https://doi.org/10.2352/EI.2024.36.2.SDA-349


# The effect of stereoscopic depth distortion on the near oculomotor response

***Eric S. Seemiller[1], Eleanor O'Keefe[2], Marc Winterbottom[1], Steven Hadley[1]***
***[1]711th Human Performance Wing, OBVA Laboratory, Wright-Patterson AFB, OH, USA; [2]KBR, Inc., Beavercreek, OH, USA***

## Abstract

*During natural viewing, the oculomotor system interacts with depth information through a correlated, tightly related linkage between convergence, accommodation, and pupil miosis known as the near response. When natural viewing breaks down, such as when depth distortions and cue conflicts are introduced in a stereoscopic remote vision system (sRVS), the individual elements of the near response may decouple (e.g., vergence-accommodation, or VA, mismatch), limiting the comfort and usability of the sRVS. Alternatively, in certain circumstances the near response may become more tightly linked to potentially preserve image quality in the presence of significant distortion. In this experiment, we measured two elements of the near response (vergence posture and pupil size) of participants using an sRVS. We manipulated the degree of depth distortion by changing the viewing distance, creating a perceptual compression of the image space, and increasing the VA mismatch. We found a strong positive cross-correlation of vergence posture and pupil size in all participants in both conditions. The response was significantly stronger and quicker in the near viewing condition, which may represent a physiological response to maintain image quality and increase the depth of focus through pupil miosis.*

## Introduction

When the two eyes must align and refocus on a near a target, they undergo a physiological response known as the near oculomotor response or near triad. The near triad consists of vergence (when the two eyes move in opposite directions to align the two visual axes), accommodation (when the intraocular lens changes shape to change the focal point of the eye), and pupil constriction (to increase the depth of focus of the eye and reduce accommodative demand). The result is a single, clear image of a target in depth. The elements of the near triad are known to be neurally linked and well-correlated with one another [1], [2], [3], insomuch that the primary visual cue that drives one can stimulate a significant response in the others [4], [5]. During natural viewing, when all the visual cues to physical depth are in good agreement, this relationship can be beneficial as it enforces a consistent and veridical oculomotor response.

However, when using a stereoscopic display, unless the parameters of the display perfectly match the users' visual system and the scene imagery, these visual cues are necessarily in conflict [6]. For example, the accommodative system and pupillary response will strive to bring the display in focus at its physical depth, while the binocular disparity in the image may drive a vergence response to a different position, a situation known as vergence-accommodation (VA) mismatch. Thus, the near triad receives conflicting stimuli, which lay lead to visual discomfort and limit the usability of the display [7], [8], [9].

Understanding the exact oculomotor disruptions elicited by depth distortions like VA mismatch may be difficult because it is cumbersome to measure all three elements of the near triad at once. Though, a small number of studies have attempted to measure vergence and accommodation simultaneously  e.g., [10], [11], or even all three [12], most studies are limited to one element, making only inferences about the others. Fortunately, most contemporary video-based eye-trackers that are used to measure vergence or accommodation must also first detect the pupil and estimate its size [13], [14], making it trivial to compare either to pupil size.

By comparing the relative temporal relationships of elements of the near triad, one can quantify the precise oculomotor deficits caused by (or adaptations to) known depth distortions of a stereoscopic display. For example, a temporal decoupling or, conversely, a tighter linkage of vergence with accommodation or pupil constriction may be indicative of visual stress that must be overcome. Alternatively, it may represent a physiological adaptation toward overcoming that stress. In this experiment, we examined the temporal coupling of two elements of the near triad (vergence and pupil constriction) while participants interacted with a stereoscopic remote vision system (sRVS) with different degrees of depth distortion.

## Methods

### *Participants*

Thirty-eight pre-presbyopic participants (including two authors, ESS and EO) were recruited and provided written informed consent before participating. The study adhered to the tenets of the Declaration of Helsinki and was approved by the Air Force Research Laboratory Institutional Review Board.

### *Apparatus*

Stereoscopic imagery was presented using a Sony VPL-GTZ280 4k laser projector (Sony, Tokyo), with a resolution of 4096 x 2160 pixels. Dichoptic presentation was achieved using Volfoni polarizing shutter glasses, synchronized to the refresh rate of the display. The Volfoni IR transmitter failed 42% of the way through the study and the remaining sessions used different stereoscopic glasses (Sony TDG-BT500A). We matched the luminance transmitted between the two sets of glasses and found that no outcome metrics were statistically different between the two types of glasses. Left and right eye imagery were temporally interlaced at 60 Hz. Crosstalk through the polarizing filters was measured at below 1% in the center of the display and below 5% in all visible regions.

Imagery was generated using a Vital 1100 image generator (IG) (Flight Safety International, Columbus, OH) for both left and right eye channels. Both channels used separate Nvidia Quadro P600 video cards (Nvidia, Santa Clara, CA). Nvidia Quadro Sync II cards synchronized the frame buffer between the two channels. No additional anti-aliasing was added beyond proprietary methods used in the IG.

We used a custom simulated telerobotic depth estimation task, developed in collaboration with Flight Safety International (Figure 1). Participants were instructed to drop a small white ball in a cup that was on the ground using a telerobotic gripper. The ball and gripper were 2.1 m above the ground in object space. The gripper was not connected to anything and appeared to hover above the ground, eliminating the use of the robotic arm as an alternative depth cue. The cup was placed on the ground of an airport tarmac, between 6 and 12 m downrange from the origin and between 1.5 m to the left and right of the origin. Lighting was controlled such that there was no shadowing in the imagery. Participants completed trials of the task as quickly and accurately as possible for 30-minute epochs.

***Figure 1.*** *A screenshot of the gripper task, presented here in 2D.*

Participants were placed in a chin rest, approximately 40 cm from an Eyelink 1000 Plus video-based eye-tracker (SR Research, Ontario). The eye-tracker was calibrated with a 13-point calibration and validation procedure. It recorded horizontal and vertical eye position binocularly at 1000 Hz, along with pupil size. Eye position data were smoothed with a rolling 50 ms average.

### *Viewing conditions*

Two viewing conditions are analyzed here (a subset of a larger study). The participant was positioned either 96 cm (Near) or 170 cm (Far) from the display screen. This led to overall changes in total field of view (measured in degrees), VA mismatch (measured in diopters) and percentage depth compression (the ratio of depth in object space to depth in image space). The separation between the two stereo cameras was set at 19 cm (approximately three times the average human interpupillary distance [15]) and the convergence distance of the cameras was set at 3.7 m by shifting the visual axes of the camera sensors (i.e., not through camera rotation). These conditions are highlighted in Table 1, with depth compression and VA mismatch calculated at 6 and 12 m of object space, the front and back of the working area in the experiment.

**Table 1: Viewing Conditions**

| | Near | Far |
|---|---|---|
| Screen Distance | 96 cm | 170 cm |
| Field of View | 49.25° | 29.02° |
| Depth Compression at 6 m | 53.8% | -13.1% |
| Depth Compression at 12 m | 116.2% | 22.1% |
| VA Mismatch at 6 m | 0.28 D | 0.16 D |
| VA Mismatch at 12 m | 0.51 D | 0.29 D |

### *Analyses*

All data were analyzed using custom MATLAB software. Missing data were interpolated with a nearest neighbor method and all data were smoothed over a rolling 100 ms windows to reduce high frequency noise. Horizontal vergence was measured as the difference between gaze positions of the right and left eye (convergence is negative). Pupil size was calibrated and computed as mm of diameter (Figure 2, top). Both were converted to velocities (Figure 2, middle). The two velocity vectors were then cross-correlated using the *xcorr* function and the normalized option. This provides a function of correlation between the two vectors as a function of temporal offset in seconds (Figure 2, bottom). The trial was summarized as the peak correlation, which is unitless, and the temporal offset of the peak correlation in milliseconds. Eye position vectors were analyzed at the individual trial level and then averaged within subject.

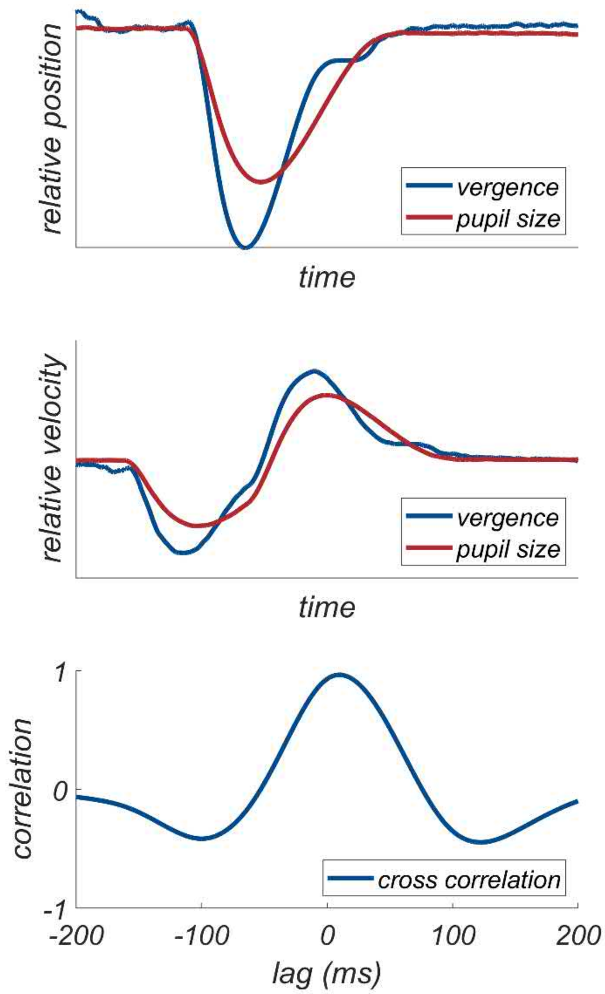


***Figure 2.*** *Raw position data of vergence and pupil size (top panel), showing a vergence eye movement and pupil constriction. This is transformed to velocity (middle panel), and the two vectors are cross-correlated (bottom panel) to yield a peak correlation and lag between the two velocity vectors.*

## Results

Subjects completed an average of 223.1 trials (SD = 59.94) in the near viewing condition and 222.4 trials (SD = 63.22) in the far viewing conditions. These were not significantly different. The peak correlations of all trials collapsed across subjects are shown in Figure 3. The mean peak correlation was 0.454 (SD = 0.136) in the near viewing condition and 0.389 (SD = 0.077) in the far viewing condition. These were significantly different ($t = 3.22$, $p = 0.003$). The relationship between the near and far peak correlations is shown in Figure 4. Data points below the unity line indicate higher peak correlations in the near viewing distance at the individual level. There was also a significant correlation between an individual's peak correlation at near and far ($r = 0.542$, $p = 0.002$).

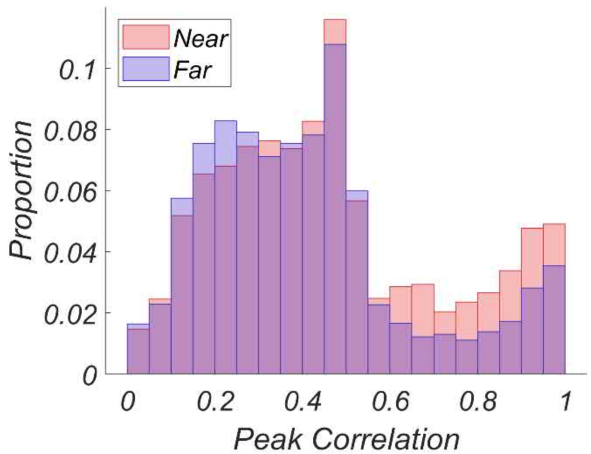


***Figure 3****. The distributions of peak correlations collapsed across subjects. Proportions for the near condition are in red and the far condition are in blue (overlapping histogram bars are purple).*

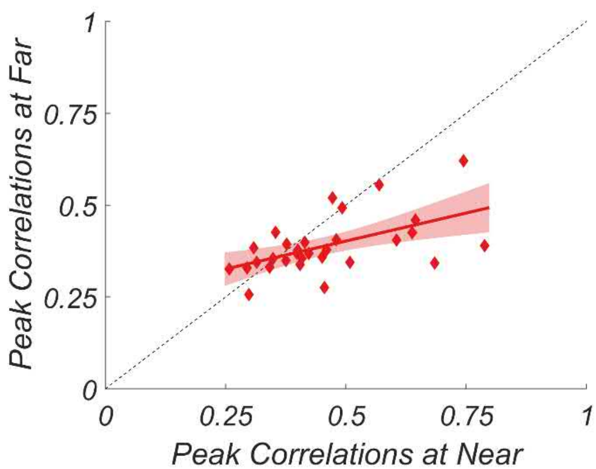


***Figure 4****. Peak correlations at far viewing distance plotted against near viewing distance. The line and shaded area are the least squares fit and 95% confidence interval. The hashed line is the unity line. Points below the unity line indicate a tighter correlation in the near viewing condition.*

There was a lag of pupil velocity relative to the vergence velocity in both conditions. The average lag was 30.08 ms (SD = 23.33) in the near viewing condition and 38.15 ms (SD = 32.94) in the far viewing condition (Figure 5). These were not significantly different. Further, there was no statistical relationship between lag and peak correlation at either far ($r = 0.070$, $p = 0.721$) or near viewing distance ($r = 0.026$, $p = 0.888$).

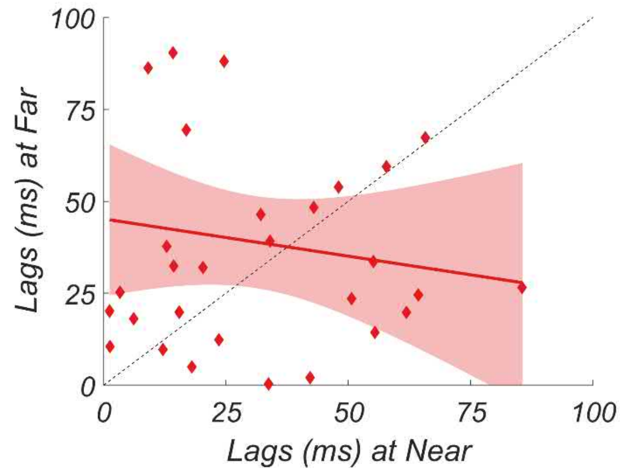


***Figure 5****. The lag of pupil velocity relative to vergence velocity (in ms) in the far viewing distance plotted against the near viewing distance. The line and shaded area are the least squares fit and 95% confidence interval. The hashed line is the unity line. Points below the unity line indicate a longer lag in the near viewing distance.*

## Discussion

The results of this experiment indicate a change in the temporal dynamics of the near oculomotor triad when the degree of stereoscopic depth distortion is changed during a depth estimation task. When the viewing distance decreases, which increases the amount of vergence-accommodation mismatch and increases the perceptual depth compression [6], the temporal relationship between vergence velocity and pupil velocity appeared to be more tightly coupled. Though the temporal lag was not statistically different, the peak cross-correlation was different, indicating a closer relationship between changes in vergence posture and pupil size. This may suggest that the near viewing distance and subsequent depth distortions create a condition where the near triad must be more responsive to eliminate extraneous stressors on the oculomotor system.

These results may suggest a physiological adaptation to depth distortion and cue conflict in the near response. The primary role of the pupil in the near response is to increase the depth of focus, reducing the accommodative demand of the intraocular lens [16], [17], [18]. The large potential VA mismatch in the near condition (as much as 0.51 D) may be too much for the intraocular lens to resolve comfortably. Thus, relying on the increased depth of focus provided by a constricting pupil may reduce the potential stress from cue conflict. The adaptation is not without cost; reduced light levels and, in the extreme, diffraction may lower aspects of image quality [19], [20]. Thus, in the far viewing condition, the pupil may be freer to respond to other stimulus elements or other biological controllers of pupil size.

Frequently the behavior of the pupil is used to estimate physiological stress, cognitive workload and other non-vision related responses in the user (for reviews, see [21], [22]). However, designers of stereoscopic displays should be wary of this approach.

As evidenced by these results, the behavior of the pupil is also a critical component of the visual system and may not necessarily betray a change in workload or stress when its role in the near triad is more urgent. Instead, the more telling pupil behavior is its tighter coupling to vergence, rather than its raw position or velocity.

Fortunately, analyses such as this one should be simple to undertake. Most contemporary video-based eye-trackers are capable of (and in fact require the measurement of) pupil position and size. Thus, simply looking at binocular eye position (or refractive state) and pupil size in terms of their temporal relationship to one another may provide display designers critical insight into the experience and stressors of the user.

## Acknowledgements

The authors wish to acknowledge Steve Nigus and FlightSafety International for their help in designing the gripper task. They also want to thank Matthew Ankrom, Margaret Hollander and Jason Russi for their assistance in data collection and management. The views expressed are those of the authors and do not reflect the official views of the United States Air Force, nor the Department of Defense. Mention of trade names, commercial products, or organizations do not imply endorsement by the U.S. Government. This research was supported by USAF contract FA8650-21-C-6277 to KBR. This work is cleared for public release: AFRL-2023-5810.

## References

[1] S. J. Judge and B. G. Cumming, "Neurons in the monkey midbrain with activity related to vergence eye movement and accommodation," *J. Neurophysiol.*, vol. 55, no. 5, Art. no. 5, May 1986, doi: 10.1152/jn.1986.55.5.915.

[2] C. Schor, "The influence of interactions between accommodation and convergence on the lag of accommodation," *Ophthalmic Physiol. Opt.*, vol. 19, no. 2, Art. no. 2, Mar. 1999, doi: 10.1046/j.1475-1313.1999.00409.x.

[3] P. J. May, S. Warren, P. D. R. Gamlin, and I. Billig, "An Anatomic Characterization of the Midbrain Near Response Neurons in the Macaque Monkey," *Invest. Ophthalmol. Vis. Sci.*, vol. 59, no. 3, pp. 1486–1502, Mar. 2018, doi: 10.1167/iovs.17-23737.

[4] T. G. Martens and K. N. Ogle, "Observations on accommodative convergence: Especially its nonlinear relationships," *Am. J. Ophthalmol.*, vol. 47, no. 1, pp. 455–463, 1959.

[5] E. F. Fincham and J. Walton, "The reciprocal actions of accommodation and convergence," *J. Physiol.*, vol. 137, no. 3, pp. 488–508, Aug. 1957.

[6] A. J. Woods, T. Docherty, and R. Koch, "Image distortions in stereoscopic video systems," *Proc SPIE 1915 Stereosc. Disp. Appl. IV*, 1993.

[7] F. L. Kooi and A. Toet, "Visual comfort of binocular and 3D displays," p. 10, 2004.

[8] D. M. Hoffman, A. R. Girshick, K. Akeley, and M. S. Banks, "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue," *J. Vis.*, vol. 8, no. 3, Art. no. 3, Mar. 2008, doi: 10.1167/8.3.33.

[9] T. Shibata, J. Kim, D. M. Hoffman, and M. S. Banks, "The zone of comfort: Predicting visual discomfort with stereo displays," *J. Vis.*, vol. 11, no. 8, Art. no. 8, Jul. 2011, doi: 10.1167/11.8.11.

[10] S. R. Bharadwaj and T. R. Candy, "Accommodative and vergence responses to conflicting blur and disparity stimuli during development," *J. Vis.*, vol. 9, no. 11, Art. no. 11, Oct. 2009, doi: 10.1167/9.11.4.

[11] E. S. Seemiller, J. Wang, and T. R. Candy, "Sensitivity of vergence responses of 5- to 10-week-old human infants," *J. Vis.*, vol. 16, no. 3, Art. no. 3, Feb. 2016, doi: 10.1167/16.3.20.

[12] L. Doyle, K. J. Saunders, and J.-A. Little, "Trying to see, failing to focus: near visual impairment in Down syndrome," *Sci. Rep.*, vol. 6, no. 1, Art. no. 1, Feb. 2016, doi: 10.1038/srep20444.

[13] K. Holmqvist and R. Andersson, *Eye tracking: a comprehensive guide to methods, paradigms, and measures*, 2nd ed. Lund: Lund Eye-Tracking Research Institute, 2017.

[14] A. Roorda, W. R. Bobier, and M. C. W. Campbell, "Geometrical theory to predict eccentric photorefraction intensity profiles in the human eye," *J. Opt. Soc. Am. A*, vol. 12, no. 8, p. 1647, Aug. 1995, doi: 10.1364/JOSAA.12.001647.

[15] N. A. Dodgson, "Variation and extrema of human interpupillary distance," *Proc. SPIE Stereosc. Disp. Virtual Real. Syst.*, vol. 5291, 2004.

[16] J. Liang and D. R. Williams, "Aberrations and retinal image quality of the normal human eye," *J. Opt. Soc. Am. A*, vol. 14, no. 11, p. 2873, Nov. 1997, doi: 10.1364/JOSAA.14.002873.

[17] R. Xu, H. Wang, L. N. Thibos, and A. Bradley, "Interaction of aberrations, diffraction, and quantal fluctuations determine the impact of pupil size on visual quality," *JOSA A*, vol. 34, no. 4, Art. no. 4, 2017.

[18] F. W. Campbell and A. H. Gregory, "Effect of Size of Pupil on Visual Acuity," *Nature*, vol. 187, no. 4743, Art. no. 4743, Sep. 1960, doi: 10.1038/1871121c0.

[19] A. B. Watson, "A formula for the mean human optical modulation transfer function as a function of pupil size," *J. Vis.*, vol. 13, no. 6, p. 18, May 2013, doi: 10.1167/13.6.18.

[20] S. G. de Groot and J. W. Gebhard, "Pupil Size as Determined by Adapting Luminance*," *JOSA*, vol. 42, no. 7, pp. 492–495, Jul. 1952, doi: 10.1364/JOSA.42.000492.

[21] B. Laeng, S. Sirois, and G. Gredebäck, "Pupillometry: A Window to the Preconscious?," *Perspect. Psychol. Sci.*, vol. 7, no. 1, pp. 18–27, Jan. 2012, doi: 10.1177/1745691611427305.

[22] A. J. Zele and P. D. Gamlin, "Editorial: The Pupil: Behavior, Anatomy, Physiology and Clinical Biomarkers," *Front. Neurol.*, vol. 11, p. 211, Apr. 2020, doi: 10.3389/fneur.2020.00211.

## Author Biography

*Eric Seemiller is the Senior Vision Scientist for the Operational Based Vision Assessment (OBVA) Laboratory at Wright-Patterson AFB, OH. He received his BS in Biological Anthropology from Penn State, his MA in Anthropology from Kent State, and his PhD in Vision Science from Indiana University. His work at the OBVA Lab has focused on binocular function and eye movements, specifically using stereoscopic remote vision systems.*