Pupillometric Workload Measurement in the 360 Degree Integrated Cueing Environment (ICE)

Amanda Hayes
Research Associate
Goldbelt Frontier, Inc.
Fort Rucker, AL, USA

Kathryn Feltman
Research Psychologist
US Army Aeromedical Research Laboratory
Fort Rucker, AL, USA

Christopher J. Aura
Sensory Scientist
Goldbelt Frontier, Inc.
Fort Rucker, AL, USA

ABSTRACT

This study is a part of an ongoing series of studies examining cueing modalities to circumvent the in-flight effects of degraded visual environments (DVEs) in a rotary wing aircraft. The suite of cueing modalities investigated include visual cueing symbology, auditory cueing, and tactile cueing. This study compared the use of combinations of these cueing modalities to find which resulted in the best performance and the least amount of workload required of the pilot. This specific paper focuses on the analysis of pupillometric data collected through video-based eye-tracking to measure cognitive workload. Results are discussed.

INTRODUCTION

Army aviation is meeting the challenges of multi-domain operations to maintain its dominance in the air with swift modernization. In 2018, the U.S. Army released a request for the next generation of the Future Vertical Lift (FVL) Future Attack Reconnaissance Aircraft (FARA). Just one year later, five prototype agreements were awarded with the expectation that two designs will compete in a "fly-off" in 2023. The Army hopes to field the winning design by 2028. The request clearly states the challenges ahead:

> *Army Aviation must operate in highly contested/complex airspace and degraded environments. The Army currently lacks the ability to conduct armed reconnaissance, light attack, and security with improved stand-off and lethal and non-lethal capabilities with a platform sized to hide in radar clutter and for the urban canyons of megacities.*

Operations in megacities need the same aviation capabilities as in today's war-space (attack, reconnaissance, assault, and medical evacuation). Megacities are simply harder to maneuver in; their dense developmental patterns create obstacles to tactical flying. Powerlines, antennas, satellite-dishes, and narrow spaces between buildings limit landing and pickup zones and make it difficult to fly close enough to ground troops to provide air support. This project tests the latest iteration of the integrated cueing environment (ICE) as part of a larger, ongoing research effort to identify ideal in-cockpit cueing (or combination of cueing) modalities that improve flight capability and safety in degraded visual environments (DVEs).

Ongoing efforts within the DVE Mitigation program (DVE-M) have resulted in a series of studies examining the use of tactile, auditory, and visual cues for flight in DVE (Ref. 1; Ref. 2; Ref. 3). The results of these studies have led to further development and refinement of a multimodal ICE package. The current iteration of this overarching ICE research effort added a tactile cueing component to updated visual symbology and cuing, and spatially-coherent 360 degree audio cues. The flight profiles were designed to challenge the pilots and force reliance on the ICE, and thus, utilized nap-of-the-earth (NOE) flight, in blackout conditions, with frequent obstacles in the flight path, that were navigated using only internal displays and external sensors for navigation.

This study collected subjective workload rating scales, subjective trust in automation, and system usability questionnaires (developed by the Combat Capabilities Development Command Data and Analysis Center; Aberdeen Proving Ground, MD), flight performance data, and biometric response data. This paper focuses on data collected through video-based eye tracking and pupil response dynamics to calculate cognitive workload using the index of cognitive activity (ICA) (Eye Tracking, Inc., Solana Beach, CA) throughout each flight.

STUDY DESIGN

Participants

Eight male UH-60 Black Hawk (UH-60) pilots evaluated the effectiveness of the ICE. All participants were Active Duty Army pilots whose military experience ranged from 8 years

Presented at the Vertical Flight Society's 76th Annual Forum & Technology Display, Virginia Beach, VA, USA, October 6-8, 2020.

to 21 years (M = 13.88 years, SD = 4.61 years). Flight hours in the UH-60 ranged from 470 hours to 2,060 hours (M = 1,390 hours, SD = 476.06 hours). Total flight hours among the pilots ranged from 540 hours to 2,060 hours (M = 1,478.75 hours, SD = 440.37 hours).

Flight Simulator

All pilots flew 10 mission-scenarios designed to stress the pilots' abilities and effectively assess the experience and performance impact of different combinations of cueing modalities for obstacle avoidance. All flights were conducted in the NUH-60FS Black Hawk simulator (NUH-60FS) at the United States Army Aeromedical Research Laboratory (USAARL) located in Fort Rucker, Alabama (Figure 1). The NUH-60FS is fully accredited by the Directorate of Simulations (DoS) and by the Program Executive Office of Simulations, Training, and Instrumentation (PEOSTRI), as a 6-degree-of-freedom (DOF), full-motion, and full-visual (Level D equivalent) NUH-60FS Black Hawk helicopter flight simulator. It uses X-IG (CATi Training Systems, Ozark, AL), an OpenGL-based visual image generator, which can simulate naturalistic flight conditions, and DVEs. The simulator also captures flight performance and simulator-state characteristics at a rate of 60 Hz.

Figure 1. USAARL UH-60 Simulator.

Visual Displays

The UH-60M Panel Mounted Display (PMD) was used to present the ICE visual avionics, and in certain configurations the visual collision avoidance symbology (ICE-CAS), overlaid onto low-latency, forward-looking-infra-red (FLIR) sensor imagery. The PMD display dimensions were 32-by-32 cm, and at a typical viewing distance of 50-70 cm, they occupied between 16 and 22 visual degrees horizontally, and between 12 and 17 visual degrees vertically. The two panels were mounted side-by-side (as shown in Figure 2). The text indicating the ground-speed, heading, and altitude were roughly 1 visual degree in size and were separated on the display by roughly 7 visual degrees. The inboard display showed aircraft position and sensor range on a topographical map.

Eye Tracking and Pupillometry.

Eye movement and pupillometric data were collected using a FOVIO FX3 camera (Seeing Machines, Mountain View, CA) and recorded using the EyeWorks™ Record (Eye Tracking, Inc., Solana Beach, CA) software package with the Scene Camera and ICA modules included. The Scene Camera Module allowed the pilot's calibrated gaze position to be overlaid onto a forward-facing video feed collected by a small camera mounted on an overhead control panel, just behind the pilot. Calibration of gaze-position to the visual scene was performed each time the pilot entered the simulator or moved the position of the seat over the course of the session. The procedure involved placing four small markers on the dashboard and asking the pilot to fixate each of those positions while the software registered the subtended angle of the eye. This procedure typically took less than 2 minutes and reported accuracy of less than 2 degrees of visual angle. This was adequate to discern on which symbology elements the pilot was fixating throughout the course of the flight.

The FOVIO FX3 cameras also collected pupil area which was adjusted for head distance and streamed to the EyeWorks™ Record software. Eye Tracking, Inc.'s trademarked (Ref. 4) ICA Module operates on this streamed pupillometric data and provides both real-time and post-collection analyses of the pilot's cognitive workload throughout the session. The ICA is expressed as a value between 0 and 1; 1 represents the highest level of cognitive workload. These data were compiled and analyzed offline to generate the figures presented in the results section.

One Fovio FX3 camera was mounted to the forward panel just below the flight instrument displays for data collection. In this configuration, the camera provided accurate and reliable eye-tracking signal quality for nearly all of the pilots tested, without interfering with the operation of the simulator or the positioning of the pilot. Data presented below detail the ICA-calculated workload from these data in six of the eight UH-60 pilots tested in the present study.

Figure 2. Cockpit with eye tracking cameras in place.

Symbology Sets and Cueing

Overview.

A multimodal obstacle avoidance cueing environment was developed; it integrated visual, spatial-auditory, and tactile elements into the ICE (Ref. 5) to provide improved situational awareness (SA) around the aircraft. The inboard (left) display contained a top-down terrain, RADAR, and obstacle map termed the Integrated Collision Avoidance Display (ICAD). The outboard (right) display presented a forward view, replicating the appearance of sensors that are currently under development, with a graphic overlay displaying critical avionics information, and additional visual collision avoidance symbology (ICE-CAS) in certain configurations. This ICE-CAS system also incorporated three-dimensional (3D) spatial audio alerts and warnings localized to the source of the cue.

For this iteration of ICE, the Tactile Situation Awareness System (TSAS) was integrated to provide an additional obstacle azimuth cueing using a 12-tactor belt. Additional tactors, placed within the shoulder harness and seat cushion, provided warnings of altitude and obstacles above or below the aircraft. Together, these four systems provide an integrated and unified cueing environment to warn of potential collisions in the vicinity of the aircraft, inside and outside the field of view. The ICE-CAS has been integrated and experimentally evaluated in previous studies at the NASA Ames System Integration Laboratory simulator (Ref. 4) and at USAARL in the same simulator used for this study. (Ref. 5; Ref. 6)

ICE-Collision Avoidance Symbology (Outboard Flight Display).

This symbology system is a tailored set of rotorcraft avionics designed specifically with the intent of allowing safe and efficient navigation and collision avoidance during NOE flight profiles in DVE conditions that are, at present, mission restricted. Figure 3 demonstrates the avionics symbology visible when the aircraft is moving at a ground speed in excess of 40 knots and is not within 0.3 nm of the landing zone. Figure 4 demonstrates the additional avionics (glideslope indicator and relative ground position indicator) that appear on the screen when within 0.3 nm of the targeted landing zone (See Table 1: Hover-to-Landing Phase). In addition to additional avionics, the ICE symbology also employed a 3D conformal landing zone, which provides an enhanced visualization of the landing point.

Mission Profile.

Ten flight vignettes were created for this study using a database based on San Francisco (See Figure 5), developed by PLW Modelworks LLC, and implemented using X-IG developed by CATi Training Systems (Ozark, Alabama). The routes involved NOE flying in which the pilots must maintain an altitude below 100ft. Failure to do so would result in a grey-out of the screen emulating low cloud cover.

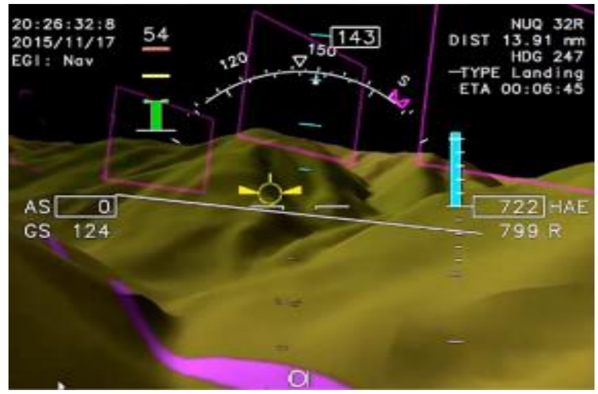

Figure 3. Baseline ICE symbology: En-route page.

Figure 4. Baseline ICE symbology: Takeoff and landing page.

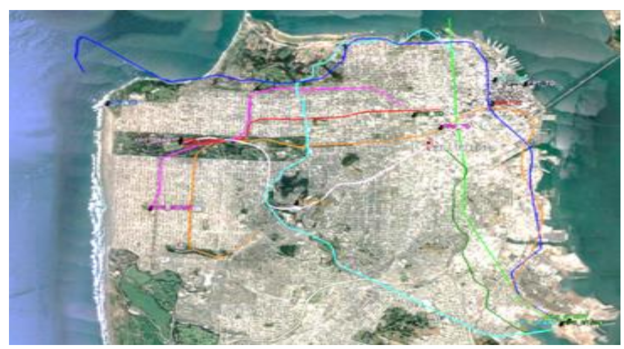

Figure 5. Ten flight routes in San Francisco Bay area.

Each route was designed and tested by USAARL Research Pilots, and involved multiple encounters with obstacles including cranes, tall buildings, and power lines. The routes were topographically separated to avoid multiple encounters with the same obstacles and prevent map familiarization. Each vignette involved the same set of maneuvers, detailed in Table 1. Each pilot was assigned a different combination of object avoidance cueing, as is shown in Table 2, for each vignette, to control for unexpected differences in the difficulty of the routes.

Table 1. Maneuver Definitions.

Maneuver	Start	End
Taxi	Aircraft wheels lift-off in the initial pre-taxi zone.	When the aircraft is wheels down in the takeoff zone.
Takeoff	When wheels lift-off the ground from the takeoff zone.	When the symbology set switches at a ground speed of 40 knots after takeoff.
En-route	When the aircraft reaches a 40 knot ground speed after takeoff.	0.8 nm from landing point.
Approach	0.8 nm from landing point.	0.3 nm from the landing point or when the symbology set switches to hover mode (GS<40kt).
Hover-to-Landing	0.3 nm from the landing point or when the symbology set switches to hover mode (GS<40kt).	Either touchdown in the landing area, or aircraft meets the conditions of a crash.

Table 2. Study Conditions.

Condition #	Cueing	Modality
1	ICE	Baseline
2	ICE-CAS	Visual
3	ICE-CAS +ARSAD	Visual + Auditory
4	ICE-CAS +TSAS	Visual +Tactile
5	ICE-CAS +ARSAD+TSAS	Visual + Auditory+ Tactile

PROCEDURES

Prior to data collection, all participants completed two training flights where USAARL Research Pilots familiarized them with the avionics and ensured they were able to complete the tasks required of the study. Figure 2 & 3 display the baseline ICE avionics and cuing symbology (Table 2, Condition 1) for the en-route phase and hover-to-landing phases of flight (Table 1). Pilots were also outfitted with biometric sensors, none of which presented any compatibility issues or design conflicts with the eye-tracking system employed in this study.

RESULTS

The ICA was calculated from the changes in the pupil diameter during each of the five flight phases and averaged across subjects and configurations. The standard error of means show a large amount of variability in mental workload between subjects, however, the highest levels of mental effort appear to occur during the phases of flight when pilots transition from air to ground and when navigating around obstacles in very close proximity to the ground, specifically in the hover-to-land phase and the takeoff phase.

Since these are regarded as the most difficult and dangerous operations, these data suggest that the ICA algorithm may be sensitive to the increased effort required to navigate these scenarios and warrants further investigation. Unfortunately, the inter-subject variability and small sample size, does not afford statistically significant distinctions between the different cueing configurations using simple group-wise statistical methods. See Appendix for figures related to ICA workload scores for individual cueing combinations and ICA workload scores by phase of flight.

CONCLUSIONS

It is important to note that the small sample size and broad set of cueing combinations used in this study makes it difficult to lend statistical meaning to our findings. However, the findings of this study support the feasibility of incorporating eye-tracking systems into the cockpit that are capable of collecting pupillometrics for the non-invasive assessment of cognitive workload during simulated flight. As would be expected, the phase of flight that induced the highest average ICA workload score was the hover-to-land phase of flight. Further, although results were not statistically significant, the average ICA workload scores indicated that the separate cueing combinations may have utility in specific scenarios. For instance, in the takeoff phase of flight, the ICA was recorded as lowest with the combination of visual and auditory cueing whereas during the hover-to-landing phase, the visual cuing alone elicited the lowest ICA workload score. The aircraft is in very close proximity to several obstacles during this phase, leading to a lot of overlapping signaling from each of the cuing sources. This abundance may actually create more information and thus require more mental effort to focus on the task.

This study is one of several working to advance the ICE as a part of the DVE-M program. The ease of integrating video-based eye-tracking into the cockpit environment, and the promising, albeit insignificant, results we have collected to date, have cemented our commitment to deploying similar non-invasive oculometric monitoring systems in future studies, including future iterations of the ICE assessments. Not only will this effort prove a critical component to further refine the effectiveness of the ICE, but this work lays the foundation for including oculometric monitoring systems in future vertical lift (FVL) aircraft and the operational environment.

APPENDIX

Note that error bars in the following depictions are standard error of means (SEM).

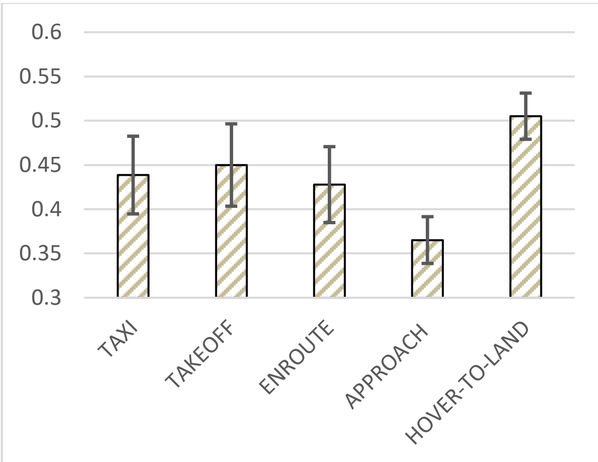

Figure A1. Mean ICA workload scores by phase of flight for baseline condition.

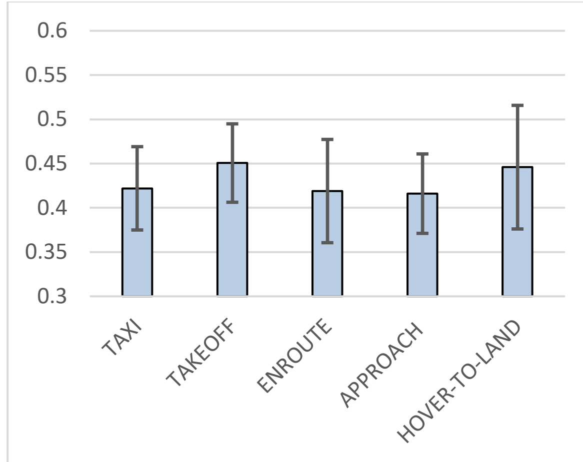

Figure A2. Mean ICA workload scores by phase of flight for visual cueing.

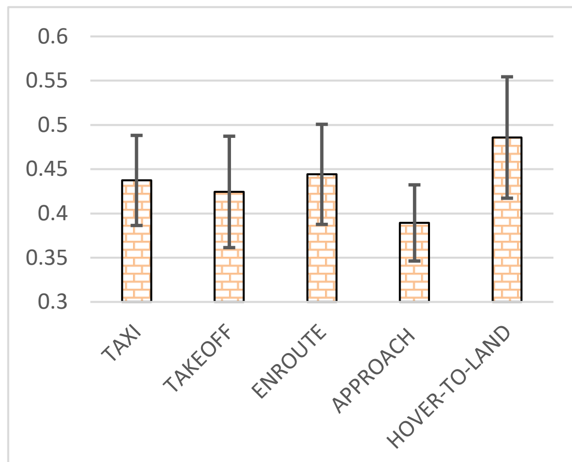

Figure A3. Mean ICA workload scores by phase of flight for visual plus auditory cueing.

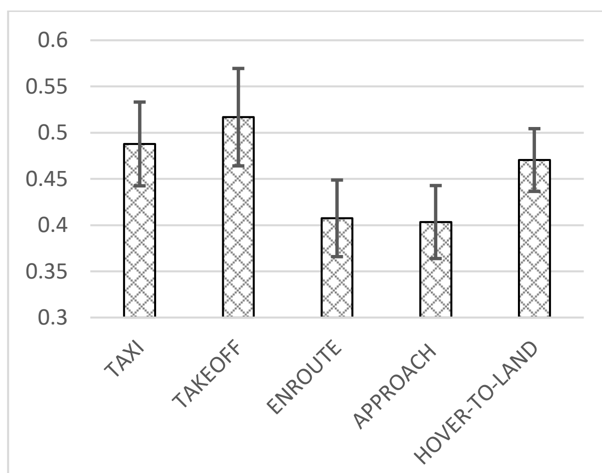

Figure A4. Mean ICA workload scores by phase of flight for visual plus tactile cueing.

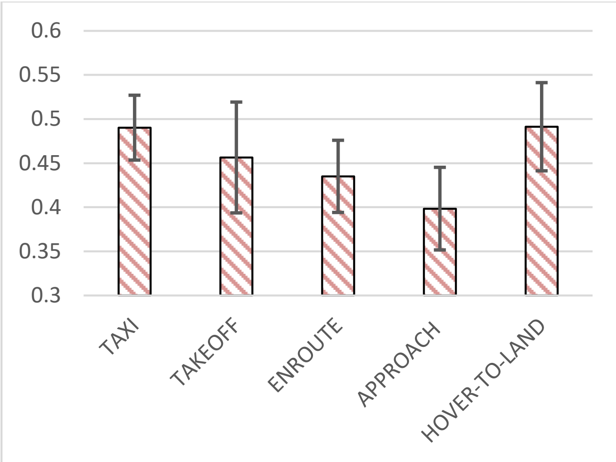

Figure A5. Mean ICA workload scores by phase of flight for visual plus auditory plus tactile cueing.

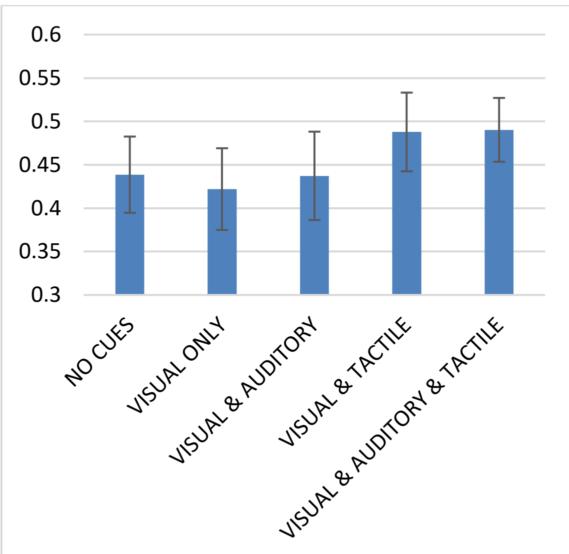

Figure A6. Mean ICA workload scores for taxi phase.

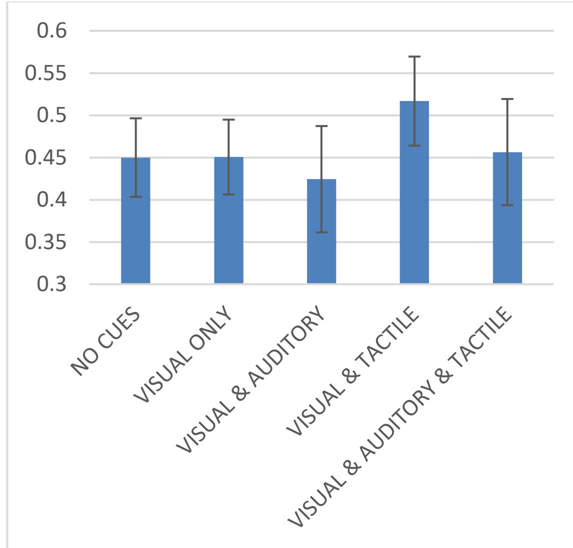

Figure A7. Mean ICA workload scores for takeoff phase.

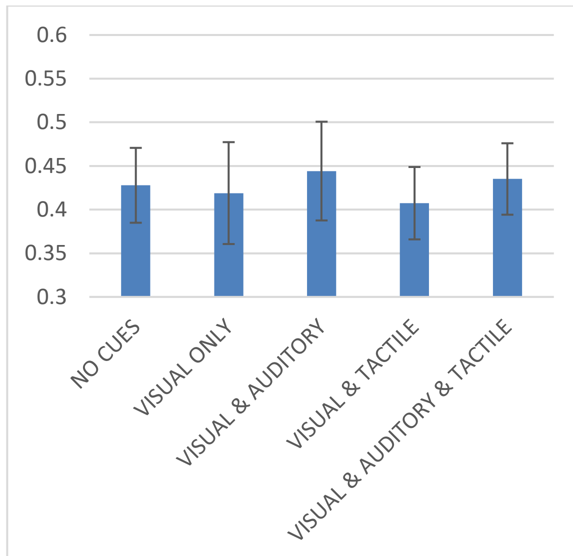

Figure A8. Mean ICA workload scores for en-route phase.

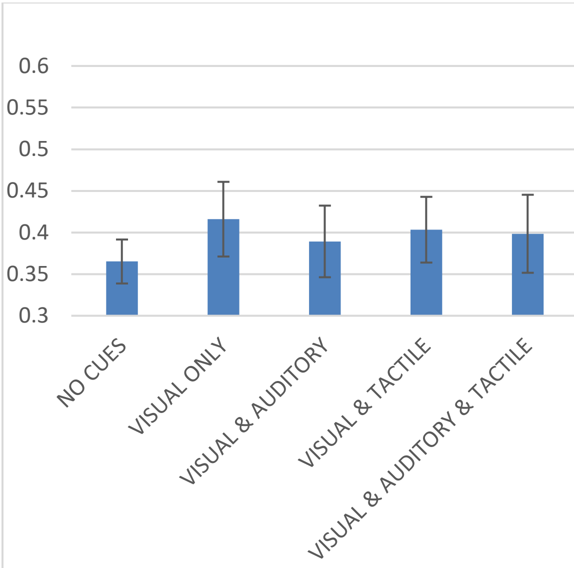

Figure A9. Mean ICA workload scores for approach phase.

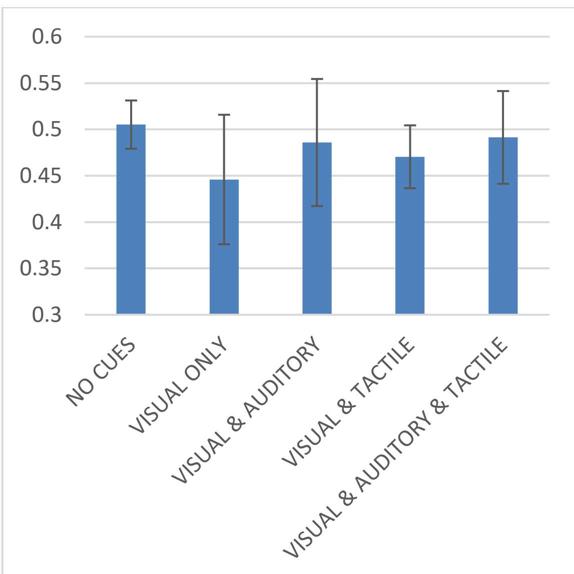

Figure A10. Mean ICA workload scores for hover to land phase.

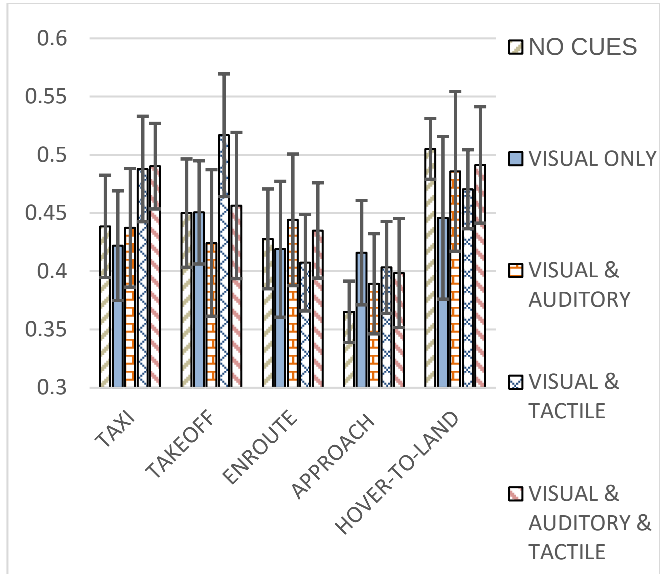

Figure A11. Mean ICA workload scores by cueing combination and phase of flight.

ACKNOWLEDGMENTS

The authors would like to acknowledge and thank Rolf Beutler and Nathan Mielcarek for all their hard work and efforts in ensuring the systems were prepared for data collection. The authors would also like to thank the simulator technicians who ensured the simulator was available and functional for use, and the technical support team who helped collect the data for the study: Raquel Goosey, Jared Basso, Jim Chiaramonte, Colby Mathews, and Melody King. Additional thanks to Dr. Kevin O'Brien, Will Irvin, and Kyle Bernhardt who assisted with the biometric data for study analyses. Most importantly, the authors would like to thank the Research Pilots in the Flight Systems Branch who aided in the flight profile design and data collection, specifically, CPT Justin Stewart. This research was supported by an appointment to the Postgraduate Research Program at the U.S. Army Aeromedical Research Laboratory administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the U.S. Army Medical Research and Development Command.

REFERENCES

1. Russell, D., Statz, J. K., Ramiccio, J., Henderson, M., Still, D., …& Estrada, A., "Pilot cueing synergies for degraded visual environments," USAARL Report No. 2016-10, 2016.
2. McAtee, A., Russell, D., Feltman, K., Swanberg, D. E., Statz, J. K., Ramiccio, J., & Harding, T. H., "Integrated cueing environment testing: Pilot cueing synergies for degraded visual environments," USAARL Report No. 2017-04, 2017.

3. Feltman, K. A., Kelley, A. M., Curry, I. P., Boudreaux, D. A., Milam, L., Mathews, C., & Russell, D., "Review of US Army aviation accident reports: Prevalence of environmental stressors and medical conditions," USAARL Report No. 2018-02, 2018.
4. Godfroy-Cooper, M., Miller, J. D., Bachelder, E., Wenzel, E. M. "Isomorphic Spatial Visual-Auditory Displays for Operations in DVE for Obstacle Avoidance," Proceedings of the European Rotorcraft Forum, Delft, The Netherlands, September 17-20, 2018.
5. Marshall S. P, "Identifying cognitive state from eye metrics," *Aviation, Space, and Environmental Medicine, Vol. 78*, 2007, pp. 165-175.
6. Miller, J. D., Godfroy-Cooper, M., & Wenzel, E. M., "ARSAD: An augmented-reality spatial auditory display for obstacle avoidance for all phases of flight," Proceedings of the 74th of the American Helicopter Society Annual Forum & Technology Display, Phoenix, Arizona, USA, May 14-17, 2018.