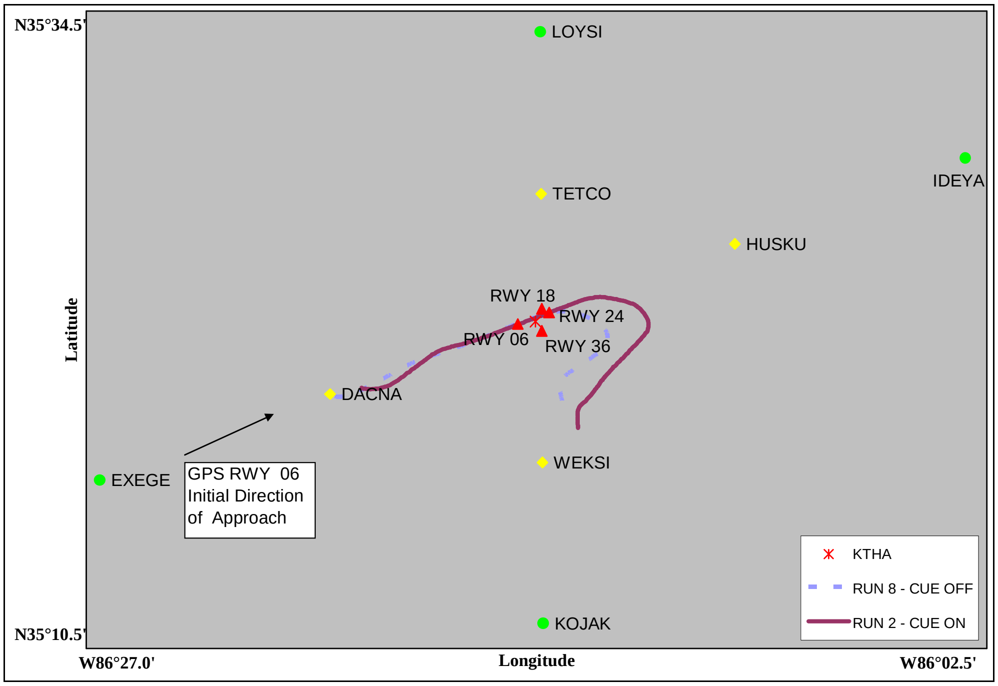


Figure 34. Aircraft Track for Evaluation Pilot 2, Run 8 & 2.

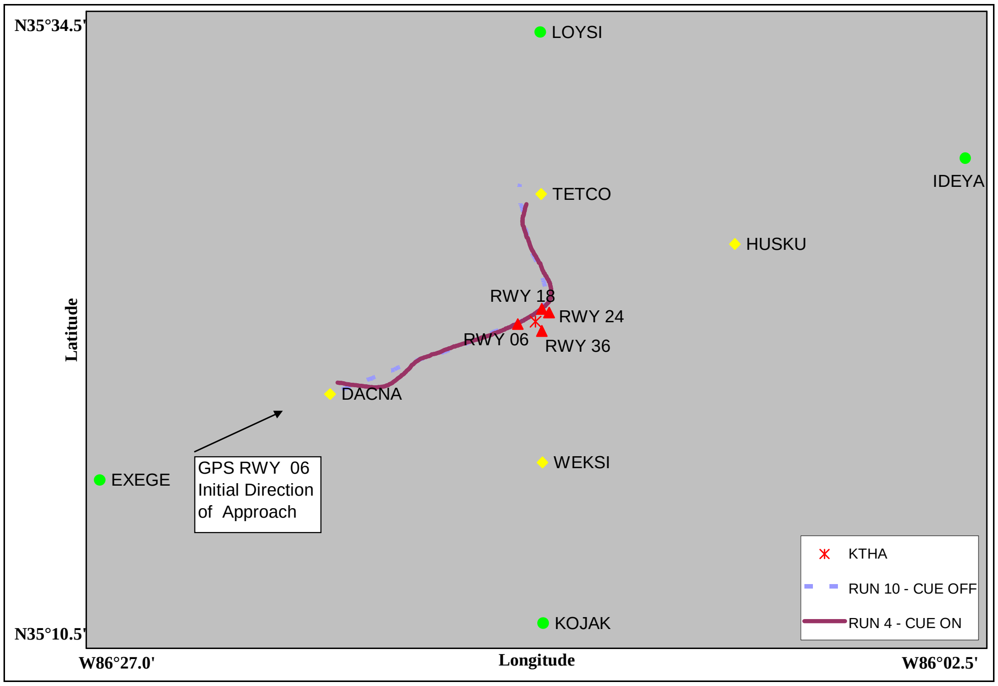


Figure 35. Aircraft Track for Evaluation Pilot 2, Run 10 & 4.

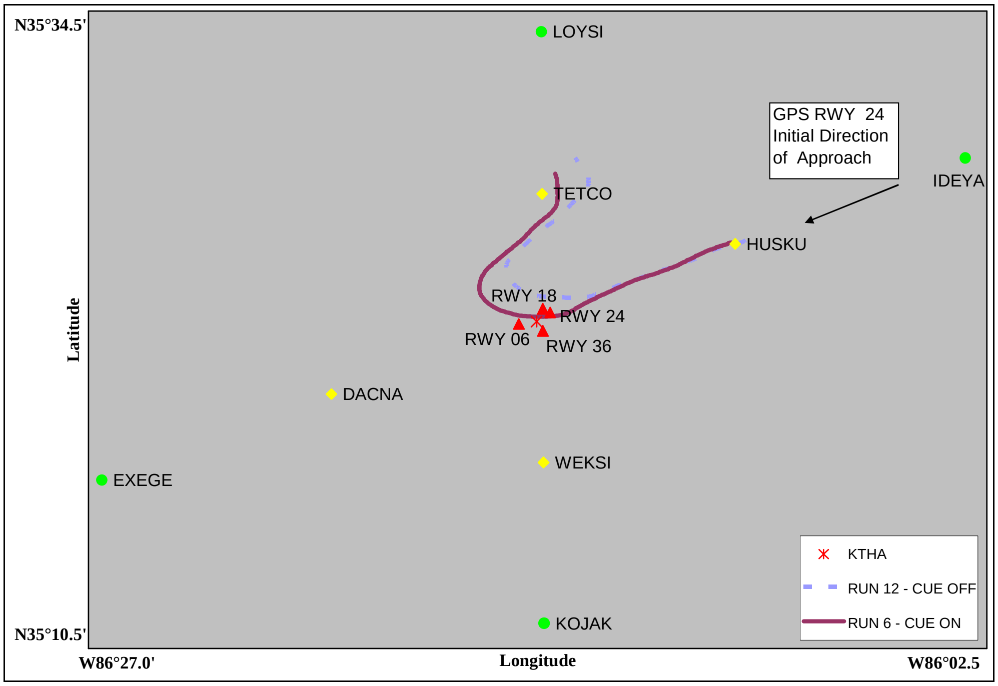


Figure 36. Aircraft Track for Evaluation Pilot 2, Run 12 & 6.

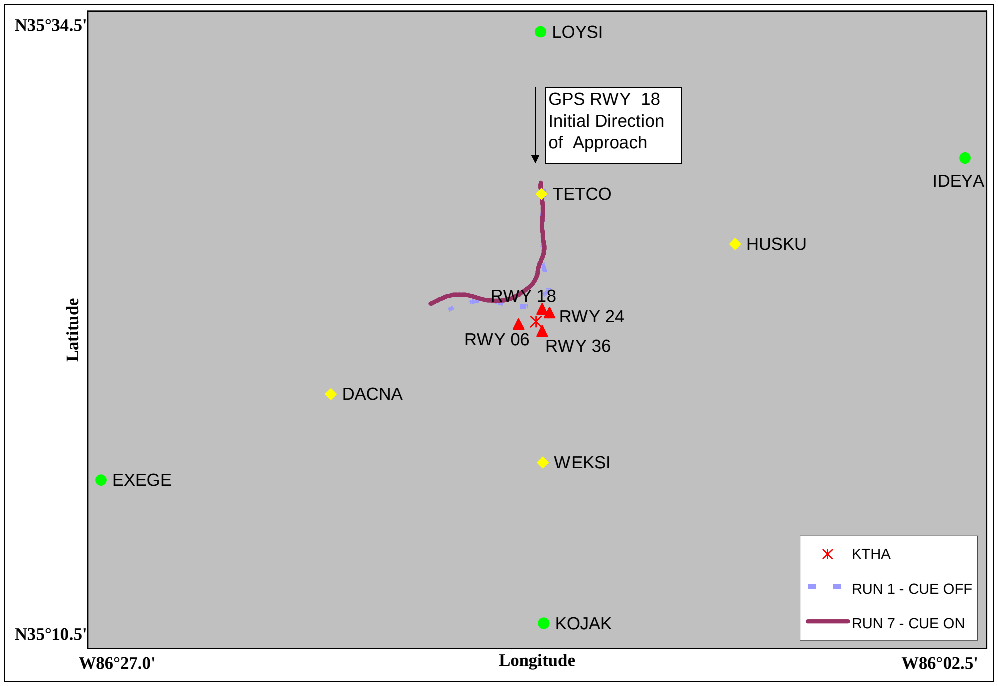


Figure 37. Aircraft Track for Evaluation Pilot 3, Run 1 & 7.

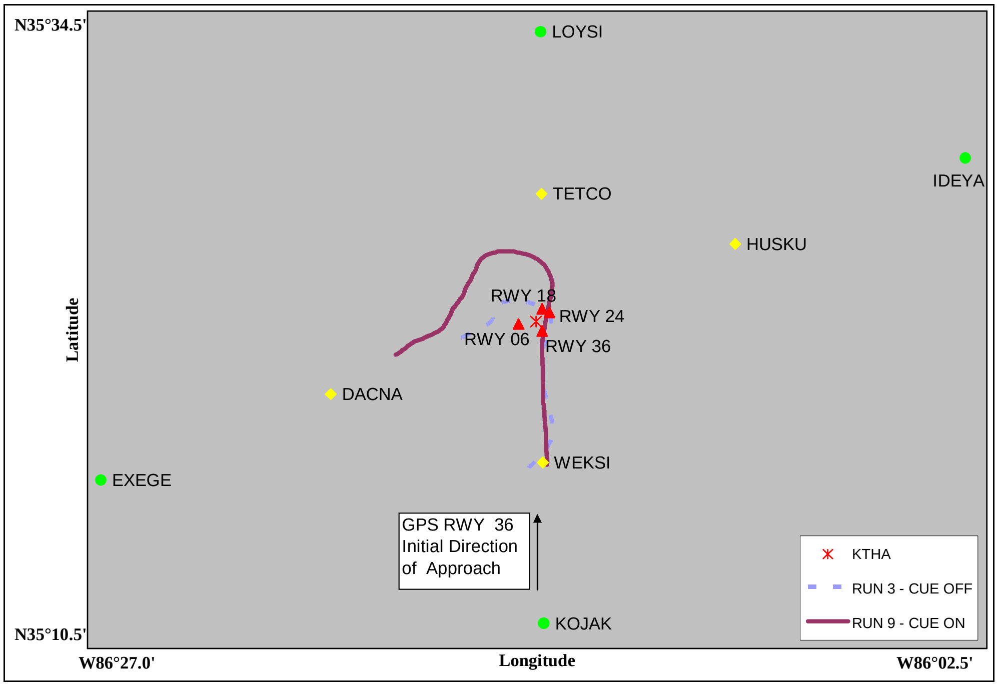


Figure 38. Aircraft Track for Evaluation Pilot 3, Run 3 & 9.

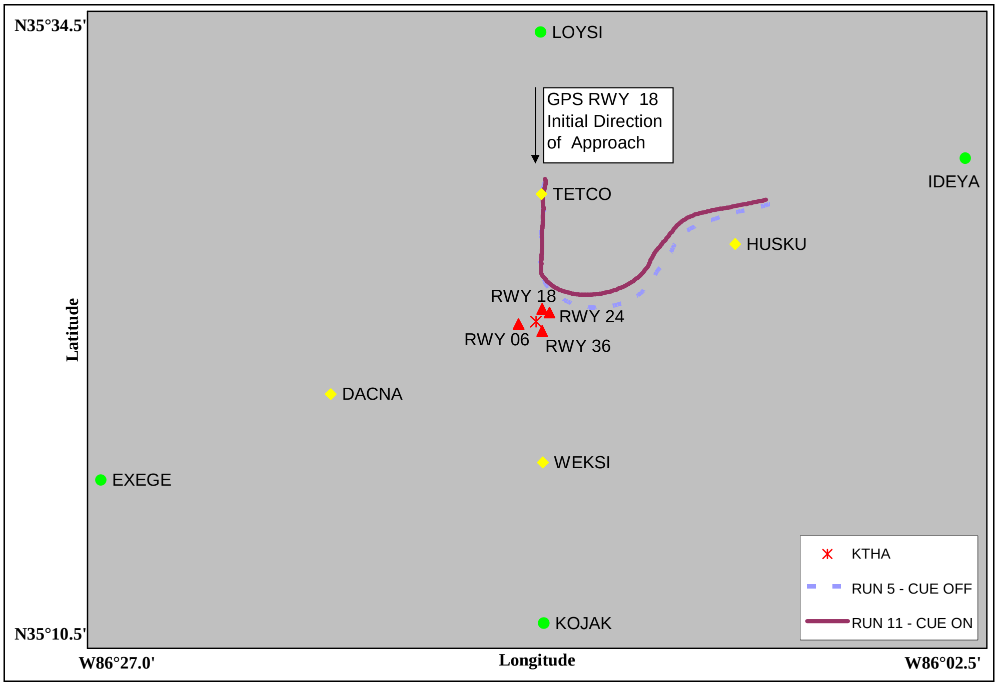


Figure 39. Aircraft Track for Evaluation Pilot 3, Run 5 & 11.

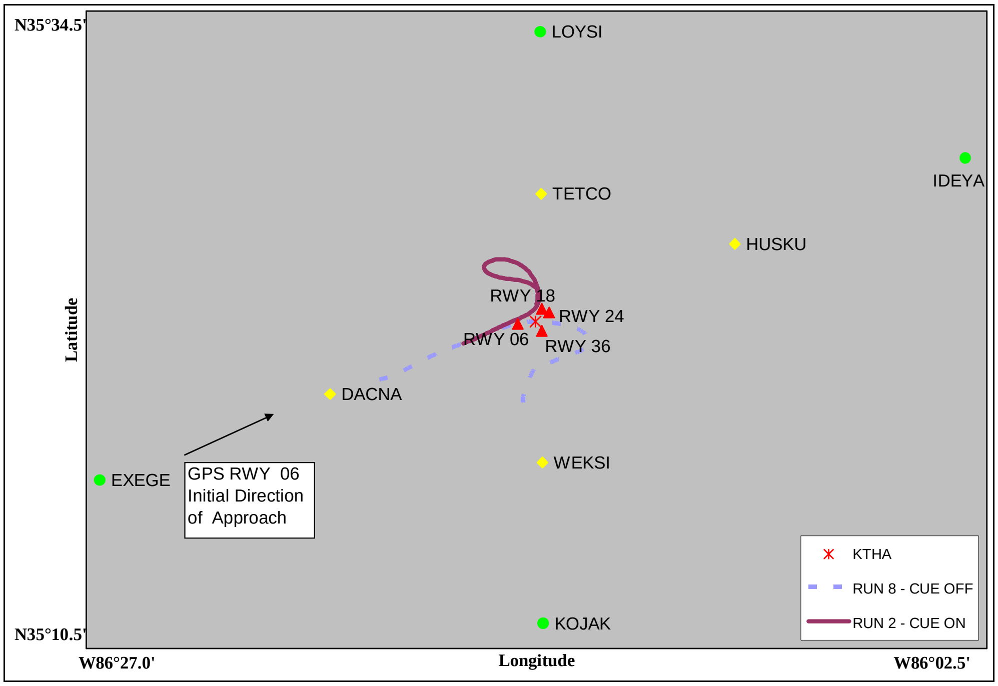


Figure 40. Aircraft Track for Evaluation Pilot 3, Run 8 & 2.

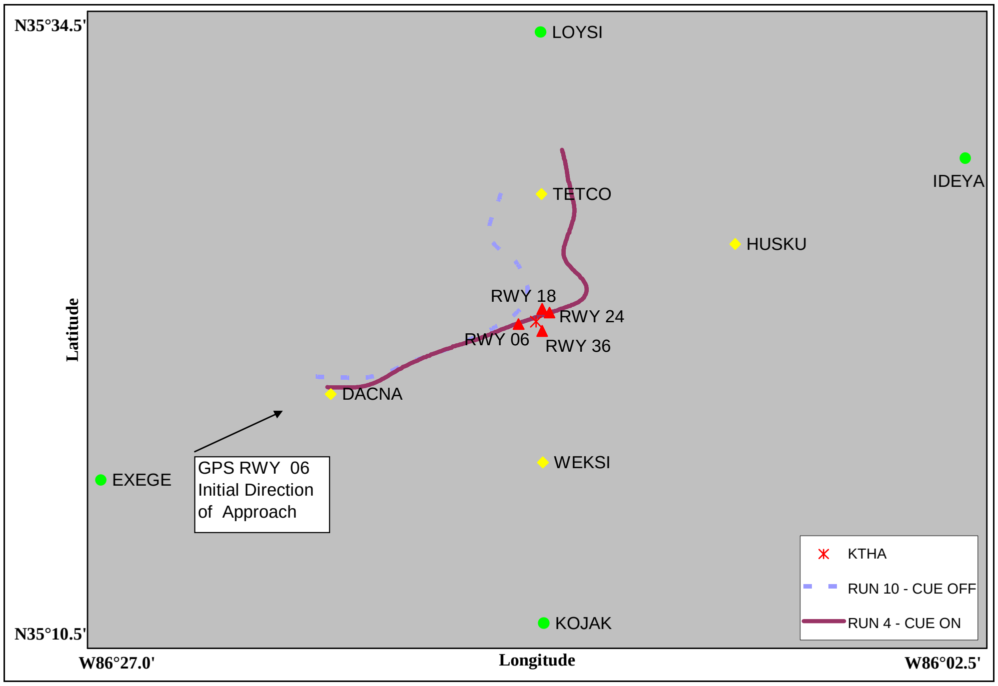


Figure 41. Aircraft Track for Evaluation Pilot 3, Run 10 & 4.

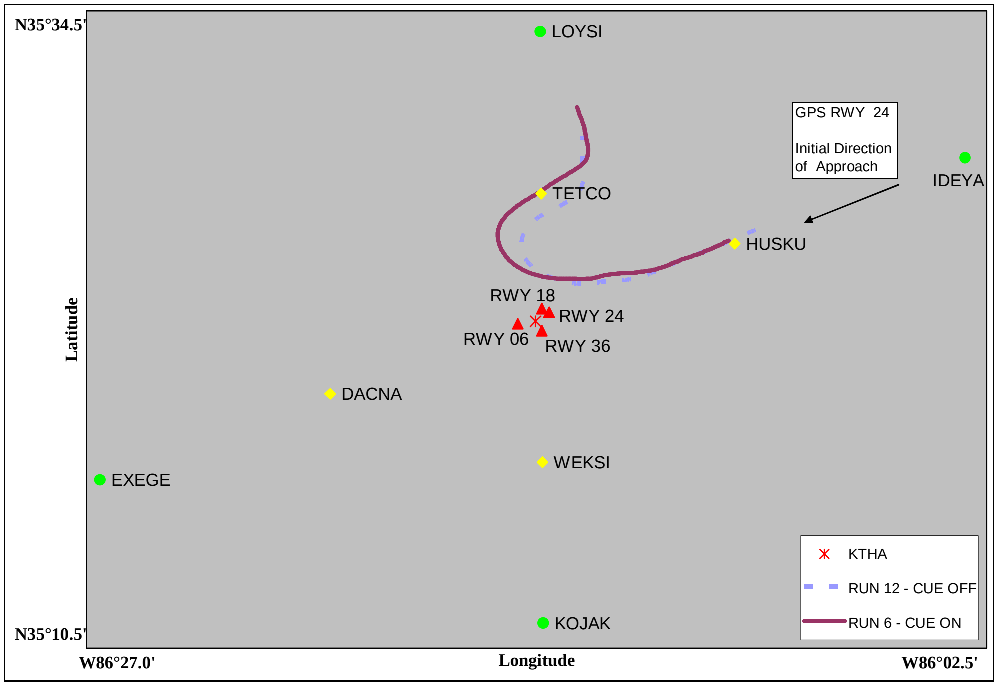


Figure 42. Aircraft Track for Evaluation Pilot 3, Run 12 & 6.

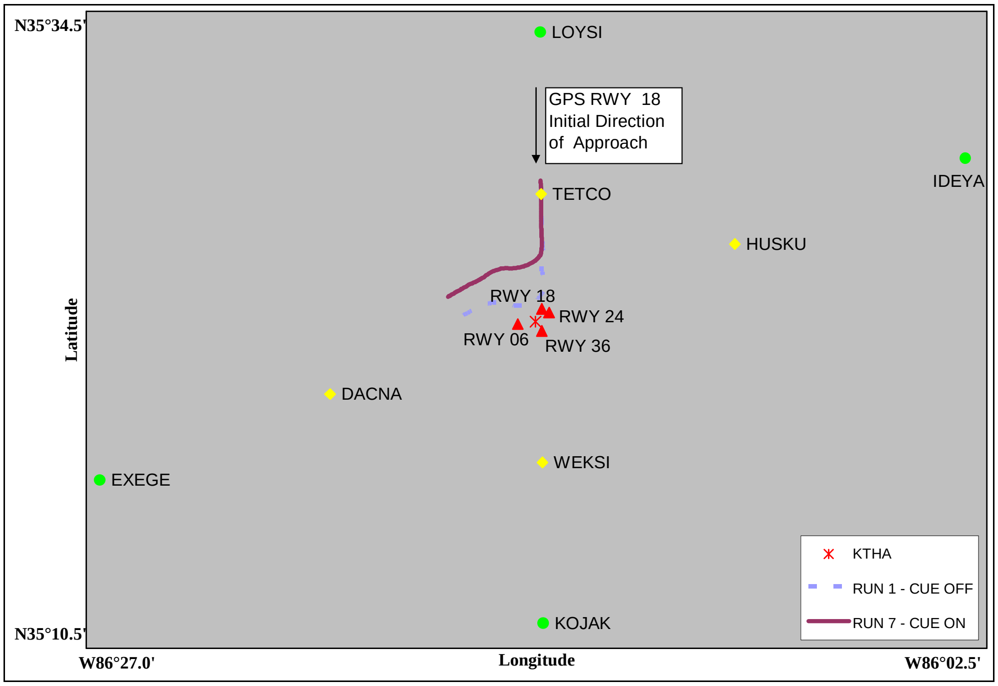


Figure 43. Aircraft Track for Evaluation Pilot 4, Run 1 & 7.

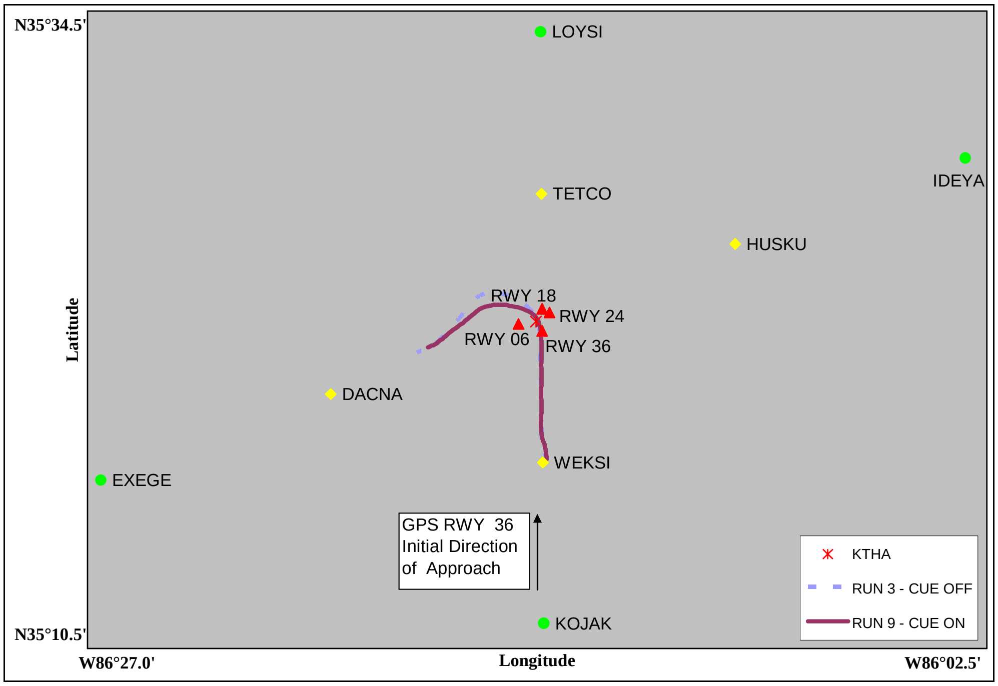


Figure 44. Aircraft Track for Evaluation Pilot 4, Run 3 & 9.

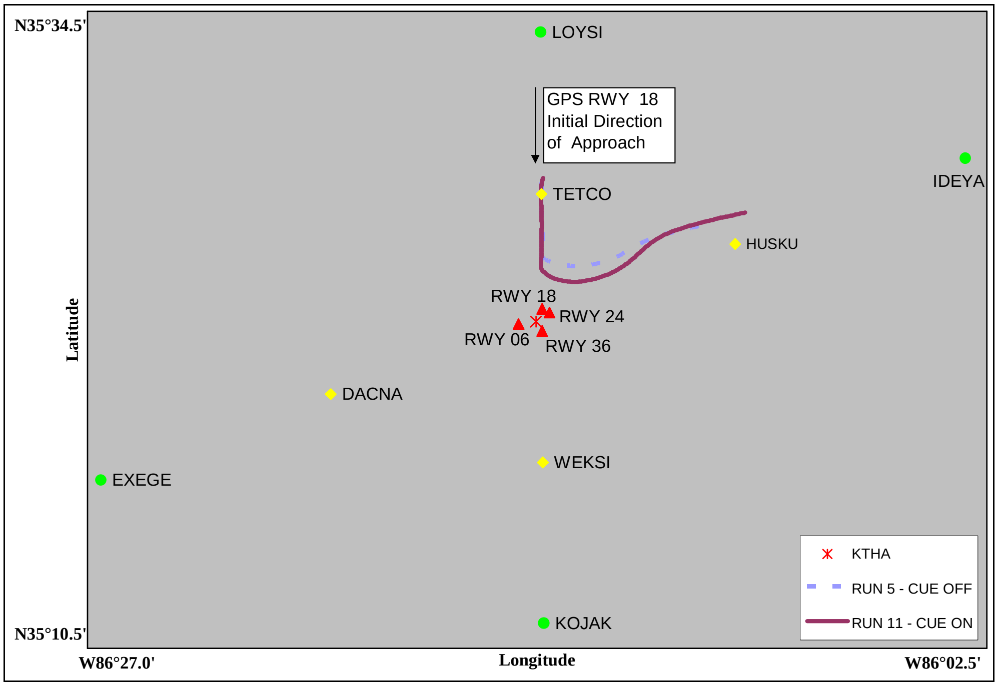


Figure 45. Aircraft Track for Evaluation Pilot 4, Run 5 & 11.

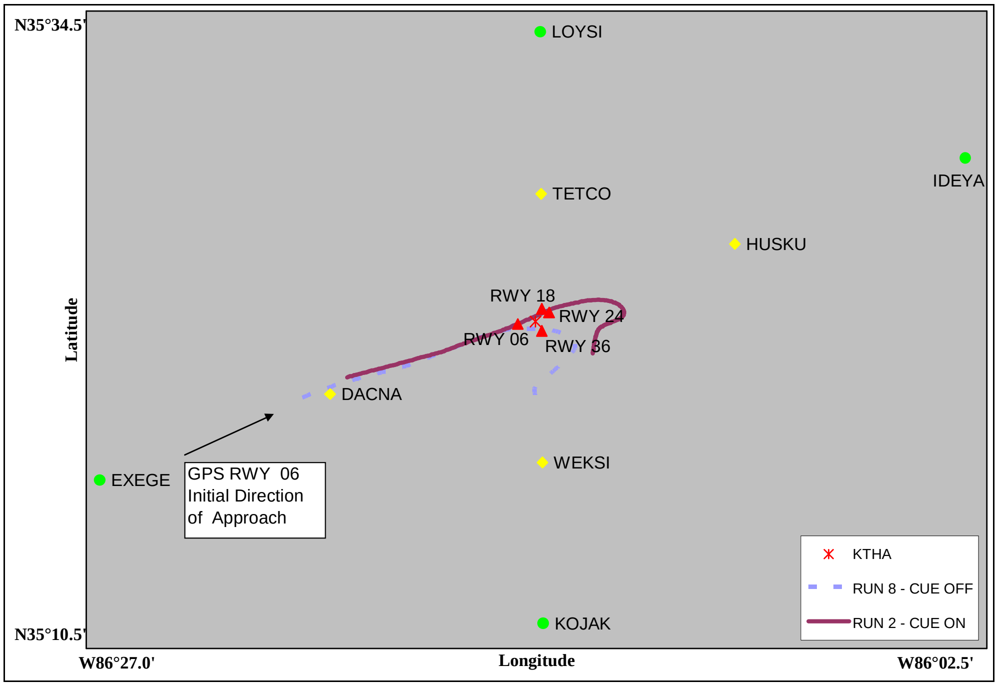


Figure 46. Aircraft Track for Evaluation Pilot 4, Run 8 & 2.

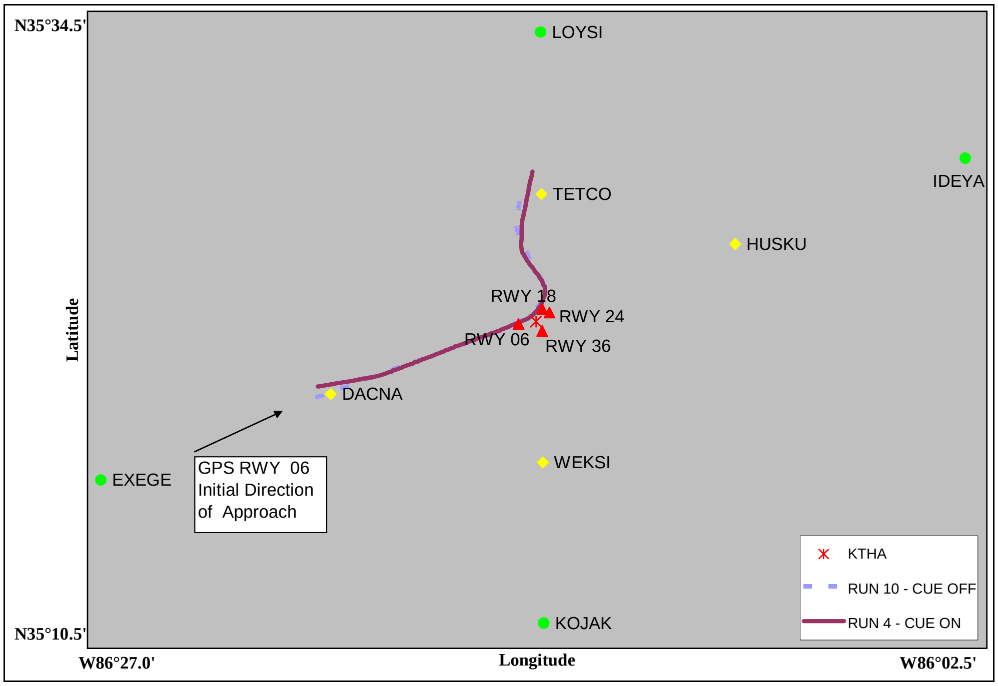


Figure 47. Aircraft Track for Evaluation Pilot 4, Run 10 & 4.

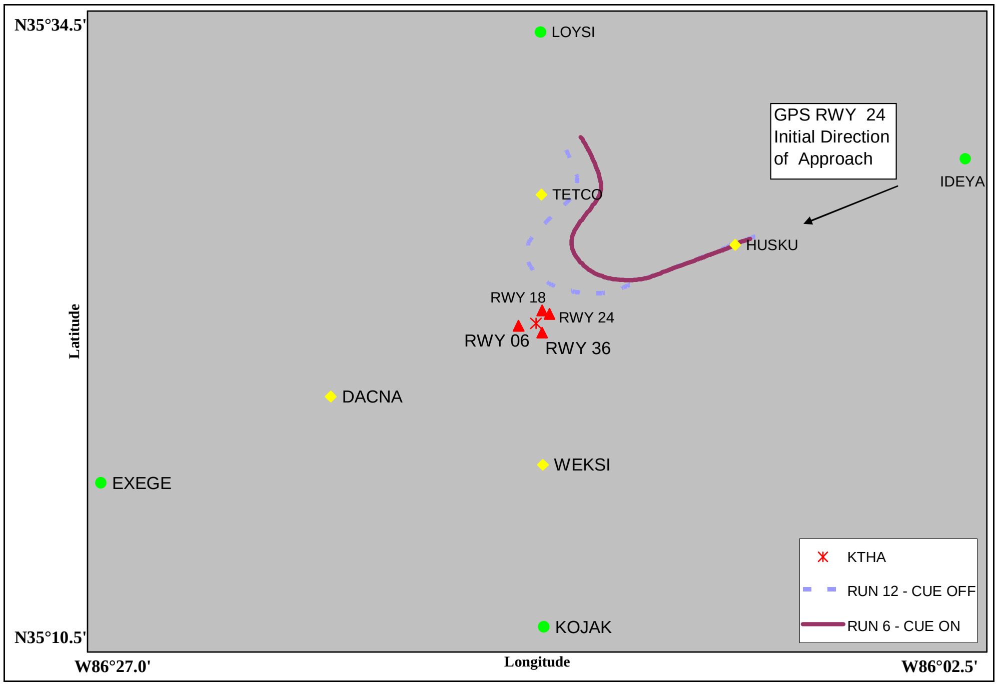


Figure 48. Aircraft Track for Evaluation Pilot 4, Run 12 & 6.

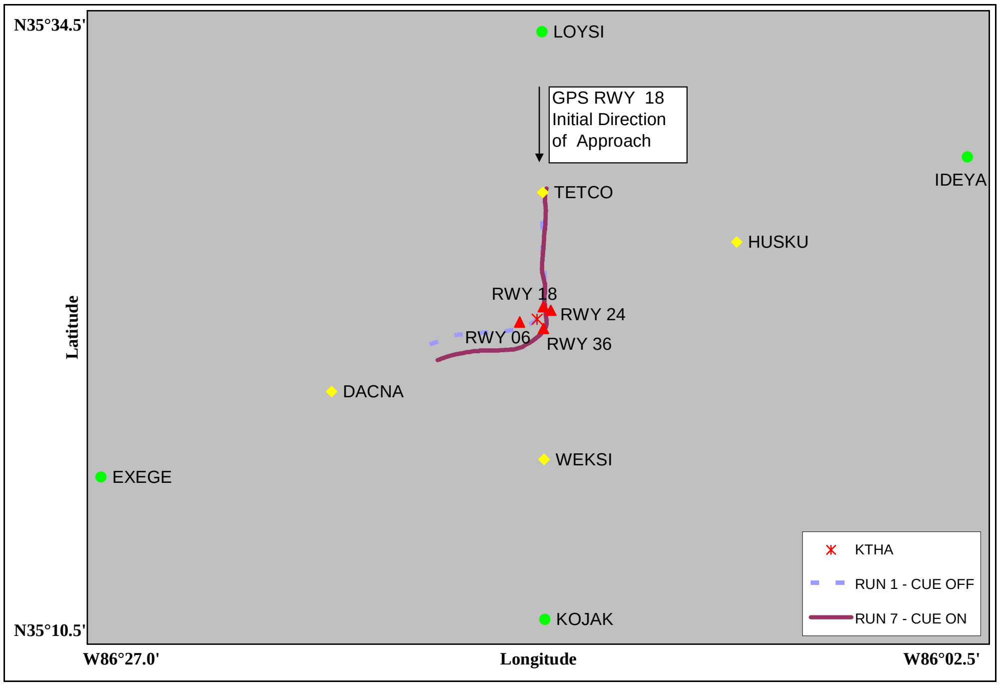


Figure 49. Aircraft Track for Evaluation Pilot 5, Run 1 & 7.

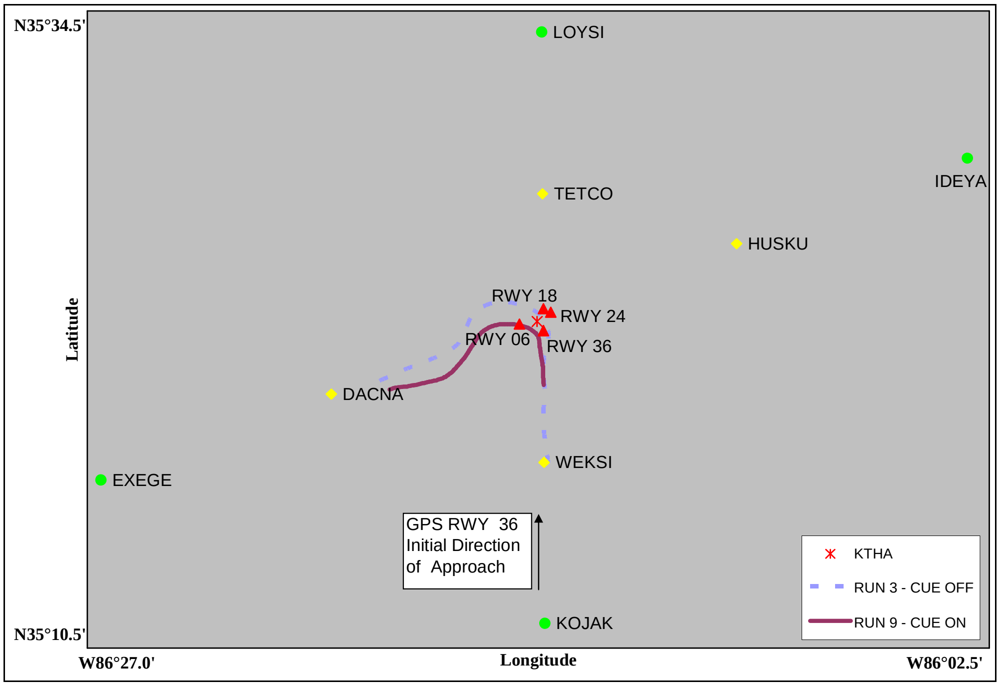


Figure 50. Aircraft Track for Evaluation Pilot 5, Run 3 & 9.

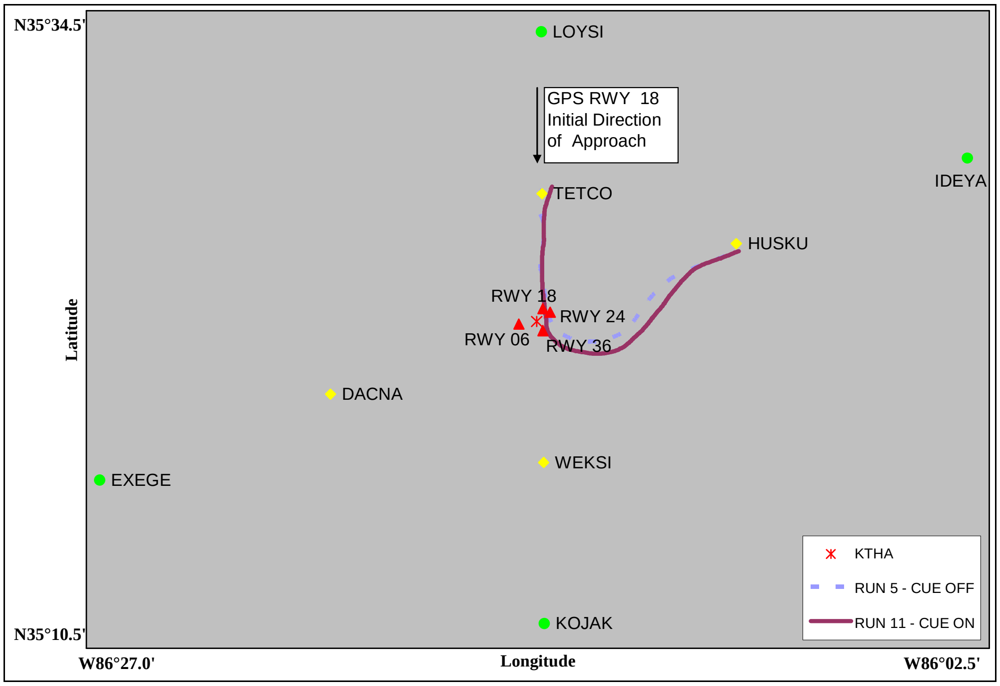


Figure 51. Aircraft Track for Evaluation Pilot 5, Run 5 & 11.

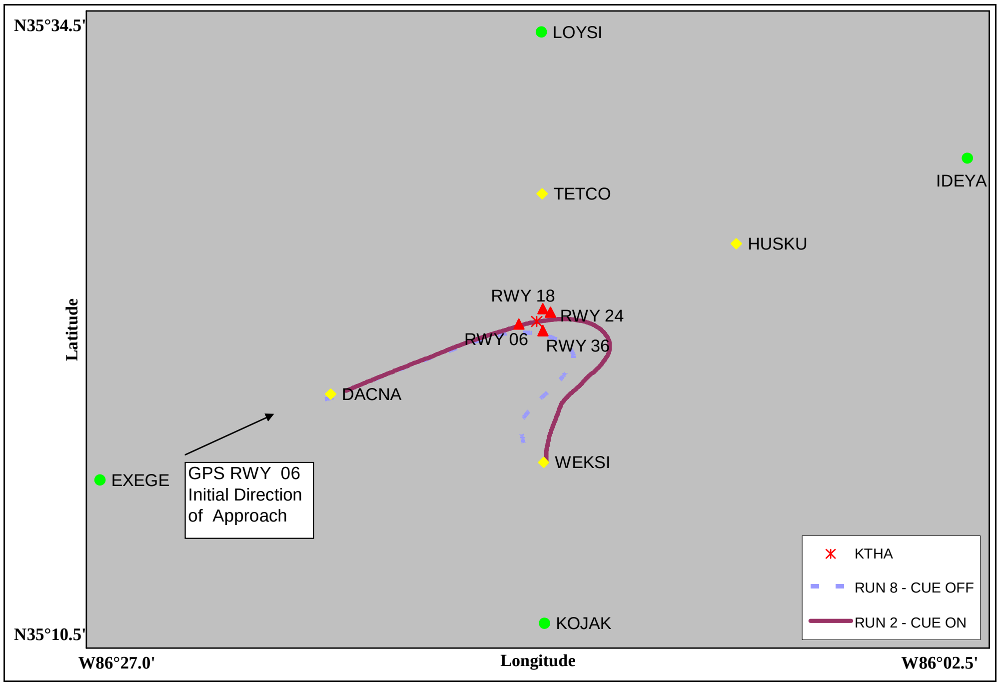


Figure 52. Aircraft Track for Evaluation Pilot 5, Run 8 & 2.

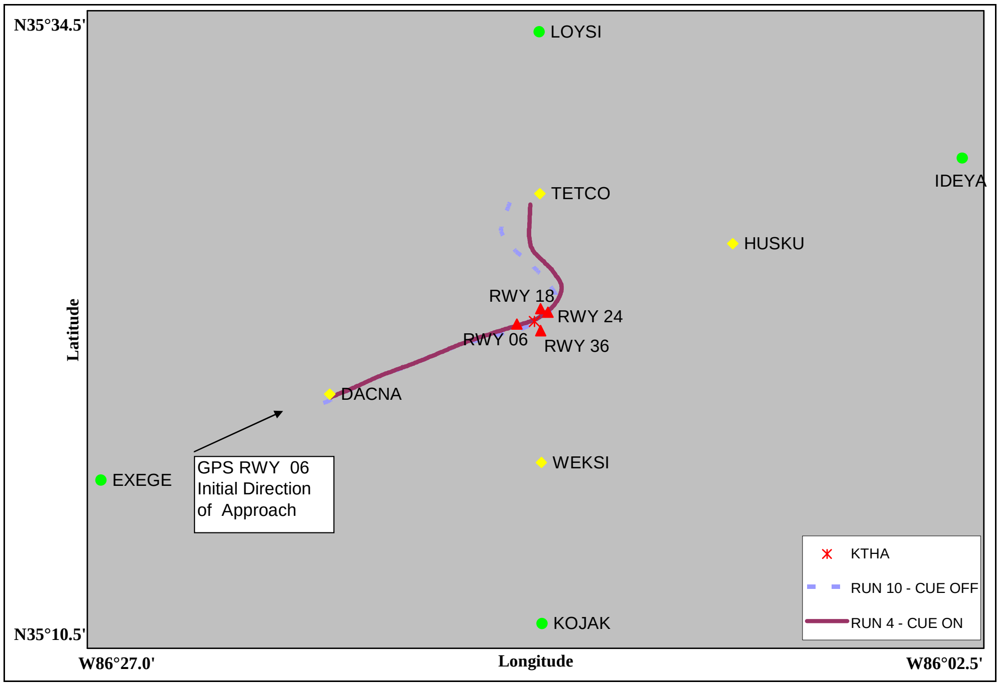


Figure 53. Aircraft Track for Evaluation Pilot 5, Run 10 & 4.

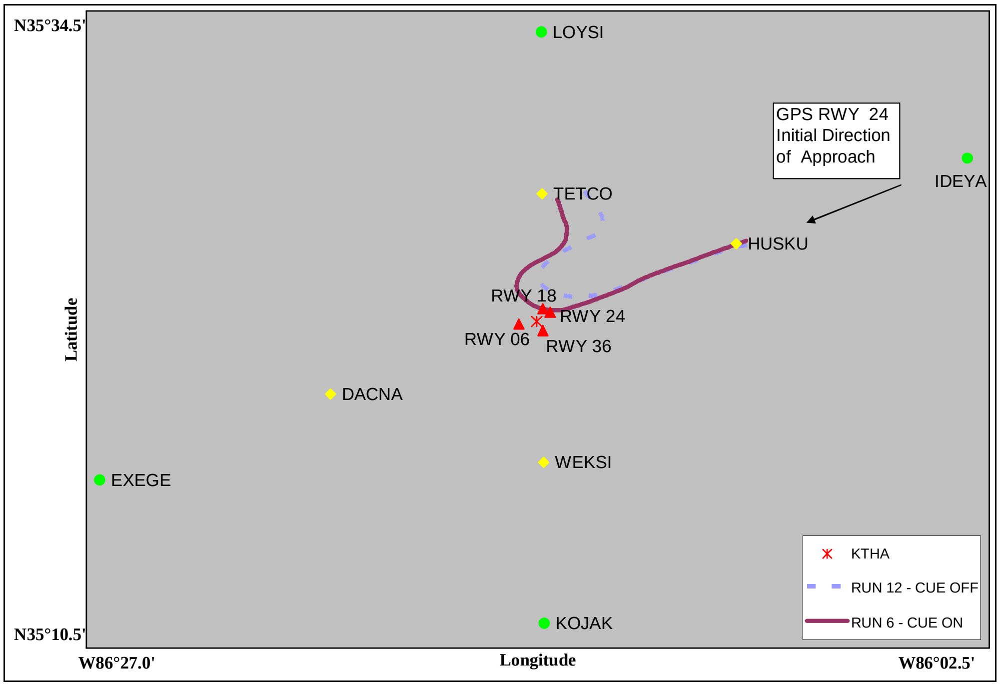


Figure 54. Aircraft Track for Evaluation Pilot 5, Run 12 & 6.

Table 1. Evaluation Pilot Demographic Data.

PILOT INFORMATION	EVAL PILOT 1	EVAL PILOT 2	EVAL PILOT 3	EVAL PILOT 4	EVAL PILOT 5
Date of Flight	April 28, 2008	April 30, 2008	May 1, 2008	May 2, 2008	April 28, 2008
Age	22	20	21	21	21
FAA Licence Held	Commercial	Commercial	Commercial	Commercial	Commercial
FAA Ratings	Multi-Engine, IFR	Multi-Engine, IFR	Multi-Engine, IFR	Multi-Engine, IFR	Multi-Engine, IFR
Total Flight Hours	940	325	359	1300	900
PIC	760	325	195	1100	900
Dual	180	N/A	164	200	N/A
Multi-Engine Hours	45	6	26	120	55
PIC	25	6	6	70	55
Dual	20	N/A	20	50	N/A
Navajo Hours	8	N/A	4.5	0	N/A
PIC	4	4	4.5	0	9
Dual	4	N/A	0	0	N/A
Instrument Hours	80	58	93.4	155	80
Actual	15	8	4.4	50	40
Simulator	20	N/A	52	55	40
Simulated	45	50	37	50	N/A

Table 2. Summary of Training and Evaluation Time for Each Evaluation Pilot.

FLIGHT INFORMATION	**EVAL PILOT 1**	**EVAL PILOT 2**	**EVAL PILOT 3**	**EVAL PILOT 4**	**EVAL PILOT 5**
Date of Training Flight	April 25, 2008	April 25, 2008	April 29, 2008	May 2, 2008	April 25, 2008
Training Time (hrs)	1.0	0.8	0.8	0.8	0.9
Date of Evaluation Flight	April 28, 2008	April 30, 2008	May 1, 2008	May 2, 2008	April 28, 2008
Evaluation Time (hrs)	2.9	2.4	2.5	1.9	2.3

Table 3. Flight Test Matrix.

Flight # / Evaluation Pilot #	Run #	3D Audio Display	Initial Approach	Initial Approach FAF	Next Approach	Next Approach Initial Fix	Direction of Turn to New Initial Fix	Engine Simulated Failed	Remarks
1 to 5	1	OFF	GPS RWY 18	TETCO	GPS RWY 06	EXEGE	Right	Left	Take off, head north
	2	ON	GPS RWY 06	DACNA	GPS RWY 36	KOJAK	Right	Left	
	3	OFF	GPS RWY 36	WESKI	GPS RWY 06	EXEGE	Left	Right	
	4	ON	GPS RWY 06	DACNA	GPS RWY 18	LOYSI	Left	Right	
	5	OFF	GPS RWY 18	TETCO	GPS RWY 24	IDEYA	Left	Right	
	6	ON	GPS RWY 24	HUSKU	GPS RWY 18	LOYSI	Right	Left	
	7	ON	GPS RWY 18	TETCO	GPS RWY 06	EXEGE	Right	Left	
	8	OFF	GPS RWY 06	DACNA	GPS RWY 36	KOJAK	Right	Left	
	9	ON	GPS RWY 36	WESKI	GPS RWY 06	EXEGE	Left	Right	
	10	OFF	GPS RWY 06	DACNA	GPS RWY 18	LOYSI	Left	Right	
	11	ON	GPS RWY 18	TETCO	GPS RWY 24	IDEYA	Left	Right	
	12	OFF	GPS RWY 24	HUSKU	GPS RWY 18	LOYSI	Right	Left	Land after data off

Table 4. Summary of NASA TLX Dimension Ratings and Workload Scores.

CONDITION		**EVAL PILOT 1**		**EVAL PILOT 2**		**EVAL PILOT 3**		**EVAL PILOT 4**		**EVAL PILOT 5**	
		CUE OFF	**CUE ON**	**CUE OFF**	**CUE ON**	**CUE OFF**	**CUE ON**	**CUE OFF**	**CUE ON**	**CUE OFF**	**CUE ON**
MENTAL DEMAND	**MEAN**	47.5	44.2	25.8	27.5	59.2	58.3	19.2	18.3	72.5	78.3
	STDEV	31.3	23.5	16.9	21.9	8.0	11.7	8.6	8.2	9.4	7.5
	MEDIAN	47.5	45.0	22.5	17.5	60.0	55.0	20.0	17.5	75.0	80.0
	RNG	70.0	55.0	40.0	50.0	20.0	30.0	20.0	20.0	25.0	20.0
PHYSICAL DEMAND	**MEAN**	29.2	27.5	28.3	32.5	72.5	70.0	25.0	24.2	80.8	79.2
	STDEV	16.9	14.1	17.2	25.0	12.5	8.9	9.5	9.7	3.8	3.8
	MEDIAN	22.5	25.0	30.0	25.0	80.0	70.0	25.0	24.2	80.2	78.8
	RNG	40.0	40.0	40.0	50.0	30.0	20.0	20.0	25.0	5.0	10.0
TEMPORAL DEMAND	**MEAN**	44.2	40.0	32.5	30.8	40.0	41.7	19.2	16.7	61.7	67.5
	STDEV	17.7	16.7	10.8	21.8	11.0	9.8	14.6	10.8	16.3	8.8
	MEDIAN	42.5	45.0	30.0	22.5	40.0	40.0	17.5	12.5	62.5	70.0
	RNG	50.0	45.0	30.0	55.0	30.0	30.0	35.0	25.0	45.0	20.0
OPERATIONAL PERFORMANCE	**MEAN**	34.2	28.3	23.3	25.8	14.2	15.8	20.0	19.2	38.3	35.0
	STDEV	25.0	21.6	8.2	17.7	4.9	5.8	12.2	11.6	15.7	10.0
	MEDIAN	32.5	22.5	25.0	17.5	12.5	15.0	15.0	15.0	32.5	30.0
	RNG	50.0	60.0	20.0	15.0	10.0	15.0	25.0	30.0	40.0	25.0
EFFORT	**MEAN**	33.3	27.5	28.3	32.5	72.5	70.0	25.0	24.2	80.8	79.2
	STDEV	24.4	14.1	17.2	25.0	12.5	8.9	9.5	9.7	3.8	3.8
	MEDIAN	22.5	25.0	30.0	25.0	80.0	70.0	25.0	22.5	80.0	80.0
	RNG	65.0	40.0	40.0	55.0	30.0	20.0	20.0	25.0	10.0	10.0

Table 4. Continued.

CONDITION		EVAL PILOT 1		EVAL PILOT 2		EVAL PILOT 3		EVAL PILOT 4		EVAL PILOT 5	
		CUE OFF	CUE ON	CUE OFF	CUE ON	CUE OFF	CUE ON	CUE OFF	CUE ON	CUE OFF	CUE ON
FRUSTRATION	**MEAN**	29.2	23.3	19.2	23.3	31.7	28.3	7.5	5.8	50.8	45.0
	STDEV	16.3	15.7	11.6	17.2	11.7	11.7	6.1	2.0	15.3	12.2
	MEDIAN	30.0	20.0	17.5	17.5	30.0	25.0	5.0	5.0	50.0	40.0
	RNG	40.0	45.0	30.0	45.0	30.0	30.0	15.0	5.0	40.0	25.0
NON-WEIGHTED NASA TLX SCORE	**MEAN**	39.4	34.4	26.9	28.9	48.8	48.1	18.2	16.7	64.0	63.9
	STDEV	20.8	16.4	13.2	19.3	6.5	5.7	10.6	9.0	7.9	6.7
	MEDIAN	38.3	31.7	26.7	19.6	48.3	47.1	16.3	14.2	63.3	62.1
	RNG	47.5	47.5	35.0	46.7	19.2	15.8	25.8	23.3	20.8	10.8
WEIGHTED NASA TLX SCORE	**MEAN**	44.1	38.0	27.9	30.4	48.3	47.2	21.3	20.0	66.7	65.7
	STDEV	23.3	17.7	13.6	19.5	6.8	5.2	11.2	10.3	7.0	7.8
	MEDIAN	43.7	32.8	28.5	20.7	49.8	47.5	19.3	17.2	67.2	64.5
	RNG	50.3	50.7	34.7	46.3	19.3	14.3	25.7	26.3	21.0	18.3

Table 5. Summary of Statistical Significance Tests.

CONDITION	WILCOXON SIGNED-RANK TEST	
	α=0.05, n=5, Tcrit=1	
	T	significance
NON-WEIGHTED NASA TLX SCORE	0	YES
WEIGHTED NASA TLX SCORE	0	YES
MENTAL DEMAND	4	NO
PHYSICAL DEMAND	-	n/a
TEMPORAL DEMAND	-	n/a
OPERATIONAL PERFORMANCE	-	n/a
EFFORT	-	n/a
FRUSTRATION	-	n/a

Table 6. General Trends in Ratings Given to the NASA TLX Dimensions

NASA TLX DIMENSION	REDUCTION IN RATING WITH 3D AUDIO DISPLAY	INCREASE IN RATING WITH 3D AUDIO DISPLAY	NEUTRAL
MENTAL DEMAND	4	1	0
PHYSICAL DEMAND	3	1	1
TEMPORAL DEMAND	2	2	1
OPERATIONAL PERFORMANCE	3	1	1
EFFORT	3	1	1
FRUSTRATION	3	0	2

Table 7. Angle of Bank Control Data for each run.

RUNS	EVAL PILOT 1			EVAL PILOT 2			EVAL PILOT 3			EVAL PILOT 4			EVAL PILOT 5		
CUE OFF	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**
1	13	5	5	9	4	7	10	4	6	15	4	4	12	3	4
3	11	2	5	11	6	7	19	4	6	12	2	4	11	4	5
5	16	3	3	12	4	5	11	7	8	1	7	15	13	4	5
8	13	3	3	17	4	5	14	4	4	12	3	4	12	4	5
10	15	2	2	13	3	3	8	16	17	14	3	4	11	3	5
12	12	4	4	11	6	7	14	3	3	11	3	5	12	3	5
MEAN	13	3	4	12	4	6	13	6	8	11	4	6	12	3	5
CUE ON	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**
7	12	4	4	7	6	10	13	3	4	12	4	5	12	4	5
9	12	4	5	11	3	5	13	3	3	16	5	5	13	2	3
11	13	3	3	14	4	4	15	3	3	14	3	4	13	6	7
2	13	5	5	13	6	6	16	4	4	14	4	4	12	4	5
4	14	2	3	13	5	5	20	6	8	16	5	5	14	3	3
6	13	4	4	11	4	6	11	5	7	12	4	4	11	5	6
MEAN	13	4	4	11	5	6	15	4	5	14	4	4	12	4	5

Table 8. Summary of Angle of Bank data for each Condition.

CONDITION	AOB		RMS	
	MEAN	**STDEV**	**MEAN**	**STDEV**
CUE OFF	12	1	5	1
CUE ON	13	1	5	1

Table 9. Indicated Airspeed Control Data for each run.

RUNS	EVAL PILOT 1			EVAL PILOT 2			EVAL PILOT 3			EVAL PILOT 4			EVAL PILOT 5		
CUE OFF	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**
1	113	2	14	110	5	12	102	4	5	103	4	6	-	-	-
3	99	6	6	103	4	6	111	5	13	96	3	4	-	-	-
5	110	4	12	101	5	6	104	7	9	100	3	3	-	-	-
8	108	3	9	111	6	14	106	6	9	103	4	6	-	-	-
10	-	-	-	106	5	8	101	5	5	102	7	7	-	-	-
12	-	-	-	102	5	6	104	4	6	102	4	5	-	-	-
MEAN	107	4	10	105	5	8	105	5	8	101	4	5	-	-	-
CUE ON	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**
7	105	3	7	110	7	13	111	8	15	104	3	6	-	-	-
9	107	4	9	105	2	6	102	4	5	99	4	4	-	-	-
11	-	-	-	101	4	4	98	5	5	100	4	4	-	-	-
2	105	3	7	104	3	6	103	5	7	101	3	4	-	-	-
4	100	4	4	105	5	7	108	5	10	110	6	13	-	-	-
6	109	4	11	106	5	9	105	4	7	102	4	5	-	-	-
MEAN	105	3	7	105	4	8	104	5	8	103	4	6	-	-	-

Table 10. Summary of Indicated Airspeed data for each Condition.

CONDITION	IAS		RMS	
	MEAN	**STDEV**	**MEAN**	**STDEV**
CUE OFF	105	3	8	2
CUE ON	104	1	7	1

Table 11. Difference in Actual and Target Intercept Heading.

RUNS	**EVAL PILOT 1**			**EVAL PILOT 2**			**EVAL PILOT 3**			**EVAL PILOT 4**			**EVAL PILOT 5**		
CUE OFF	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**
1	1	2	2	4	4	6	28	9	29	15	3	7	4	3	5
3	0	3	3	7	2	7	1	2	2	2	2	3	6	2	6
5	7	2	7	6	1	27	3	2	4	42	1	42	5	1	5
8	3	3	4	1	3	3	1	2	2	0	2	2	1	3	4
10	4	4	6	0	4	4	4	4	5	0	2	2	2	3	3
12	9	4	6	7	5	8	10	8	13	2	3	4	1	5	5
MEAN	4	3	5	4	3	9	8	4	9	10	2	10	3	3	5
CUE ON	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**	**MEAN**	**STDEV**	**RMS**
7	2	2	3	0	4	4	0	3	3	16	1	16	0	2	2
9	2	3	4	2	4	4	0	3	3	4	2	5	4	1	4
11	2	4	4	1	2	2	2	4	4	4	1	4	4	2	5
2	5	3	6	3	3	5	3	1	3	2	2	3	1	4	4
4	5	3	6	5	3	6	1	2	2	1	2	2	0	3	3
6	9	4	10	9	5	10	7	3	7	8	2	8	2	1	2
MEAN	4	3	5	3	4	5	2	3	4	6	1	6	2	2	3

Table 12. Summary of Difference in Actual and Target Intercept Heading.

CONDITION	**HDG DIFF**		**RMS**	
	MEAN	**STDEV**	**MEAN**	**STDEV**
CUE OFF	4	2	6	3
CUE ON	3	1	4	1

Table 13. Expeditious Turns toward the Revised Missed Approach Waypoint.

EVAL PILOT	NUMBER OF EXPEDITIOUS TURNS WITH 3D DISPLAY	NUMBER OF EXPEDITIOUS TURNS WITHOUT 3D DISPLAY	NO DIFFERENCE
1	3	3	0
2	3	3	0
3	3	3	0
4	2	3	1
5	1	5	0
TOTAL	12 (40%)	17 (57%)	1 (3%)

VITA

M. Hassan Naqvi is currently the Officer-in-Charge of Fixed Wing Evaluations for the Canadian Forces' Aerospace Engineering Test Establishment. He has been directly involved with numerous developmental and engineering flight test programs associated with modifications and new aircraft acquisitions. Hassan completed his undergraduate degree from the Aerospace Engineering program at Carleton University in Ottawa, Ontario, Canada in 1994 while under the Canadian Forces' Regular Officer Training Program. He graduated from the United States Naval Test Pilot School as a flight test engineer in 2002. He also has twenty years of general aviation experience and holds a commercial pilot license with multi-engine, instrument and instructor ratings.