Investigating the role of binocular depth cues in altitude judgements with real and simulated imagery

Prepared by: Lesley M. Deas
Department of Psychology, Centre for Vision Research, 4700 Keele St., York University, Toronto, Ontario, M3J 1P3
And

Robert S. Allison, Department of Psychology, Centre for Vision Research, York University[1]

Brittney Hartle, Department of Psychology, Centre for Vision Research, York University

Elizabeth L. Irving, School of Optometry and Vision Science, University of Waterloo

Laurie M. Wilcox, Department of Psychology, Centre for Vision Research, York University

Technical Report prepared under PWGSC Contract Number: W7714-145967
"CIMVHR Task 15: Operational Requirements for Stereopsis and Binocular Processing in Military Aviation"

Scientific Authority:

Mackenzie G. Glaholt
Defence Research and Development Canada
1133 Sheppard Avenue W.
Toronto, ON, M3K 2C9
416 635 2000 x2137
Mackenzie.Glaholt@forces.gc.ca

Contract Report
DRDC-RDDC-2016-C352
October, 2016

Abstract

Canadian Forces Health Services (CFHS) has an ongoing requirement to evaluate CAF aircrew vision status both at initial recruitment and over the course of their career. Vision status is derived from a set of clinical tests. Stereopsis, while not presently an aircrew requirement, is suspected to be important for certain aircrew tasks. Stereopsis is the ability to perceive depth based on binocular disparity and individual differences in stereoscopic ability (e.g. stereoacuity) are substantial. This feature of functional binocular vision may provide, for example, additional precision in estimating altitude by eye when flying close to the ground (e.g. during rotary wing operations). Accordingly, DRDC is undertaking a project to investigate the link between clinical tests of stereopsis (e.g. those measured by CFHS) and aircrew operational performance. In this series of experiments, we investigate the importance of stereopsis in call-to-landing during rotary-wing operations, typically performed by the Flight Engineer. Six types of naturalistic still imagery (three viewing angles, two landing locations) were captured at different altitudes ranging from 10 to 100 ft. The images were displayed stereoscopically and observers were asked to estimate the distance to the ground relative to a reference image. All observers participated in stereoscopic 3-D and 2-D viewing conditions. There was no consistent significant difference between the 3-D and 2-D viewing conditions, and therefore we found no evidence of a strong requirement for stereopsis during call-to-landing scenarios in which the viewer makes estimates of absolute distance to ground at altitudes in the 10 to 100 ft range. Further work is underway to investigate the role of stereopsis in altitude estimation during rotary wing low-hover (e.g. 0-5 ft).

INTRODUCTION

Binocular vision, the ability to see with two eyes, provides a number of important advantages to interactions with the world around us. In addition to the simple benefit of having two light sensors, binocular vision provides enhanced depth perception, a sense of the three-dimensional layout of the world that is not available from either eye's image along. This sense of 'solid sight' or stereopsis, is based on registration of slight positional offsets (retinal disparity) between points in the two eye's images (for review see Howard, 2012). While it is widely known that relative depth judgements from stereopsis are very precise at short distances (<2m) it has been demonstrated that stereopsis can be used to estimate depth at much larger distances, up to 248m (Palmisano, Gillam, Govan, Allison, and Harris, 2010).

Given the potential advantages afforded by binocular vision to obstacle avoidance, it is not surprising that the contribution of stereopsis has been the focus of aviation research. For instance, stereopsis has been identified as a relevant depth cue for operations such as landing, formation flying, aerial refueling, helicopter operations and ground operations (see Karlsberg, Karlsberg & Rubin, 1971; Acromite, 1999). Early investigations of the contribution of stereopsis to aerial tasks concluded that stereopsis provided little advantage to aircraft landing. Most of these studies were conducted in physical environments, and evaluated pilots' landing maneuvers under binocular vs monocular viewing (Jongbloed, 1935; Pfaffmann, 1948; Lewis & Krier, 1969; Grosslight et al, 1978; Diepgen, 1993). However, it is arguable that these results speak to how well pilots can adjust landing strategies to rely on monocular depth information, not to the use of stereopsis in general. In fact, it has been documented that although monocular landings were successful, when relying on one eye only, pilots engaged in steeper approaches, flares were initiated too high and sink rates at touchdown were higher (Grosslight et al, 1978; Krier & Lewis, 1969; Pfaffmann, 1948). Furthermore, under monocular conditions, pilots exhibited increased head movements (Lewis & Krier, 1969), reported higher cognitive workload (Perry, Dana & Bacon, 1967; Krier & Lewis, 1969) and increased apprehension (Jongbloed, 1935). Moreover, Pfaffmann (1948) reported that monocular pilots underestimated the size of objects from altitudes 500-800ft. In sum, it is clear that forced monocular viewing in pilots accustomed to binocular vision has significant negative consequences. It is likely that some of these issues could be ameliorated with extended monocular flight, however, the point remains that stereopsis may not be required to perform aerial tasks, but could provide important performance benefits. The advantages afforded by stereopsis are difficult to assess in field studies because they typically involve experienced pilots under conditions where multiple depth cues were available and are difficult to control. Moreover, the reduced field of view which accompanied monocular occlusion in such studies makes it even more difficult to evaluate the role of stereopsis.

More recent research has attempted to isolate stereopsis in controlled settings, using flight simulators or head mounted displays. For instance, Lloyd & Nigus (2012) showed a strong performance advantage (by a factor of 2.9) in a remote vision system refueling task for stereoscopic over 2-D viewing. In a subsequent study Winterbottom, Gaska, Wright, & Hadley (2016) confirmed this advantage. Furthermore, they report that individuals with lower scores on acuity tests performed more poorly on the stereoscopic in-flight fueling task and reported higher rates of discomfort. It is clear that stereopsis confers advantages in performing refueling tasks using remote vision systems, both in terms of precision and comfort. However, the contribution of stereopsis to aircraft landing remains uncertain given the issues raised above regarding real-world testing. The aim of the series of studies described here is to bridge the gap in the literature by using stereoscopic display systems to present real-world and simulated imagery.

In the experiments presented here, we focus on rotary-wing call-to-landing, a task where depth and distance perception are of critical importance. In this scenario, the aircraft descends from 100ft and the Flight Engineer estimates distance to the ground by eye (supplemented by pilot call-outs). The objective is to determine if, and to what extent, stereoscopic depth provides a benefit for call-to-landing during rotary wing operations. More specifically, we assessed the impact of stereopsis on estimation of absolute distance using natural and simulated stereoscopic-3D imagery.

Experiment 1A Magnitude Production using Natural Images

In this first experiment, we captured stereoscopic images in a real-world setting and observers used a magnitude production procedure to estimate altitude.

Method

Stimuli The stimuli consisted of stereoscopic still images taken in February 2015 at CFB Borden from a CH-146 Griffon helicopter, during six confined space landings. The photographer and assistant sat in the gunner position and photographed out the side of the helicopter (door open) during the landing events. The pilot was instructed to maintain hover at each of the pre-specified altitudes during the descent, and once the photos were taken the pilot was instructed to proceed to the next altitude. A Fujifilm FinePix REAL 3D W3 stereoscopic camera was used to capture the images. The stereo pairs were captured at a resolution of 3584 × 2016. See Figure 1 for subset of images. There were two image sets (different landing locations) for each viewing angle, resulting in a total of six image sets. Images were captured at 3 viewing angles (10°, 45°, 90°). At 10° and 45°, images were taken at 6 altitudes (10, 20, 40, 60, 80, 100 ft). At 90°, 3 altitudes were tested: 20, 40 and 60ft in one set, 40, 60 and 80 ft in a second (exclusions made due to poor image quality). Two 'image types' were created: stereoscopic-3D and the same set in 2D (left image in both eyes). In the S3D set, disparity varied across images and within each image – the average separation on the projection screen was 63mm. A black oval aperture was added to each image at zero disparity to frame the images.

Figure 1. Sample images from Experiment 1A, of a natural scene taken from a helicopter. From left to right, the image is taken at three different viewing angles (10°, 45° and 90°) at the maximum altitude for that set (100ft, 100ft and 60ft respectively).

Apparatus The experiment was conducted in a screening room with multiple viewers tested simultaneously. Images were rear projected onto a cinema screen (300x168cm) using a Christie Digital 3D Mirage projector (resolution 1920x1080) via Stereoscopic Player (Wimmer, 2005). Viewers wore LC shutter glasses that alternately blocked the left and right eye view at 120 Hz in synchrony with the display of the left and right images, respectively. This provided a time-multiplexed stereoscopic display and the percept of the content in stereoscopic 3-D. Viewers were positioned in two rows, at viewing distances of 325cm and 420cm from the screen. At the closest distance, the screen subtended 50° of visual angle and one pixel subtended 0.03°.

Observers Seventeen undergraduate students were recruited from the York university participant pool (8 female, mean age = 20.3 years (sd 3.3); 9 male, mean age = 21.1 (sd 2.8)). These observers were inexperienced in visual psychophysics experiments. All observers were screened for stereopsis, and all testing followed the tenets of the Declaration of Helsinki and the protocol was approved ethics review boards at York University, The University of Waterloo, and DRDC Toronto Research Center (Protocol 2015-040).

Procedure Viewers performed a magnitude production task[1] (Stevens, 1956) to estimate the altitude in the scenes. For each set of images, a reference image depicting an altitude of 40ft was presented as the standard. Viewers were instructed to write down a value to represent the altitude or “distance to the ground” and they then judged all other altitudes relative to this modulus. That is, if they chose a modulus of 10, and the distance to the ground seemed twice as large, then they would assign that altitude ‘20’.

[1] A pilot study was conducted with a subset of observers to evaluate the relative utility of magnitude estimation (fixed modulus) and magnitude production (free modulus) methods with 4 observers per group. Estimated depth

increased with altitude using the free modulus method, but not with the fixed modulus method.

Each image was presented for 10 seconds, followed by a blank response screen (5s). Responses were made in paper booklets. Observers completed a short practice session to familiarize themselves with this method (using 2-D lines as stimuli). Conditions were blocked by viewing angle (10°, 45°, 90°) and image set (location of landing). There were 6 blocks, corresponding to 6 image sets. Image type (2-D / 3-D) was interleaved within each block and each image was shown twice within a block. For viewing angles 10° and 45°, there was 24 trials per block (6 altitudes x 2 image types x 2 repetitions), and there were 12 trials for blocks with 90° VA (3 altitudes x 2 image types x 2 repetitions).

Analysis Normalized estimates [average (mean individual estimate / individual modulus)] were computed for each image set, and compared statistically using conventional repeated measures ANOVA and tests of means. the analysis was done separately for each image set/viewing angle.

Results & Discussion: Experiment 1A

Six different images sets were tested. For all sets of images perceived magnitude increased with increasing altitude. There were no significant interactions (Image Type x Altitude) and the slopes were not significantly different. See Appendix 1 for statistical details. The effect of Image Type (2-D / 3-D) varied across image sets. Three out of six images sets showed a difference between estimates in 3-D versus 2-D, while others did not. Table 1 shows the P-values associated with the main effect of Image Type for each of the six image sets. Figure 2 below illustrates the results for Image Sets that showed significant differences between 2-D and 3-D estimates.

Viewing Angle	Image example (max Alt.)	Main Effect Image Type * Significant effect
10		$F_{(1,16)}$ = 6.69, p = 0.03 *
10		$F_{(1,16)}$ = 7.2, p = 0.016 *
45		$F_{(1,16)}$ = 1.5, p = 0.238
45		$F_{(1,16)}$ = 3.36, p = 0.086
90		$F_{(1,16)}$ = 2.75, p = 0.117
90		$F_{(1,16)}$ = 9.58, p = 0.007 *

Table 1. Statistics of 2-D vs 3-D comparisons for each image set. For three viewing angles (10°, 45° and 90°), two sets of images were captured at different locations. The 2nd column shows a sample still from the set.

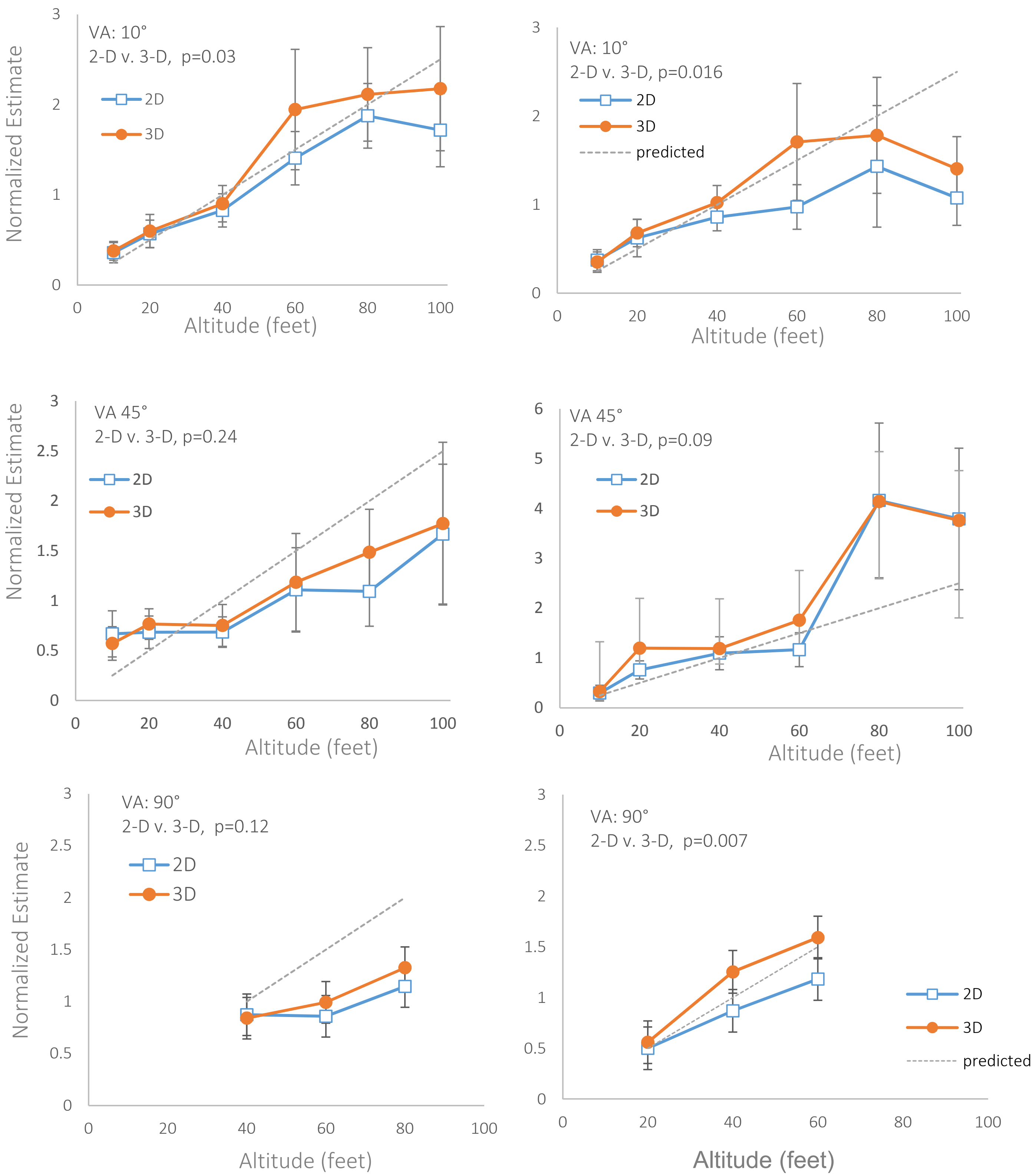

VA: 10°
2-D v. 3-D, p=0.03
2D
3D
Normalized Estimate
Altitude (feet)
VA: 10°
2-D v. 3-D, p=0.016
2D
3D
predicted
Altitude (feet)
VA 45°
2-D v. 3-D, p=0.24
2D
3D
Normalized Estimate
Altitude (feet)
VA 45°
2-D v. 3-D, p=0.09
2D
3D
Altitude (feet)
VA: 90°
2-D v. 3-D, p=0.12
2D
3D
Normalized Estimate
Altitude (feet)
VA: 90°
2-D v. 3-D, p=0.007
2D
3D
predicted
Altitude (feet)

Figure 2. The figures show results from 6 image sets with p-values corresponding to 2-D and 3-D comparisons. In the figure titles, VA refers to Viewing Angle. Note only 3 altitudes were tested at 90deg angle due to image artefacts and in one case (centre right column) the y-axis scale was doubled. In all graphs the orange, blue and dashed lines indicate 3-D, 2-D and predicted depth, respectively. Error bars represent 95% CI.

The results of Experiment 1A show no *consistent* significant difference between the 3-D and 2-D altitude estimates, though there appears to be a trend for the 3-D estimates to be larger than the 2-D. It is possible that this was a bias caused by the presence of a disparity offset between the oval and the 3-D imagery. To evaluate this possibility, we conducted a follow-up experiment using sub-set of the images from Experiment 1A (one image set at 10° and one set at 45°) and added a '2-D control' condition where the same image was presented to both eyes and was offset in depth relative to the oval aperture (which had zero disparity). The disparity offset was constant across conditions at 1.1 deg – the average disparity across the test conditions. Eleven new observers participated using the procedure described in 1A and the results are shown in Figure 3. Repeated measures analyses of variance were conducted separately for each viewing angle (with two factors: Image type and Altitude) using Greenhouse-Geisser to correct for violations of sphericity. These analyses showed a significant effect of image type for the 10° image set ($F_{(1,13)} = 4.38$, $p = 0.047$), but no significant effect in the 45° condition ($F_{(1,11)} = 0.81$, $p = 0.401$).

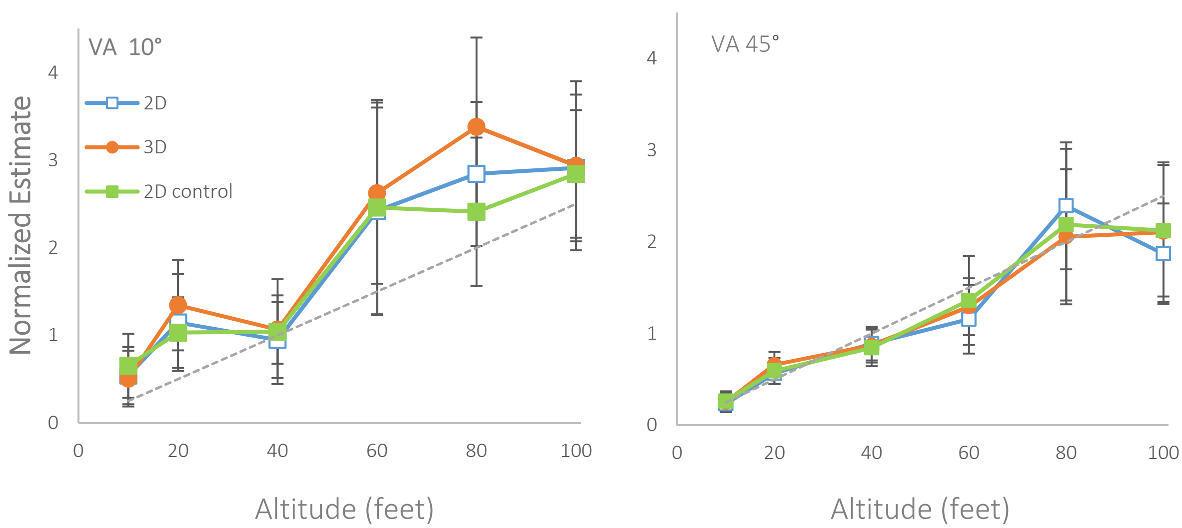


Figure 3. Normalized altitude estimates are plotted as in Figure 2, for the 11 observers that completed control Experiment 1A. Results for the 10° viewing angle are on the left, and the 45° are on the right. Orange circles, blue open squares and green filled squares represent 3-D, 2-D and 2-D control data respectively. The dashed line indicates the predicted estimates. The error bars represent 95% CI.

Conclusion Experiment 1A The results suggest that stereopsis does not provide a consistent benefit to judgements of absolute altitude for the range of heights tested here. Where the statistics do suggest a difference between 2-D and 3-D estimates, the p-values tended to be close to 0.05 (the threshold for significance). We also note that there is considerable variability at the higher altitudes which may be in part due to the fact that the images were taken in situations making it difficult to keep all capture conditions consistent. The follow-up study which included a 2D-control condition showed that the slight elevation in 3D estimates in Experiment 1A was not due to the fact that these images were offset from the fixation plane. However, it is difficult to draw strong conclusions given the variability of these data. In subsequent experiments we evaluate the potential advantage provided by stereoscopic viewing using simulated imagery in an effort to control some of the parameters which may have been responsible for the variable results.

Experiment 1B   Magnitude Production using simulated imagery

In Experiment 1A we used images captured in a real-world setting. This provided a degree of external validity, however as a result there was considerable variability between images and poor control over specific image parameters. To evaluate potential differences between 2D and 3D viewing with more precision we replicated Experiment 1A using computer software to simulate 3-D helicopter point-of-view (POV) and manipulate altitude.

Methods

Stimuli

Stereoscopic still images were created using commercial gaming software (ARMA II ™). Each image in the stereopair was at resolution 1920x1080. Two terrains, featuring different monocular texture cues were used which we refer to as desert and runway (see Figure 4). Images were captured at one viewing angle (10°) and six altitudes (10, 20, 40, 60, 80, 100 ft). Consistent with Experiment 1A, a black oval aperture was added to each image, positioned at the screen plane. Three conditions were generated: 2-D, 3-D (orthostereoscopic for front row) and 2-D control (as described above, left image in both eyes, offset in depth relative to the black oval).

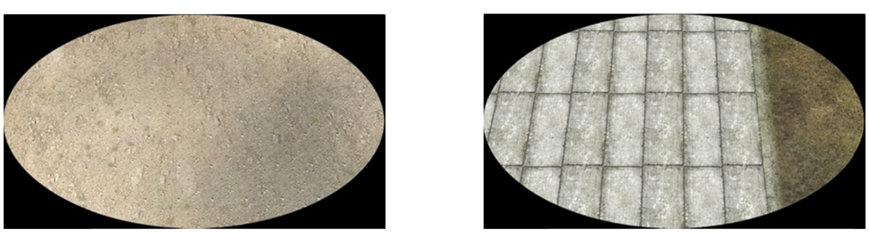


Figure 4. Sample images from Experiment 1B show a simulated desert (left) and runway (right) at 100ft. Altitudes of 10 to 100ft were tested at one viewing angle (10°).

Observers 9 undergraduate students with no previous experience in psychophysical tasks.

Apparatus As described in Experiment 1A.

Procedure The same procedure was used as described for Experiment 1A, with the 40ft altitude used as the standard. Image sets (desert and runway) were blocked, with 2-D and 3-D trials interleaved. As in Experiment 1A the 2-D and 3-D conditions were tested in separate blocks, and within a block, 6 altitudes were assessed for each of two terrain types. Observers evaluated each condition twice for a total of 24 trials per block.

Results and Discussion Experiment 1B

For simplicity, the results obtained using the two terrains (desert and runway) are discussed separately (see Figure 5).

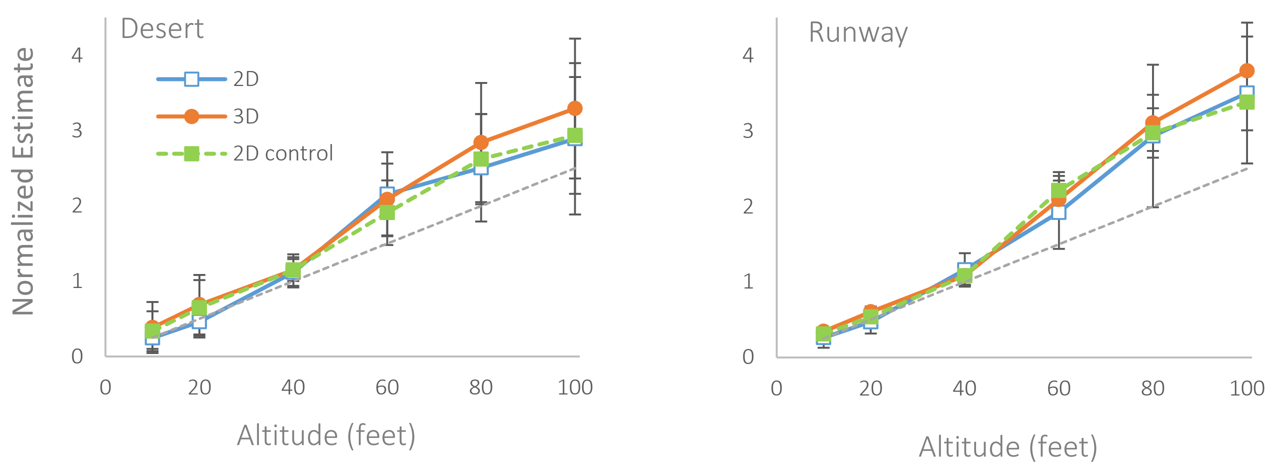


Figure 5. Normalized altitude estimates are plotted for 9 observers for Experiment 1B. Results using Desert imagery are on the left, and Runway imagery on the right. Orange circles, blue open squares and green filled squares represent 3-D, 2-D and 2-D control date respectively. The dashed line indicates the predicted estimates. Error bars represent 95% CI.

First consider the Desert scenario results (Figure 5, left). There was a significant effect of Image Type ($F_{(2,16)} = 5.65$, $p = 0.014$). Pairwise comparisons (with Bonferonni correction) revealed that estimates for the 2-D images were significantly different from the 3-D images ($p=0.003$). There was also a significant effect of Altitude ($F_{(5,40)} = 25.41$, $p < 0.0001$, Obs. power = 1.0), but no significant Image Type x Altitude interaction ($F_{(10,80)} = 1.36$, $p = 0.22$). In the runway condition, there was no significant effect of Image Type ($F_{(1,10)} = 1.86$, $p = 0.21$). There was a significant effect of Altitude ($F_{(1,10)} = 32.76$, $p < 0.0001$) but no significant Image Type x Estimate interaction ($F_{(2,19)} = 2.15$, $p = 0.14$).

Casual comparison of the results of Experiment 1A and 1B (Figures 2 and 5 respectively) shows that there is less variability in the pattern of results in Experiment 1B. This may be a consequence of the increased control over capture parameters and conditions (as expected). It may also reflect the fact that the imagery used in Experiment 1B contained more reliable 2D depth cues, for instance from texture, perspective and/or familiar size. The potential contribution of 2D cues is highlighted by the results obtained from the different terrain types. The desert terrain contained sparse texture (less 2D depth information) and we find that there was a slight but significant increase in altitude estimates when the images were viewed stereoscopically. However, the runway scene contained more structure and stronger perspective cues, and in this case there was no difference between the 2D and 3D conditions. As in Experiment 1A it appears that, while there are some suggestions of a 3D advantage, the impact is not substantial and is modulated by the availability of other depth cues.

Experiment 1C Altitude discrimination

In the preceding experiments, we used a magnitude production task to determine if the presence of stereoscopic 3D information influenced observers' ability to judge altitude. The range of altitudes was selected to be representative of real-world helicopter landing scenarios. While there were some instances where the presence of stereoscopic depth did seem to increase estimates, this effect was inconsistent, and highly dependent on the imagery used. Psychophysical studies of stereopsis have shown that while stereopsis can be used effectively to make depth judgements over long distances (Allison et al 2009; Palmisano et al, 2010) it is limited in its ability to support absolute distance judgements (Foley, 1980). It is possible that stereopsis might provide a clearer advantage if observers were asked to judge *changes* in altitude (rather than estimate their altitude). In Experiment 1C we conducted an altitude discrimination task, with conditions similar to those used in Experiment 1B to permit comparison of the two studies.

Methods

Stimuli The Experiment 1B image sets, created using the ARMA II™ gaming software, were used here. Desert and Runway terrains were both assessed under 2-D and 3-D viewing conditions. An altitude of 60ft was selected as the reference image and 7 test images were generated (including 60ft), with a step size of 3ft. In addition, to dissuade observers from using monocular features a lateral displacement 'jitter' was added to each image (2 and 4 ft offsets to the left and right).

Observers 9 observers who all had prior experience with psychophysical tasks were recruited.

Apparatus As described in Experiment 1A.

Procedure	A two interval forced choice (2IFC) task was used. On each trial, two images were presented sequentially for 2s each. The first interval always contained the reference stimulus, corresponding to the image type being presented in that block (2-D or 3-D). Observers indicated in which interval they appeared to be more distant from the ground, i.e. in which interval was the helicopter higher? Image sets (desert and runway) and image type (2-D and 3-D) were blocked. Each condition was repeated 20 times. There were 560 trials in total (2 image sets x 2 image types x 7 altitudes x 20 repetitions). This was separated into 16 blocks of 35 trials (7 altitudes x 5 repetitions), corresponding to 4 blocks per image set and type.

Results and Discussion Experiment 1C

Cumulative normal psychometric functions were fit to the data for each condition, using the number of times the 2nd interval was reported 'higher'. The Just-Noticeable Difference (JND) and Point of Subjective Equality (PSE) were calculated for each observer, and the averages are plotted in Figure 6. For reference, individual JND and PSE values are shown in Appendix 2. Note that given the Experimental design used here, the PSE reflects the degree of response bias.

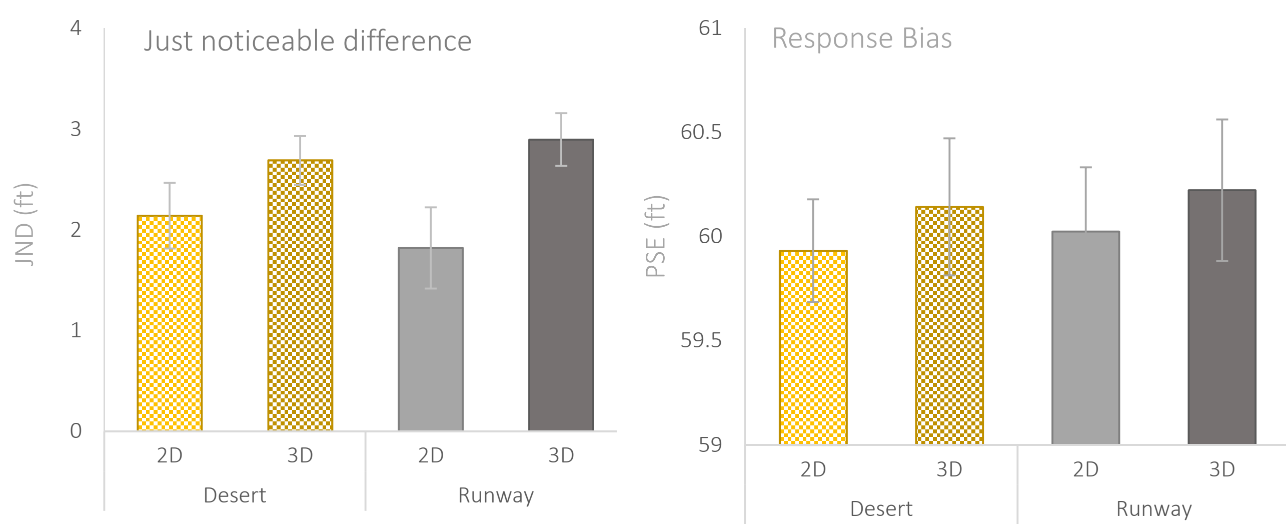


Figure 6. The average Just Noticeable Difference (left) and Response Bias (right) are shown here for 9 observers, in 2-D (light shades) and 3-D (darker shades) conditions for the two test scenarios. Error bars represent the SEM.

There was a significant difference between the JNDs (just noticeable difference) when comparing 2-D to 3-D viewing for the runway only ($t_{(8)} = -2.97$, $p=0.018$) with the average JND larger in the 3-D condition. There was no significant difference between the PSEs for any condition or terrain.

Observers can perform this altitude discrimination task; on average, thresholds are near 3ft. In the runway scenario we find that thresholds are significantly higher in the 3-D condition suggesting that performance is *poorer* in the presence of stereoscopic depth. It is not clear why this occurred; the 2-D cues are strongest in that condition, but are not in conflict with stereopsis. However, it is important to note that the statistical power associated with this comparison is low. This, combined with the fact that this pattern is not evident when the desert scenario is tested, undermines its impact.

General Discussion

The series of experiments described here suggest that stereopsis provides, at best, a very weak input to altitude estimation for still images in the range of 10 – 100ft. In the magnitude production tasks (Experiments 1A and 1B), most conditions do not show a positive impact of the presence of disparity on altitude estimations, and where there are trends it appears that distance estimates are larger for the 3-D conditions. Note that where we do see statistically significant effects, the p-values are relatively high and are not consistent across different observer groups. Further, Experiment 1C shows that observers can discriminate altitude using these stimuli, and that thresholds are as low as 3ft. However, in this experiment, as in previous experiments, there was no clear advantage provided by stereoscopic 3-D. In sum, in our experiments, binocular vision provides very weak input to absolute distance judgements at large altitudes (10-100ft).

It is well established that binocular disparity provides accurate and precise estimates of the relative separation of objects in space. However, to extract absolute distance from binocular disparity the visual system requires information concerning where the observer is fixating. This can be obtained by monitoring the vergence state of the two eyes (for review see Howard & Rogers, 2012). That is, when an observer changes their fixation to a different distance the vergence angle between the eyes changes; by monitoring this angle it is theoretically possible to estimate the absolute distance to the fixated location. Vergence eye movements have been shown to influence absolute distance judgements but only at near distances (< 6.5ft). At longer distances, 2-D cues such as relative size and texture, are relied on. The fact that there is no consistent improvement in altitude estimation with the addition of stereopsis in our studies suggests that in the conditions tested here, viewers relied on monocular cues to perceive depth. From an operational perspective, the dependence on 2-D information for landing highlights the importance of situation-specific training. For instance, aircrew might be trained to use average tree height as a 2D cue during altitude estimation. However, average tree high can change dramatically with changing latitude, and therefore this cue might not be reliable across flight situations. If the flight crew make errors in the assumed tree height, their altitude estimates will also be in error; a potentially disastrous mistake. However, over a range of distances, relative distance information from binocular disparity would be independent of such changes. Thus, in an ongoing series of experiments we are currently evaluating the contribution of stereopsis to relative distance estimation in another common operational scenario: low hover (0-5ft). As predicted, our results show that, in contrast to the experiments reported here, binocular viewing provides significant improvement in the accuracy of depth estimation. Specifically, when observers are asked to estimate the distance to the ground relative to a helicopter skid, estimates are accurate under binocular viewing, even when no 2-D cues are available. Conversely, accuracy is low when only 2-D depth cues are available. Taken together, the results of this series of studies underscore the need to assess the operational relevance of depth cues over a wide range of scenarios, and task demands.

References

Acromite, M.T. (1999) U.S. Navy Aeromedical Reference and Waiver Guide. *Naval Aerospace Medical Institute*. http://www.med.navy.mil/sites/ navmedmpte/nomi/nami/. [12.4 on stereopsis].

Allison, R. S., Gillam, B. J., & Vecellio, E. (2009). Binocular depth discrimination and estimation beyond interaction space. *Journal of Vision*,*9*(1), 10-10.

Palmisano, S.A., Gillam, B.J., Govan, D., Allison, R.S., and Harris, J.M. (2010) Stereoscopic perceptions of real depths at large observation distances. *Journal of Vision*, 10(6):19, 1–16.

Foley, J. M. (1980). Binocular distance perception. *Psychological review*,*87*(5), 411.

Grosslight, J. H., Fletcher, H. J., Masterton, R. B., & Hagen, R. (1978). Monocular vision and landing performance in general aviation pilots: cyclops revisited. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *20*(1), 27-33.

Howard, I. P., & Rogers, B. J. (2012). Perceiving in depth. Vol. 2: Stereoscopic vision. New York: Oxford University Press.

Jongbloed, J. (1935). Landing carried out by experienced aviators with the use of one eye only. *Acta Brevia Neerland*, *5*, 123-125.

Lewis Jr, C. E., & Krier, G. E. (1969). Flight research program. XIV. Landing performance in jet aircraft after the loss of binocular vision. *Aerospace medicine*, *40*(9), 957.

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.

Karlsberg, R. C., Karlsberg, F. S., & Rubin, M. (1971). Aerospace physiological optics. I. Depth perception. *Aerospace medicine*, *42*(10), 1080-1085.

Palmisano, S., Gillam, B., Govan, D. G., Allison, R. S., & Harris, J. M. (2010). Stereoscopic perception of real depths at large distances. *Journal of Vision*,*10*(6), 1-16.

Pfaffmann, C. (1948). Aircraft landings without binocular cues: a study based upon observations made in flight. *The American journal of psychology*, *61*(3), 323-334.

Stevens, S.S. (1956) The direct estimation of sensory magnitudes – loudness. *American Journal of Psychology,* 69, 1, 1-25.

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

Winterbottom, M., Lloyd, C., Gaska, J., Wright, S. and Hadley S. (2016) Stereoscopic Remote Vision System Aerial Refueling Visual Performance. *Stereoscopic Displays and Applications XXVII*, 437 1-10.

This space is intentionally left blank

Appendix 1

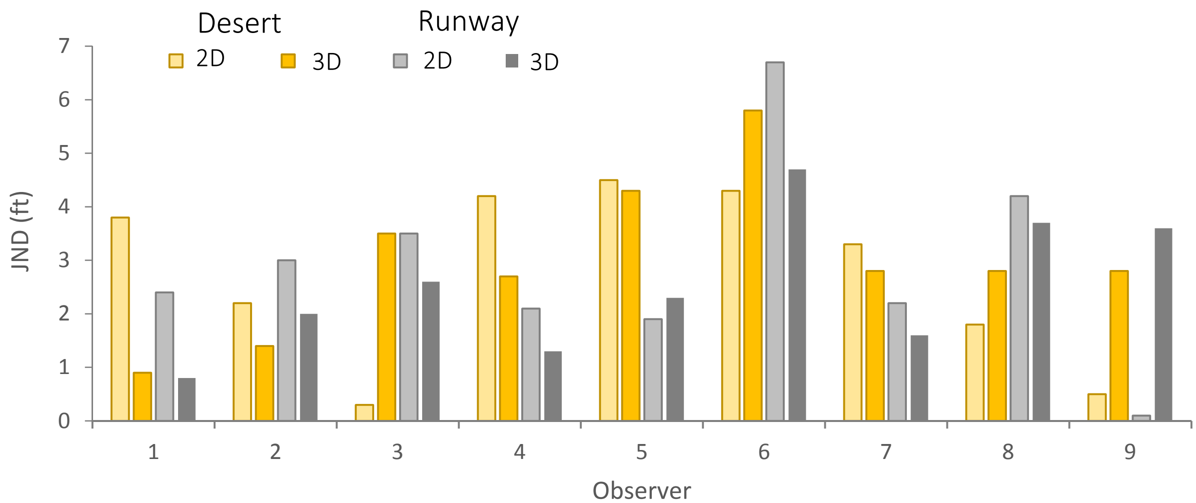
Additional Statistics for Experiment 1A

Viewing Angle	Image example (max Alt.)	Main Effect Altitude	Altitude x Image Type Interaction
10		F(3,41) = 27.45, p < 0.0001	F(2,37) = 1.9, p = 0.16
10		F(2,29) = 7.91, p = 0.002	F(2,31) = 3.26, p = 0.053
45		F(1,23) = 7.36, p = 0.007	F(2,33) = 0.73, p = 0.493
45		F(1,20) = 14.53, p = 0.001	F(2,31) = 0.46, p = 0.63
90		F(2,32) = 2.70, p = 0.083	F(2,32) = 0.95, p = 0.397
90		F(1,18) = 5.33, p = 0.03	F(2,28) = 2.75, p = 0.088

Appendix 2

Individual JNDs (A) and PSEs (B) for Experiment 1C. Conventions as for Figure 6.

A



B