Cueing Research by the US Army Aeromedical Research Laboratory

LECTURE SERIES CHAIR

Dr. Arthur ESTRADA
P.O. Box 620577
Fort Rucker, AL 36362
UNITED STATES
E-mail: arthur.estrada.civ@mail.mil

UNITED STATES

Dr. John CROWLEY
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577
Fort Rucker, AL 36362

Dr. Kathryn FELTMAN
Oak Ridge Institute for Science and Education
MC-100-44
P.O. Box 117
Oak Ridge, TN 37831

Dr. Thomas HARDING
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577
Fort Rucker, AL 36362

Mr. Mike HENDERSON
Science Application International Corporation
1710 SAIC Drive
McLean, VA 22102

Mr. Aaron McATEE
Oak Ridge Institute for Science and Education
MC-100-44
P.O. Box 117
Oak Ridge, TN 37831

Mr. John RAMICCIO
U.S. Army Aeromedical Research Laboratory
6901 Farrel Road
Fort Rucker, AL 36362

Dr. Bethany RANES
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577
Fort Rucker, AL 36362

Ms. Deborah RUSSELL
Wyle CAS Group
100 Quality Circle NW
Huntsville, AL 35806

Mr. J. Keegan STATZ
Oak Ridge Institute for Science and Education
MC-100-44
P.O. Box 117
Oak Ridge, TN 37831

Dr. David STILL
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577
Fort Rucker, AL 36362

Mr. Donald E. SWANBERG
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577
Fort Rucker, AL 36362

Dr. Leonard TEMME
U.S. Army Aeromedical Research Laboratory
P.O. Box 620577
Fort Rucker, AL 36362

ABSTRACT

Degraded visual environment (DVE) is the term used to describe the condition when visibility outside the aircraft is severely degraded or nonexistent. Ever since human flight began, vision has been the primary source of orienting information. Aviation pioneers and engineers have relied heavily on visual instruments to provide pilots with somatogyral. somatagravic, audio, and spatial information. This has resulted in very high visual and cognitive workloads, control reversal errors, poor perception, delayed reaction times, and ultimately, loss of situational awareness. In recent years, interest in multimodal interfaces and multisensory cueing has increased for complex, event-driven domains that are at risk for sensory overload due to their overreliance of visual displays. The U.S. Army Aeromedical Research Laboratory has conducted several multimodal studies in support of and sponsored by the U.S. Army Aviation and Missile Research Division and Engineering Center of the U.S. Army Research Development and Engineering Command. This manuscript reports on USAARL's DVE research capabilities and the general results of two experiments conducted in simulated flight which compared various multisensory cueing displays and their combinations.

1.0 INTRODUCTION[1]

Degraded visual environment (DVE) is the term used to describe the condition when visibility outside the aircraft is severely degraded or nonexistent. Of primary concern is brownout, which occurs when the visual environment is obscured by recirculated dust, dirt, or sand due to rotor downwash as a helicopter takes off, hovers, or lands. A similar phenomenon occurs in snow and is called whiteout [1]. In order to successfully operate in DVE, pilots must be able to detect and perceive drift, height above terrain, descent rate, ground speed, attitude, ground slope, terrain features, landing point location, obstacle clearance, and moving obstacles [1].

Ever since human flight began, vision has been the primary source of orienting information. Aviation pioneers and engineers have relied heavily on visual instruments to provide pilots with somatogyral. somatagravic, audio, and spatial information. This has resulted in very high visual and cognitive workloads, control reversal errors, poor perception, delayed reaction times, and ultimately, loss of situational awareness. Current instrumentation suffers from two major limitations. First, available displays do not contain sufficient information, i.e., drift, ground slope, terrain features, landing point location, obstacle clearance, and moving obstacle detection. Second, information bandwidth is insufficient to communicate the necessary information in a timely manner. During an approach to landing in a normal visual environment, the pilot will rely on outside visual references for information regarding ground speed, lateral drift, landing point location, and the landing zone environment. However, once the pilot enters DVE, he can no longer access those outside visual cues. Switching to flight instruments does not solve the problem because they do not provide these key parameters. Thus, the lack of necessary information in DVE increases the pilot's risk of crashing due to unrecognized excessive descent rates, unintended drift, and ground obstacle collisions [2].

Most information is presented visually in modern cockpits; thus, the visual channel can become overloaded while operating in high-workload conditions such as DVE [3]. Overreliance on any one sensory channel, especially during periods of high workload, can cause cognitive tunneling and sensory bottleneck [4]. Visual channels are often overburdened by cluttered visual displays and complex symbology, rendering pilots susceptible to cognitive tunneling. Cognitive tunneling is a phenomenon of focusing so intently on a display that

---

[1] Content, when not attributed to another source, is excerpted from USAARL Report 2016-10, Pilot Cueing Synergies for Degraded Visual Environments by Russell, D., Statz, J.K., Ramiccio, J., Henderson, M., Still, D., Temme, L., Ranes, B. Crowley, J., and Estrada, A.

the pilot loses focus of the environment as a whole. As more visual attention is required, the visual sense may become overloaded and critical information may be missed or misinterpreted [5]. There is also a temporal cost to cluttered visual displays: Displays can distract or slow down the pilot from obtaining necessary information. Longer search times can negatively impact performance and increase workload [6]. Sensory overload associated with congested displays and complex symbology can actually cause pilots to see and comprehend less as more information is provided.

USAARL's NUH-60 Aeromedical Research Flight Simulator provides a versatile platform from which to perform cueing research. It is a full-motion, full-visual, 6 degrees of freedom. It is equipped with seven Dell X-IG Image Generators, one of which is dedicated to simulating sensor images, through which visual databases created from worldwide satellite imagery are viewed. It possesses enhanced brownout/whiteout models to more closely simulate real-world blowing dust/snow characteristics. In addition, the NUH-60 has the capability to collect synchronized flight performance and biomedical data. The cockpit is configurable to the displays and instrumentation of A, L, V, and M model UH-60s, allowing comparison studies of existing and novel displays and technologies.

Figure 1-1: Evaluation of visual displays.

2.0 MULTISENSORY APPROACH[2]

Flying in DVE presents pilots with the potential for high workload and sensory overload. Single modality solutions can increase the already high workload and provide an incomplete picture of the outside world, resulting in a negative effect on performance. Sensory overload and increased workload can lead to missed cues, loss of situation awareness (SA), and adversely affect overall safety. To overcome the risks associated with DVE, effective and efficient use of pilot resources are required. In recent years, interest in multimodal interfaces has increased for complex, event-driven domains that are at risk for sensory overload due to their overreliance of visual displays [7]. Wickens' Multiple Resource Theory (MRT) predicts that performance can be improved by distributing information across sensory channels.

According to MRT, humans are capable of processing information from multiple sensory sources in parallel. Thus, pilots are capable of processing visual, sound, and tactile inputs simultaneously using multiple sensory resources [8]. Tasks using compatible resources that allow parallel processing may usually be performed simultaneously. Multimodal systems support time-sharing and attention management. Based on MRT, a multimodal approach that utilizes visual, audio, and tactile senses may provide pilots with the information required for safe DVE operations and prevent overreliance on the visual sense. Many bimodal research studies in which auditory and tactile cues have been introduced to provide directional and navigational guidance have

---

[2] Content, when not attributed to another source, is excerpted from USAARL Report 2016-10, Pilot Cueing Synergies for Degraded Visual Environments by Russell, D., Statz, J.K., Ramiccio, J., Henderson, M., Still, D., Temme, L., Ranes, B. Crowley, J., and Estrada, A.

supported this theory. A meta-analysis of more than 600 studies investigated the effectiveness of tactile cues versus visual cues versus visual-tactile cues found that a multisensory approach using complementary visual and tactile cues increased performance for orientation, task information, and alerts [9]. Research conducted by Sklar and Sarter [10] investigated response times for uncommanded changes of an automatic flight deck system using tactile, visual, and tactile-visual cueing. It was found that response times for tactile and tactile-visual conditions were significantly better than response times for a visual only condition, demonstrating the advantage to bimodal presentations. Research that explored the efficacy of audio-tactile systems found that well-designed audio-tactile displays have the potential to result in more resilient systems that enable the operator to receive the necessary information, even when one modality is compromised [5].

While there is great promise with a multisensory approach to solving the DVE problem, it is important to consider certain limitations of the multimodal approach. Moving from uni- or bi- to multimodal displays involves certain tradeoffs. Multimodal systems may aid in time-sharing, but there is also a potential increase in interface management and monitoring demands. Due to limited capabilities in regard to human information processing, multisensory cueing could overload the pilots' cognitive abilities, resulting in increased workload and missed cues.

An additional cautionary note to consider is The Principle of Inverse Effectiveness in Multisensory Integration which states:

As the strength of multisensory integration responses increase, the strength of responsiveness to individual sensory stimuli decreases. Consequently, multisensory cueing indices will naturally serve to improve associated performance when compared to individual stimuli. This improved degree of performance may be illusory to a certain degree, merely by the nature of the multisensory inputs [11].

3.0 USAARL STUDY 1: PILOT CUEING SYNERGIES FOR DEGRADED VISUAL ENVIRONMENTS[3]

There is a multitude of research exploring bimodal systems, but trimodal systems using visual, audio, and tactile cueing research is limited. It was the goal of this study to determine: 1) if the different symbology/cueing sets will be compatible, 2) if combining symbology/cueing sets would improve flight performance and/or reduce workload/stress, 3) if the effectiveness of different combinations of the symbology/cueing sets would be reflected in their subjective evaluations, observed flight performance, and pilot workload/stress, 4) if the effectiveness of different combinations of symbology/cueing sets would vary with the flight task. The sponsor specified that eight formally trained and rated rotary wing pilots serve as evaluation pilots for this study. These pilots were selected by the sponsor. The eight evaluation pilots performed the flight tasks and provided subjective estimations of workload and cue utility, biometric measures of workload and stress, and objective measures of their flight performance. Demographics of the evaluation pilots' flight experiences were collected with a questionnaire.

3.1 Test Equipment

All testing was conducted in the USAARL's NUH-60FS research flight simulator. Infrared (IR) scene and symbology information was shown on the primary flight display of the UH-60M instrument panel emulation. Tactile cues were presented via the Tactile Situation Awareness System (TSAS) belt, shoulder harness, and seat

---

[3] Content, when not attributed to another source, is excerpted from USAARL Report 2016-10, Pilot Cueing Synergies for Degraded Visual Environments by Russell, D., Statz, J.K., Ramiccio, J., Henderson, M., Still, D., Temme, L., Ranes, B. Crowley, J., and Estrada, A.

cushion tactors. Audio cues were presented via speakers inside HGU-56/P rotary-wing aircrew helmets. The windscreen and chin bubble exterior views were clouded with realistic dust and obscured to match the aircraft integration test configuration currently utilized by AFDD, NASA-AMES Research Center.

3.2 Cueing (Visual, Aural, Tactile)

3.2.1 Forward-looking Infrared (FLIR) and Visual Symbology Sets (Legacy Head Up Display [HUD AN/AVS-7], Brownout Symbology Set [BOSS], and Forward-Looking Integrated Systems for Helicopters [FISH])

FLIR refers to an infrared sensor mounted on the aircraft (in this context, the nose of the aircraft) that provides pilots with a view of the external world even when operating at night or in obscured environments. This capability provides pilots with cues that traditional instruments cannot, such as drift, ground slope, terrain features, obstacle clearance, and landing point location. Having an image of the external visual environment provides pilots with pertinent information necessary for safe operations in DVE; however, FLIR imagery has its limitations. The quality of the FLIR image can be distorted due by sensor restrictions. Very fine dust particles can obscure the sensors and prevent them from detecting objects. FLIR imagery relies on differences in heat to detect objects and produce images. While useful, FLIR imagery is not infallible—rain-soaked terrain on a cloudy day may present little thermal contrast. Also, depending on the climate, twice a day objects are likely to have very small heat differences making object detection difficult and FLIR imagery less useful; this phenomenon is known as thermal crossover. For the purpose of this study, a panel-mounted display with real time FLIR imagery was paired with the three visual symbology sets described below. The FLIR display presented the sensor's 60 degree field of view. The real time FLIR images were over-layed with the advanced symbologies tested with the intent to increase situation awareness during DVE.

The Legacy HUD symbology is the baseline control display system currently used in many U.S. Army helicopters. The system's cues include attitude, heading, waypoint bearing and distance, altitude, performance, and velocity vector. The system was designed for use in conjunction with a HUD system; however, it may also be used with a head-down IR scene display.

The BOSS provides "visual quality" landing capabilities in zero visibility and horizontal and vertical speed guidance to the landing zone [12]. BOSS utilizes 3D conformal symbols which allow a 3D view of the landing zone added to the 2D symbology set. Conformal symbology facilitates the mental integration of information outside the helicopter and the symbology presented on the display [13]. BOSS was developed using symbology tailored for rotorcraft during brownout conditions and presents critical flight information necessary for safe take-off, hovers, and landings in DVE. A 2009 study of BOSS showed that pilots were able to effectively use the symbology to land helicopters in brownout conditions; however, the system did not indicate lateral drift at a level considered to be suitable for safe DVE operations [14].

FISH uses pictorial pursuit guidance symbology enroute, and prior to landing switches to forward-looking landing guidance symbology. FISH symbology maintains a forward looking perspective which is consistent with how pilots naturally fly when in visual flight conditions. FISH attempts to reduce visual search times by integrating all cues in the center of the screen. Clutter is also reduced as the system switches from an Enroute page to Hover page format prior to landing. The display also possesses the ability to reroute the trajectory in-flight, which might be needed due to unforeseen events such as obstacles, traffic, or an unexpected change in the direction of approach to the landing site. In both simulator and in-flight testing of FISH, pilots were able to achieve improved tracking performance, fewer missed waypoints, and lower workload [15].

3.2.2 Aural Cueing

The initial aural cues are designed to emulate an easily fielded system that has already been approved for flight use. To this end, this study utilized SwiftTalker voice audio symbology system to monaurally alert pilots of altitude and unintentional drifts. Subsequent studies will incorporate more complex aural cues (e.g., 3D, earcons). SwiftTalker verbal alerts are provided through aircrew helmets. SwiftTalker uses Text-to-Speech (TTS) technology to create the verbal alerts. These alerts are easy to manipulate, nonreliant on human participation, low lifecycle cost, and have rapid prototype-delivery phases and iterations. In this study, SwiftTalker provided pilots with altitude alerts during approach to hover, verbally announcing “250 FEET,” “40 FEET,” “30 FEET,” “20 FEET,” and “10 FEET.” Altitude alerts were linked to the radar altimeter to ensure accurate feedback was provided to the pilots. Drift detection occurred at the 30 foot (ft.) hover condition and the alerts were via verbal cueing. If a lateral drift was detected the system announced “DRIFTING LEFT” or “DRIFTING RIGHT,” while longitudinal drifts announced “DRIFTING FORWARD” or “DRIFTING AFT.” Additional cues alerted pilots when to accept approach guidance via the voice command “Assume Guidance” or notified pilots if their heading or airspeed did not match desired parameters via the voice commands “Check Heading” and “Check Speed.” SwiftTalker cues provided pilots with altitude information that was redundant with all symbology sets.

3.2.3 Tactile

As visual and aural channels have become overwhelmed in the cockpit, there is new interest in utilizing the sense of touch with tactile cues [16]. A primary reason for utilizing tactile displays is that their use results in minimal interference with visual and aural channels. Tactile cueing was selected based on its potential ability to aid pilots operating in DVE; specifically, tactile cueing was used to provide redundant drift, course, and altitude information. The tasks pilots performed (approach to landing, approach to hover, hover, and sidestep) were selected because they test a pilot’s ability to enter and operate in DVE conditions. Additionally, it was necessary that the tactile display ensemble be comfortable, compatible with aircraft operation, and allowed for emergency egress. Finally, the tactile display system’s software and hardware was integrated into USAARL’s NUH-60FS.

The Tactile Situation Awareness System (TSAS) was found to meet all selection requirements. Specifically, it used noncontinuous tactor stimulation to preclude sensory habituation, met airworthiness requirements, and was compatible with the NUH-60 simulator software requirements. In flight tests, TSAS was shown to be capable of providing altitude, attitude, velocity, navigation, acceleration, threat location, and target location data [17]. Additionally, TSAS has shown to be effective at reducing tracking errors and improving situational awareness during landings in both degraded and good visual environments [18]. Further, Kelley, Grandizio, Estrada & Crowley [19] found that aviator performance with vibro-tactile displays was not adversely affected by adaptation or habituation following 12 continuous hours of simulated flight. The TSAS system selected for used in this experiment consisted of eight tactors along the belt that correspond to direction of drift or course deviation. Shoulder and seat tactors reported altitude deviations. On course, forward flight was cued via the center tactor.

3.3 Summarized Results

Detailed results can be found in USAARL Report 2016-10, including physiological . Flight performance data (i.e., flight path, speed, heading, altitude, position) were evaluated for pilot performance for Approach to Landing, Approach to Hover, Hover, and Sidestep Maneuvers. Subjective assessments included results by maneuver for Cooper-Harper, Bedford Workload, and Visual Cue Index ratings.

3.3.1 Approach to Landing

To determine the optimized cueing display configurations used to facilitate helicopter approach to landing in DVE, all combinations of cueing were recorded and analyzed. Between visual symbologies, BOSS resulted in significantly better performance than the Legacy HUD and FISH in position and speed, and better than Legacy HUD in heading. FISH provided significantly better altitude control over BOSS and Legacy HUD. The data indicate that pilot position accuracy improved when BOSS was paired with either the TSAS or Aural cueing displays than when using BOSS paired with only a FLIR scene. Subjective assessment overall ratings by the test pilots indicated pilot preference for BOSS over Legacy HUD symbology. Bedford workload data indicate that pilots perceived workload to be significantly lower using BOSS than using either Legacy HUD or FISH. Additionally, a cueing effect was found for workload. Pilots reported that workload was significantly lower when symbology was paired with the TSAS and Aural cueing display than when paired with the Aural cueing display only.

3.3.2 Approach to Hover

For the approach to a hover maneuver, BOSS symbology resulted in significantly better performance in maintaining position than the Legacy HUD and FISH. BOSS also resulted in better altitude control over Legacy HUD. FISH provided significantly better results over Legacy HUD for position, altitude and speed. Subjective ratings showed the test pilots significantly preferred BOSS symbology over the Legacy HUD symbology. The best overall approach to hover performance was attained by the BOSS symbology with supplemental TSAS and Aural cueing and was rated as preferred by the test pilots.

3.3.3 Hover

In the Hover maneuver, significant differences were found between the visual symbology sets for position, heading, and altitude. BOSS resulted in significantly better performance over Legacy HUD and FISH in position and altitude, although FISH was significantly better than Legacy HUD in heading maintenance. Likewise, the test pilots subjectively preferred the BOSS over the Legacy HUD. Additionally, the BOSS symbology was rated better than the FISH symbology on three of the five subjective questionnaires. For the Hover maneuver, the BOSS symbology combined with supplemental Aural cueing or TSAS cueing produced the best performance and was preferred over other combinations.

3.3.4 Sidestep

Although the Sidestep maneuver began and ended with twenty second hovers only the sidestep segment was analyzed. The data indicate performance with both BOSS and FISH was significantly better than with Legacy HUD. Subjectively, the TSAS cueing display was ranked as the easiest to fly during sidestep, followed by the TSAS and Aural cueing combination. The Aural cueing display was ranked the most difficult to fly. The results indicate that for the Sidestep maneuver, the BOSS symbology paired with the TSAS cueing display produced the best performance and was preferred over other displays.

3.4 General Conclusions

1. Pilots performed better using advanced visual symbologies (BOSS and/or FISH) when combined with a supplemental form of cueing (aural and/or tactile).

2. Advanced visual symbologies outperformed Legacy symbology for almost all maneuvers.

3. Test pilots' preferred supplemental cueing modality was dependent on the type of visual symbology and/or flight maneuver.
4. As configured in this study, aural cueing degraded flight performance in some test pilots when using either Legacy or FISH visual symbology sets due to pilot-induced oscillation during the hover and sidestep maneuvers.
5. Overall, subjective and flight performance measures indicated that the BOSS symbology was the preferred visual symbology set.
6. Pilots preferred aural cues that provided situational information over aural cues that demanded corrective action to satisfy a required performance measure.
7. In general, test pilots preferred the TSAS cueing display over the aural cueing display.

4.0 USAARL STUDY 2: INTEGRATED CUEING ENVIRONMENT TESTING: PILOT CUEING SYNERGIES FOR DEGRADED VISUAL ENVIRONMENTS[4]

DVE has driven the development of new display technology which in turn presents new challenges, including the integration of scene imagery, visual symbology, tactile cues, and aural cueing. Visual symbology must also be studied to determine the best modality to be presented, panel mounted display (PMD) or helmet mounted display (HMD).

This effort was another important step in the U.S. Army Aviation and Missile Research Division and Engineering Center's (AMRDEC) development of an integrated cueing environment for future aviation applications. As such, experienced test pilots were used in AMRDEC's effort to evaluate the optimization of the integration and to establish display requirements before the testing of resulting integrations are tested on less experienced pilots whose inexperience could bias for or against novel technologies.

The primary test objective was to evaluate the Integrated Cueing Environment's (ICE) visual symbology overlaid over imagery from a forward-looking infrared radar (FLIR) sensor. During simulated night flight, the imagery was displayed on a UH-60M PMD or on a SA Photonics high definition (HD), wide field-of-view (FOV), binocular HMD. During simulated day flight, composite imagery was displayed on both the PMD and HMD. Additionally, the synergistic effects of aural and tactile cues were assessed. All conditions were tested with and without a distraction task.

This was a complex test plan with specific sequences that were accomplished in order for the test to be successful. These sequences included the following plans and procedures:

1. Integrate ICE's visual symbology, the selected PMD or HMD, the cueing sets, and the FLIR and head tracked imagery into the United States Army Aeromedical Research Laboratory (USAARL) immersive, full-motion, enhanced brownout Black Hawk simulator.
2. Implement the selected operational flight tasks for evaluation.

---

[4] Content, when not attributed to another source, is excerpted from USAARL Report 2017-04, Integrated Cueing Environment Testing: Pilot Cueing Synergies for Degraded Visual Environments by McAtee, A., Russell, D., Feltman, K., Swanberg, D.E., Statz, J.K., Ramiccio, J., and Harding, T.H.

3. Familiarize the evaluation pilots with the selected flight tasks, FLIR sensor display, and ICE cueing package.
4. Conduct test flights of the selected tasks under different cueing configurations.
5. Evaluate the utility of the selected PMD and HMD and cueing sets using the evaluation pilots' subjective ratings, flight performance, and psychophysiological measures.

The compatibility and effectiveness of each combination of FLIR sensor imagery, selected PMD and HMD, and cueing set were measured with quantitative measures of flight performance, pilot psychophysiological measures, and the pilot's subjective reports. The Modified Multi-Attribute Task Battery [20] was integrated on a kneeboard tablet and acted as a distraction task to increase pilot workload to a comparable operational level.

ICE symbology test configurations were evaluated three ways: 1) flight performance metrics which track deviations from an ideal flight path, 2) workload metrics, and 3) pilot subjective assessments. Simulator data documented the symbology sets' effect on flight performance. Psychophysiological data were collected as measures of the configurations' effect on workload and stress. Cooper-Harper Handling Qualities Ratings Scale, National Aeronautics and Space Administration Task load Index (NASA-TLX) workload assessment, and Situational Awareness Rating Technique (SART) data, along with free reports documented the evaluation pilots' assessments for each of the configurations.

Three results were obtained from these tests. 1) The relative efficacy of the ICE cueing package when teamed with the selected PMD and/or HMD. 2) The effect of each condition on flight performance, workload, and situational awareness. 3) Recommendations for managing the integration of the ICE cueing package technologies into helicopter operations.

4.1 Methods

All testing was conducted in the USAARL NUH-60FS research flight simulator.

4.1.1 Imagery/Cue Display

FLIR scene and symbology information were shown on the PMD or the HMD when testing each system separately. When the displays were tested simultaneously, ICE symbology was present on both PMD and HMD, but the FLIR imagery was only displayed on the PMD.

Tactile cues were presented via belt, shoulder harness, and seat cushion tactors. Audio cues were presented via HGU-56/P rotary wing aircrew helmets. The windscreen and chin bubble exterior views were clouded with realistic dust during takeoff, hover, and landing for all conditions. Night flights were conducted using a starless night environment. Day flights were conducted in day visual meteorological conditions (VMC) environments.

4.1.2 Evaluation Pilots

Seven experienced test pilots performed the flight tasks and provided subjective estimations of workload and situational awareness. The evaluation pilots were formally trained and rated rotary-wing Experimental Test Pilots medically fit to fly. These pilots were selected by the sponsor and USAARL, and they had sufficient flight experience to enable them to provide expert guidance in the establishment of display requirements. Accordingly, this test plan was specifically designed to provide a structured environment to capture these experts' flight performance and subjective assessment of display, symbology, and cueing characteristics. Objective measures included psychophysiological measures of workload and stress, and objective measures of flight performance.

4.1.3 Flight Maneuvers and Metrics

The ICE cueing package and selected displays were evaluated while performing operational flight tasks; specifically, enroute flight, approach to hover, hover, landing, and takeoff. The flight tasks were flown in the order presented for each condition. Each maneuver represents a logical ordering of events of a simulated operational mission. Flight performance was measured by speed, heading/flight path, and altitude deviations from ICE guidance.

4.1.4 Enroute

This task initiated with the aircraft traveling on an established nap of the earth flight path moving at 80 knots toward the approach point. Pilots received ICE enroute guidance while attempting to maintain 80 knots. Metrics for this task included deviations from an ideal flight path (altitude, speed, and position).

4.1.5 Approach to Hover

This task started with the aircraft at 250 feet above ground level (AGL) and moving at 80 knots toward the hover point 0.8 nm away. Descent from 250 feet AGL began at 0.8 nm from the hover point. The pilot attempted to approach the hover point in a straight line while following approach guidance until hover guidance was activated. Metrics for this task included deviations from an ideal approach path and heading, as well as deviations from ideal collective and cyclic inputs.

4.1.6 Hover

This task started when the symbology cycles from approach to hover. The pilot attempted to navigate to the hover point and maintain a 30 foot hover for 1 minute. Metrics included drift, altitude, and heading deviations.

4.1.7 Landing

The pilot attempted to lower the aircraft into the confined space of the compound and touchdown with minimal drift. Metrics for this task included maximum velocity, heading maintenance, and position maintenance when the aircraft touched down and activated the weight on wheels switch, and aircraft was at zero speed.

4.1.8 Takeoff

This task started with the aircraft parked in the compound. The pilot attempted to lift off the ground while maintaining heading with minimal lateral drift and ascend to an altitude of 30 feet AGL. Once the aircraft reached 30 feet and cleared obstacles, the pilot attempted to follow ICE guidance and accelerate to 80 knots. Metrics for this task included altitude, heading maintenance, and position maintenance.

4.2 Psychophysiological Metrics

The instrumentations' effects on workload were measured using four metrics: heart rate (HR), heart rate variability (HRV), respiratory rate (RR), and electroencephalogram (EEG). These metrics were selected for their responsiveness to workload, minimal invasiveness, compatibility with the testing paradigm, tolerance of the simulator environment, and potential utilization in aircraft. The heart rate and respiratory rate data were synchronously collected with BIOPAC's BioNomadix® and BHAPI® instrumentation/software. The EEG data were collected separately using the B-Alert X-24 EEG and was synchronized with the HR/HRV and RR measures post-data collection. This instrumentation was also selected for its tolerance of the simulator testing environment and potential utilization in aircraft. The biometric data were analyzed with BIOPAC's AcqKnowledge® and IBM's SPSS software. (See USAARL Report 2017-04 for the results of psychophysiological measures.)

4.3 Qualitative Subjective Metrics

Following completion of each scored test run, pilot subjective impressions were captured with: the Cooper-Harper grading criteria, NASA TLX workload scale, and SART. Questionnaire responses and free reports were also collected from each pilot.

4.4 Test Conditions

Test flights were flown with a single unassisted (minimal crew coordination) evaluation pilot at the controls with wind and turbulence turned off. The out-the-window views, including the chin bubbles, were obscured with blowing sand and dust below 100 feet AGL for all flights. The FLIR scene imagery within the display was unobscured.

4.5 Test Order

Training and testing required two days for each of the evaluation pilots. The training phase consisted of a safety and risk briefing and 4 simulator flight hours of socialization/training. The testing phase consisted of three test runs for each of the Night DVE conditions and two test runs for each of the Day DVE conditions.

4.6 Analysis

The outcome measures from the flight test results included a quantitative assessment of flight performance. These quantitative results of the display and cueing configurations were tabulated and compared. Similarly, the pilots' psychophysiological results and qualitative reports were also summarized, and agreement/disagreement of the findings tabulated. Qualitative, psychophysiological, and quantitative results were compared to check for agreement. Descriptive statistics were generated for the dependent measures.

4.7 Instrumentation

4.7.1 Evaluation Pilots

The test instrumentation worn/sat upon by the evaluation pilots included: 1) HGU-56/P rotary-wing aircrew helmets with HMD, 2) Tactile Situational Awareness System (TSAS) tactor belt, shoulder harness, and seat cushion, and 3) psychophysiological instrumentation.

4.7.2 ICE, Tactile, and Aural Instrumentation

Tactile cues were presented via TSAS belt, shoulder harness, and seat cushion electromechanical tactile stimulators (tactor) instrumentation and supporting software algorithms. Aural cues were presented via HGU-56/P rotary wing aircrew helmets.

4.7.2.1 *Tactile Cues*

The study utilized belt, shoulder harness, and seat cushion tactors operated under TSAS algorithms for speed, drift, and altitude control. Tactile cues provided feedback predicated on the maneuver being flown. Using data from existing aircraft sensors, TSAS provided touch cues via an array of tactors on the skin. The system incorporated belt, shoulder harness, and seat cushion tactors to provide drift and altitude cueing.

4.7.2.2 *Aural Cues*

ICE symbology provided aural alerts for altitude, heading, and speed/drift via HGU-56/P rotary wing aircrew helmets. Aural cues provided feedback predicated on phase of flight.

4.7.2.3 *ICE Visual Display*

Two visual systems were evaluated: SA Photonics Low Cost Augmented Reality system (LARS) HMD and the UH-60M PMD. The displays were tested in three configurations; PMD, HMD, and both PMD and HMD. ICE visual symbology and FLIR sensor imagery were always present on the PMD configuration and the HMD configuration. When the ICE system was tested on the PMD and HMD displays together, FLIR imagery was only present on the PMD. FLIR scene imagery with overlaid visual symbology was displayed on an emulated UH-60M instrument panel installed in the USAARL NUH-60FS. This emulated panel was manufactured by SGB Enterprises. The emulated panel's extended graphics array (XGA) display used to present the FLIR scene imagery measures 8.06 x 6.11 inches and is positioned 28 to 30 inches from the aircraft's design eye point. Thus, the display screen visible area subtends 15.8 x 12.0 degrees. The screen has a resolution of 1024 x 768 pixels (with 997 x 756 viewable) and a maximum viewing angle of 85°. The display is capable of 700:1 contrast and brightness greater than 50 FL.

Figure 4-1: UH-60M instrument panel emulation.

The red circled screens were used to display the FLIR scene imagery and visual symbology.

The HMD tested was the SA Photonics LARS HMD provided by Army Research Lab (ARL). The LARS HMD was selected because it provides high resolution (1920 x 1200 pixels), wide FOV (76° Horizontal, 33° Vertical) imagery via see-through binocular optics with almost no peripheral obscurations. The LARS HMD was also selected based on the ease to integrate ICE visual symbology.

Figure 4-2: SA Photonics LARS HMD.

ICE utilized a 3D Conformal landing zone to provide a 3D perspective view of the landing point. This symbology system was a tailored set of rotorcraft symbology with guidance to allow for safe landings in brownout. It was developed to present critical flight information to enable safe landing, hover, and take-off while in zero visibility conditions. In order for the aircraft to land safely at the landing point, the pilot was required to concurrently manage three profiles: vertical (altitude) profile, lateral (cross-track) profile, and longitudinal (speed) profile.

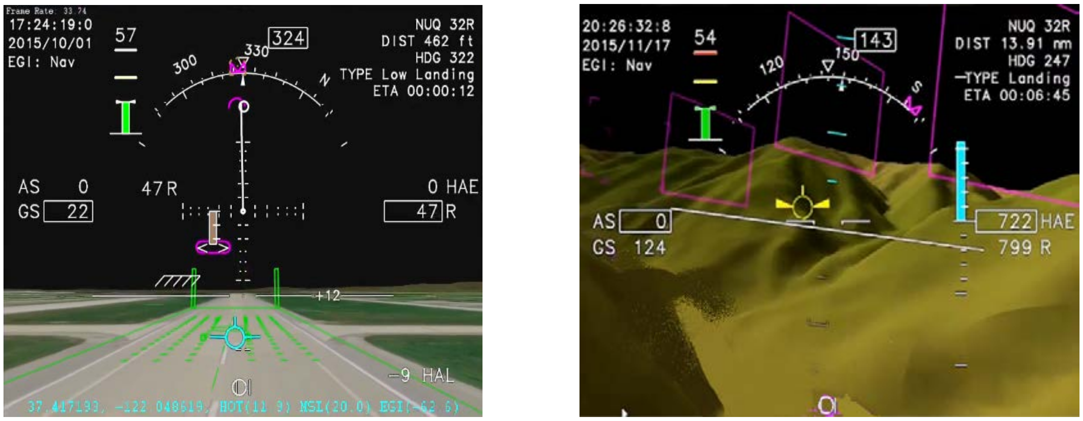

Figure 4-3: ICE symbology Enroute and Hover/Approach/Take-off pages.

The design of the symbology started with the AH-64A. The plan-view velocity vector, acceleration cue (ball), and target hover position (doghouse) symbols are from the AH-64 design. Other aircraft that use these symbols include the UH-60M, CH-47F, and OH-58D. What was missing in the AH-64 design was guidance to the landing point. ICE explicitly showed what the current horizontal speed should have been at every instant during the approach to hover. If the pilot followed the horizontal speed guidance, the aircraft would come to near zero speed at the intended hover/landing point.

Altitude information was improved upon with ICE. Rather than an AH-64 type altitude tape, which grew smaller and further from the center of the screen as the aircraft neared the ground, the ICE symbol used a rising ground symbol which moved closer to the center of the screen as the aircraft neared the ground. ICE provided vertical speed guidance missing on the AH-64 display. The ICE symbology explicitly showed the pilot what the vertical speed should have been at all times, to touch the ground, or to stabilize at a hover at the intended hover/landing point. It was up to the pilot to close the control loop and move toward the required state as determined by the guidance.

ICE included a perspective view landing pad, which most pilots called three-dimensional. This was not a binocular-viewed landing pad (true 3D), but rather a perspective-viewed image of a 3D model. The 3D conformal addition included head tracked 3D symbology which added the other pilot's line of sight, 3D Wingmen, and 3D landing grid.

Figure 4-4: ICE Display with 3D symbology.

4.8 Modified MATB

The Modified MATB is a Windows© based computer program designed to evaluate operator performance and workload. The Modified MATB tasks presented to the pilot included only system monitoring. The monitoring task required pilots to monitor simulated system instruments. If a system instrument light came on, the pilot had to respond within 10 seconds by touching the screen of the tablet. The Modified MATB was programmed to a high workload setting.

4.9 Summarized Results

Detailed results can be found in USAARL Report 2017-04.

The ICE configurations were evaluated using flight performance metrics that tracked deviations from an ideal flight path and pilot subjective assessments that measured the adequacy of the ICE systems for completing the assigned tasks.

For each phase of flight, multivariate linear regression models were developed to determine if the observed differences in flight performance metrics were statistically significant. The flight performance metrics used to determine deviations from an ideal flight path were dependent on the flight phase and, in many cases, a transformation of the metrics was required prior to model selection.

The Cooper-Harper HQR scale was the primary tool used to subjectively assess the adequacy of the ICE systems during individual phases of flight. The Cooper-Harper HQR scale is a decision tree that uses adequacy for selected task, aircraft characteristics, and demands on the pilot to reach a pilot rating that ranges from 1 to 10. For this study, the HQR scale was considered to be an ordinal scale; therefore, multinomial logistic regression models were developed for each phase of flight to determine if either display type or cueing had a statistically significant effect on the scores.

Overall subjective assessments of the ICE system, across all phases of flight, were conducted using the NASA TLX, SART, and a pilot questionnaire.

During the enroute phase, the type of display utilized had a statistically significant effect on flight performance of night DVE flights. The pilots were better able to maintain an ideal flight path with the PMD than with the HMD. However, pilot subjective comments revealed no difference in handling quality ratings for the two displays. During all other phases (approach to hover, hover, landing, and takeoff), there were no observed differences in flight performance or handling quality ratings.

The NASA TLX is a multi-dimensional rating technique designed to measure workload in terms of the following six subscales: mental demand, physical demand, temporal demand, performance, effort, and frustration level. After each test run the pilots provided a rating, on a scale from 1 to 10, for each of the subscales. An overall test score was calculated for each run by giving each subscale equal weighting. During night DVE flights, there was no difference in NASA TLX overall scores when using HMD vs using PMD (p-value > 0.99) and there was no difference in NASA TLX overall scores for Cues on vs Cues off (p-value = 0.67). During day DVE flights, there was no difference in NASA TLX overall scores for Cues on vs Cues off (p-value = 0.34).

The SART is a multi-dimensional technique designed to measure the pilot’s situational awareness in terms of the following three subscales: demands on attentional resources, supply of attentional resources, and understanding of the situation. After each test run the pilots provided a rating, on a scale from 1 to 10, for each of the subscales. An overall SART (SA) score was calculated for each run using the following formula: $SA = U - (D - S)$. During night DVE flights, there was no difference in overall SART scores when using HMD vs using PMD (p-value = 0.58) and there was no difference in overall SART scores for Cues on vs Cues off (p-values > 0.00). During day DVE flights, there was no difference in overall SART scores for Cues on vs Cues off (p-value = 0.57).

After each run, pilots were asked to rate the overall effectiveness of the following ICE subsystems: PMD symbology, PMD imagery, HMD symbology, HMD imagery, aural cueing, and tactile cueing. A five category rating system, ranging from 1: very effective to 5: very ineffective, was used for each subsystem. During night DVE flights, pilots considered symbology very effective on the HMD and PMD. The imagery, aural cueing, and tactile cueing were all rated as effective. During day DVE flights, when symbology was provided on both the HMD and PMD concurrently, the PMD symbology was given a better rating. The difference in ratings of PMD and HMD symbology during day DVE flights was statistically significant (p-value < 0.0001).

4.10 General Conclusions

Recall that this effort was an important step in AMRDEC's development of an integrated cueing environment for future aviation applications. As such, experienced test pilots were used in to evaluate the optimization of the integration and to establish display requirements. The primary objective of this study was to evaluate the ICE's visual symbology overlaid over imagery from a FLIR sensor. The compatibility and effectiveness of FLIR sensor imagery, displays (PMD and/or HMD), and cueing set were assessed using quantitative measures of flight performance, pilot psychophysiological measures, and pilot subjective reports. These results document advantages and disadvantages of the panel and helmet mounted displays, the aural and tactile cues, and the interaction effects of various combinations of displays and cueing systems. The results are based on data collected from seven highly experienced test pilots that performed flight maneuvers of relative short duration. The results are not likely representative of a more general population of Army Aviators.

4.10.1 Displays

The type of display utilized had a statistically significant, but not operationally significant, effect on flight performance during the enroute phase of night DVE flights. The pilots were minimally better able to maintain an ideal flight path with the PMD than with the HMD. Pilot subjective comments revealed no difference in handling quality ratings for the two displays. There was no difference in flight performance or handling quality ratings during any other phase of flight when using HMD versus PMD. There was no difference in the NASA TLX Overall scores or the SART scores when using HMD versus PMD.

During night DVE flights, pilots considered symbology very effective on both the HMD and the PMD. During day DVE flights, when symbology was provided on the HMD and PMD concurrently, pilots gave the PMD symbology a very effective rating and the HMD an effective rating. The difference in ratings was statistically significant.

4.10.2 Cues

There was no difference in flight performance, handling quality ratings, the NASA TLX Overall scores or the SART scores during any phase of flight with Cues on versus Cues off. During night DVE flights, pilots considered imagery, aural cueing, and tactile cueing as effective.

4.11 Future ICE Research

Plans are underway to conduct the next phase of testing in the next few months in which line pilots who have not been previously exposed to the ICE cueing system will be used as test subjects. This will allow for the assessment of whether the ICE cueing system is easily learned and able to assist in maintaining performance in DVE for those who have had less or no experience with the systems. The overall testing objectives will be the same as the previous study, such that the symbology will be assessed on both a PMD and HMD, and the synergistic effects of aural and tactile cues will be examined.

5.0 REFERENCES

[1] NATO Task Group HFM-162. (2012). Rotary-wing brownout mitigation technology and training. RTO/NATO.

[2] Colucci, F. (2007, Spring). Digging out from brownout. Vertiflite, pp. 50-55.

[3] Veltman, J. A., Oving, A. B., & Bronkhorst, A. W. (2009). 3D audio in the fighter cockpit improves task performance. The International Journal of Aviation Pyschology, 14(3), 239-256. Retrieved from http://dx.doi.org/10.1207/s15327108ijap1403_2.

[4] Allan, K., White, T., Jones, L., Merlo, J., Haas, E., Zets, G., & Rupert, A. (2010). Getting the buzz: Whats next for tacile information delivery? Preceedings of the Human Factors and Ergonomics Society Annual Meeting 2010, 1331-1334.

[5] Mateo, J., Simpson, B., Gilkey, B., Iyer, N., & Brungart, D. (2012). Spatial multisensory cueing to support visual target-acquisition performance. Proceedings of the Human Factors and Ergonomics Society 56th Annual Meeting, 1312-1316.

[6] Curtis, M. T., Jentsch, F., & Wise, J. (2010). Aviation Displays. In E. Salas, & D. Maurino, Human Factors in Aviation (2nd ed., p. 453). Burlington: Elsevier.

[7] Sarter, N. (2007). Multiple resource theory as a bases for mutlimodal interface design: Success stories, qualifications, and research needs. In A. F. Kramer, D. Weigmann, & A. Kirlik, Attention: From Theory to Practice. New York: Oxford Publishing.

[8] Vidulick, M., Wickens, C., Tsang, P., & Flach, J. (2010). Information Processing in Aviation. In E. Salas, & D. Maurina, Human Factors in Aviation (pp. 175-207). Burlington: Elsevier.

[9] Elliott, L., Coovert, M., Prewett, M., Walvord, A., Saboe, K., & Johnson, R. (2009). A review and meta analysis of vibrotactile and visual information displays. Aberdeen Proving Ground: Army Research Laboratory.

[10] Sklar, N. and Sarter, N. (1999). Good vibrations: Tactile feedback in support of attential allocations and human-automation coordinationin event-driven domains. Human Factors: The Journal of the Human Factors and Ergonomics Society, 543-552.

[11] Holmes, N. (2009). The principle of inverse effectiveness in multisensory integration: some statistical considerations. Brain Topograpy, 21, 168-176. doi:10.1007/s10548-009-0097-2.

[12] Thompson, B. (2011). Symbology Collaboration Helps Pilots "See" During Brownouts. Retrieved from Wright-Patterson Air Force Base: http://www.wpafb.af.mil/news/story.asp?id=123271120.

[13] Wickens, C. (2003). Aviation Displays. In P. Tsang, and M. Vudilich, Principles and Practice of Aviation Psychology (pp. 147-200). Mahwah, Ney Jersey: Lawrence Erlbaum Associates.

[14] Szoboszlay, Z. P., Turpin, T. S., and McKinley, R. A. (2009). Symbology for Brown-Out Landings: The First Simulation for the 3D-LZ Program. In proceeding of: American Helicopter Society 65th Annual Forum.

[15] Moralez, E., Shively, R. J., Grumwald, A. J., and Hovev, M. (2011). Forward-looking integrated symbology for 4-D reroutable helicopter approach-to-landing. American Helicopter Society 66th Annual Forum.

[16] Lu, S., Wickens, C. D., Sarter, N. B., and Sebok, A. (2011). Informing the design of multimodal displays: A meta-analysis of Empirical studies comparing auditory and tactile interruptions. Proceedings of the Human Factors and Ergonomics Society 55th Annual Meeting, 1170-1174.

[17] McGrath, B., Estrada, A., Braithwaite, M., Raj, A., and Rupert, A. (2004). Tactile situational awareness system flight demonstration final report. f: USAARL.

[18] Craig, G., Jennings, S., Cheung, B., Rupert, A., and Schultz, K. (2008). Flight-test of a tactile situational awareness system in a land-based deck landing task. Proceedings of the Human Factors and Ergonomics Society, 142-146.

[19] Kelley, A. M., Grandizio, C. M., Estrada, A., & Crowley, J. S. (2014). Tactile Cues in Continuous Operations: A Preliminary Study. Aviation, Space, and Environment Medicine, 85(2), 172-176.

[20] Cline, J., Arendt, D., Geiselman, E., & Blaha, L. M. (2015). Web-based implementation of the Modified Multi-Attribute Task Battery. Poster presented at the Fourth Annual Midwestern Cognitive Science Conference, Dayton, OH.