https://doi.org/10.2352/EI.2022.34.2.SDA-289
This is a work of the U.S. Government.

The Relationship between Vision and Simulated Remote Vision System Air Refueling Performance

***Eleanor O'Keefe*[1]*, Matthew Ankrom*[1]*, Eric S. Seemiller*[1]*, Tommy Bullock*[2]*, Marc Winterbottom*[2]*, Jonelle Knapp*[2]*, and Steve Hadley*[2]**
[1]*KBR; Beavercreek, OH, USA;* [2]*Airman Systems Directorate, OBVA Laboratory, Wright-Patterson AFB, OH, USA*

Abstract

United States Air Force (USAF) vision screening tests have remained largely unchanged since WWII and it is unclear whether current standards are applicable for users of new human-machine interfaces (e.g., stereoscopic remote vision system (RVS) in the KC-46 refueling tanker). This study examined the relationships between several vision tests, including a set of automated vision tests (AVT) developed by the Operational Based Vision Assessment (OBVA) Laboratory, an electronic version of the standard Titmus stereo test, and the current Armed Forces Vision Tester (AFVT), and simulated air refueling task performance using a stereoscopic RVS. Additionally, the relationships between vision test scores and subjective measures of visual fatigue were analyzed. Results showed that AVT measures of disparity discrimination and horizontal fusion correlated with simulated air refueling performance. AVT measures of acuity, contrast sensitivity, disparity discrimination, and radial motion sensitivity were significantly associated with subjective measures of discomfort and visual fatigue. Notably, neither the electronic Titmus results, nor the AFVT measures were associated with either air refueling task performance or subjective measures of visual fatigue. Adjustments to the vision standards and test methods used for USAF aeromedical vision screening should therefore be considered.

Introduction

One of the aims of the Operational Based Vision Assessment (OBVA) Laboratory is to modernize current vision tests and establish a quantitative relationship between vision and to operational performance, particularly for Airman and Guardians using modern digital displays and vision enhancement devices. A recent example of this issue is the incorporation of a stereoscopic remote vision system (RVS) in the KC-46 refueling tanker. Previously, as in the KC-10 and KC-135, in-flight refueling specialists, or boom operators, viewed receiver aircraft directly through a window in the rear of the tanker aircraft. However, in the KC-46, the out-the-window workstation has been relocated to the front of the aircraft and the boom operator now views refueling operations through the RVS, consisting of cameras and stereoscopic three-dimensional (S3D) display. This design evolution may require a change to aeromedical selection and certification requirements for boom operators.

Previous RVS research showed that several computer-based vision tests were predictive of simulated RVS refueling performance (stereo acuity, contrast sensitivity) while standard USAF tests generally were not [1-2]. This research also showed that measures of ocular alignment were correlated with simulated RVS refueling performance and reported discomfort. Additional research found a similar correlation between ocular alignment and performance on other S3D systems. For instance, a positive correlation between horizontal fusion range and performance on a simulated helicopter landing alignment task using a binocular head-mounted device (HMD) was shown [3]. Others have also found a connection between vision tests and 3D performance, including inferior performance on the Wilkin's rate of reading test (WRRT) in 3D by those with poor binocular vision status, as defined by 10 common clinical optometric tests [4], correlation between both near fusion range and stereoacuity scores with performance on an S3D object placement task [5], and worse performance on 3D simulated surgical tasks by surgeons with poorer stereoacuity [6]. Thus, there is substantial evidence that binocular function affects user performance when using a 3D display.

During S3D system use, some individuals are more likely to experience subjective discomfort that may be related to their visual characteristics or ocular health. Those characterized as having moderate/poor binocular vision status were more likely to report significant discomfort in a 3D condition of the WRRT [4]. Similarly, phoria and the zone of clear singular binocular vision predicted ratings of eye tiredness and eye strain when viewing an S3D display [7]. The association between an automated near phoria test and self-reported visual fatigue on a misaligned binocular HMD task also suggests that ocular health is important for the use of S3D and binocular HMD technologies [8]. Additionally, participants with slow fusional response and smaller fusion ranges have found to be more sensitive to visual fatigue [9]. However, particularly when examining the relationship between vision and self-reported discomfort, the results are often mixed. For example, previous research showed significant heterophoria changes after 20 minutes of S3D display viewing, but there was no correlation with self-reported visual comfort [10]. In another study, neither fusion range nor stereoacuity scores correlated with responses on the Simulator Sickness Questionnaire (SSQ) [5]. Others have found that those with superior vision experience more fatigue. Kooi and Toet proposed that observers with good stereo acuity and visual acuity were bothered more by image misalignments than those with poor vision [11], and Kim et al reported that observers with good stereopsis reported more discomfort (dizziness, headache, eye fatigue) after S3D use [12]. Finally, surgeons with poorer stereo acuity reported lesser symptoms on both the SSQ and lower headache/dizziness ratings after completing 3D surgical simulations [6]. Thus, there appears to be no consensus on whether specific visual traits predict subjective discomfort when using an S3D display.

Given the transition to RVS refueling and the critical need to ensure that boom operators are medically qualified to safely carry out the refueling mission, a comprehensive reassessment of the necessary visual traits to optimize performance is needed. Current stereoacuity requirements for aircrew and Special Warfare Airmen is 40 arcseconds (arcsec) of disparity on the Armed Forces Vision Tester (AFVT) stereo test [13]. However, a waiver can be obtained with a score of 120 arcsec or better on the AO Vectograph or a score of 30mm on the Howard-Dolman (HD) depth test. The HD test uses a method of adjustment depth estimating procedure, which may mask specific deficiencies in stereopsis [14]. In fact, Serrano-Pedraza et. al. [15] write "…in all current stereo tests, monocular artifacts are a potential issue to a greater or lesser extent." This limitation of existing standard stereoacuity tests has been known for many years [16]. Thus, individuals with poor stereoacuity (as much as 30% of the adult population [17]) could potentially pass current stereoacuity tests.

The purpose of this study was to assess the relationship between measures of spatial vision (visual acuity, contrast sensitivity), binocular function (stereoacuity, horizontal and vertical fusion range, horizontal and vertical phoria), and motion perception with air refueling task performance on a stereoscopic RVS simulation representative of the original KC-46 aerial refueling operator station delivered to the USAF in 2019. This study also assessed the relationship between visual function and subjective measures of discomfort and visual fatigue. This study used both computer-based tests (OBVA AVT battery) and standard vision test methods (AFVT and eTitmus), thereby allowing comparison of the performance of various tests relative to the study outcome measures. The results of this study are intended to inform the development of evidence-based vision standards for KC-46 boom operators.

Method

Participants

Participants consisted of 27 individuals with no USAF air refueling experience ("non-booms") and 13 current Air Force Boom Operators ("booms"). Their ages ranged from 21 to 61 years, with a median of 29.5 years, which is generally representative of USAF active-duty booms. Three of the participants were over the age of 50, which would be unusual for active-duty (though not for the USAF Guard and Reserve). Twenty-four participants (60%) were male, and 20 participants (50%) wore vision correction in the form of either glasses or contact lenses during experimentation.

Non-boom participants were recruited based on pre-existing AVT stereoacuity and horizontal fusion range scores with the goal of obtaining an equal number of participants with poor (scores worse than one standard deviation (SD) from the mean), average (scores within ±1 SD of the mean), and good scores (scores better than one standard deviation from the mean). Booms were recruited through coordination with the Air Force Life Cycle Management Center KC-46 System Program Office (SPO), which as a group tended to have better scores than non-boom participants. All participants provided written informed consent before participating. The study was conducted under a research protocol approved by Wright-Patterson Air Force Research Laboratory Institutional Review Board. Booms and other USAF employees volunteered their time, remaining participants were compensated in accordance with the IRB approved protocol.

Apparatus

This study employed a KC-46 RVS simulation developed by the OBVA Laboratory based on consultation with boom operator subject matter experts (SMEs) to replicate key design elements of the KC-46 AROS (Figure 1) and air refueling task. The workstation dimensions and viewing distances were designed to be similar to that found in the aircraft. The simulation used a FlightSafety Vital 1100 5-channel image generator, which provided 3D receiver aircraft models. Two Vital 1100 channels driving the 3D imagery used an NVidia Quadro P6000 video card, and NVidia Quadro Sync II cards synced framebuffers. The image generator for each channel fed the video signal into a Westar EZwindow Video Combiner box to spatially interlace the video into a single stereo image at 60 frames per second. The video output fed into a hardened-DVI (HDVI) converter to reformat the video signal into an HDVI signal, which was compatible with the KC-46 aircraft 3D display provided by Boeing and Collins Aerospace. The display resolution was 1920 x 1200 pixels and spatially interlaced with the micro-polarizers blocking every other horizontal line. The display was viewed from a distance of 0.89 m. The luminance presented to each eye through the 3D glasses was approximately 95 cd/m^2. Stereoscopic display crosstalk was low, less than 3% across most of the display measured at the design eye-point.

Three additional monitors, HP Pavilion 21.5-in, 1920 x 1080, In-Plane Switching (IPS) LED HDs, displayed a panoramic view of the airspace behind the simulated KC-46 aircraft. These monitors were configured such that only a portion of each display was viewed to simulate the lower resolution panoramic displays used in the KC-46. Two Saitek X52 flight controllers were incorporated to control movements of the boom and boom nozzle.

The virtual cameras were rotated inwards, or toed-in. This camera/display configuration produces several artifacts common to toed-in hyper-stereoscopic systems including significant vertical misalignment, or dipvergence (> 15 arcmin at the corners of the display), depth plane curvature, and depth plane compression [18, 13]. The virtual cameras were also spaced much farther apart than an observer's typical interpupillary distance, representative of the hyperstereoscopic design of the KC-46 RVS.

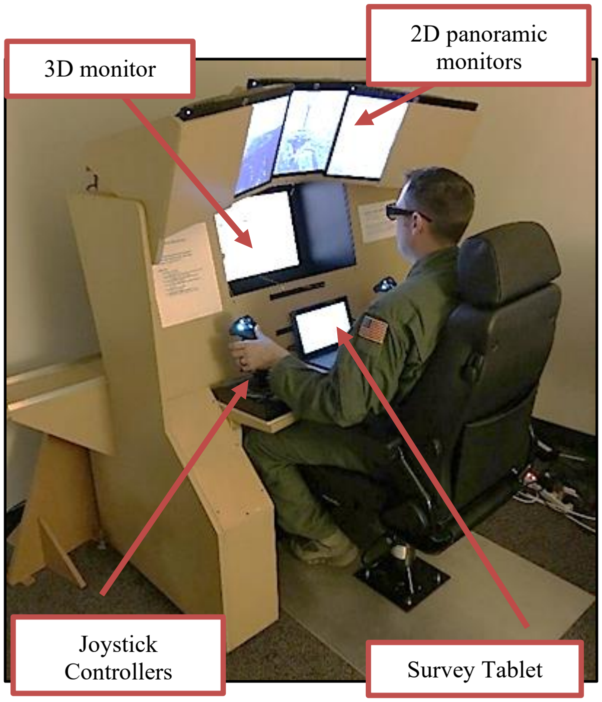

Figure 1. *Air Refueling Operator workstation simulation*

Experimental Tasks and Metrics

Air Refueling Task

During the simulated refueling task, each receiver aircraft started from the right side behind the refueling aircraft from the operator's perspective (from the port side of the simulated tanker aircraft). At this position, it was only visible in the far-right panoramic monitor. From this initial position, the receiver accelerated into pre-contact position then to the contact position. Once the receiver aircraft was in contact position (Figure 2), the participant used the two joysticks to maneuver the boom into the receiver aircraft receptacle. The right joystick controlled the azimuth and elevation of the boom, and the left joystick controlled the extension and retraction of the boom nozzle.

Figure 2. *RVS display with overlay*

Participants were instructed to attempt to make contact with the receiver aircraft as quickly as possible while avoiding hitting the receiver outside the receptacle. If the boom nozzle hit outside the receptacle, a red ‘X’ appeared on screen, the trial ended, and the receiver aircraft flew away. If instead the connection between the boom nozzle and the receptacle was successful, the participant monitored for the occurrence of any of three of the following events: (1) transfer of the planned quantity of fuel, which was determined by viewing a fuel information box in the lower left of the primary display (see Figure 2), (2) fuel spray, which was evident from a change in color of the receiver fuselage behind the receptacle, and (3) the receiver drifting outside the refueling envelope. Once the planned amount of fuel was reached, or if either fuel spray or an out-of-envelope movement occurred, the participant disconnected the boom nozzle from the receiver and the aircraft flew away. This task was repeated for seven minutes.

Six measures were collected from the tracking task: number of contacts per block (*Number of Contacts*), the time to make contact in seconds (*Time to Contact*), the number of collisions per block (*Number of Collisions*), the velocity of the boom nozzle when a contact occurred (*Boom Velocity*), the velocity of the boom nozzle when a collision occurred (*Collision Velocity*), and the absolute difference in the planned versus actual fuel delivered (*Fuel Error*). An additional derived task measure, *Air Refueling Performance*, was calculated by subtracting the *z*-score of *Number of Collisions* from the *z*-score of the *Number of Contacts*.

Vision Metrics

The OBVA Laboratory previously developed and validated a computer-based Automated Vision Test (AVT) battery that is documented elsewhere [20]. All tests used the psi adaptive thresholding procedure [19]. Included in the battery were the following measures: acuity (*AVT Acuity Test*), contrast sensitivity, stereo acuity, horizontal and vertical fusion (*AVT H Fusion Test* and *AVT V Fusion Test)*, and motion perception (*AVT Radial Test* and *AVT Rotational Test).* The acuity test and contrast sensitivity tests were administered at 4 meters using Landolt C stimuli. The contrast sensitivity test Landolt C’s had gap sizes of 6.67, 2.5, and 1.25 arcmin yielding the measures: *AVT CS Test 16.67*, *AVT CS Test 2.5*, and *AVT CS Test 1.25*. A derived measure, *AVT CS Test Cumulative*, was calculated by summing the three thresholds [21]. There were two stereoacuity tests in the battery: the dual ring and the stereo search test (*AVT SST*). The dual ring test was presented at 1 meter and 4 meter viewing distances, yielding the following measures: *AVT Dual Ring Test 1* and *AVT Dual Ring Test 4*.

Stereoacuity was also measured using the Optec 2300 Armed Forces Vision Tester, and the eTitmus test, an electronic version of the traditional Titmus booklet stereo test. On the Optec, a participant was presented with an image of a series of rings in rows and was asked to indicate which ring appeared closest in each row. The degree of stereoacuity was assessed using the provided scoring key and reported in arc seconds, yielding the single measure, *AFVT Depth Test*. The eTitmus test was taken on an ASUS 3D VG248 monitor with 1920 x 1080 pixels using the same ring stimuli and disparity levels as the Titmus stereo test, reconfigured for a viewing distance of 1 meter. The degree of stereoacuity was reported in log arcseconds, yielding the single measure, *eTitmus*.

Horizontal and vertical phorias were measured using the AFVT. To measure horizontal phoria, a participant was presented with dichoptic images: one eye saw a series of numbers and the other eye saw an arrow (Figure 5). The participant was asked to report the number the arrow appeared to point toward. The degree of phoria was assessed using the scoring key and reported in prism diopters (PD). To measure vertical phoria, a participant was presented with an image of a staircase and asked to identify which step was aligned with a dotted line. Again, the degree of phoria was assessed using the scoring key and reported in PD. The AFVT provided horizontal and vertical phoria measurements at both near (16 inches) and far (20 feet) distances, yielding a total of four measures: *AFVT Near V Phoria Test*, *AFVT Far V Phoria Test*, *AFVT Near H Phoria Test*, and *AFVT Far H Phoria Test*.

Subjective measures

Subjective measures were obtained using a 7-item questionnaire based on a similar questionnaire from Shibata et al. [3] and the NASA TLX [23] that addressed ocular issues associated with the use of stereoscopic displays. The questionnaire was administered using a tablet and each question was presented followed by an unnumbered visual analog scale divided into 20 increments with two anchors. A value of 0-100 in steps of 5 was obtained for each question based on the participant’s response using the visual analog scale. The output of the survey comprised seven measures: *Tired Eyes*, *Discomfort*, *Difficult Adjustment*, *Double Vision*, *Hard Work*, *Headache*, and *Eye Strain*.

Procedure

Participants completed four sessions on separate days. All AVT and AFVT tests took place on days 1-2, for approximately two hours each session. Non-boom participants completed a 1.5-hour training session to familiarize themselves with the experimental apparatus and task prior to day 3. On days 3 and 4, participants first completed the eTitmus test, the AVT horizontal and vertical fusion range tests, and the AVT SST. Experienced booms completed a shorter 20-minute practice session. All participants completed a sequence of eight 10-minute experimental blocks comprised of the air refueling task and survey. At the completion of the experimental blocks, participants repeated the AVT horizontal and vertical fusion range tests, and the AVT SST. The third and fourth day of testing took approximately 3-3.5 hours to complete for each subject.

Results

All AVT measures had an N of 40, except the AVT Radial Test (N = 36), as some participants were unable to successfully differentiate between expansive and contractive motion. All AFVT measures had an N of 40, except the AFVT Depth Test (N = 32). The test does not measure stereopsis above 40 arcsec, therefore if participants were unable to successfully discriminate that disparity, they received an ‘X’ as their score and removed from analysis. The eTitmus test was implemented after four participants had already completed the study, therefore this measure had an N of 36.

Table 1 provides Pearson product-moment correlation coefficients for associations between vision measures and air refueling task performance measures. Asterisks identify significant associations as hypothesized at $p < 0.01$. Significant associations were observed between the AVT Dual Ring Test 1 and 4, the AVT SST, the AVT H Fusion Test, and several air fueling task measures.

Table 1. Pearson Correlation Coefficients for associations between vision measures and air refueling task measures

	Air Refueling Task Measure						
Vision measure	**Number of contacts**	**Time to contact**	**Number of collisions**	**Boom velocity**	**Collision velocity**	**Fuel error**	**Air refueling performance**
AVT Dual Ring Test 1	-0.37	0.13	**0.47***	0.05	0.01	-0.14	**-0.45***
AVT Dual Ring Test 4	-0.39	0.21	**0.41***	-0.03	-0.06	-0.15	**-0.43***
AVT SST	**-0.42***	0.23	**0.48***	0.16	0.15	-0.16	**-0.48***
AVT H Fusion Test	**0.58***	-0.34	**-0.57***	0.03	-0.01	-0.02	**0.62***
AVT V Fusion Test	-0.24	0.18	0.19	-0.03	-0.02	-0.24	-0.23
AVT CS Test 1.25	-0.15	0.10	0.13	-0.08	-0.12	-0.07	-0.09
AVT CS Test 2.5	-0.07	0.10	-0.01	-0.15	-0.23	-0.14	-0.03
AVT CS Test 16.67	-0.07	0.09	-0.02	0.02	-0.08	-0.07	-0.03
AVT CS Test Cumulative	-0.11	0.1	0.06	-0.09	-0.16	-0.11	-0.09
AVT Acuity Test	-0.25	0.17	0.22	-0.10	-0.14	-0.05	-0.25
AVT Radial Test	0.05	0.03	-0.09	-0.11	-0.13	-0.34	0.07
AVT Rotational Test	-0.07	-0.09	0.18	0.05	0.05	-0.22	-0.14
AFVT Far V Phoria Test	-0.24	0.28	0.05	0.05	-0.02	-0.08	-0.15
AFVT Far H Phoria Test	-0.01	0.19	-0.15	0.23	0.23	0.12	0.07
AFVT Depth Test	-0.28	0.18	0.30	0.25	0.21	-0.32	-0.32
AFVT Near V Phoria Test	-0.17	0.14	0.10	-0.11	-0.19	0.13	-0.14
AFVT Near H Phoria Test	-0.18	0.24	0.09	0.02	-0.10	0.08	-0.14
eTitmus	-0.44	0.23	0.44	0.03	0.01	-0.08	-0.48

Table 2 provides Spearman’s rho correlation coefficients for the associations between vision measures and subjective measures. Asterisks identify significant associations as hypothesized at $p < 01$. Significant associations were observed between the AVT Dual Ring Test 1, the AVT Dual Ring Test 4, the AVT CS Test 1.25, the AVT CS Test Cumulative, and the AVT Acuity test and various subjective measures.

Table 2. Pearson Correlation Coefficients for associations between vision measures and subjective measures

	Subjective Measures						
Vision Measures	**Tired Eyes**	**Headache**	**Eye Strain**	**Discomfort**	**Difficult Adjust-ment**	**Double Vision**	**Hard Work**
AVT Dual Ring Test 1	-0.39	-0.32	-0.49*	-0.33	-0.40	-0.28	0.01
AVT Dual Ring Test 4	-0.47*	-0.31	-0.52*	-0.42*	-0.46*	-0.34	-0.07
AVT SST	-0.35	-0.14	-0.40	-0.25	-0.37	-0.20	0.02
AVT H Fusion Test	0.36	0.23	0.24	0.10	0.12	0.22	-0.18
AVT V Fusion Test	-0.16	0.00	-0.12	0.05	0.00	0.12	0.02
AVT CS Test 1.25	-0.58*	-0.25	-0.56*	-0.48*	-0.48*	-0.30	-0.51*
AVT CS Test 2.5	-0.31	-0.18	-0.33	-0.26	-0.27	-0.22	-0.35
AVT CS Test 16.67	-0.18	0.02	-0.17	-0.09	-0.06	-0.12	-0.04
AVT CS Test Cumulative	-0.47*	-0.20	-0.46*	-0.40	-0.39	-0.28	-0.43*
AVT Acuity Test	-0.45*	-0.34	-0.51*	-0.45*	-0.49*	-0.37	-0.41*
AVT Radial Test	-0.42	0.07	-0.28	-0.34	-0.40	-0.22	-0.15
AVT Rotational Test	-0.14	0.02	-0.20	-0.13	-0.17	-0.18	0.04
AFVT Far V Phoria Test	-0.10	0.20	0.00	0.12	0.12	0.27	0.03
AFVT Far H Phoria Test	0.23	0.15	0.21	0.21	0.22	0.17	0.09
AFVT Depth Test	-0.34	-0.13	-0.32	-0.23	-0.29	-0.13	-0.02
AFVT Near V Phoria Test	-0.10	0.02	-0.10	-0.06	-0.06	0.02	0.10
AFVT Near H Phoria Test	-0.08	0.12	-0.07	0.18	0.14	0.21	0.17
eTitmus	-0.18	-0.13	-0.15	-0.17	-0.23	-0.12	0.13

Paired *t*-tests showed no significant change in before and after scores for AVT SST or H Fusion Test, while AVT V Fusion Test scores significantly increased after the refueling task (Table 3), at a Cohen's *d* effect size of 0.35. One participant's scores on the AVT H Fusion Test could not be recovered.

Table 3. Statistics of vision measures taken before and after the refueling task

	Statistics						
Vision Measures	**Pre-Refueling Mean**	**Pre-Refueling SD**	**Post-Refueling Mean**	**Post-Refueling SD**	***t*-value**	***df***	***p*-value**
AVT SST	1.30	0.54	1.30	0.52	0.07	39	0.94
AVT H Fusion Test	2.33	0.51	2.35	0.52	-0.56	38	0.58
AVT V Fusion Test	1.33	0.14	1.35	0.14	-2.24	39	0.03*

Discussion

This study examined the relationship between a range of vision attributes, as measured by multiple vision tests (OBVA AVT, AFVT, and eTitmus), and both objective and subjective air refueling task performance using a stereoscopic RVS currently employed on the KC-46. AVT measures of stereoacuity and horizontal fusion range were significantly associated with air refueling task performance; as binocular function improved, task performance improved. AVT measures of visual acuity, contrast sensitivity, stereoacuity, and radial motion sensitivity were significantly associated with self-reported measures of discomfort and visual fatigue. Notably, the AFVT and eTitmus were not associated with either air refueling task performance or subjective measures. The eTitmus was designed to simulate the Titmus booklet test, and, as expected, resulted in noisier results, with significant floor and ceiling effects, which accounts for the lack of significance despite an r value similar to that of the other stereoacuity measures. Neither AVT horizontal fusion range nor SST scores changed following the refueling task, but there was a significant change in vertical fusion range, suggesting that visual fatigue may affect the control of vertical binocular alignment. It's possible that this effect on vertical fusion is driven by the significant amount of dipvergence inherent in the RVS 1.0 design.

While all three AVT stereoacuity tests and the AVT H Fusion Test were associated with overall air refueling performance, the AVT V Fusion Test and the two AVT motion tests were not, which was unexpected. However, there was not much variation in these tests in the sample population used. It may be worth noting that both motion tests used in the present study were somewhat different than those used in previous research where a correlation with performance was reported [1, 24]. The new radial motion test in particular is more difficult, and a significant proportion of subjects were unable to discriminate expanding from contracting motion even at high levels of motion coherence.

None of the AFVT vision tests were associated with any air refueling task measures. These results may be due to limitations of these tests described above, including monocular depth cues in the stereo test. In addition, those who could not detect the 40 arcsecond depth disparity (that is, is their stereoacuity threshold was larger than 40 arcseconds) on the AFVT stereoacuity test were simply given an 'X' as their score instead of a true stereoacuity threshold. The eight participants in this study who were given an 'X' were not included in the correlation analyses, limiting the variance used in a correlation calculation. It should be noted that for the USAF vision screening, patients are given another test, the AO Vectograph, with larger disparities if they fail the AFVT. This was not done for the present study. However, previous research failed to show a correlation between a combination of the AFVT and AO Vectograph and simulated air refueling performance [2].

The negative correlation between vision test scores and subjective measures, and lack of correlation with fusion range, was an unexpected result that contradicts previous research that showed that individuals with poorer, not better, quality of vision tend to report more discomfort with S3D displays [2, 4, 7, 9]. However, this was consistent with other research that has shown that individuals with good vision will be more bothered by image distortions than individuals with poor vision, potentially because they are simply better able to see those distortions [6, 11, 12].

Across several (unpublished) OBVA RVS studies, the subjective questionnaires have been found to be highly unreliable. During data collection, experimenters documented numerous occasions where subjects commented on significant headache and eyestrain, but then failed to provide questionnaire responses consistent with their comments. This may provide an alternative explanation to the differing results between the current study and previous studies examining the relationship between vision and self-reported physiological symptoms. In our future research with S3D displays, we will incorporate more objective, methods to document potential fatigue/discomfort (e.g., eye-tracking, heart rate variability, etc.).

Conclusion

The results of this study suggest that the AFVT currently used for aeromedical vision screening does not predict boom operator performance in a simulated RVS refueling task representative of the KC-46 RVS. Thus, the current Flying Class III vision standard, which applies to boom operators, will very likely pass operators who may experience reduced performance and/or significant discomfort (e.g., headache, eyestrain) when using the system. Additionally, the current standard is likely to disqualify operators who would perform well with the new RVS. The OBVA Laboratory AVT includes several vision measures that predict simulated refueling task performance when using a stereoscopic RVS. Adjustments to the vision test battery used to screen and medically qualify KC-46 Boom Operators should be considered to improve aeromedical decision making and to optimize human performance with the use of new vision enhancement technologies such as the KC-46 RVS.

Disclaimer and Acknowledgments

The views expressed are those of the authors and do not necessarily reflect the official policy or position of the Air Force, the Department of Defense, or the U.S. Government.

The authors would like to thank the Air Force Life Cycle Management Center (AFLCMC), Mr. Steve Nigus, Mr. Dan Myers and Flight Safety International for their support of this research.

References

[1] M. Winterbottom, "Individual differences in the use of remote vision stereoscopic displays," (Doctoral dissertation), Wright State University CORE Scholar, Theses and Dissertations, 2015. Retrieved from: https://corescholar.libraries.wright.edu/cgi/viewcontent.cgi?article=2449&context=etd_all.

[2] M. Winterbottom, C. Lloyd, J. Gaska, S. Wright & S. Hadley, "Stereoscopic remote vision system aerial refueling performance," Electronic Imaging Stereoscopic Displays and Applications XXVII, pp. 1-10, 2016.

[3] M. Winterbottom, C. Lloyd, J. Gaska, L. Williams, E. Shoda, & S. Hadley, "Investigating aircrew depth perception standards using a stereoscopic simulation environment," Electronic Imaging, Stereoscopic Displays and Applications XXVIII, pp. 29-40, 2017.

[4] M. Lambooij, M. Fortuin, W. A. IJsselsteijn, B. J. Evans, & I. Heynderickx, "Susceptibility to visual discomfort of 3-D displays by visual performance measures," IEEE Transactions on Circuits and Systems for Video Technology, vol. 21, no. 12, pp. 1913-1923, 2011.

[5] J. P. McIntire, S. T. Wright, L. K. Harrington, P. R. Havig, S. N. Watamaniuk, & E. L. Heft, "Optometric measurements predict performance but not comfort on a virtual object placement task with a stereoscopic three-dimensional display," Optical Engineering, vol. 53, no. 6, 2014.

[6] A. Alhusuny, M. Cook, A. Khalil, J. Treleaven, A. Hill, & V. Johnston, "Impact of accommodation, convergence and stereoacuity on perceived symptoms and surgical performance among surgeons," Surgical Endoscopy, vol. 35, pp. 6660-6670, 2021.

[7] T. Shibata, J. Kim, D. M. Hoffman, & M. S. Banks, "The zone of comfort: Predicting visual discomfort with stereo displays," Journal of Vision, vol. 1, no. 8, pp. 1-29, 2011.

[8] L. Williams, E. Palmer, J. Gaska, M. Winterbottom, E. O'Keefe, E. Shoda & Hadley, S, "Ocular Fatigue in Biocular Helmet-Mounted Displays," in the Aerospace Medical Association Annual Meeting, Las Vegas, Nevada, 2019.

[9] M. Emoto, Y. Nojiri, & F. Okano, "Changes in fusional vergence limit and its hysteresis after viewing stereoscopic TV," Displays, vol. 25, pp. 67-76, 2004.

[10] E. Karpicka & P. A. Howarth, "Heterophoria adaptation during the viewing of 3D stereoscopic stimuli," Ophthalmic and Physiological Optics, vol. 33, no. 5, pp. 604-610, 2013.

[11] F. L. Kooi & A. Toet, "Visual comfort of binocular and 3D displays," Displays, vol. 25, pp. 99-180, 2004.

[12] S. H. Kim, Y. W. Suh, C. Yun, E. J. Yoo, J. H. Yeom & Y. A.Cho, "Influence of stereopsis and abnormal binocular vision on ocular and systemic discomfort while watching 3D television," Eye, vol. 27, no. 11, pp. 1243-1248, 2013.

[13] United States Air Force Medical Standards Directory. Retrieved from: AF Special Warfare, afspecialwarfare.com, 19 Mar 2021.

[14] W. L. Larson, "Does the Howard-Dolman really measure stereoacuity?" American journal of optometry and physiological optics, vol. 62, no. 11, pp. 763-767, 1985.

[15] I. Serrano-Pedraza, K. Vancleef, & J. Read, "Avoiding monocular artifacts in clinical stereotests presented on column-interleaved digital stereoscopic displays," JOV, vol. 16, no. 14, pp. 1-14, 2016.

[16] J. Cooper, & J. Warshowsky, "Lateral displacement as a response cue in the Titmus Stereo test," American journal of optometry and physiological optics, vol. 54, no.8, pp. 537-541, 1977.

[17] R. F. Hess, L. To, J. Zhou, G. Wang, & J. R. Cooperstock, "Stereo vision: the haves and have-nots," i-Perception, vol. 6, no. 3, 2015.

[18] D. B. Diner & M. von Sydow, "Stereo Depth Distortions in Teleoperation," JPL Publication, vol. 87, no. 1, 1988.

[19] N. Prins, "Applying the model-comparison approach to test specific research hypotheses in psychophysical research using the Palamedes Toolbox;" Frontiers in psychology, vol. 9, pp. 1-18, 2018.

[20] E. A. Shoda, M. Winterbottom, A. van Atta, & J. Gaska, J., "Operational based vision assessment automated vision test collection user guide," U.S. Air Force School of Aerospace Medicine, Aeromedical Research Department, Wright-Patterson AFB, Defense Technical Information Center. Retrieved from http://www.dtic.mil/docs/citations/AD1034796, 2017.

[21] E. S. Seemiller, J. Gaska, M. Winterbottom, E. Palmer, E. O'Keefe, & E. Shoda, "Dimension reduction reveals underlying factors of performance within a vision screening battery," Investigative Ophthalmology and Vision Science, vol. 62, no. 8, 2021.

[22] K. H. Britten, M. N. Shadlen, W. T. Newsome, & J. A. Movshon, "Responses of neurons in macaque MT to stochastic motion signals," Visual neuroscience, vol. 10, no. 6, pp. 1157-1169, 1993.

[23] S. G. Hart, "NASA-task load index (NASA-TLX); 20 years later;" Proceedings of the human factors and ergonomics society annual meeting, vol. 50, no. 9, pp. 904-908, 2006.

[24] J. P. Gaska, M. Winterbottom, R. Gray, "Operational Based Vision Assessment: Visual Motion Sensitivity and Operational performance," Presentation at the Aerospace Medicine Association Annual Meeting, 2016.

Author Biography

Eleanor O'Keefe received her BS in biopsychology from the University of California, Santa Barbara (2010) and her PhD in experimental psychology from the University of Louisville (2017). Since then she has worked in the Operational Based Vision Assessment Lab at Wright-Patterson Air Force Base. Her work has focused on remote vision system image space and stereoscopic 3D performance.