https://doi.org/10.2352/ISSN.2470-1173.2017.5.SD&A-354
This is a work of the U.S. Government.

Investigating Aircrew Depth Perception Standards Using a Stereoscopic Simulation Environment

Marc Winterbottom[1]*, Charles Lloyd[2], James Gaska[1], Logan Williams[1], Elizabeth Shoda[3], and Steven Hadley[1]

[1]Operational Based Vision Assessment (OBVA) 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]KBRWyle, Dayton, OH

Abstract

Due to concern that current U.S. Air Force depth perception standards and test procedures may not be adequate for accurately identifying aircrew medically fit to perform critical depth perception tasks during flight, the U.S. Air Force School of Aerospace Medicine developed a stereoscopic simulation environment to investigate depth perception vision standards. The initial results of this research showed that while the use of stereoscopic displays clearly improved performance for a helicopter landing task involving depth judgments, an individual's stereo acuity was not predictive of performance. However, landing task performance could be predicted when stereo acuity was used together with binocular fusion range. However, motion perception was a better predictor of performance than stereo acuity. Potential implications for medical vision standards and the potential complexities involved in predicting real-world performance based on performance in a stereoscopic flight simulation are discussed.

Introduction

A depth perception standard has been enforced for aviators since the early years of aviation. For example, Wilmer and Berens noted that "the value of stereoscopic vision . . . is of great value in judging distance and landing . . . The importance of this qualification seems to grow greater as our experience increases" [1]. Howard developed one of the first tests of depth perception for screening purposes and, on the basis of his research, believed that "to possess normal judgment of distance one's binocular parallactic angle should not be greater than 8.0 arcsec [2]. However, the debate concerning the utility of depth perception has also been ongoing since the early 1900s. Howard, in 1919, noted "some examiners have questioned the absolute necessity of binocular single vision as a preliminary requirement" [2]. Although a 1996 Delta MD-88 crash at LaGuardia was partly attributed to defective stereopsis, some researchers have concluded that stereopsis is not required for flight safety, owing to the fact that other cues to depth are sufficient [3], [4]. According to U.S. Air Force (USAF) medical policy, good stereo acuity and ocular alignment are both considered to be critical for pilots and also for non-pilot aircrew (Flying Class III, or FCIII, aircrew) involved in certain tasks such as clearing aircraft for landing [5]. An FCIII depth perception standard has been enforced for USAF aircrew since 1998, following a fatal accident involving two H-60 helicopters where defective stereopsis was identified as a contributing factor [6]. However, a similar standard is not maintained for Army personnel in similar aircrew positions, and many other countries do not maintain a depth perception standard, even for pilots.

Research examining the importance of either stereo acuity or the use of stereo displays has generated mixed results. In a systematic review of 71 experiments, previous researchers found that although about 67% showed a benefit of three-dimensional (3D) displays, the remaining 33% either did not show a benefit or had mixed results [7]. Similarly, a review of the importance of depth perception in aviation showed that not only is it difficult to clearly identify the importance of good stereo acuity, traditional methods used to measure stereo acuity may be lacking, which likely contributes to confusion concerning the utility of stereopsis and stereo displays [8]. In simulation and training applications, the use of stereo displays has been very limited. This may be due to several factors. Conventional knowledge has held that stereo is not useful beyond a few meters. Previous studies using electronic displays [9], [10] found stereo acuity thresholds of ~140 arcsec (i.e., many times higher than reported for real objects). Thus, previous experience with inadequate displays may have led to the conclusion that stereo cues would be ineffective for larger distances. Previous attempts to incorporate stereo displays into training systems proved difficult to implement [11], and two previous efforts to demonstrate the effectiveness of stereo displays for boom operator training in the USAF were cancelled. Difficulties with the use of stereoscopic displays are well known and may be attributable to a number of different factors such as vergence-accommodation mismatch, image distortion/misalignment between the left and right eye images, use of differing filters in the left/right eye (e.g., red/green filtering), conflicting depth cues (e.g., blur vs. disparity, lack of appropriate motion parallax), etc. [12]–[18].

As noted briefly above, a major limitation for many studies examining the utility of depth perception for performance of real-world tasks is that the measures of depth perception are often coarse and suffer from significant floor effects. If stereo acuity or other clinical metrics relevant to binocular health are actually obtained, they are often limited to, for example, a 40- or 60-arcsec minimum threshold, or simply "fly positive," meaning that subjects could see the 3D fly on a commonly available near stereo acuity test. Thus, part of the confusion concerning the utility of stereopsis may stem from the use of limited measures of binocular health. Although the potential limitations of some commonly used stereo acuity tests have been discussed [19]–[21], these tests are still frequently used. Our own research suggests that a more carefully designed computer-based stereo acuity test, although correlated with the USAF standard Armed Forces Vision Tester (AFVT) and AO Vectograph stereo acuity tests, differs substantially in outcome (Figure 1). As shown, there is a substantial floor effect on the standard test, and further, individuals obtaining the best score of 15 arcsec on the standard test may score anywhere from approximately 5 arcsec to 250 arcsec on the adaptive, threshold-based test. These results are consistent with

---

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

Cleared, 88PA, Case # 2016-6379, 8 Dec 2016.

previous research that suggests that the standard stereo acuity tests may actually test something other than stereo acuity. For this reason, we used our computer-based stereo acuity test in the research presented here rather than rely only on the more commonly available chart-based methods.

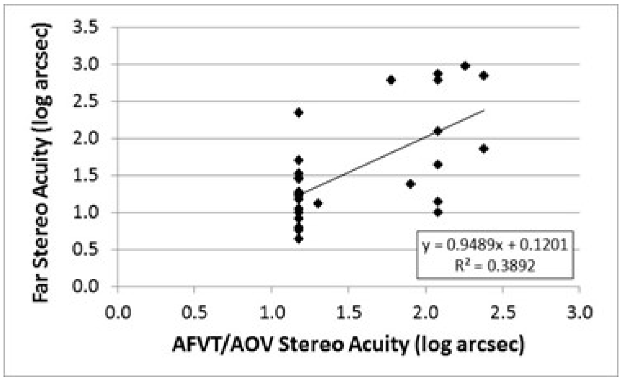

Figure 1. Relationship between chart-based AFVT/AO Vectograph stereo acuity tests and OBVA lab, computer-based adaptive stereo acuity test.

The objectives of the research presented here were to 1) examine the effect of stereo displays on the performance of a helicopter landing task, and 2) examine whether current USAF FCIII depth perception requirements are in fact relevant to job performance. Although the USAF depth perception standard is fairly strict (stereo acuity of 25 arcsec or better), the USAF waiver policy is much less stringent (see Table 1). This creates a situation where although many aircrew fail the standard, they are then placed on a special waiver status. In waiver status, aircrew are approved for flight, but their performance must be tracked throughout their career, and they must be re-tested on an annual basis. Thus, the current system is costly to enforce.

Table 1. U.S. Air Force, Navy, and Army Depth Perception Standards.

		USAF (Air Force Waiver Guide)		USN (NAVMED P-117)	USA (Army Regulation 40-501)
		FCI/IA/II/III (Scanners)	FCIII	Class I, Class II (except Fixed Wing Aircrew), Class III (including UAV Operators, Critical Flight Deck Personnel)	Class 1/2/2F/3/4
Stereo-acuity (Arc Second)		≤25	N/A	VTA-DP or OVT (≤25) or Randot or Titmus (≤40) or Verhoeff: 8/8 on 1st trial; or 16/16 on the 2nd and 3rd trials	≤40
Phoria	Eso	< 10 PD	< 15 PD	< 6	<8
	Exo	< 6	< 8	< 6	<8
	Hyper	< 1.5	< 2	< 1.5	<1
Tropia		0	< 15	0	0

VTA-DP: Vision Test Apparatus-depth perception
OVT: the Optec Vision Tester

For this research, a representative helicopter landing task was selected in which subjects were required to discriminate the distance between the rear wheel of the aircraft and the top of an object over which the aircraft hovered. To initiate this line of research, we have broken down a very complex helicopter call-to-landing task into sub-components, beginning with the hover task presented here. Future research is intended to examine time to contact and height estimation prior to researching performance in a full combat landing simulation and in collaboration with researchers in Canada and Australia. The simulation was constructed using X-Plane® software running on a pair of Windows PCs and viewed using a head-mounted display (HMD). A relatively unique aspect of this research is that each observer's stereo acuity, fusion range, and motion sensitivity were thoroughly evaluated using computer-based vision tests developed in our laboratory prior to participation in the simulated helicopter landing task.

It is important to note that depth perception involves much more than binocular disparity. A wide variety of monocular cues, such as optic flow [22], [23], motion parallax [15], [16], [24], relative size, and occlusion [25], all contribute to depth perception. In this research, the use of a head-tracked, wide field-of-view (FOV) HMD and highly detailed simulated environment preserved many of the cues to depth that would normally be encountered in a natural environment. Thus, the research described here should be relevant for examining the contribution of stereo displays and quality of vision to the performance of a highly complex task such as a helicopter call-to-landing.

Four experiments were devised to accomplish different objectives. Experiment 1 was designed to ensure that observers with good stereo acuity could indeed make fine discriminations in depth using the apparatus developed for this research. A significant effort was made to ensure that binocular disparities displayed in the HMD were accurate. Experiment 2 was designed to examine the effect of viewing condition (stereo or no-stereo) and the effect of height above the obstacle over which the simulated helicopter hovered. Experiments 3 and 4 were designed to examine individual differences in performance, effect of visibility (low vs. high contrast), and whether stereo acuity, fusion range, or sensitivity to motion affected performance on this task related to clearing rotary wing aircraft for landing.

Experiment 1

Comparison to OBVA Stereo Acuity Test

Experiment 1 was designed to compare stereo acuity thresholds obtained using the HMD in a simulated environment to those obtained using the OBVA stereo acuity test.

Methods

Subjects

Three subjects with good stereo acuity participated in this experiment. All subjects provided informed consent and the experimental protocol was approved by the Wright-Patterson Institutional Review Board (IRB).

Apparatus

The OBVA stereo test isolates stereo cues by employing constant-size concentric ring stimuli on a flat-panel electronic display (ASUS VG278). The OBVA stereo test is similar to the Freiburg stereo acuity test [19], in that it uses antialiasing when displaying the concentric rings to obtain sub-pixel shifts in disparity. The concentric ring stimuli are similar to the well-known Titmus/Randot graded circle test. The OBVA stereo test, with concentric ring stimulus, is illustrated in Figure 2.

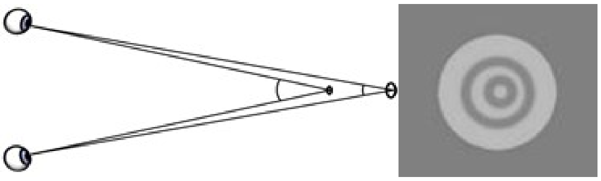

Figure 2. Geometry (left) of the OBVA stereo test battery with concentric ring stimuli (right).

The OBVA stereo test was modified to present the concentric ring stimuli on a binocular SA Photonics SA-55 HMD shown in Figure 3. This HMD uses two 1920x1200 opaque organic light-emitting diode displays with 100% overlap, forming a 55° horizontal FOV. A 4-meter virtual image distance was used. This HMD was selected for this research in part because it is one of the highest resolution, largest FOV HMDs currently available. Previous work [26] suggested that a relatively high resolution was needed to adequately display stereo imagery.

Figure 3. SA-55 HMD configured with 3D printed infrared (IR) reflective rigid body. Photo provided by SA Photonics, used with permission.

The native optics of the SA-55 HMD produce significant pincushion distortion; therefore, the Brown-Conrady model for radially symmetric image warping [27] was implemented to minimize optical distortion while adding negligible latency. The model parameters were determined empirically by subjective evaluation of the final image geometry. This correction reduced the active horizontal FOV to 44° due to the loss/deactivation of pixels near the image borders, as shown in Figure 4. However, the modified FOV is still within the 40° FOV typical of night vision goggles used routinely by USAF aircrew.

For both the OBVA stereo test and the HMD stereo test, the Ψ (psi) method [28] was adopted to estimate a psychometric function using a simple two-alternative forced-choice experiment for stereo acuity. For this test the observer is simply asked to repeatedly discriminate whether the inner circle is in front of, or behind, the larger outer circle. In previous research, we showed that this stereo acuity test was correlated with performance on a simulated stereoscopic remote vision system aerial refueling task [29], while the standard AFVT stereo test was not.

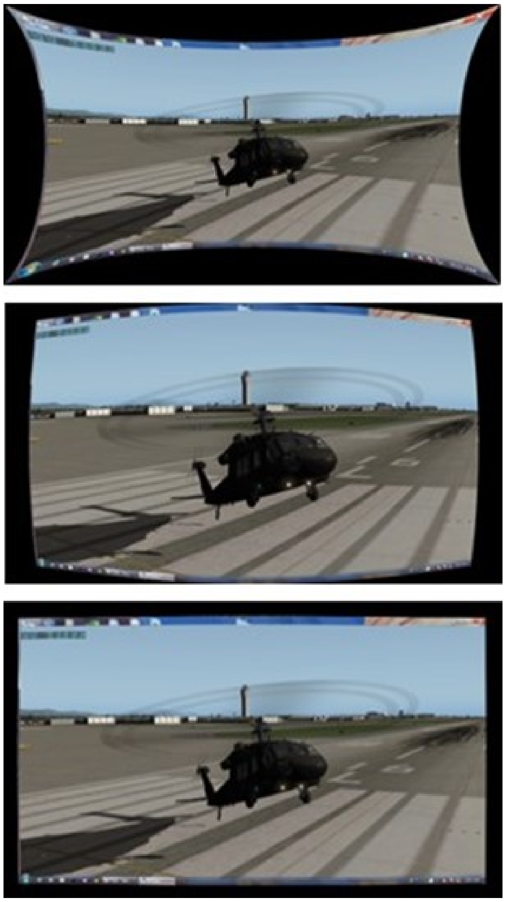

Figure 4. Simulation of SA-55 HMD image distortion and correction: native distortion (top), Brown-Conrady correction (center), and corrected image (bottom).

Results

Table2 below summarizes the stereo acuity results for each display and each of the three observers.

Table 2. Comparison of OBVA Stereo Test on Flat Panel and HMD for Three Observers.

Observer	1	2	3
Subject IPD	66 mm	65 mm	63 mm
Stereo acuity (flat panel)	3.7″±1.7″	6.0″±1.4″	3.0″±1.2″
Stereo acuity (HMD)	5.0″±1.6″	5.1″±1.7″	6.2″±1.4″

Cleared, 88PA, Case # 2016-6379, 8 Dec 2016.

Depth Judgments of Objects in a Virtual Environment

There are several differences between the measurement of stereo acuity using the OBVA stereo acuity test and similar measurements performed in a rendered scene. First, the OBVA stereo acuity test presents stimuli of a constant size, regardless of depth, in an effort to isolate only the stereo cue. Second, because the scene is not rendered in perspective, there is no need to account for an individual observer's interpupillary distance (IPD). Third, as noted above, the OBVA stereo acuity test makes specific use of antialiasing algorithms to achieve sub-pixel shifts in the concentric rings to obtain very small changes in disparity. However, in the rendered scene the level of antialiasing may be set for primarily aesthetic or computational performance reasons, without regard to specific stereo cues, which may affect the achievable stereo acuity limit in a realistically rendered scene. Therefore, it becomes necessary to test the stereo acuity limit of the complete apparatus in the as-built configuration, including the intended virtual environment and rendering settings.

Methods

Subjects

Three subjects with good stereo acuity participated in this experiment. Stereo acuity was tested using the OBVA stereo acuity test described above.

Apparatus

The SA Photonics HMD and X-Plane rendering system described above were used for this evaluation.

Procedure

To quantify the stereo threshold of the display in a realistic simulated environment, experimental stimuli were created that could be placed in a virtual environment to simulate the additional size and geometric perspective cues, which will always be present in a realistic scene. The experimental stimuli consisted of two static cross-shaped objects placed d =10 m from the observer (within the virtual environment), with one object slightly closer than the other, as shown in Figure 5. The target objects were separated by 3° (center to center).

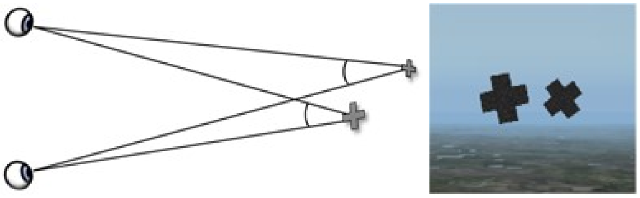

Figure 5. Geometry of the targets used to determine stereo thresholds (left). Cross-shaped test targets in X-Plane environment (right). Note that the relative sizes reveal which object is closer in the absence of stereo cues.

Each cross was randomly rotated after each trial. The Ψ method was again implemented with 30 trials per block, with at least 10 blocks per antialiasing condition, across 5 available antialiasing conditions, under both stereo and non-stereo rendering conditions, and with identical Brown-Conrady image warping implemented across all trials (to correct native HMD pincushion distortion). Stereo images were rendered using each observer's measured IPD, whereas non-stereo images were rendered with IPD = 0, which produces identical centered images for each eye.

Results

The results of this evaluation are summarized in both Table 3 and Figure 6.

Table 3. Comparison of Stereo Acuity Test Results and Depth Discrimination Thresholds for the Cross-Shaped Targets When Viewed with Stereo (in arcsec).

Observer	**1**	**2**	**3**
IPD (mm)	66	65	63
Stereo Acuity	3.7	6.0	3.0
Depth Threshold at 0x	8.0	14.3	8.7
Depth Threshold at 2x	7.6	12.2	9.8
Depth Threshold at 4x	7.3	11.2	10.4
Depth Threshold at 8x	6.8	11.8	7.6
Depth Threshold at 16x	6.2	10.3	8.9

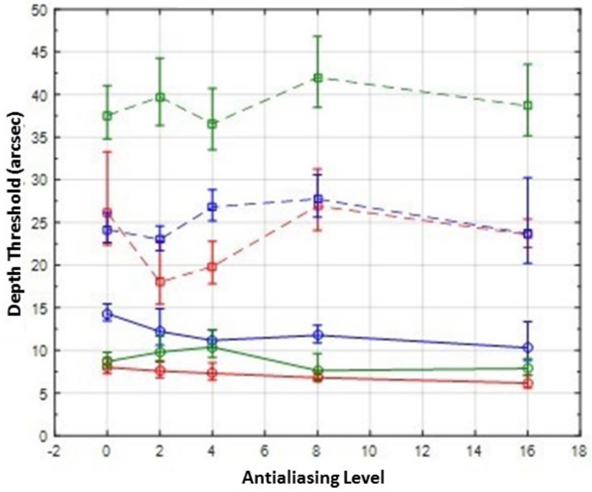

Figure 6. Stereo/depth thresholds of three observers at various antialiasing multiples, with (solid) and without (dashed) stereo enabled.

Discussion

The results of Experiment 1 provide evidence that the simulation environment developed for this research was capable of providing eye-limited binocular disparity. This experiment also provided preliminary evidence that manipulating the stereo viewing conditions affected the performance of subjects with good stereo acuity.

Experiment 2

Methods

Subjects

Five subjects with stereo acuity better than 30 arcsec participated in this experiment. All subjects provided informed consent and the experimental protocol was approved by the Wright-Patterson IRB.

Apparatus

Similar to Experiment 1, the 3D virtual environment was generated using two instances of Laminar Research's X-Plane to stereoscopically render a head-tracked, out-the-window visualization. The two instances of X Plane were implemented using two separate PCs, each incorporating Intel i7 processors and Nvidia Quadro K4200 video cards with Quadro Sync to ensure 60-Hz rendering synchronization between each channel of the stereoscopic visualization. The SA Photonics SA-55 binocular HMD was again used to display the virtual environment. However, head-tracking was also implemented to update the imagery based on head position, thus simulating motion parallax. Head-tracking was performed using a NaturalPoint OptiTrack (Motive:Tracker) IR tracking system, utilizing seven Flex-13 cameras, with tracking latency of approximately 0.4 ms. A custom 3D printed rigid body constellation, containing five IR reflectors, was mounted to the HMD for integration with the OptiTrack system as shown in Figure 3. The structure built by OBVA lab personnel to configure the head-tracking cameras is shown in Figure 7.

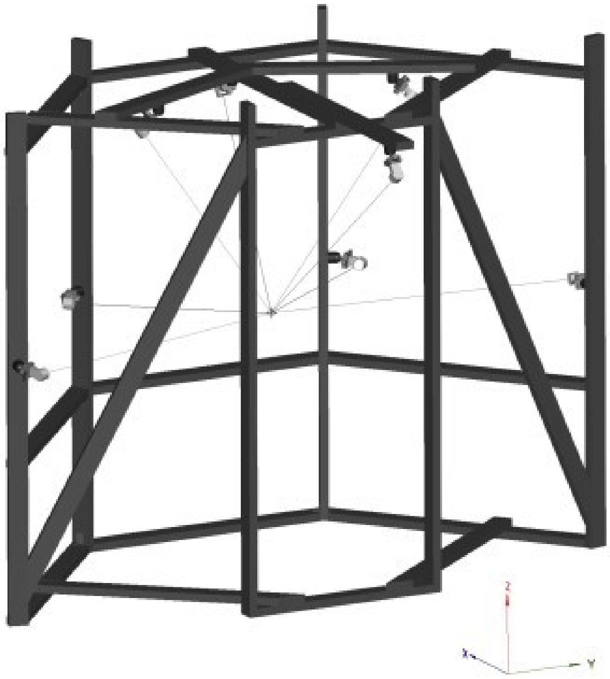

Figure 7. Structure and head-tracking cameras constructed by OBVA lab personnel.

Noticeable jitter did require the use of significant smoothing, which introduced a noticeable delay with rapid head movements. However, this artifact is likely inconsequential, since subjects tended to remain relatively still during the distance estimation task. The simulation control host was written in Matlab (MathWorks, 2014) by the OBVA team and operated from a third, separate PC. Communication between the Matlab host and the X-Plane rendering machines was implemented using UDP multicast packets containing the relevant control parameters (e.g., aircraft position, etc.). The apparatus has also been described in greater detail in a previous publication [30].

Procedure

Subjects were first fitted with the HMD, ensuring that the binocular imagery was properly aligned. Subjects were instructed to indicate in which trial the tail wheel of the H-60 was most closely aligned vertically with the windsock pole. At the beginning of each trial, the aircraft descended, or hovered, into place from a higher position and then remained stationary for the remainder of the trial. In different blocks of trials, the aircraft height above the tail wheel was varied from 0.1 m to 2 m and stereo imagery was turned on or off. A small degree of random variation was applied to each flight path (between trials) to both enhance realism and eliminate frequent repetition of rendering artifacts that might be used as unintended cues by the observer during the call-to-landing experiment. For the purposes of this experiment, the subjects' viewing position (the eye-point) was placed such that they could easily view the tail wheel and windsock, as shown in Figure 8, while remaining comfortably seated within the eye-tracking structure shown in Figure 5. For each viewing condition, the Ψ method [28] was used to estimate the smallest displacement in depth relative to the windsock that each subject could reliably detect.

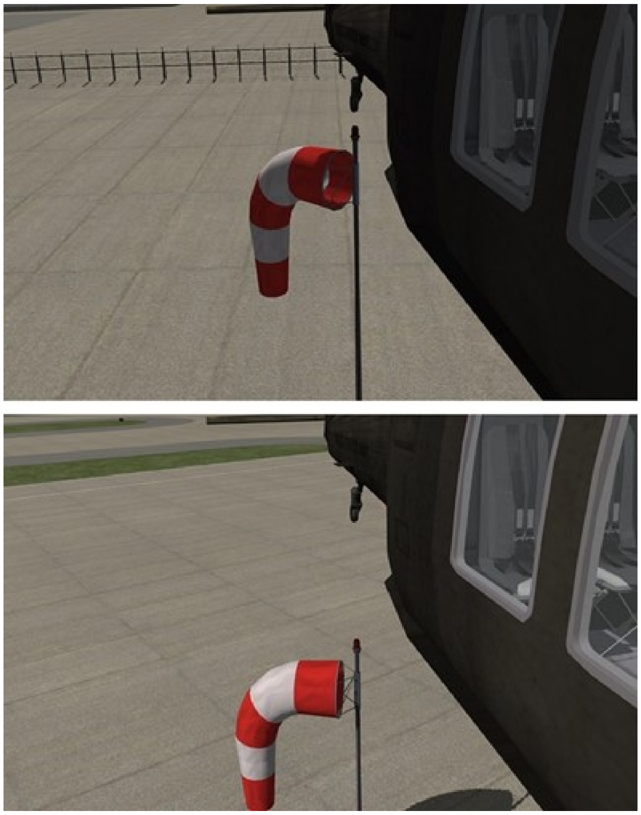

Figure 8. Subject's view of the tail of the aircraft and the rear wheel at two different heights (0.2 and 2.0 m). Image captured by OBVA lab personnel.

Results

Figure 9 shows the results of Experiment 2. As shown, smaller thresholds are obtained under stereoscopic viewing conditions in comparison to monoscopic viewing conditions ($t = -6.21$, $p \ll 0.001$; averaged over tail wheel height). However, this difference diminishes as the separation between the tail wheel and pole is increased.

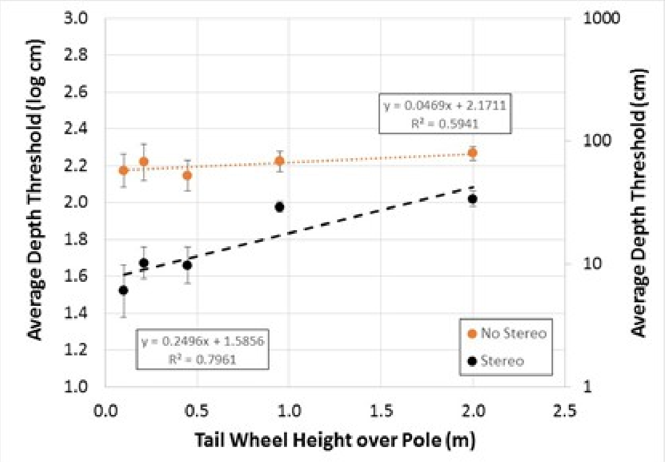

Figure 9. Experiment 2 results. Threshold depth judgments for each aircraft height and for each viewing condition (stereoscopic vs. monoscopic).

Discussion

The results of Experiment 2 provide evidence that stereoscopic viewing improved performance on this depth judgment task relevant to helicopter landing for a small group of subjects with good stereo acuity. However, the results also suggest that, even for subjects with good stereo acuity, the advantage gained by stereo viewing diminished with increasing separation between the two objects, in this case, the helicopter tail wheel and the pole.

Experiment 3

Methods

Subjects

Forty subjects volunteered to participate in Experiment 3. All subjects provided informed consent and the experimental protocol was approved by the Wright-Patterson IRB.

Apparatus

The apparatus for the helicopter landing simulation used in Experiment 3 was the same as that described for Experiment 2. A Dell Precision T7610 with Nvidia GeForce GTX 680 graphics card was used to administer the stereo acuity, fusion range, and motion tests. The tests were displayed on an Asus VG278HE 3D monitor (1920x1080 pixels) using active shutter glasses that was compatible with Nvidia 3D Vision2. At a 1-m viewing distance, the angular pixel size was 1.1 arcmin.

Procedure

All subjects were first administered the stereo acuity, fusion range, and motion perception tests developed by the OBVA laboratory. Figures 10 and 11 show examples of the test stimuli. For the stereo acuity tests, subjects indicated whether the smaller inner circle appeared popped out or receded in depth relative to the larger reference circle using two buttons on a game controller (Figure 10, top). The Ψ method was used to estimate threshold stereo acuity. The vertical fusion range test required that participants indicate when a circle viewed at a distance of 1 m on the Asus stereo monitor resulted in double vision (i.e., binocular fusion was broken) using the game controller as the circles (displayed separately to the left and right eyes) moved apart in the vertical direction. The direction of motion then reversed, and the participant next indicated when the circles returned to a single "fused" image using the game controller. This task was repeated several times. The amount of separation between the left and right eye images was recorded at the time the subject pressed the button on the game controller for each trial. A similar procedure was used for horizontal fusion range, except that the two circles moved apart in the horizontal direction, requiring subjects to either cross their eyes or uncross their eyes to maintain binocular fusion (Figure 10, bottom).

The motion sensitivity task is designed to estimate sensitivity to optic flow and consisted of a field of black and white dots (Figure 11) that could move in one of four directions: clockwise, counter-clockwise, expanding, or contracting. The Ψ method was used to estimate the minimum coherence (the proportion of dots moving in the same direction) required to detect motion for each of the four directions of motion. The stimuli and procedure were similar to that described in previous research [31], except that a 4AFC task was used instead of 2AFC. Subjects used the arrow keys (right, left, up, down) to indicate the direction of motion.

For the helicopter hover task, the procedure was similar to that described in Experiment 2, except that only two heights above the windsock (0.2 and 2 m) were used.

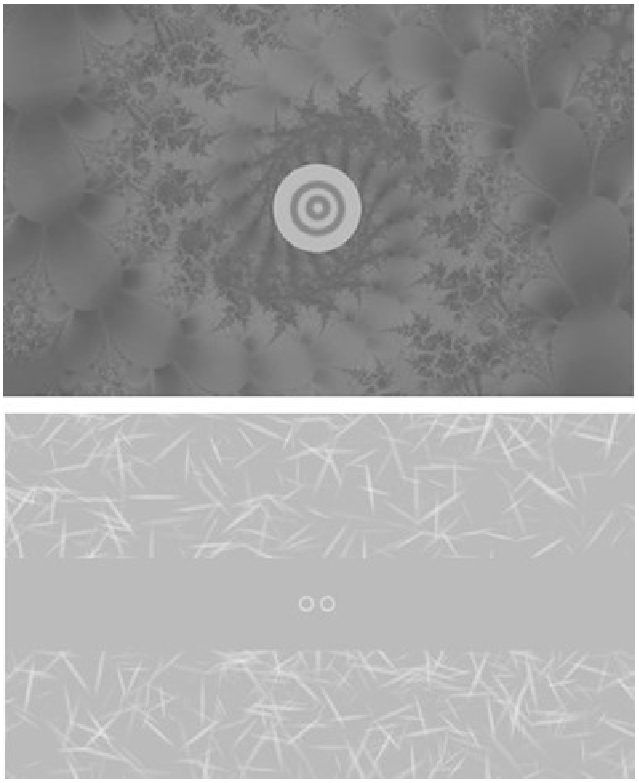

Figure 10. Stereo acuity test image (top) and fusion range test image (bottom).

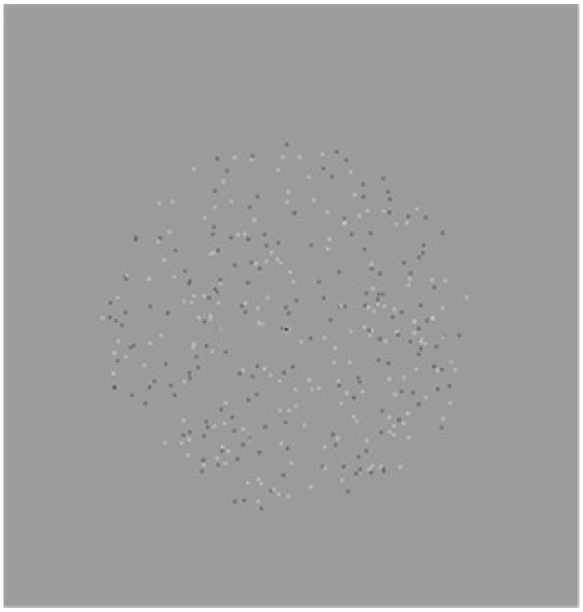

Figure 11. Motion sensitivity test stimuli.

Results

Figure 12 shows the distribution of stereo acuity (average of near and far), fusion range, and motion sensitivity scores. Fusion range was computed by taking the sum of crossed and uncrossed fusion recovery values in log prism diopters (PD), then subtracting the vertical fusion recovery range in log PD (i.e. the ratio of total horizontal fusion range to vertical fusion range). As shown, subjects vary substantially on both dimensions. For stereo acuity, the best observers obtained scores better than 10 arcsec, while the worst were 300 arcsec or greater.

Figure 13 shows the distribution of depth thresholds for each viewing condition and aircraft height (0.2 m, top graph; 2 m, bottom). Similar to Experiment 2 the use of a stereoscopic display clearly decreases depth thresholds. Shown in Figure 14 are the average depth thresholds for each viewing condition and aircraft height. The average depth thresholds were 114 cm, 198 cm, 171 cm, and 299 cm for the 0.2 m stereo, 0.2 m no-stereo, 2 m stereo, and 2 m no-stereo viewing conditions, respectively. A repeated measures analysis of variance reveals that stereoscopic viewing and height both had significant effects [$F(1, 156) = 25.7$, $p << 0.001$; $F(1, 156) = 13.5$, $p < 0.001$], but no significant interaction (see Figure 14).

Shown in Figure 15 are the relationships between individual average stereo acuity scores (top), individual fusion range scores (bottom), and depth thresholds averaged over aircraft height. Notably, the correlation between stereo acuity and depth accuracy was not significant. However, the correlation between fusion range and stereoscopic depth accuracy was significant ($r = 0.46$, $p = 0.003$). The correlation between fusion range and monoscopic depth accuracy was not significant ($r = 0.29$, $p = 0.07$). Shown in Figure 16 is the relationship between the combination of stereo acuity and fusion range and depth accuracy. The surface shown in this figure is a fitted multiple regression model ($R^2 = .513$, $p = 0.00012$, 30 df). Motion sensitivity was predictive of both stereoscopic and monoscopic depth accuracy ($r = 0.34$, $p = 0.03$; $r = 0.45$, $p = 0.003$).

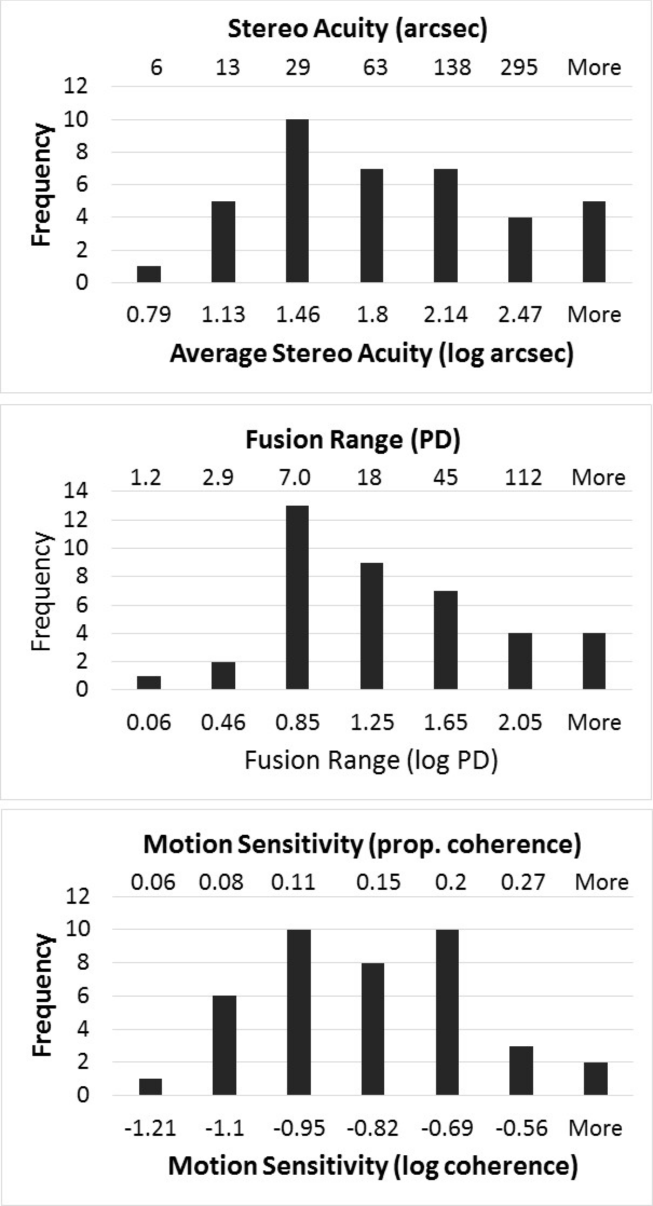

Figure 12. The distribution of stereo acuity (average of near and far), fusion range, and motion sensitivity scores (top, middle, and bottom, respectively).

Cleared, 88PA, Case # 2016-6379, 8 Dec 2016.

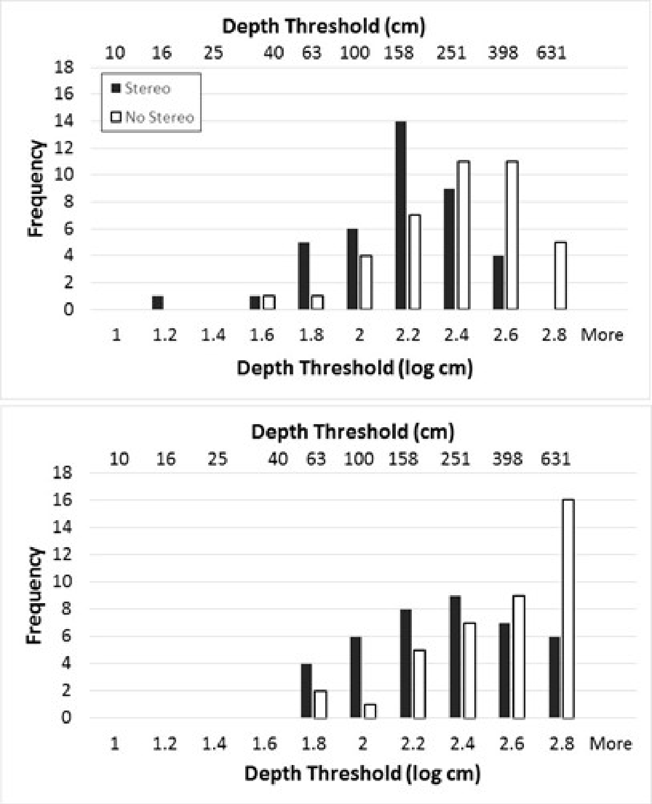

Figure 13. Distribution of depth thresholds for the hover task in Experiment 3 at a height of 0.2 m above the pole (top) and at 2 m above the pole (bottom).

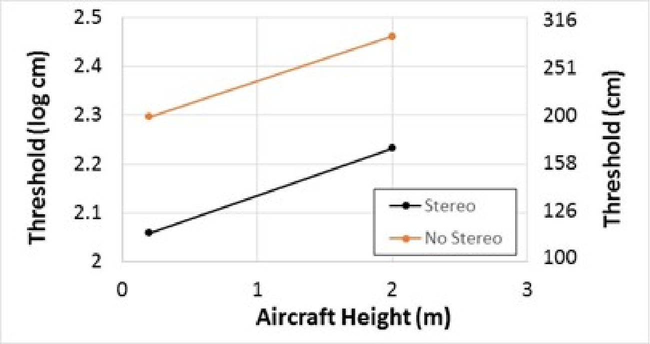

Figure 14. Average depth thresholds in Experiment 3 for each viewing condition and aircraft height.

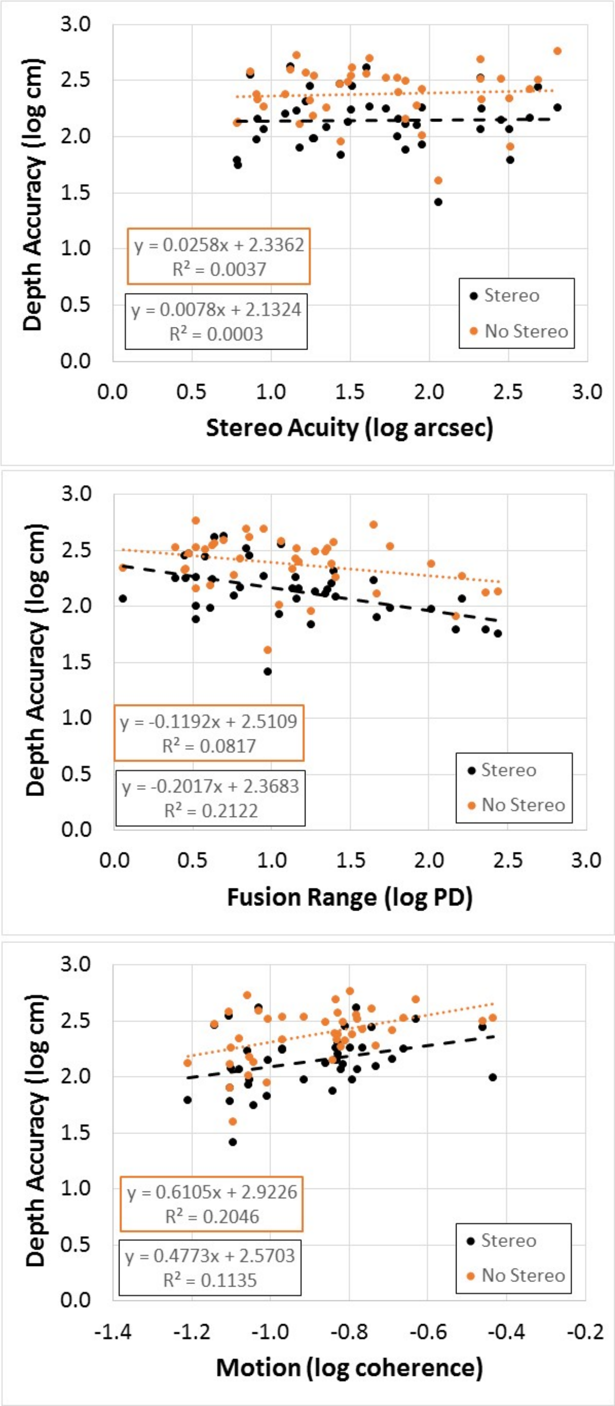

Figure 15. Scatterplots showing relationships between depth thresholds averaged over aircraft height and individual stereo acuity scores (top), fusion range scores (middle), and motion sensitivity scores.

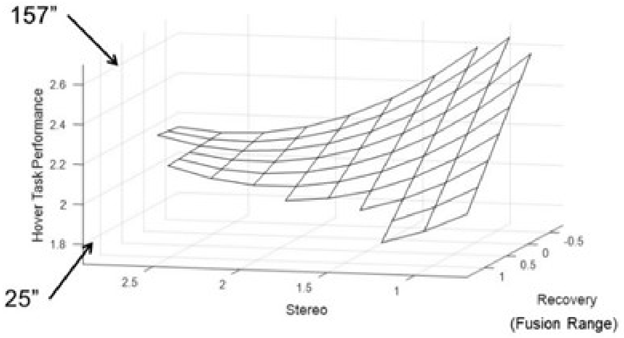

Figure 16. The relationship between the combination of stereo acuity and fusion range and depth accuracy.

Table 4 summarizes the correlations between each vision test and helicopter tail wheel depth thresholds for stereo and no-stereo viewing conditions (averaged across two altitudes).

Table 4. Correlations between Vision Tests and Helicopter Tail Wheel Depth Thresholds for Stereo and No-Stereo Viewing Conditions (averaged across two altitudes). Significant correlations are enlarged and bolded.

	Viewing Condition			
	Stereo		No-Stereo	
Vision Test	**Correlation**	**Significance**	**Correlation**	**Significance**
Stereo Acuity	r = 0.017	p = 0.91	r = 0.06	p = 0.71
Fusion Range	**r = 0.46**	**p = 0.002**	r = 0.29	p = 0.07
Motion Sensitivity	**r = 0.34**	**p = 0.03**	**r = 0.45**	**p = 0.003**

Experiment 4

Methods

Subjects

Thirty-one subjects volunteered to participate in Experiment 4. All subjects provided informed consent and the experimental protocol was approved by the Wright-Patterson IRB. However, because a few subjects either did not return to complete all conditions or were missing data, not all 31 subjects were included in subsequent analyses. Twenty-two subjects completed all eight viewing conditions.

Apparatus

The apparatus for the helicopter landing simulation and vision tests used in Experiment 4 were the same as that described for Experiment 3. A low contrast/low visibility condition was added using a pixel shader. The contrast level used was 10%, which was set using a Minolta LS-100 luminance meter.

Procedure

The procedure used in Experiment 4 was similar to that used in Experiment 3, except that an additional viewing condition was included – a low-contrast, or reduced visibility, condition. Thus, there were eight viewing conditions total: two stereo conditions (on or off) × two altitudes (0.21 and 2 m) × 2 levels of visibility (low and high). An average contrast of 10% between the tail wheel and pole and the concrete tarmac background was used in the low visibility condition. Figure 17 shows the low visibility condition at the 2-m altitude.

Figure 17. Subject's view of the tail of the aircraft and the rear wheel at 2.0-m height and low visibility. Image captured by OBVA lab personnel.

Results

Shown in Figure 18 are the average depth thresholds for each viewing condition in Experiment 4. After averaging across height, a repeated measures analysis of variance reveals that stereoscopic viewing again had a significant effect on depth thresholds [F(1, 84) = 14.8, p < 0.001], but that neither the overall effect of contrast [F(1, 84) = 1.78, p = 0.19] nor the interaction was significant [F(1, 84) = 2.33, p = 0.13]. However, performance was significantly better for the high-contrast, 0.21-m height with stereo compared to the low-contrast, 0.21-m stereo condition [t(21) = -4.27, p <0.001].

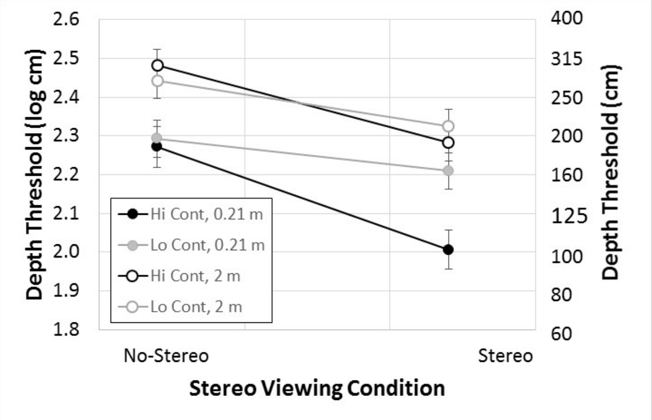

Figure 18. Average depth thresholds for each viewing condition in Experiment 4.

Table 5 summarizes the correlations between each vision test and performance, in terms of depth thresholds, for Experiment 4. Similar to Experiment 3, stereo acuity was not correlated with performance, but motion sensitivity was correlated with performance. With fewer subjects, the relationship between fusion range and performance was not significant, but one component of the fusion range measure, vertical recovery, was correlated with depth thresholds.

Table 5. Correlations between Each Vision Test and Hover Task Depth Thresholds.

Vision Test	Correlation	Significance
Stereo Acuity	r = 0.22	p = 0.22
Fusion Range	r = 0.33	p = 0.07
Vertical Fusion	**r = 0.36**	**p = 0.047**
Motion	**r = 0.43**	**p = 0.017**

General Discussion

The results of Experiment 1 provide evidence that stereo judgments were not likely to have been limited by either the selected display system or an aspect of the simulation environment. Subjects with very good stereo acuity were able to make very accurate depth judgments when provided with suitable stimuli such as the concentric rings shown in Figure 2 and were able to make very accurate judgments using the cross-shaped stimuli in the X-Plane virtual environment.

The results of Experiments 2, 3, and 4 clearly show that stereoscopic viewing improves performance when judging the depth of a simulated H-60 tail wheel relative to a landing zone obstacle, in this case a windsock pole. Thus, it might be expected that subjects with better stereo acuity should perform better than subjects with poor stereo acuity on this relative depth task. However, the results of this series of experiments show that stereo acuity considered in isolation does not predict performance. However, the results indicate that good ocular alignment, which we assessed using a fusion range test, may be important for accurate depth judgments. Somewhat surprisingly, given that there was relatively little motion in this simulation, increased sensitivity to motion was more consistently correlated with improved depth judgment in comparison to either stereo acuity or fusion range. This suggests that motion parallax was used as an important cue to depth by subjects in this experiment. However, as shown in Figure 16, fusion range and stereo acuity test scores may need to be considered together to predict performance on this simulated call-to-landing task. Subjects with a combination of good ocular alignment and good stereo acuity may perform better on this depth judgment task.

Conclusion

This research clearly shows that the use of stereo displays improves performance on a depth judgment task – in this case a simulation of a task relevant to clearing rotary wing aircraft for landing. This finding may be an important consideration for ground-based training for flight engineers tasked with clearing aircraft for landing. As noted in the introduction, numerous visual cues contribute to depth perception, such as relative size, motion parallax, texture gradient, etc.; thus, observers may be able to use a combination of these cues to perform the task adequately despite weak stereo acuity. Nonetheless, it is still somewhat surprising that observers with superior stereo acuity were not clearly superior in performing this operationally relevant task involving depth estimation. This finding may have implications for minimum vision standards for this career field. Currently, flight engineers are held to the same depth perception standard as fighter pilots. These results suggest that the stereo acuity standard could potentially be less stringent without adversely affecting safety, and which would have the desirable effect of enlarging the potential pool of qualified candidates. The results also suggest that the waiver policy concerning stereo acuity could be adjusted to significantly reduce costs. Based on these results, tests of ocular alignment and motion sensitivity deserve more research for potential use in aircrew vision screening. In future research we will continue to examine simulation requirements for stereoscopic displays, as well as the effect of stereo displays and quality of vision on time to contact and height above terrain judgments, as well as performance in clearing the aircraft for landing in a complex combat landing simulation.

References

[1] Wilmer WH, Berens C Jr. The eye in aviation. In: Aviation medicine in the A.E.F. Washington, DC: Government Printing Office; 1920:165-9.

[2] Howard HJ. A test for the judgment of distance. Trans Am Ophthalmol Soc. 1919; 17:195-235.

[3] Diepgen R. Do pilots need stereopsis? [article in German]. Klin Monbl Augenheilkd. 1993; 202(2):94-101.

[4] Entzinger JO. The role of binocular cues in human pilot landing control. Paper presented at the Thirteenth Australian International Aerospace Congress; 2009 Mar 9-12; Melbourne, Australia.

[5] U.S. Air Force. Medical standards directory (MSD). 2016. [Available to those with access].

[6] U.S. Air Force. United States Air Force HH-60G Helicopters (88-26105/91-26359). Nellis AFB, NV; 1998. Accident Investigation Board Report AIB-HH60G 19980903, Nellis AFB.

[7] McIntire JP, Havig PR, Geiselman EE. What is 3D good for? A review of human performance on stereoscopic 3D displays. In: Marasco PL, Havig PR, Desjardins DD, Sarma KR, eds. Head- and helmet-mounted displays XVII; and Display technologies and applications for defense, security, and avionics VI. Bellingham, WA: SPIE Press; 2012. SPIE Proceedings Vol 8383.

[8] Winterbottom M, Gaska J, Wright S, Lloyd C, Gao H, et al. Operational based vision assessment research: depth perception. J Aust Soc Aerospace Med. 2014; 9:33-41.

[9] Tidwell RP. Stereopsis takes off in flight simulation. IEEE Proceedings on Southeastcon, New Orleans, LA. 1990; 2:578-582.

[10] Yeh YY, Silverstein LD. Limits of fusion and depth judgment in stereoscopic color displays. Hum Factors. 1990; 32(1):45-60.

[11] Couvillion WC, Roberts AJ, Newell BD. KC-135 boom operator training hi-fidelity stereoscopic display technology demonstrator. Proceedings of the Interservice/Industry Training, Simulation and Education Conference, 2003.

[12] Held RT, Hui TT. A guide to stereoscopic 3D displays in medicine. Acad Radiol. 2011; 18(8):1035-1048.

[13] Hoffman DM, Girshik AR, Akeley K, Banks MS. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. J Vis. 2008; 8(3):33.1-33.30.

[14] Kooi FL, Toet A. Visual comfort of binocular and 3D displays. Displays. 2004; 25(2–3):99-108.

[15] Lloyd CJ, Nigus SG. Effects of stereopsis, collimation, and head tracking on air refueling boom operator performance. Proceedings of IMAGE 2012 Conference; 2012 Jun; Scottsdale, AZ.

[16] Patterson R. Human factors of stereoscopic displays. New York, NY: Springer; 2015.

[17] Shibata T, Kim J, Hoffman DM, Banks MS. The zone of comfort: predicting visual discomfort with stereo displays. J Vis. 2011; 11(8):11.

[18] Wann JP, Rushton S, Mon-Williams M. Natural problems for stereoscopic depth perception in virtual environments. Vision Res. 1995; 35(19):2731-2736.

[19] Bach M, Schmitt C, Kromeier M, Kommerell G. The Freiburg Stereoacuity Test: automatic measurement of stereo threshold. Graefes Arch Clin Exp Ophthalmol. 2001; 239(8):562-566.

[20] Fawcett SL, Birch EE. Validity of the Titmus and Randot circles tasks in children with known binocular vision disorders. J AAPOS. 2003; 7(5):333-338.

[21] Larson WL. Does the Howard-Dolman really measure stereoacuity? Am J Optom Physiol Opt. 1985; 62(11):763-767.

[22] Gibson JJ. The optical expansion pattern in aerial locomotion. Am J Psychol. 1955; 68(3):480-484.

[23] Lee DN. The optic flow field: the foundation of vision. Philos Trans R Soc Lond B Biol Sci. 1980; 290(1038):169-179.

[24] Gibson JJ, Olum P, Rosenblatt F. Parallax and perspective during aircraft landings. Am J Psychol. 1955; 68(3):372-385.

[25] Gibson J. The ecological approach to visual perception. Hillsdale, NJ: Lawrence Earlbaum Associates; 1986.

[26] Lloyd CJ. On the utility of stereoscopic (3D) displays for simulation training. Proc. Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC); 2012 Dec 3-6; Orlando, FL.

[27] Brown DC. Decentering distortion of lenses. Photogramm Eng. 1966; 32(3):444-462.

[28] Kingdom FA, Prins N. Psychophysics: a practical introduction. Cambridge, MA: Academic Press; 2009.

[29] Winterbottom M, Lloyd C, Gaska J, Wright S, Hadley S. Stereoscopic remote vision system aerial refueling visual performance. Proceedings of the IS&T International Symposium on Electronic Imaging 2016 Stereoscopic Displays and Applications XXVII; San Francisco, CA; 2016:1-10.

[30] Williams L, Bullock T, Winterbottom M, Gaska J, Hadley S. 3D immersive environment using X-Plane for depth perception research. Proceedings of the Interservice/Industry Training, Simulation, and Education Conference; 2015; Orlando, FL.

[31] Morrone MC, Burr DC, Di Pietro S, Stefanelli M. Cardinal directions for visual optic flow. Curr Biol. 1999; 9(14):763-766.

Author Biographies

Dr. Marc Winterbottom is a senior research psychologist supporting the 711th Human Performance Wing, U.S. Air Force School of Aerospace Medicine, Aeromedical Research Department, Operational Based Vision Assessment Laboratory. His most recent research focuses on U.S. Air Force vision standards and modernization of vision screening practices. He was awarded a DoD SMART scholarship in 2010 and completed his PhD in Human Factors Psychology at Wright State University in 2015. He received his MS in Human Factors Psychology from Wright State University in 2000. From 2000-2011 he conducted research concerning visual displays, human performance, and flight simulation training technologies for the 711th Human Performance Wing, Warfighter Readiness Research Division.

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 30 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 a 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 U.S. Air Force School of Aerospace Medicine at Wright-Patterson AFB.

Dr. Logan Williams is a senior research engineer at the U.S. Air Force School of Aerospace Medicine at Wright-Patterson AFB, Ohio, and currently leads technology development for the Operational Based Vision Assessment Laboratory. Previously, he has led multiple lines of research in various fields such as human effectiveness, immersive environments, visual display system design, and distributed simulation for aircrew training. He has served as the lead systems engineer for F-16, A-10, and KC-135 aircrew training systems and has two decades of experience in analog and digital circuit design, networked control systems, optical & electro-optical system design, computer programming, and physics-based modeling and simulation. He has earned a PhD in Electro-Optics, ME and BS degrees in Electrical Engineering, as well as a BS in Physics.

Elizabeth Shoda is a research specialist who, after completing a Pathways Internship with the U.S. Air Force, 711th Human Performance Wing, accepted a position with KBRwyle to provide contract research support to the Operational Based Vision Assessment Laboratory. She received her MS in Industrial/Organizational Psychology from Wright State University in 2013 and is currently a doctoral candidate at the same institution. Her research interests include human performance, workload, and selection and measurement.

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 U.S. Air Force Command Pilot/Pilot Physician with over 4,000 pilot flight hours in 8 different military aircraft during 9 combat deployments with 74 combat missions including 6 rescues behind enemy lines. He received his BS in Aerospace Engineering at West Point, completed his MD at Uniformed Services University of the Health Sciences, Bethesda, MD, and his ophthalmology residency at Madigan Army Medical Center, Tacoma, WA. He retired as a colonel after 34 years active duty service.

Cleared, 88PA, Case # 2016-6379, 8 Dec 2016.

Acknowledgments

The authors would like to thank the 66 RQS for their support of this research and valuable input on the performance and training required for the call-to-landing task. Thanks also to Alex van Atta and Tommy Bullock for supporting the software development involved with the computer-based vision tests and development of the helicopter landing simulation. 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. This document has been approved for public release (Case # 2016-6379, 8 Dec 2016). 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.