DOI: 10.2352/ISSN.2470-1173.2016.5.SDA-437

# Stereoscopic Remote Vision System Aerial Refueling Visual Performance

*Marc Winterbottom[1]*, Charles Lloyd[2], James Gaska[1], Steven Wright[3], and Steven Hadley[1]*

*[1]Operational Based Vision Assessment Laboratory, Operational Health and Performance Research Division, Aeromedical Research Department, U.S. Air Force School of Aerospace Medicine, 711th Human Performance Wing, Wright-Patterson AFB, OH*

*[2]Visual Performance LLC, Ellisville, MO*

*[3]Aerospace Ophthalmology Branch, Aerospace Medicine Department, U.S. Air Force School of Aerospace Medicine, 711th Human Performance Wing, Wright-Patterson AFB, OH*

## Abstract

*The performance and comfort of aircrew using stereoscopic displays viewed at a near distance over long periods of time are now important operational factors to consider with the introduction of aerial refueling tankers using remote vision system technology. Due to concern that the current U.S. Air Force vision standards and test procedures may not be adequate for accurately identifying aircrew medically fit to operate this new technology for long mission durations, we investigated performance with the use of a simulated remote vision system and the ability of different vision tests to predict performance and reported discomfort. The results showed that the use of stereoscopic cameras generally improved performance but that individuals with poorer vision test scores performed more poorly and reported greater levels of discomfort. In general, newly developed computer-based vision tests were more predictive of both performance and reported discomfort than standard optometric tests.*

## Introduction

With the introduction of the next generation of aerial refueling tankers (e.g., Boeing KC-46 for the U.S. Air Force (USAF), Airbus KC-30 for the Royal Australian Air Force, Boeing KDC-10 for the Royal Netherlands Air Force, and KC-767 for the Japan Air Self-Defense Force), in which aerial refueling operators (AROs), or boom operators, will use relatively recently developed indirect view stereo displays in place of direct view crew stations, existing vision standards for boom operators may not be adequate [1]–[5]. In particular, the level of stereo acuity and oculomotor capabilities required to maintain stereo fusion with these new stereoscopic remote vision displays in rested and fatigued states are generally unknown. Mild anomalies (currently allowed by USAF vision standards) in binocular alignment may permit stereopsis, but may also predispose those individuals to visual complaints such as asthenopia (eyestrain) or headaches [6]. These visual complaints may not occur under normal viewing conditions, but may arise under more unnatural viewing conditions, such as viewing stereoscopic content on 3D displays. Although three-dimensional (3D) displays have been in use for many years, their popularity has grown in recent years, and the sales of 3D displays for television, movies, and gaming have grown considerably. Additionally, head-mounted display (e.g., Joint Strike Fighter) and remote view display applications (e.g., tele-robotic surgery, remotely controlled air and ground vehicles, remote view aerial refueling) have drawn attention to the need for more research on the use of stereoscopic imagery.

Stereoscopic displays offer a number of potential benefits, such as aiding in the encoding of large amounts of complex information (e.g., 3D modeling of complex structures) or improving identification of important details, especially in noisy/complex scenes [7]. However, there are also serious drawbacks associated with 3D displays. Inconsistent cues may cause discomfort/eyestrain and 3D perception may be inaccurate or totally disrupted. In fact, reports of serious discomfort are very common. A recent study found that 55% of respondents reported discomfort after viewing a 3D movie versus 14% for a two-dimensional (2D) movie [8]. Potential sources of problems and/or discomfort include binocular asymmetry (differences in left and right image quality or position); perceptual inconsistencies (e.g. vergence-accommodation mismatch, motion parallax-convergence mismatch resulting from the depth plane differing from the image plane); distortions due to hyper-stereo; and crosstalk (resulting from incomplete separation of left/right eye images). In fact, some commercially available 3D displays come with warnings that users could experience altered vision, lightheadedness, confusion, or nausea.

Several of these distortions, in particular vergence-accommodation mismatch, hyper-stereo, vertical misalignment, and crosstalk, may be present in the remote vision system (RVS) for the aerial refueling task. And, due to demanding mission requirements, viewing periods could be quite lengthy, so there is the potential for significant levels of discomfort. However, because experimentation with lengthy and repeated viewing is time consuming and difficult, the extent to which increased exposure duration when viewing stereoscopic displays affects discomfort is not well known. Research examining reports of discomfort often use relatively short viewing periods but very large misalignment, disparities, or other distortions to more efficiently examine factors affecting stereoscopic 3D display performance [7], [9]. Other researchers have commented that amusement park stereoscopic displays are viewed for relatively short periods of time (30 minutes or less) and are therefore “visually bearable” [10] and go on to note that most studies examining discomfort involve only relatively short viewing periods. However, several studies have used fairly long view periods (20 – 45 minutes) to examine 3D displays and discomfort/fatigue [10]–[12]. It stands to reason that lengthier viewing periods could lead to increased fatigue and discomfort, and previous research does, in fact, show that symptom severity increased over the course of a 20-minute viewing period [12].

A potentially significant problem for the wider use of stereoscopic displays, particularly in applications where operational task performance depends on information conveyed through a 3D display, is that a significant portion of the population may have binocular vision anomalies [9]. Some anomalies prevent normal stereopsis, such as strabismus or amblyopia. Other anomalies permit stereo, but may predispose patients to visual


*Corresponding author: marc.winterbottom@us.af.mil

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.5.SDA-437

complaints (i.e., asthenopia, or eyestrain). In previous research [9], 50 volunteers were evaluated on their quality of ocular alignment and classified as having either 1) good binocular status (GBS) or 2) moderate binocular status (MBS) based on standard optometric vision tests such as visual acuity, stereo acuity, phoria, etc. Subjects were then asked to perform a reading task for 60 seconds using either a 2D or a 3D display (crossed disparity only). In the stereo condition, the relative disparity between text and frame in 3D condition was 1.5 degrees, which is beyond the 1-degree limit typically rated as comfortable, to stress subjects' visual systems. A questionnaire was also administered to assess comfort. The results of this study reveal that both the GBS and MBS participants reported significantly more eyestrain for the 3D condition. However, the MBS participants tended to report a greater degree of eyestrain and other symptoms compared to GBS subjects, and MBS subjects performed significantly worse on the 3D display relative to the 2D display on the reading task.

For real world objects, accommodation and vergence are coupled whereas for a typical stereo display, best-focus is at the screen and optimal vergence changes with disparity. This can result in Vergence-Accommodation (VA) mismatch which has been suspected as a source of discomfort for many years [13]. However, a definitive link between VA mismatch and performance and discomfort was not established until more recently using unique displays that allowed better control over vergence and accommodative distances [11], [12]. Previous research has shown that subjects' reaction time to discriminate grating orientation substantially increases as the difference between accommodation/vergence distance increases, thresholds for spatial frequency increase as the difference between accommodation/vergence distance increases, and that subjects report a significantly greater degree of discomfort when accommodation and vergence cues are inconsistent. These data provide definitive evidence that VA mismatch contributes to discomfort [11]. Additional research has also shown that binocular status, or individual phoria measurements, may be related to comfort ratings when viewing stereo displays [12].

The same technology can be used to generate hyper-stereoscopic viewing conditions. This is achieved by placing the pair of cameras, or another pair of image sources, generating the left and right eye images at a greater distance apart than the normal interpupillary distance. Hyper-stereo magnifies the perceived disparity, increasing the sense of depth. The KC-46 RVS, and similar systems, use a hyper-stereoscopic configuration. The use of hyper-stereo can potentially reduce viewing comfort [7]. However, a separate but equally important issue is the effect of hyper-stereo on performance. The use of hyper-stereo is becoming an important issue since it is being employed in a variety of applications, such as with helmet-mounted displays and aerial refueling. Increased disparity associated with hyper-stereo can result in distorted cues to distance/depth and can lead to visual illusions that may affect flight performance (e.g., crater or bowl effect, distorted perception of slope, distorted time to contact estimates, etc.) [14]. Hyper-stereo may lead to an underestimation of time to contact and an increase in the perceived slope of a textured surface [15]–[17]. However, in flight tests with a TopOwl helmet-mounted display, pilots seemed to adapt to the hyper-stereoscopic conditions after about five sorties [14].

A hyper-stereoscopic display system can result in depth plane curvature, depth non-linearity, shear distortion, and keystone distortion and vertical parallax (dipvergence). The distortions that result from different hyper-stereo camera configurations can be predicted based on a set of equations developed in previous work [18]. Camera toe-in in particular can lead to vertical disparity and may also reduce perceived depth. However, little work has been done to examine depth distortions caused by camera toe-in [19]. A key aspect of the research presented here is to replicate, to the greatest extent feasible, any of these distortions that may result from the particular configuration of the hyper-stereoscopic KC-46 RVS.

### *Aerial Refueling Task*

The Boeing KC-46 ARO crew station uses two cameras mounted at the rear of the aircraft to generate a hyper-stereoscopic view of aircraft approaching the KC-46 for refueling. The imagery is presented to the ARO using a stereoscopic display. Additional cameras (panoramic cameras) and displays allow the ARO to view the airspace around the KC-46 to monitor approaching aircraft and aircraft waiting to refuel. The ARO controls the boom using two joysticks mounted on the console. When a receiver aircraft approaches to within about 1.6 km (1 mile), the ARO takes control of the receiver aircraft (i.e., becomes the air traffic controller) and relies primarily on the panoramic cameras to guide the receiver to pre-contact position. When the receiver aircraft approaches to a close enough distance, the ARO transitions to the 3D display. The pre-contact position is at approximately 15 meters (50 feet). An overlay display (analogous to a heads-up display, or HUD), showing boom position and other information can also be used by the ARO to aid the refueling task. The ARO calls out distances and issues verbal commands or uses the indicator lights to adjust the position of the receiver aircraft. Figure 1 shows the view of a KC-10 boom operator during refueling.

*Figure 1. View from the KC-10 ARO crew station.*

The simulated aerial refueling task for this research was designed based on in-depth interviews with experienced KC-135 and KC-10 boom operators. The simulated refueling task is similar to that of a "fighter drag" refueling scenario where a boom operator repeatedly refuels several aircraft while on a long-haul flight. The characteristics of the RVS were simulated based on detailed and proprietary data provided by the Boeing Co.

DOI: 10.2352/ISSN.2470-1173.2016.5.SDA-437

Based on our review of the research literature, interviews with boom operators, input from the KC-46 program office, and in anticipation of fielding and selecting aircrew to use the new hyper-stereoscopic aerial refueling RVS, the objectives of the research presented here were to 1) compare RVS refueling performance for several different types of viewing conditions (2D, normal stereo, and hyper-stereo); 2) examine individual differences in performance while using a simulated RVS for a long period of time (2 hours); and 3) determine whether vision screening could effectively predict which individuals might experience difficulty when using the RVS.

## Methods

### *Apparatus and Vision Tests*

#### *Remote Vision System Aerial Refueling Simulation*

A 5-channel PC-based image generator (Flight Safety International Vital X) was used to present the boom model and receiver aircraft model over ground terrain. A standard flight database provided by Flight Safety consisting of desert terrain was used. In this simulation the simulated KC-46 aircraft orbited a location over the western United States. The tanker traveled at 593 km/h (320 knots), at an elevation of 9,100 meters (30,000 feet) and completed the circuit in approximately 20 minutes. Two of the IG channels drove the left and right eye views of the simulated boom camera viewpoints. Two Black Magic DVI Extenders and an AJA Video Multiplexer were used to combine the two video channels into one top-and-bottom stereoscopic image, which was presented using a ViewSonic V3D231 23-inch 3D monitor. Three additional channels drove the left, center, and right panoramic camera viewpoints, which were presented on three HP monitors (HP Pavilion 21.5-inch IPS LED HD monitors). A separate PC served as the host emulator, which allowed the experimenter to control the simulated refueling scenario. The ViewSonic monitor uses passive polarizing glasses to present the left and right eye views. Thus, each eye viewed a 1920- x 540-pixel image (approx. 1.1 arcmin/pixel) at 24 Hz when viewing the 3D display. The 3D monitor was mounted in front of the observer and viewed at a distance of 33 inches. The panoramic displays were mounted above the 3D display. Subjects controlled the boom using 2 Saitek X52 flight controllers mounted on either side of the 3D display. The controller on the right provided control over the boom left, right, up, and down position. The controller on the left provided control over the boom extension. Light turbulence was simulated using a sum of sines. The relative motion of the two aircraft occurred in the left-right direction with a peak amplitude of approximately 2 feet. The range of temporal frequencies present in the relative motion was approximately 0.1 to 1 Hz.

An HP Envy 15t Quad laptop computer running the Windows 8.1 OS was also used to control an aspect angle recognition task that was displayed on the 2D HP monitors and the ViewSonic stereoscopic monitor. For this task, participants entered their responses using a 6-button response box (Cedrus Corp. model RB-730). A Microsoft Surface Pro 2 was used to administer a questionnaire. Figure 2 shows the refueling task apparatus.

#### *Vision Tests*

**Contrast Sensitivity**

A specially modified contrast sensitivity (CS) test was provided by Adaptive Sensory Technology (AST). For this test, participants viewed band-pass filtered Landolt C stimuli at a distance of 4 meters and responded to the position of the gap (Figure 3). The AST CS test uses a sophisticated adaptive procedure to rapidly estimate the contrast sensitivity function [20]. Each participant completed the CS test three times: 1) binocularly (OU), 2) monocular left eye (OS), and 3) monocular right eye (OD). Diffusing filters were used to equalize the luminance across the two eyes for the monocular conditions. For the binocular condition, neutral density filters were used to maintain the same luminance across all three conditions. No feedback concerning accuracy of responses was provided for this test. An NEC MultiSync P463 flat panel display was used to present the imagery, which was generated by a Unix-based Shuttle PC. Participants entered responses using a Microsoft X-box game controller.

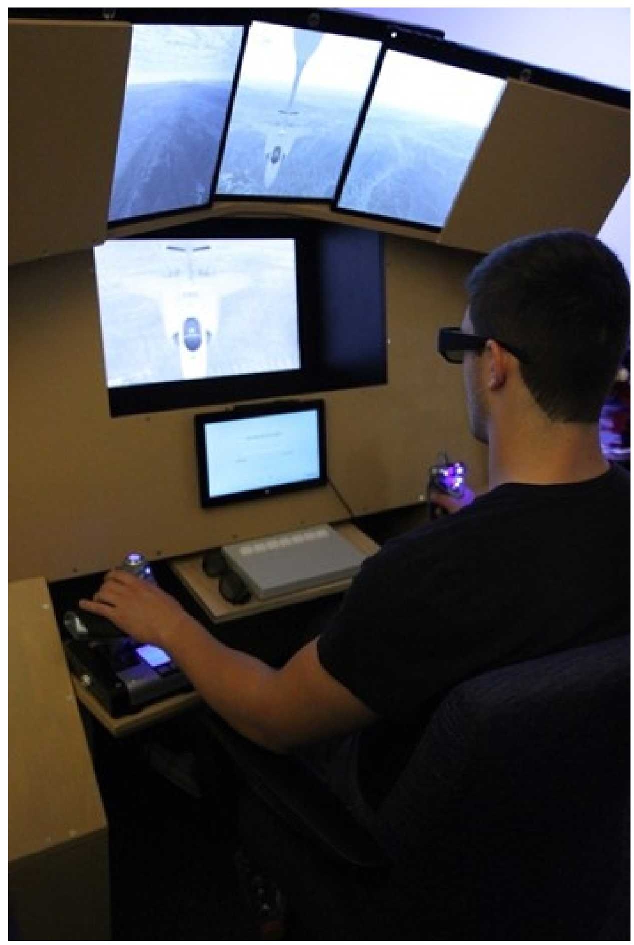

*Figure 2. Simulated remote vision system aerial refueling task apparatus.*

*Figure 3. Band-pass filtered Landolt C stimuli. Figure provided by AST. Image used with permission.*

### Near and Far Stereo Acuity

This test involved depth discrimination of a center circle relative to a larger outer circle presented on a stereo-display (Figure 4) at a viewing distance of either 1 meter or 4 meters. For this test, the circles were presented for 2 seconds. The participants' task was simply to indicate whether the small inner circle appeared to be in front of or behind the larger outer circle using the game controller. For this test, the luminance of the gray background was approximately 38 cd/$m^2$ and the luminance of the center disc was approximately 81 cd/$m^2$. However, the luminance through the 3D glasses was substantially reduced. An adaptive procedure was used to estimate the stereoscopic depth threshold [21], with the number of trials fixed at 35. Prior to the actual test, each participant completed several practice trials to become familiar with the test procedure. Auditory feedback (tone: correct; buzzer: incorrect) was provided. This stereo acuity test is similar to a computer-based test described in previous research [22].

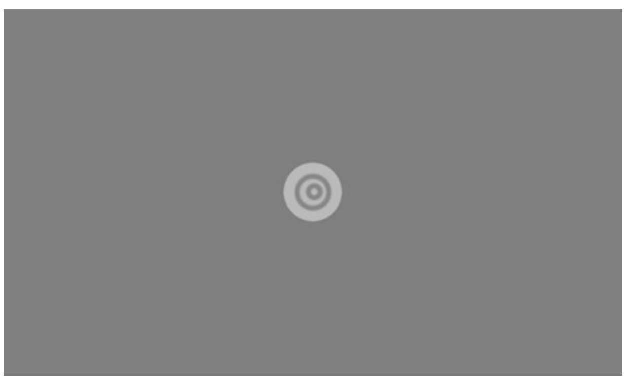

*Figure 4. Near stereo test stimulus.*

### Horizontal and Vertical Fusion Range

This task required that participants indicate when they could no longer fuse a circular target that was moved in depth in either the crossed or uncrossed directions (i.e., they could no longer cross/uncross their eyes and the left and right eye images became doubled). For this test the display was viewed at a distance of 1 meter and the game controller was used to respond. After the fusion break point was recorded, the direction of motion reversed, and the participant indicated when the circles returned to a single "fused" image (recovery). This task was repeated several times. The test images viewed by the participant (for left and right eye images) at one instant in time are shown in Figure 5. A similar procedure was used to record vertical fusion range, with stimuli moving in the vertical direction rather than horizontal.

A Dell Precision T7610 with NVidia GeForce GTX 680 graphics card was used to administer the stereo acuity and fusion range tests. Participants used a Logitech game controller to enter responses. The test software was developed using Visual Basic, Matlab (www.mathworks.com), and Octave (http://www.gnu.org/software/octave/). Tests were displayed on an Asus VG278HE 3D monitor with 1920 x 1080 pixels that was compatible with NVidia 3D Vision2 using active shutter glasses. At a 1-meter viewing distance, the angular pixel pitch was 1.1 arcmin.

*Figure 5. Horizontal fusion range stimuli (left and right eye images).*

### *Stereo Display Luminance and Crosstalk*

For most of the computer-based vision testing, the participants donned active shutter NVidia 3D vision glasses. The brightness of the Asus display was noticeably lower in 3D mode compared to 2D mode. To increase the luminance, the NVidia brightness boost was enabled and set to the highest level.

To evaluate crosstalk, black and white test images were generated using a software application that allowed different images to be displayed to the left and right eyes in 3D view and measured using a Minolta LS-100 through the left and right lenses of the glasses. Crosstalk was calculated according to an equation described in previous research [23]:

$$OCT_{RL} = (O_{GL} - O_{BL})/(O_{WL} - O_{BL})$$

where $OCT_{RL}$ is observed crosstalk- right to left, $O_{GL}$ is observed ghost image-left, OBL is observed black-left, and $O_{WL}$ is observed white-left. Based on this procedure, a crosstalk level of 0.2% was estimated for the Asus 3D display using the active shutter glasses.

For the aerial refueling task, the participants donned passive, polarized 3D glasses to view the ViewSonic 3D display. The luminance of the F-35 receiver aircraft without glasses was approximately 20.3 cd/$m^2$. The luminance of the receiver aircraft through the glasses was approximately 7.4 cd/$m^2$. Thus, the glasses reduced luminance by approximately 64%. A crosstalk value of 1.4% was obtained based on the same procedure described above. Although no ghosting was apparent when viewing the refueling simulation, this value is greater than the 0.3% recommended for comfortable viewing [7]. Figure 6 shows the two different kinds of 3D glasses used in this research.

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.5.SDA-437

*Figure 6. NVidia 3D Vision active shutter glasses (left) and passive polarizing 3D glasses supplied with the ViewSonic 3D monitor (right).*

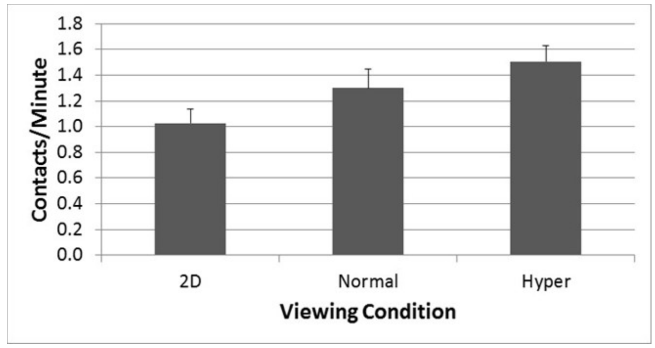


*Figure 7. Number of contacts/minute for each viewing condition.*

## *Experiment 1*

### *Participants*

Thirteen participants ranging in age from 18 to 57 were recruited to participate in Experiment 1. Nine of the participants in this experiment were male and 4 were female. Two experienced KC-135 boom operators also completed the refueling task but did not participate in the vision testing.

### *Procedure*

In this experiment, the viewing conditions switched between hyper-stereoscopic (KC-46 configuration), normal stereo (camera separation set to an interpupillary distance of 65 mm), and 2D. The order of viewing conditions was randomized for each participant. Participants viewed only the central stereoscopic display for nearly the entire duration of the experiment. The refueling duration was 5 minutes. Following each 5-minute refueling period, the experimenter changed the viewing condition according to the randomized order. An aerial refueling metric defined as Success Rate = #Contacts⁄Time (average number of successful contacts per minute) was used to compare performance across viewing conditions. To assess individual refueling performance a sum of Z-transformed metrics was used: number of contacts, number of collisions, number of attempts, duration, and distance at attempt.

All participants had previously engaged in a period of training prior to beginning Experiment 1; additionally, prior to beginning Experiment 1, participants engaged in 10-30 minutes of additional practice. Participants completed 3 viewing conditions x 4 repetitions = 12 blocks of aerial refueling. The first block for each viewing condition was excluded from subsequent analysis to further reduce practice effects.

### *Results*

Figure 7 shows the difference in aerial refueling success rate for each viewing condition. Paired sample t-tests revealed that the difference in aerial refueling performance between the 2D and Normal viewing conditions was significant [$t(14) = -3.6$, $p = 0.003$], between the 2D and Hyper viewing conditions was significant [$t(14) = -5.1$, $p < 0.001$], and between the Normal and Hyper viewing conditions was significant [$t(14) = -2.4$, $p = 0.03$].

The effect of quality of vision was also examined. Figure 8 shows the difference in aerial refueling performance for participants categorized as either "good" or "poor" based on combined stereo acuity, fusion range, and contrast sensitivity scores. Participants were simply categorized as above or below the mean (i.e., positive vs. negative combined z-scores). According to these criteria, 7 participants were categorized as "good" and 6 as "poor." As shown, participants with better vision appear to benefit from the use of hyper-stereo, while participants classified as poor appear to perform worse in the hyper-stereo viewing condition, and an analysis of variance revealed a significant interaction [$F(2,22) = 4.9$, $p = 0.036$]. These results suggests that some individuals may have more difficulty with stereo viewing conditions relative to normal 2D displays, or at least may not benefit from the use of hyper-stereo.

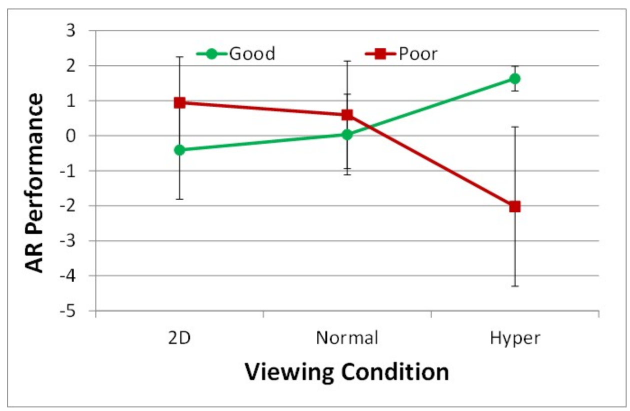


*Figure 8. Aerial refueling performance for participants categorized as either "good" or "poor" based on their combined quality of vision scores.*

## *Experiment 2*

### *Participants*

Twenty-eight participants (23 male, 3 female) were recruited for Experiment 2. Eight participants were over the age of 40 and exhibited varying levels of presbyopia. Two subjects were between the ages of 30 and 40, and 17 subjects were between the ages of 18 and 30. Participants over the age of 40 were included since aerial refueling operators could also be over the age of 40.

©2016 Society for Imaging Science and Technology
DOI: 10.2352/ISSN.2470-1173.2016.5.SDA-437

***Procedure***

Experiment 2 was designed to simulate conditions a KC-46 boom operator might experience during an extended “fighter drag” refueling mission. Thus, Experiment 2 was designed to require participants to switch back and forth between the 2D panoramic and 3D stereoscopic displays and to make repeated refuelings over a 2-hour time period – conditions that were anticipated to induce visual fatigue or discomfort for some individuals. To simulate the approach phase, an aircraft aspect angle task was employed where each participant was asked to identify the aspect (orientation) of a fighter aircraft viewed through the panoramic displays. Each participant then repeated the aspect task using the 3D display. The Psi method [21] was employed to vary the receiver aircraft position and estimate a threshold recognition range for each participant and for each of six different positions (left panoramic, center panoramic, right panoramic, and left, right, and center in the 3D display). In a given block of trials, participants completed 35 trials while viewing the panoramic display and another 35 trials in the 3D display, which took approximately 5 minutes. Figure 9 shows two images of the target aircraft on the 3D display (i.e., the boom camera view) at different distances and positions (center vs. left).

Following completion of the aspect angle task, the refueling task began. At the beginning of the task, the experimenter moved the receiver aircraft into pre-contact position and deployed the boom. The participant was required to recognize when the receiver began moving into the contact position and to then fully deploy the boom and prepare for refueling. Participants were instructed to make contact with the receiver as quickly as possible, while avoiding striking the fuselage with the boom. In the event of a strike, the 3D display flashed to indicate a strike, and the experimenter retracted the boom to indicate a failed trial. The experimenter then repositioned the aircraft into the pre-contact position, then back to contact position and deployed the boom (thus imposing a time penalty). To lock the refueling boom to the receiver aircraft’s receptacle, the participant had to successfully maneuver the boom to within 3 inches and depress the trigger button on the right-hand controller. If successful, the boom locked to the receptacle and the participant was instructed to leave the boom in place to simulate fuel off-load for several seconds. The experimenter then repositioned the receiver to prepare for the next trial. If unsuccessful, the boom would retract, and the participant could try again to make a connection.

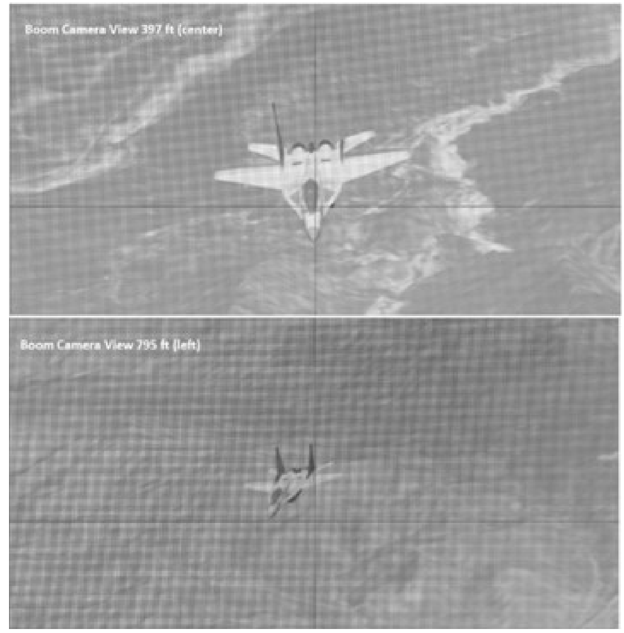


*Figure 9. Images of the aspect angle task target aircraft as displayed on the 3D display (boom camera view) at 2 distances and 2 positions (center and left).*

Following each 7-minute refueling block, the participant completed a short questionnaire to estimate fatigue/discomfort during the course of the experiment. The questionnaire items were adopted from previous research [12]. Participants repeated this procedure for 2 hours. The receiver aircraft position was varied by as much as 1.5 m (5 feet) left, right, fore, and after and 0.3 (1 foot) up, and down across refueling attempts. Varying the position of the receiver aircraft from trial to trial was intended to prevent participants from simply learning a series of steps to complete the task. Twelve different receiver aircraft refueling positions were used. Figure 10 shows 4 of the 12 receiver aircraft positions. In the pre-contact position (upper left), the receiver aircraft is about 60 feet away from the tanker and out of reach of the boom. In the other three positions, the receiver aircraft is about 40 feet from the tanker. However, as shown, the receiver varies noticeably fore/aft and/or horizontally/vertically.

The same combination of metrics described in Experiment 1 (number of contacts, number of collisions, number of attempts, duration, and distance at attempt) was used to assess the performance of all 28 participants individually.

DOI: 10.2352/ISSN.2470-1173.2016.5.SDA-437

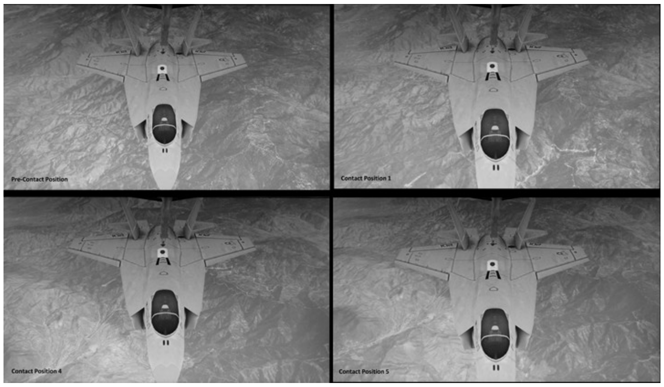


*Figure 10. Receiver aircraft positions viewed through the simulated RVS. Pre-contact position (top left), contact position 1 (top right), contact position 4 (lower left), and contact position 5 (lower right).*

## Results

### *Stereo Acuity*

A frequency histogram for near stereo acuity test results is shown in Figure 11. As shown, participants varied widely on stereo acuity test scores. Since near and far stereo acuity test scores were highly correlated ($r = 0.62$, $p < 0.001$), near and far scores were combined to compute a single stereo acuity value for each participant.

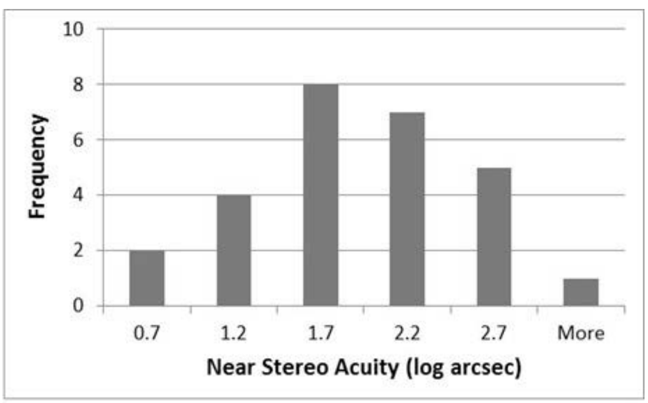


*Figure 11. Near stereo acuity histogram showing the distribution of scores.*

Figure 12 shows the log transformed horizontal and vertical fusion ranges for two subjects. Subject 87 has a relatively small horizontal fusion range while subject 76 has a relatively larger fusion range. In subsequent analyses, log horizontal and vertical fusion range values were combined (log vertical range subtracted from log horizontal). Horizontal and vertical fusion ranges were combined in this way based on discussions with U.S. Air Force School of Aerospace Medicine optometrists and ophthalmologists. A larger horizontal range is indicative of good ocular alignment and ocular motility. For vertical fusion range though, a smaller range is indicative of good ocular alignment.

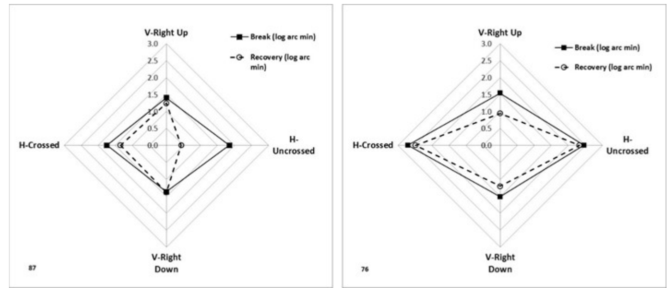


*Figure 12. Horizontal and vertical fusion ranges for two subjects (log arc minutes).*

### *Contrast Sensitivity*

Figure 13 shows binocular and monocular contrast sensitivity test results for two subjects in Experiment 2. As shown, one subject has a decreased contrast sensitivity in one eye, while the other subject has approximately equal contrast sensitivity in each eye. To assess the relationship between contrast sensitivity and RVS refueling performance, the area under the log contrast sensitivity function for the weakest eye was used.

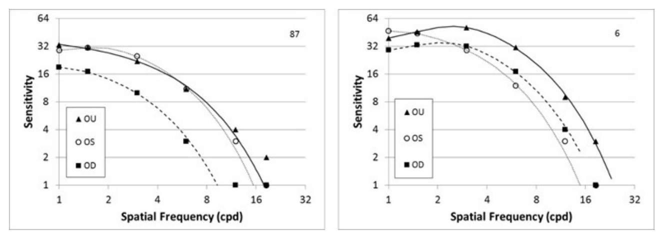


*Figure 13. Contrast sensitivity (CS) test results for two subjects. CS functions are shown for binocular (OU), left eye only (OS), and right eye only (OD).*

### *Reported Discomfort*

Reports of discomfort generally increased over the course of the 2-hour refueling task. Figure 14 shows the average level of eye-tiredness, difficulty maintaining focus/vergence, and viewing comfort throughout the 2-hour task (average value at the end of each 7-minute refueling session). However, although several subjects took one or more rest breaks during the task, no one became so uncomfortable that they quit the task. Although most subjects only reported mild to moderate discomfort, several subjects reported being very uncomfortable for a substantial proportion of the task duration. Figure 15 shows the number of subjects reporting different levels of eye-tiredness.

### *Simulated Aerial Refueling Performance*

Figure 16 shows the range of values obtained for number of contacts per block, number of collisions per block, number of attempts (average number misses prior to a successful connection), mean duration to make contact, mean distance across attempts, and the composite aerial refueling performance score (AR performance, based on combined z-scores).

DOI: 10.2352/ISSN.2470-1173.2016.5.SDA-437

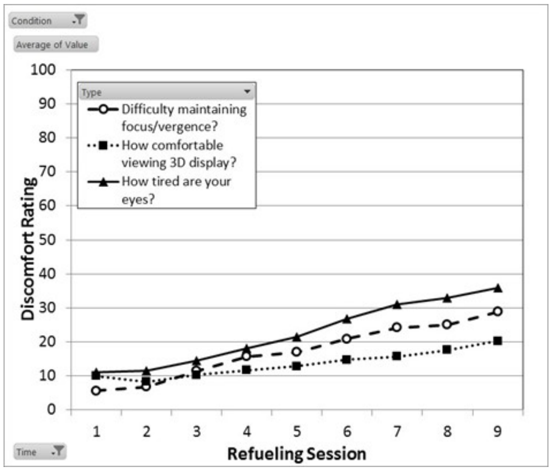


*Figure 14. Average values for focus/vergence difficulty, viewing comfort, and eye-tiredness throughout 2-hour refueling task.*

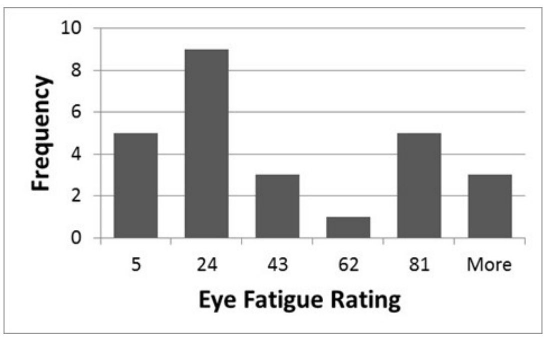


*Figure 15. Histogram showing the number of subjects reporting different levels of eye-tiredness at the conclusion at the end of the 2-hour refueling task.*

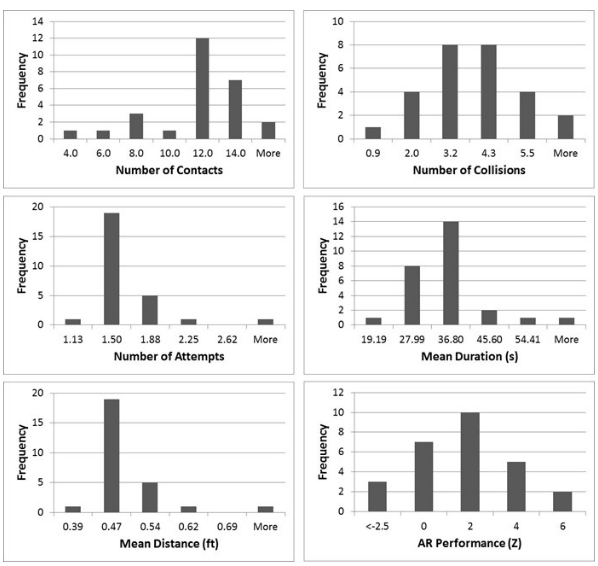


*Figure 16. The range of values obtained for number of contacts per block (top left), number of collisions per block (top right), average number of attempts (middle left), mean duration (middle right), mean distance (bottom left), and composite aerial refueling (AR) performance (bottom right).*

### *Vision Tests and Aerial Refueling Performance*

Table 1 summarizes the correlations between each vision test and simulated RVS aerial refueling performance. Each vision test was highly correlated with aerial refueling performance.

| Vision Test | Correlation | Significance |
|---|---|---|
| Stereo Acuity | 0.56 | 0.002 |
| Fusion Range | 0.7 | < 0.001 |
| Contrast Sensitivity | 0.68 | < 0.001 |

*Table 1. Correlation between each vision test and simulated aerial refueling performance.*

Similarly, Table 2 summarizes the correlations between each vision test and discomfort ratings. The correlation between fusion range and discomfort ratings was highly significant, but the relationship between stereo acuity and contrast sensitivity was not as strong.

| Vision Test | Correlation | Significance |
|---|---|---|
| Stereo Acuity | 0.4 | 0.04 |
| Fusion Range | 0.55 | 0.003 |
| Contrast Sensitivity | 0.39 | 0.04 |

*Table 2. Correlation between each vision test and average discomfort ratings reported during the refueling task.*

Interestingly, age appears to be a significant factor affecting reported discomfort with the use of the RVS. Figure 17 shows average discomfort ratings for two groups – over the age of 30 and 30 years of age or younger. A one-tailed t-test shows that the effect of age on discomfort ratings was significant [$t(25) = 2.05$, $p = 0.03$]. Re-examining the discomfort ratings reveals that the correlation between minimum contrast sensitivity and aerial refueling performance ($r = 0.8$, $p < 0.001$) is highly significant for participants 30 years of age and younger. For these younger participants, the correlation between discomfort ratings and RVS refueling performance is also high ($r = 0.89$, $p < 0.001$).

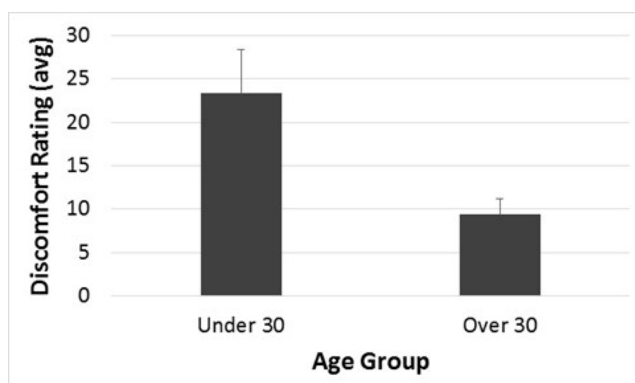


*Figure 17. Average discomfort ratings by age group.*

## Discussion

Several important conclusions can be drawn from the results of this research. The use of stereo and particularly hyper-stereo clearly improved simulated refueling performance. Although stereo/hyper-stereo was beneficial for most participants, performance either did not improve or actually worsened for a few participants in Experiment 1 with poorer ocular health. We investigated individual differences in vision and refueling performance more thoroughly in Experiment 2, which showed that performance with the use of a hyper-stereoscopic remote vision system is clearly dependent on ocular health, or what other authors have termed good binocular status. We showed that stereo acuity, fusion range, and/or weakest eye contrast sensitivity are all highly predictive of performance when using a hyper-stereoscopic display system to perform a simulated aerial refueling task. Individuals with poor scores on one or more of these vision tests not only tended to perform more poorly, but also tended to report higher levels of discomfort. However, it is important to note that most participants performed well and generally reported low levels of eyestrain/discomfort. Thus, it is not the case that a high proportion of individuals experienced difficulty using the hyper-stereoscopic RVS. Another interesting result was that presbyopic (older) participants reported very little eyestrain/discomfort and generally performed well on this simulated RVS task. This suggests that vergence-accommodation mismatch may be a source of discomfort.

These results provide evidence that carefully designed vision screening tests could be used to identify a minority of observers likely to have difficulty with stereoscopic display tasks such as the RVS aerial refueling task described here. As we note in a previous review paper, the type of vison tests selected for research involving stereoscopic displays and for vision screening is probably due for improvement [24]. Although not reported here, we found that most existing tests of stereo acuity, phoria, etc. were not predictive of either performance or discomfort on this simulated RVS refueling task.

Very few studies have been published regarding human performance with the use of hyper-stereoscopic displays. Although there are published studies examining eyestrain and discomfort with the use of stereoscopic displays, nearly all of these studies rely on existing clinical vision screening methods in an effort to predict which individuals may be susceptible to adverse symptoms when viewing these types of displays. While these tests may be appropriate for use in clinical settings, these tests may not have the sensitivity and reliability to accurately predict differences in performance and comfort with the use of newly developed electronic displays. Further, very few studies examine performance over long periods of time. Thus this study, examining hyper-stereoscopic displays, the comparison of standard clinical vision tests vs. newly designed computer-based vision tests, and lengthy viewing periods is a valuable contribution to this area of research. The results suggesting that a hyper-stereo RVS may improve refueling performance and that relatively few individuals experienced degraded performance and/or discomfort are also important since they may help address concerns that indirect view technology was not a suitable replacement for the direct view approach that has been in use for many years (e.g. KC-10, KC-135).

## Conclusion

Aerial refueling is a challenging and specialized task that requires extensive and costly training, and for which safety is obviously a significant concern. The results of this research suggest that appropriate vision screening can be applied to identify training candidates who are more likely to be able to tolerate the unique demands of using hyper-stereoscopic RVS technology for long periods of time. Thus, continued research on this topic could contribute to reduced training costs for aerial refueling operators and improved safety during refueling operations.

## References

[1] J. Gooch, “Stereopis standards for U.S. Air Force Boom Operators,” in *Proceedings of the Aerospace Medical Association Annual Meeting*, 2012.

[2] T. Konishi, “Stereo vision test for JASDF aviators,” *Presentation at the Aerospace Medical Association Annual Meeting*, 2012.

[3] B. Singh, “Boom enhanced vision system of KC-30 tanker aircraft,” *Presentation at the Aerospace Medical Association Annual Meeting*, 2012.

[4] T. Smart and B. Singh, “Stereo visual acuity standards for air refueling operators,” *Presentation at the Aerospace Medical Association Annual Meeting*, 2012.

[5] A. Smith, “Stereopsis of air refueling operators in the Royal Australian Air Force: the initial cohort,” *Presentation at the Aerospace Medical Association Annual Meeting*, 2012.

[6] M. Lambooij, M. Fortuin, W. a. Ijsselsteijn, B. J. Evans, and I. Heynderickx, “Susceptibility to visual discomfort of 3-D displays by visual performance measures,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 21, no. 12, pp. 1913–1923, 2011.

[7] F. L. Kooi and A. Toet, “Visual comfort of binocular and 3D displays,” *Displays*, vol. 25, no. 2–3, pp. 99–108, 2004.

[8] A. G. Solimini, “Are There Side Effects to Watching 3D Movies? A Prospective Crossover Observational Study on Visually Induced Motion Sickness,” *PLoS One*, vol. 8, no. 2, 2013.

[9] M. Lambooij, M. Fortuin, W. Ijsselsteijn, B. Evans, and I. Heynderickx, “Measuring visual fatigue and visual discomfort associated with 3-D displays,” *J. Soc. Inf. Disp.*, pp. 931–943, 2010.

[10] M. Emoto, T. Niida, F. Okano, and S. Member, “Repeated Vergence Adaptation Causes the Decline of Visual Functions in Watching Stereoscopic Television,” *J. Disp. Technol.*, vol. 1, no. 2, pp. 328–340, 2005.

[11] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, “Vergence-accommodation conflicts hinder visual performance and cause visual fatigue.,” *J. Vis.*, vol. 8, no. 3, pp. 33.1–30, 2008.

[12] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks, “The zone of comfort: Predicting visual discomfort with stereo displays.,” *J. Vis.*, vol. 11, no. 8, p. 11, 2011.

[13] J. P. Wann, S. Rushton, and M. Mon-Williams, “Natural

DOI: 10.2352/ISSN.2470-1173.2016.5.SDA-437

problems for stereoscopic depth perception in virtual environments," *Vision Res.*, vol. 35, no. 19, pp. 2731–2736, 1995.

[14] G. W. Stuart, S. a. Jennings, M. E. Kalich, C. E. Rash, T. H. Harding, and G. L. Craig, "Flight performance using a hyperstereo helmet-mounted display: adaptation to hyperstereopsis," *Proc. SPIE, Head- Helmet-Mounted Displays XIV Des. Appl.*, vol. 7326, 2009.

[15] P. Flanagan, G. W. Stuart, and P. Gibbs, "Hyperstereopsis in helmet-mounted NVDs: Time to contact estimation," in *Proc. of SPIE Vol. 6557, Head- and Helmet-Mounted Displays XII: Design and Applications*, 2007, vol. 6557, p. 65570O–65570O–7.

[16] P. Flanagan, G. W. Stuart, and P. Gibbs, "Hyperstereopsis in helmet-mounted NVDs: Absolute distance perception," in *Proc. of SPIE Vol. 6557, Head- and Helmet-Mounted Displays XII: Design and Applications*, 2007, vol. 6557.

[17] G. W. Stuart, P. Flanagan, and P. Gibbs, "Hyperstereopsis in helmet-mounted NVDs: slope perception," in *Proc. of SPIE, Head- and Helmet-Mounted Displays XII: Design and Applications*, 2007, vol. 6557.

[18] A. Woods, T. Docherty, and R. Koch, "Image Distortions in Stereoscopic Video Systems," *SPIE's Symp. Electron. Imaging Sci. Technol.*, vol. 36, no. February 1993, pp. 36–48, 1993.

[19] M. S. Banks, J. C. a Read, R. S. Allison, and S. J. Watt, "Stereoscopy and the Human Visual System," *SMPTE Motion Imaging J.*, pp. 24–43, 2012.

[20] L. A. Lesmes, Z.-L. Lu, J. Baek, and T. D. Albright, "Bayesian adaptive estimation of the contrast sensitivity function: the quick CSF method.," *J. Vis.*, vol. 10, no. 3, pp. 17.1–21, 2010.

[21] F. Kingdom and N. Prins, *Psychophysics: a practical introduction*. 2010.

[22] M. Bach, C. Schmitt, M. Kromeier, and G. Kommerell, "The Freiburg Stereoacuity Test: automatic measurement of stereo threshold.," *Graefes Arch. Clin. Exp. Ophthalmol.*, vol. 239, no. 8, pp. 562–566, 2001.

[23] M. a. Weissman and A. J. Woods, "A simple method for measuring crosstalk in stereoscopic displays," *Displays*, vol. 7863, pp. 786310–786310–11, 2011.

[24] M. Winterbottom, J. Gaska, S. Wright, C. Lloyd, F. Tey, and J. Mcintire, "Operational Based Vision Assessment Research: Depth Perception," *J. Australas. Soc. Aerosp. Med.*, vol. 9, pp. 33–41, 2014.

## Author Biography

*Dr. Marc Winterbottom is a senior research psychologist supporting the Operational Based Vision Assessment Laboratory. His most recent research focuses on U.S. Air Force vision standards and modernization of vision screening practices. He received a BA degree in Psychology from Purdue University (1996), an MS in Human Factors Psychology from Wright State University (2000), and recently completed a PhD in Human Factors Psychology at Wright State University.*

*Dr. Charles J. Lloyd is president of Visual Performance LLC, where he addresses research and development challenges relating to training display system design, requirements, metrics, and measurements. He has 28 years of experience in display systems and applied vision research at such organizations as Honeywell's Advanced Displays Group, The Lighting Research Center, BARCO Projection Systems, Flight Safety International, and the Air Force Research Laboratory. Charles has published more than 75 papers in this arena.*

*Dr. James P. Gaska received a PhD degree in Biopsychology from Syracuse University. From 1981 to 2000 he worked at UMASS Medical School where he researched the cortical visual processes of primates. From 1995 to 2005, Dr. Gaska worked as contractor to the Air Force supporting research on the design of camouflage, laser eye protection, and flight simulation technologies. He is currently Senior Scientist for OBVA at the USAF School of Aerospace Medicine at WPAFB.*

*Dr. Steve Wright is a clinical and research optometrist with the U.S. Air Force School of Aerospace Medicine, Ophthalmology branch, Wright-Patterson AFB, OH. He has worked in this capacity since 2004 and spent the prior nine years in private practice. He received a B.S. (1989) and O.D. degree (1995) from Pacific University, Forest Grove, OR and an M.S. (1990) from the University of Washington.*

*Dr. Steven C. Hadley is the Operational Based Vision Assessment Program Lead. He is a board certified Ophthalmologist, Operational Pilot-Physician, Aeromedical Specialist, and Special Operations Consultant. He has served as a chief flight surgeon and command USAF helicopter instructor pilot with over 4000 flight hours and is qualified in eight different military aircraft. He is a West Point graduate and completed his MD at Uniformed Services University of the Health Sciences (USUHS), Bethesda, MD.*

## Acknowledgments

The authors would like to thank Scott Humphrey and Kelly Bishop for their help with recruiting volunteers and many, many hours of data collection, and Alex Van Atta for supporting the software development involved with the computer-based vision tests. Thanks also to The Boeing Company and Pat Corbett who provided critical technical details; Mr. John Luu and Joe Rich at the KC-46 Program Office (AFLCMC) for their support of this research; Flight Safety International and Dan Myers for providing the Vital X image generator; and to Dr. Lu Lesmes and Adaptive Sensory Technology for providing the specially modified contrast sensitivity test used in this research.

The views expressed are those of the author and do not necessarily reflect the official policy or position of the Air Force, the Department of Defense, or the U.S. Government. This document has been approved for public release (Case # 2015-6232). This work was supported by U.S. Air Force contract FA8650-12-D-6280 to Wyle Laboratories and was funded by the 711 HPW/USAFSAM.