IMAGE SIZE AND VISUAL COMFORT IN 3D VIEWING

A thesis presented to the graduate faculty of
New England College of Optometry in partial fulfillment
of the requirements for the degree of Master of Science.

Manshul Nagpaul

May 2024

©Manshul Nagpaul,
All rights reserved.

The author hereby grants New England College of Optometry permission to reproduce and to distribute publicly paper and electronic copies of the thesis document in whole or in part.

DocuSign Envelope ID: 587FB165-9020-4107-89ED-00FF69FF8AB8

IMAGE SIZE AND VISUAL COMFORT IN 3D VIEWING

Manshul Nagpaul

This manuscript has been read and accepted by the Thesis Committee in satisfaction of the thesis requirement for the degree of Master of Science.

5/14/2024
Date

DocuSigned by:
Glen McCormack
C299988CF8094C2...
Glen McCormack, OD, Ph.D.
Graduate Faculty Advisor

DocuSigned by:
8C96C99C286D4CA...
Christopher Patrick Taylor , Ph.D.
Thesis Committee Member

DocuSigned by:
Elizabeth Pallante
3C03C07182C846E...
Elizabeth Pallante , OD
Thesis Committee Member

DocuSigned by:
Peter Bex
3E92398B601647F...
Peter Bex ,Ph.D.
Thesis Committee Member

5/14/2024
Date

DocuSigned by:
78A8AD978899447...
Thanasis Panorgias, MS, Ph.D.
Director of Graduate Studies

Abstract

IMAGE SIZE AND VISUAL COMFORT IN 3D VIEWING

Manshul Nagpaul

New England College of Optometry, 2024

This thesis explores the relationship between fixation target size and perceived visual comfort during 3D viewing. The specific aim was to ascertain whether small, medium, and large 3D fixation targets, moving within the same depth range, evoke distinct comfort ratings.

The study tested twenty-five subjects with normal binocular vision who were subjected to three main experimental sessions utilizing fixation targets of varying sizes. Participants' comfort levels were assessed through a symptom survey. Results indicated no association between fixation target size and visual comfort. This research holds practical implications for creators of 3D content.

Table of Contents

General Introduction 1

Methods 6

Subjects 6

Apparatus 7

Experimental Setup 9

Procedure 10

Data Analysis 15

Results 17

Discussion 34

Conclusion 41

References 42

List of Figures and tables

Figure 1 - Vergence range as a function of vergence type and image size. Inset shows our Maltese cross right eye fixation target. From McCormack and Kulowski 4

Figure 2 - Symptom survey used by Shibata et al. 8

Figure 3 - *Hypothetical relative symptom severity functions of fusional vergence innervation for small and large targets*.....13

Figure 4 - Relative vergence limits in this study (in prism diopters (Δ)) are plotted as functions of vergence type and target size. 18

Figure 5. Graphs of posttest symptoms versus pretest symptoms for the tired and headache symptoms for 3 target sizes: small, medium, and large. 28

Table A - Friedman Analysis of 3D Symptoms by Target Size, based on ‘difference” symptom ratings. 20

Table B - Difference eyestrain scores for 25 subjects 21

Table C - Skew values of distributions of symptom data by symptom, test type, and target size. 23

Table D - Summary of a repeated measures ANOVA of pretest, posttest, and difference symptom values for the eyestrain symptom, by target size, with F and p values. 24

Table E - Mean pretest symptoms for all symptom types and target sizes, including ANOVAs 25

Table F - Summary of a repeated measures ANOVA of pretest, posttest, and difference symptom values for the backache symptom, by target size, with F and p values. 27

Table G - Mean posttest symptoms for all symptom types and target sizes, including statistical analysis 31

Table H - The distribution of target size presentations presented to 25 subjects during experimental sessions 3,4, and 5.....33

General Introduction

Fixated 3D images[a] moving in depth are less comfortably viewed than their naturalistic counterparts [1,2,3]. 3D discomfort occurs in those with clinically normal binocular vision[4], so 3D discomfort cannot be dismissed as a problem of clinically - abnormal binocularity. Hoffman et al.[1] and Shibata et al. [2] observed that discomfort occurring during dynamic 3D convergence in their study is partly connected with Accommodation/Convergence Mismatch (ACM). ACM is an artificial stimulus condition inherent to most 3D displays, in which the convergence stimulus changes with respect to the video screen, but the accommodative stimulus does not. Maintaining clear and single vision on a target moving in 3D depth requires the oculomotor system to change convergence while holding accommodation constant. 3D ACM truncates the range of clear vision relative to naturalistic vision because strong convergence drags ocular focus off of the display screen, leading to blur. It is generally assumed that the struggle to maintain focus on the screen, while changing convergence, gives rise to ACM discomfort symptoms[2].

Head-mounted displays (HMDs)[5] are on the verge of revolutionizing the fields of education, entertainment, and communication. However, there are challenges related to HMD optics that have yet to be solved. Such near-to-eye displays can provide a natural and comfortable visual experience only if the displays accurately simulate natural accommodation and convergence. Soomro and Urey[6] found a minor reduction in visual acuity (20/20 to 20/25), when the clear vision of the user was covered with a retro-reflective

[a] In this paper a "3D image" is a binocular image viewed in a stereoscopic display such as a 3D movie, 3D TV, or VR HMD.

HMD screen. In current HMDs, used in VR applications, fixed lens and prism optics shift the viewing screen to optically distant vision so as to reduce near vision strain. However, when the images are adjusted to stimulate convergence, the accommodation stimulus remains fixed to the screen, sometimes causing ACM with blur and discomfort. It has also been observed that screen distance has a minimal effect on clarity when large objects are displayed on an HMD screen. McCormack and Kulowski[7] observed that large images strengthen fusional vergence and increase the neural depth of focus. These findings provide guidelines for maintaining clear images while designing virtual interfaces. A potential solution for ACM is to use an adaptive focus lens to adjust the accommodative stimulus with the convergence stimulus. However, using such lenses adds significantly to the cost of the device. Adaptive focus lenses are also not a good simulation of optical reality because they move the optical focus of the entire visual field and not just the focus of the fixated object. It appears that the problem of 3D ACM may not be solved soon.

Shibata et al.[2] showed the likelihood that their subjects would experience 3D discomfort correlated with well-known clinical criteria for patient comfort - Sheard's[8] and Percival's[9] criteria. Meeting these criteria is dependent on phoria size and fusional vergence strength. Shibata et al.[2] employed prism vergences to quantify fusional vergence strength to calculate whether subjects met Sheard's and Percival's criteria. A prism vergence is a standardized clinical test in which prisms stimulate convergence using binocular disparity, but the accommodative stimulus and image size do not change. Prism vergence testing causes ACM similar to 3D viewing. Because of the similarity of prism vergence to 3D vergence, we would expect weakness of fusional vergence to impair both prism vergence responses and 3D

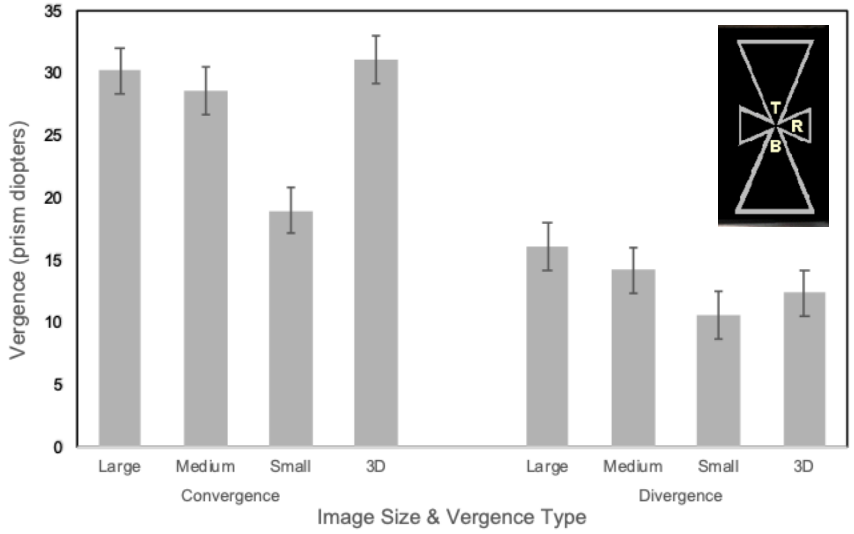
motion-in-depth responses. Shibata et al. suggested that their positive correlation of 3D discomfort with the clinical comfort criteria might predict which viewers are uncomfortable during 3D viewing, making it possible to clinically identify uncomfortable 3D viewers based on clinical prism vergence tests. However, there is at least one concern with this suggestion: Their "3D" images had fixed sizes in all conditions, unlike most 3D imagery and more like prism vergences. Hence, their correlations might not translate to individuals' visual performance and comfort in typical 3D vision.

McCormack and Hogan[4] tested prism vergences and 3D vergences in the same subjects. 3D vergences are analogous to prism vergences in that convergence and divergence targets are moved in depth to the limits of clear and single vision while watching a static accommodative stimulus. However, they are executed in a 3D display in which image size covaries with binocular disparity. McCormack and Hogan reasoned that if prism vergences equaled 3D vergences in the same subjects, then 3D image size might not affect clarity and comfort, and clinical prism vergences might then correctly identify those who are susceptible to discomfort caused by ACM. They found instead that 3D convergence was much larger than prism convergence and that 3D divergence was smaller than prism divergence. These finding questions whether an individual's 3D discomfort can be predicted by applying Sheard's and Percival's criteria based on clinical prism vergences.

McCormack and Kulowski[7] suspected that 3D fixation target image size might account for McCormack and Hogans' 3D findings. It was previously shown by Feldman et al.[10] that larger fixed-size video targets stimulate more vergence than small video targets, an effect explained by the fact that larger targets add more fusional textures to peripheral vision,

which increases fusional vergence strength.[11] The 3D image size hypothesis is supported by part of McCormack and Kulowski's[4] findings, reproduced in Fig.1 below. The "small", "medium", and "large" responses were stimulated by targets with fixed small, medium, or large size but all having changing disparity, like the images in Feldman et al. It is evident that McCormack and Kulowskis' small, medium, and large targets replicated Feldman et als' findings. During "3D" viewing, responses were stimulated by targets whose size and disparity were changing congruently as in typical 3D viewing. It is evident in Fig. 1 that the 3D vergence value in convergence (with its growing image size) approximates the vergence value of the fixed-size large target, while the 3D vergence

Figure 1. Vergence range as a function of vergence type and image size. Inset shows our Maltese cross right eye fixation target. From McCormack and Kulowski[4]



value during divergence (with its shrinking image size) approximates the vergence value of the fixed-size small target. These size/vergence correlations led McCormack and Kulowski to

suggest that fixation target size dynamically influences vergence range during 3D vergence similar to how various static image sizes influence vergence, which explains the observations of McCormack and Hogan[4].

Hofstetter[12] argued that prism vergence limits are reached when fusional vergence innervation is depleted. Presumably, Hofstetter's argument would also apply to 3D vergences. The fixed image size conditions in Fig. 1 suggest that fusional convergence strength should be low when viewing the small target, increase significantly when viewing the medium target, and increase a little more when viewing the large target. This size effect is a marker for relative fusional vergence strength because the strength of other forms of near innervation, such as accommodation (and therefore accommodative vergence) do not appear to increase significantly with target size. For instance, Labishetty et al.[13] observed that accommodative response did not increase as peripheral texture was added to foveal texture in conditions similar to ours (the "no conflict" conditions in their Fig. 6). Because fusional vergence strength correlates to both image size and to Sheard's and Percival's criteria (as noted above). It follows that large 3D images should be more comfortable than small 3D images.

McCormack and Kulowski[7] did not study the relationship of visual comfort with 3D vergence ranges, and Shibata et[2] al did not study the effect of fixation target size on comfort in their study, so that leaves open the question of whether 3D image size impacts visual comfort in 3D viewing. We hypothesize that for variously sized 3D fixation targets moving over the same range of depth, with image size changing congruently with disparity, visual comfort will be related to target size.

The significance of this study is that if our hypothesis is true, then makers of 3D content will have to consider the size, convergence demand[1,2], ACM[1,2], and velocity[12] of intended fixation targets when calculating the maximum range of disparities compatible with comfort in 3D viewing. This knowledge could help the 3D industry produce more acceptable technology and content and may eventually help optometrists better understand the relationship between their patients' complaints about 3D vision and commercial 3D content.

Methods

Subjects

Subjects read and signed a human subject's consent form before tests began. If not emmetropic, subjects wore their spectacle or contact lens correction or a trial-framed correction during all sessions. The great majority of 3D viewers complaining of 3D discomfort have clinically normal binocular vision, so we selected subjects with normal binocular vision using conventional clinical tests in a screening session. Twenty-five subjects, male and female, completed the study and met the following criteria: 18-35 years old, no strabismus or amblyopia, 20/20 or better corrected visual acuity, local and global stereo thresholds less than or equal to 40", normal general and ocular health by self-report, 40cm Von Graefe phorias from ortho to 6 exo, Maddox far phorias from 1 eso to 3 exo, 40cm divergence no less than 9Δ, 40cm convergence no less than 12Δ, distance divergence no less than 4Δ, distance convergence no less than 5Δ, normal amplitude of accommodation for age

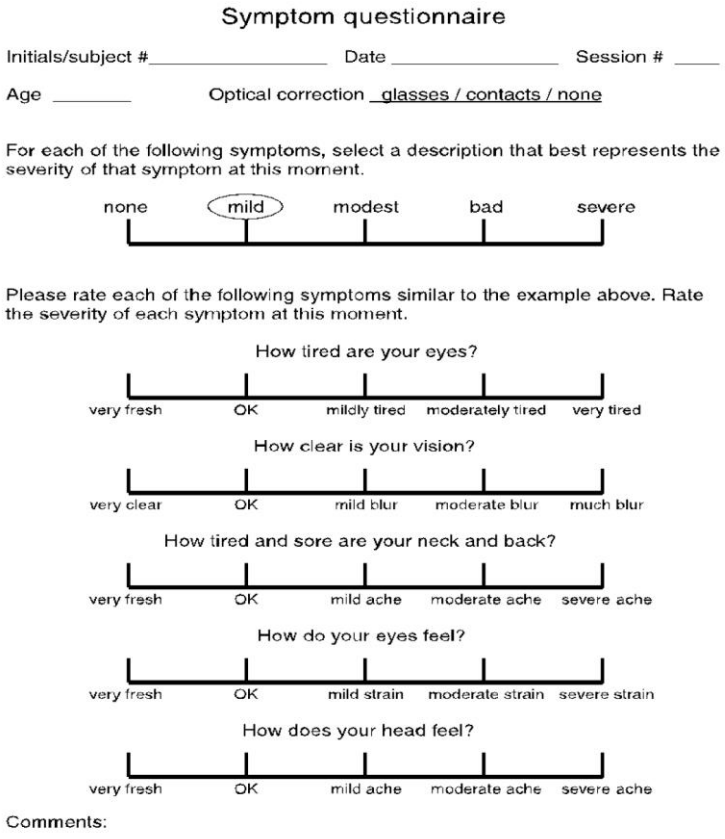
based on Hofstetter's minimum accommodation equation, normal near-point of convergence (<7 cm), and no reported visual discomfort in naturalistic viewing conditions. Subjects were asked if they had a history of chronic headaches (e.g., migraine), and if so, were excluded.

Subjects had to identify the onset of blur during vergences measured in the second session, so blur was demonstrated by having subjects view, a row of Snellen letters one line higher than their acuity level while wearing their refractive correction at six meters, and then blurring those letters with a half diopter sphere.

Apparatus

Experimental trials were run on a Dell Windows 3D computer and 27" 3D monitor. Images were viewed with the monitor 53.5cm from the eyes (2'/pixel). 3D images were controlled with MatLab and PsychToolBox software on the Dell PC.

We used an edited version of the symptom survey described by Shibata et al. (Fig 2), which was based on a similar version previously developed by Sheedy and Bergstrom[14] to assess discomfort symptoms. Subjective discomfort was measured by asking the subject to report three symptoms


Symptom questionnaire

Initials/subject #_____ Date _____ Session # ____

Age _____ Optical correction _glasses / contacts / none

For each of the following symptoms, select a description that best represents the severity of that symptom at this moment.

none mild modest bad severe

Please rate each of the following symptoms similar to the example above. Rate the severity of each symptom at this moment.

How tired are your eyes?

very fresh OK mildly tired moderately tired very tired

How clear is your vision?

very clear OK mild blur moderate blur much blur

How tired and sore are your neck and back?

very fresh OK mild ache moderate ache severe ache

How do your eyes feel?

very fresh OK mild strain moderate strain severe strain

How does your head feel?

very fresh OK mild ache moderate ache severe ache

Comments:

Fig. 2. The symptom survey used by Shibata et al.[3]

tired eyes, eye strain, and headache. These symptoms are well-known to be related to visual activity. Like Shibata et al., we used a neck/back comfort question as a validity check for subjects who automatically responded at the same level for all questions. Viewing targets of various sizes could not reasonably be expected to cause differences in neck or back pain between experimental target size conditions. We did not use Shibata's clarity question as it is a sight symptom (like diplopia or low contrast), not a discomfort symptom, and because it is unlikely to be reported in this study because our experimental design limits vergence to disparities perceived as clear and single. Subjects were asked to rate their discomfort by marking the position on a horizontal symptom scale that best represented their discomfort.

The symptom scale delivered numbers ranging from 0 (left end), to 100 (right end). The numbers were not visible to the subject. The textual descriptions on the scale guided subjects in selecting the position of their discomfort on the scale. Symptom surveys were administered in all three experimental sessions. Symptoms were assessed immediately before and immediately after each experimental session.

Experimental Target

The convergence target was a dark gray Maltese cross (14.5cd/m^2) on a light gray background (24.7 cd/m^2) (spatially similar to the insert in Fig.1). The cross had three different base sizes when imaged at the screen depth: Small (1° height), Medium (3° height), and Large (9° height). The target width was set at 45% of the height. Target sizes were scaled with nearness, as in typical 3D viewing, by the formula: size=(53.5*distance^-1)*(base size), where size is in degrees and distance from the observer is in centimeters. The target size was clamped at a minimum of .66°, which is required to prevent the cross from shrinking to zero size in large divergence. The distance value of the Maltese cross (required for size computation) is derived from its disparity: distance=100*PD/disparity where distance and PD (interpupillary distance) are in cm and disparity is in prism diopters.

Target depth motions were linear at 2Δ/sec during 3D vergence tests in the second session. Depth oscillations during experimental sessions were composed of alternating divergence and convergence cosine functions over time, with a peak velocity of 2Δ/sec. Kim, Kane, and Banks found that viewing rapidly oscillating motion-in-depth is more fatiguing than slowly oscillating motion-in-depth[15]. Therefore, we used a consistent 2Δ/sec peak

vergence velocity in all experimental trials to avoid introducing a confounding velocity variable.

Procedure

A subject's completed experiment was composed of five sessions run on five different days: (1) a screening session for selecting subjects (30 minutes), (2) a one-hour "vergence session" in which the 3D convergence and divergence limits were measured in our experimental apparatus, and (3) three 1-hour symptom induction sessions defined by image size ("Small", "Medium", and "Large"). Sessions were no more frequent than 1/day, and no less frequent than 1/week.

In the screening session, we obtained informed consent and run the clinical tests which check for normal binocularity. Potential subjects who did not pass the clinical screening session were excluded.

The purpose of the vergence session (the second session) was to establish the limits of vergence oscillation to be used in the experimental sessions. We measured the relative convergence and divergence limits – how much the eyes can converge or diverge from the screen surface while maintaining clear and single vision. Each vergence test began with the cross target at the screen distance of 53.5 cm. The "Convergence" limit was defined as the amount of convergence (in prism diopters) that first induced blur or diplopia, whichever came first. The "divergence" limit was defined as the amount of divergence (in prism diopters) which first induces blur or diplopia, whichever came first. Target relative disparity was incremented linearly from zero disparity at the screen surface, at 2Δ/sec, until the subject

reported the onset of blur and then of diplopia. The size of the cross varied with its disparity as it moved in depth, as described above. The vergence was measured for each image size (small, medium, and large). Each condition (3 sizes X con/div) was repeated three times, and the repeated values were averaged.

The experimental sessions created the 3D symptoms, which were the primary responses, and these symptoms can be variable. The optimal range of 3D motion had to be chosen for each subject because the range of clear and single vision measured by the vergences varies significantly between normals. The experimental session vergence range met these standards in each subject:

(1) The range of 3D motion-in-depth, within each subject, had to be one and the same value in all three experimental sessions, so that vergence motion magnitude did not vary by target size. This standard is critical because it is already known that convergence per se, without ACM, can cause symptoms in 3D viewing[1,2], so we must prevent the convergence stimulus from covarying with target size.

(2) Vision was clear and single throughout the entire range of 3D motion. We did not want blur and diplopia to confound discomfort reports.

(3) The peak vergence velocity was the same for all stimuli, so differences in velocity were controlled for.

(4) We stimulated as much fusional vergence as possible, subject to the first three standards, so that sufficient discomfort could be measured.

Choosing the vergence value for stimulating a difference in comfort between conditions is explained in Fig. 3. In this illustration, we show hypothetical functions of symptom severity

versus fusional vergence innervation for small and large targets. Zero prism diopters on the abscissa is the convergence distance at which the subject uses no fusional vergence. For subjects viewing our screen at 53cm, zero prism diopters was probably very close to the computer screen surface (±2Δ). It was assumed that symptoms rise monotonically with fusional vergence innervation, initially rising slowly when innervation is small, and leveling off at high fusional vergence levels when symptoms will have saturated. These shapes are monotonic S-shaped functions, but their exact shape should not significantly affect the results. Fig. 3 shows discomfort functions for one vergence direction only (divergence in this case). Symptom severity of 1.0 in Fig. 3 represents the upper limits of discomfort recorded on the symptom survey described above. The right end values of the functions are borrowed from the mean divergence limits of 11Δ and 17Δ, respectively, for the small and large target values displayed in Fig.1 above. Because of the small target's smaller vergence range and steeper slope, symptoms are expected to rise more rapidly with small target fusional vergence innervation.

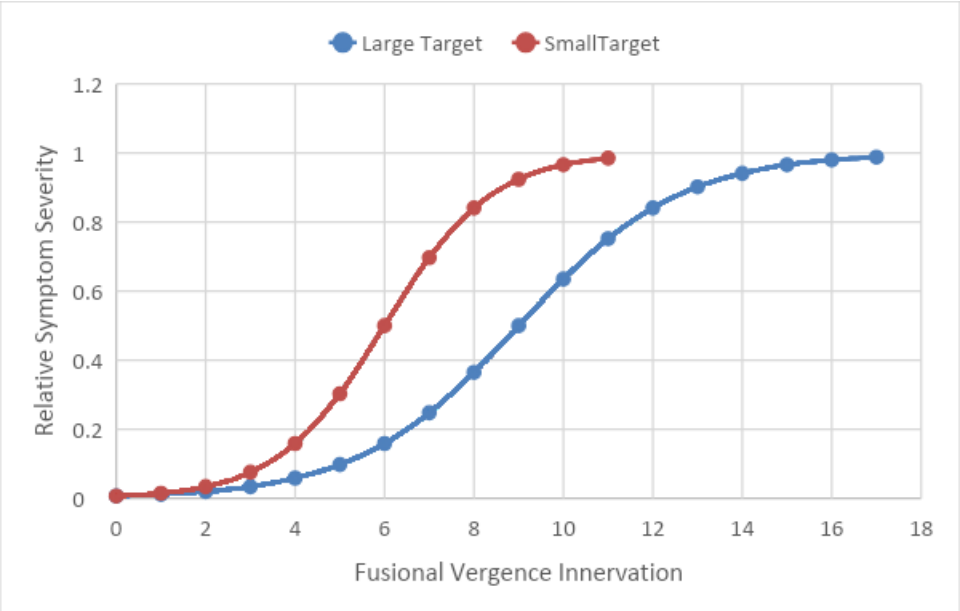


Fig. 3. Hypothetical relative symptom severity functions of fusional vergence innervation are shown for small and large targets. The right-hand ends of the functions on the fusional vergence innervation axis represent the limits of clear and single vision for the small and large targets.

The strategy for choosing the optimal symptom induction stimulus was to identify the one fusional vergence innervation that activated the greatest discomfort difference between the small and large target conditions. In Fig. 3 that would be approximately 8Δ on the fusional vergence axis, which is 47% of the large target vergence limit and 73% of the small target vergence limit. The small target should generate stronger symptoms because 8Δ stimulates a much higher percentage of small target vergence than of large target vergence. Based on these functions, we oscillated our experimental targets between 73% of the subjects' small

target divergence limit and 73% of their small target convergence limit. These limits were chosen because they would stimulate as much fusional vergence as possible without pushing the limits of blur and diplopia. According to Sheard's criterion, for the subject in Fig. 3 viewing the small target, they would spend 55% of their time viewing the target when its disparity is beyond the Limits of Sheard's criterion (disparities greater than 3.6Δ on the fusional vergence axis), and should therefore be symptomatic. Conversely, for large targets the subject will spend a much smaller percentage of their vergence oscillation time beyond the limits of Sheard's criterion because of the larger fusional vergence range of large targets.

We alternately stimulated convergence and divergence within each target oscillation, in experimental sessions, to minimize the confounding effects of prism adaptation, which become significant during unidirectional vergence stimulation and could alter comfort. This method choice means that symptoms induced during experimental sessions cannot be attributed separately to divergence or convergence.

All three experimental sessions used the same peak velocity of depth motion (in prism diopters per second). The only difference between the experimental sessions was the fixation target size. The order of the three symptom induction sessions was randomized. Only one experimental session was run on any given day and at approximately the same time, preferably in the morning. Another variable that influences symptom severity is the duration of viewing the experimental stimulus because symptoms take time to develop. Pilot studies found that 20 minutes was sufficient to induce measurable discomfort.

In each experimental session, the subject executed a vigilance task that required stereoscopic single binocular vision for success. This task is necessary because if subjects

frequently relaxed their eyes and failed to follow the targets with convergence, our stimulus conditions would not induce discomfort. The task was also intended to "entertain" the subject during experimental sessions. Four 8' diameter rings were simultaneously presented for 1.5 seconds of time above, below, left, and right of the center of the Maltese cross at a random time within divergence and then within convergence. The size and eccentricity of the rings were constant, regardless of cross size, to maintain the same level of "difficulty" at all target depths. One of the four rings, chosen randomly, had an 8' crossed disparity with respect to the other rings. After each divergence and convergence half of the target oscillation, the subject had two seconds to enter their guess of which ring stood out in depth. After their guess, the subject was given immediate auditory feedback (a beep) whose pitch depended on success. On average, approximately 50 vigilance task guesses were answered in a 20-minute experimental session. At least 70% of the guesses had to be correct for the session to count. If subjects failed the vigilance task in a session, they were asked to repeat the failed session on another day. If they failed a repeat session, then they were excluded from the remainder of the study, and their data was not processed.

Data Analysis

Our data analyses were based on methods used in three earlier studies. Kuzi and Ukai[16] analyzed display-induced motion symptoms, which included visual fatigue and eyestrain symptoms like our study, and other symptoms as well. They used a method of subtracting pretest symptoms from posttest symptoms – the "difference" method. This

method is intended to measure only stimulus-induced discomfort and not preexisting discomfort. Their analyses used parametric factor analysis and repeated measures ANOVAs. They did not comment on the normality of their data. Shibata et al. and Sheedy and Bergtrom[14] both assumed that their symptom data were not normally distributed. That is, they were assumed to be skewed. They both adopted nonparametric analyses such as the Freidman analysis, which is a nonparametric alternative to the repeated measures ANOVA. Neither of the latter two studies used the "difference" method for calculating stimulus-induced symptoms, but rather, only used post-test surveys, which were equivalent to our post-test symptom surveys. In light of these methodological differences, we elected to take a middle path of using parametric tests where the data were reasonably normally distributed, and Freidman analysis where the data were not normally distributed. Our normality decisions were based on the skew parameter of a tested data distribution as determined by StatPlus statistics software. A skew of zero is purely normal while Keim[17] suggests that a skew within the range of -1 to +1 may be considered sufficiently normal for parametric analysis. Skew values outside that range necessitate nonparametric analysis. We used a Freidman analysis when the data contained one or more variables skewed out of the range of normality, otherwise we used repeated measures ANOVAs.

Results:

In summary, we found that the magnitude of 3D vergences increased with fixation target size, similar to vergences stimulated by different fixed-size targets. While discomfort was caused by sustained fusion of our 3D experimental targets moving in depth, that discomfort was not significantly influenced by target size, contrary to our hypothesis.

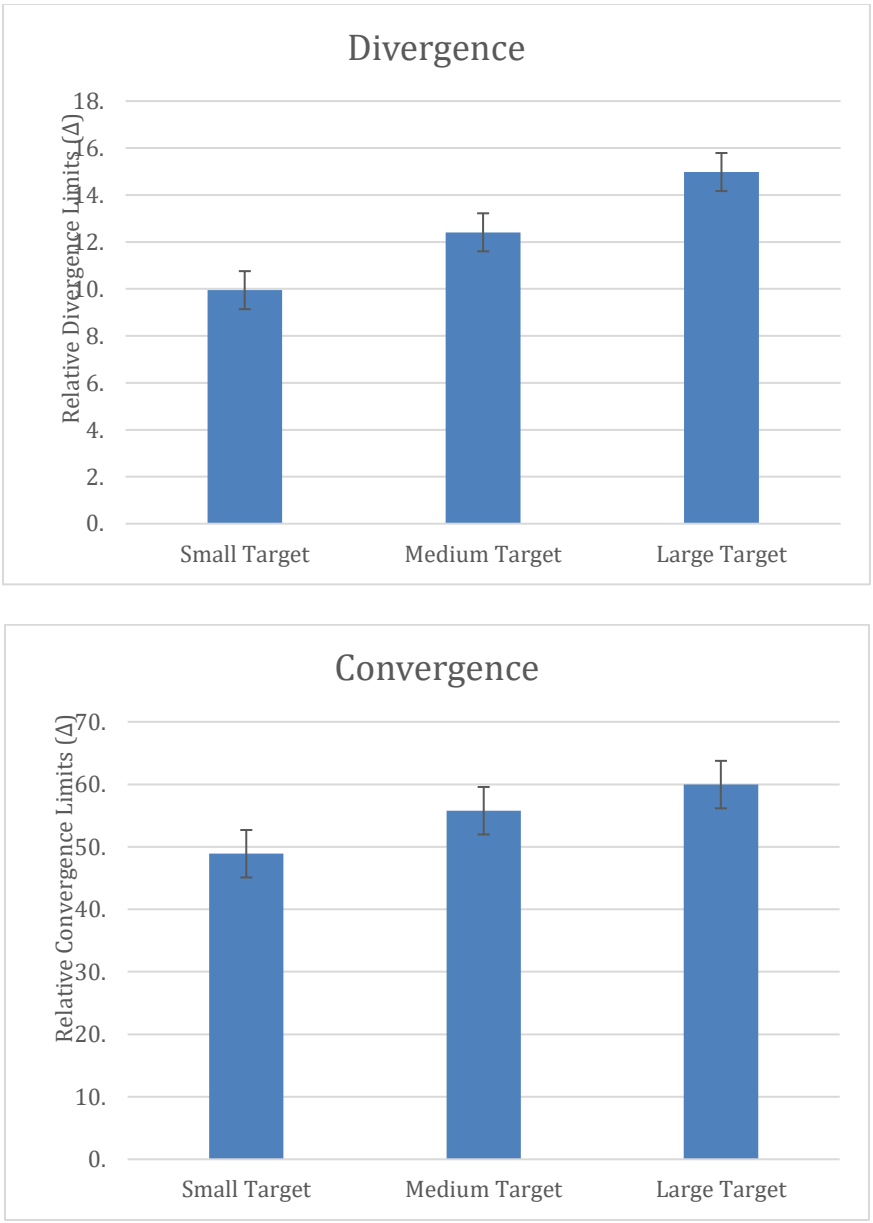
A. 3D vergences and Target Size

Analysis of our 3D convergence and divergence magnitudes among 25 subjects (14 females and 11 males) supported the hypothesis that these data were normally distributed and that there were no significant differences in standard deviation by target size within each type of vergence. Separate repeated measures ANOVAs for convergence and divergence were applied because convergence variability is much greater than divergence variability. Vergence limits varied significantly with target size in convergence ($F=16.8$, $p=0.0000028$) and divergence ($F=84.3$, $p=0$) (Fig. 4). These vergence magnitude differences indicate that there should be differences of fusional vergence gain operating between the experimental conditions. That is, fusional vergence will be stronger in the large target experimental conditions than in the small.

Comparison of Fig. 4 with Fig. 1 shows that the impact of target size on vergence range is very similar whether the targets are fixed size (Fig. 1) or size is changing congruently with disparity (Fig. 4). The divergences are very similar between the studies. In

contrast, the convergence difference between small and large targets is somewhat less in this study.

Figure 4. Relative vergence limits in this study (in prism diopters (Δ)) are plotted as functions of vergence type and target size. Error bars are ±1 standard error from the mean.



B. Stereo Vigilance Task Performance during Experimental Sessions

Pilot studies found that subjects who frequently reported double vision (and were not entered into the study) performed below 70% correct on our stereo vigilance task, so it was assumed that the stereo task was a good monitor of vergence accuracy. As expected by our protocol, all accepted subjects exceeded 70% correct in all conditions. Most subjects had the same stereo performance during convergence and divergence, while some had more difficulty in convergence or in divergence. Mean and median stereo performances in divergence were 96.7% and 100%, respectively, and ranged from 75% to 100%. Mean and median stereo performances in convergence were 97.4% and 100%, respectively, and ranged from 83% to 100%. The distributions of percent correct values were skewed toward 100% in both convergence and divergence. Because stereo performance scores were identically skewed in convergence and divergence, we felt comfortable analyzing the impact of target size and experimental session number on stereo performance by repeated measures ANOVA. There was no significant difference in stereo performance by target size (divergence: $F=0.38$, $p=0.68$; convergence: $F=0.0007$, $p=0.999$) or by session number (divergence: $F=0.63$, $p=0.53$; convergence: $F=1.69$, $p=0.19$).

C. 3D Discomfort

We initially applied Friedman ANOVAs to subjects' "difference scores" because the difference scores were skewed. The mean scores and p values are shown in Table A. While mean values suggest that symptoms were worse for the small target, Friedman p values did not support a significant effect of target size on any symptoms. The reason for the nonsignificant Friedman test result can be visualized by scanning the difference eyestrain scores of all 25 subjects in Table B. A significant Friedman test supporting our hypothesis would occur if symptoms were greater for the small target and lesser for the medium and large targets in most subjects. However, only 10 of 25 subjects showed that pattern of scores in Table B.

Table A

Friedman Analysis of 3D Symptoms by Target Size, based on 'difference" symptom ratings.

	Tired	Backache	Eyestrain	Headache
small	20.08	15.12	25.32	13.08
medium	15.44	10.4	16.32	6.44
large	14.24	10.6	13.68	8.64
p value	0.555	0.137	0.648	0.957

Table B. Difference eyestrain scores for 25 subjects

Small	Medium	Large
28	31	10
-1	7	8
45	12	1
9	5	10
21	-5	23
22	1	-1
47	-4	3
70	42	11
44	16	8
1	24	-2
-1	6	8
11	23	14
-1	-1	24
49	23	0
76	0	48
72	75	47
17	26	27
22	0	28
20	43	2
22	25	26
3	-1	2
-1	17	45
15	5	-1
27	32	-4
16	6	5

The higher mean symptom scores for small targets in Table A suggests that the difference between the small target symptoms and those of the larger targets might be significant if tested by parametric statistics. We tested the effect of target size on symptoms using a “within-subjects” repeated measures parametric ANOVA on our skewed symptom difference values, obtaining p values of 0.41 (tired), 0.36 (backache), 0.04 (eyestrain), and 0.18 (headache). The eyestrain finding is “significant” at the 0.05 level.

The marginally significant eyestrain results from our parametric ANOVA, taken in context with the lack of significant target size effects on eyestrain by Friedman analysis, led us to look for the source of our supposedly “significant” eyestrain ANOVA. We ran a parametric ANOVA on the log of our eyestrain data. Our hypothesis for this test is that the initial low eyestrain p value was an artifact caused by skew in the difference data, and that the log transformation should reduce the skew and therefore raise the eyestrain p value to a less significant level. We observed that the skew parameter dropped below 1.0 for all eyestrain log target size variables, normalizing the eyestrain difference data. p was elevated to a less significant value of 0.158 by the log transform. Skew in the linear data was at least partly responsible for the “significant” (p=.04) result in our “difference” eyestrain data.

The presence of skew was not limited to the eyestrain data. We have tabulated the skew of all linear pretest, posttest, and difference symptoms in Table C below.

Table C. Skew values of distributions of symptom data by symptom, test type, and target size.

Target		Tired			Backache			Eyestrain			Headache	
	Pretest	Posttest	Difference	Pretest	Posttest	Difference	Pretest	Posttest	Difference	Pretest	Posttest	Difference
Small	0.67831	0.74186	0.28755	0.99855	0.67091	0.83519	0.70861	0.51101	0.82522	0.91051	1.06679	1.63551
Medium	0.33166	0.12083	1.13649	-0.12532	0.35995	1.08552	0.07963	0.28852	1.3037	0.89132	0.61141	1.4156
Large	0.69348	0.24569	0.48793	0.66817	0.52336	0.58514	0.24813	0.38805	0.99904	0.91807	1.04921	0.77305

Difference symptom data were significantly skewed for all four symptoms. On the other hand, the pretest and posttest symptom measures were not skewed for the Tired, Backache, and Eyestrain data and were less skewed for Headache. Therefore, the Posttest-Pretest subtraction operation skewed most difference measures for all four symptoms. It therefore seems that the skew effect will falsely reduce the p values of all parametric tests of difference scores, for all types of symptoms. Therefore, we cannot use parametric analysis of symptom difference scores to test our main hypothesis that target size affects symptoms.

The skew problem is not the only problem in our symptom data. A casual review of our raw pretest and posttest symptom data suggested that pretest symptoms might be much lower for the small target sessions. Therefore, we ran a repeated measures ANOVA on the pretest and posttest eyestrain data to verify these casual observations. The parametric ANOVAs of the pretest and posttest eyestrain data are justified by the acceptable pretest and posttest symptom skew values noted in Table C. Table D compares these pretest and posttest symptom survey ANOVAs to the ANOVA of the difference eyestrain data. The pretest symptoms for the small target were significantly less than those of the medium and large targets and there was no difference between the three posttest symptom survey values.

Table D. Summary of a repeated measures ANOVA of pretest, posttest, and difference symptom values for the eyestrain symptom, by target size, with F and p values.

Target Size	Mean pretest symptoms	Mean posttest symptoms	Mean of difference symptoms
small	16.28	42.56	25.32
medium	25.88	42.2	16.32
large	28.8	42.48	13.68
	F=3.86; P=0.028	F=0.00256, P=0.997	F=3.37, P=0.042

We claim that the significant differences between the pretest eyestrain symptoms for different target sizes must be a target-related subject bias because the pretest symptoms were measured before subjects saw the experimental session targets, and therefore the pretest symptoms cannot be attributed to our subjects having seen the visual stimuli in experimental sessions. We have labeled this bias a small target bias because the medium and large target pretest symptom scores were similar, as expected, while the small target pretest symptom score was out of line with the others. In principle, all three pretest symptoms should be the same. It cannot be denied that it is possible that the small target pretest scores are correct and that the medium and large target pretest scores are both biased. While some random variation of pretest symptoms within and between subjects is to be expected, this result appears to be beyond random variation.

Table E. Mean pretest symptoms for all symptom types and target sizes, including ANOVAs.

	Tired	Backache	Eyestrain	Headache
Small	17.76	16.67	16.28	12.25
Medium	26.72	23.67	25.88	23.33
Large	29.96	24.54	28.80	20.5
ANOVA	F=6.67, p=0.0027	F=2.66, p=0.080	F=3.863, p=0.0278	F=3.99, p=0.025

The bias evident in the pretest eyestrain data also significantly affected the other symptoms, with the exception of a backache trend (Table E). The bias is remarkably consistent across symptom types, suggesting that it is not the symptoms themselves that the pretest surveys are measuring, but rather, knowledge of the experimental session (small, medium, or large) that will follow a pretest survey. This will be addressed in the Discussion section.

In comparing the mean pretest, posttest, and difference scores in Table D it is evident that the mean pretest symptom levels, when subtracted from the mean posttest symptom levels, very closely predicts the mean difference symptom levels for each target size. In particular, the small target pretest symptoms are low by the amount that the small target difference symptoms are high. In other words, it appears that the small target symptom bias almost completely explains the relatively high small target difference symptoms.

It might be suggested that the small target bias in the eyestrain data caused the skew of the eyestrain difference data. However, the eyestrain data for the other target sizes does

not support that hypothesis. For instance, the medium target eyestrain data does not have any apparent bias in its pretest data, and the skew values of the medium target pretest data (.08) and posttest data (.39) were normally distributed. Nonetheless, when the pretest medium target symptoms were subtracted from the posttest medium target symptoms, the medium target difference symptoms were skewed with a skew value of 1.3. The skew in this example was caused by the mathematical differencing operation, not the small target bias.

The above eyestrain data raises the question of how our subjects could report the same posttest eyestrain symptoms across target size after reporting significantly lower pretest eyestrain symptoms for the small target size. We suggest two hypotheses on how this could happen: (1) We first assume that subjects' pretest symptoms for each target size are remembered and added to stimulus-induced symptoms to yield the posttest symptoms. The differences between the pretest symptoms are almost exactly offset by differences of stimulus-induced symptoms between target sizes to produce the same posttest symptom reports we observed. (2) We assume that biases are volatile and can be dropped in the face of discomfort experience contrary to the bias. Subjects have the same posttest target-induced discomfort across target size and answer accordingly after discarding their pretest biases. The law of parsimony would urge that the only reasonable choice between these hypotheses is the second hypothesis. We conclude that the posttest symptom values are unbiased and represent how subjects felt after completing the experimental sessions.

The backache data, which was a control variable in this study, support the hypothesis that our subjects ignored pretest symptom biases when they answered posttest questions.

Table F shows our backache symptom data with the same layout as the eyestrain data in Table D.

Table F. Summary of a repeated measures ANOVA of pretest, posttest, and difference symptom values for the backache symptom, by target size, with F and p values. The difference symptom ANOVA values are inaccurate because of data skew.

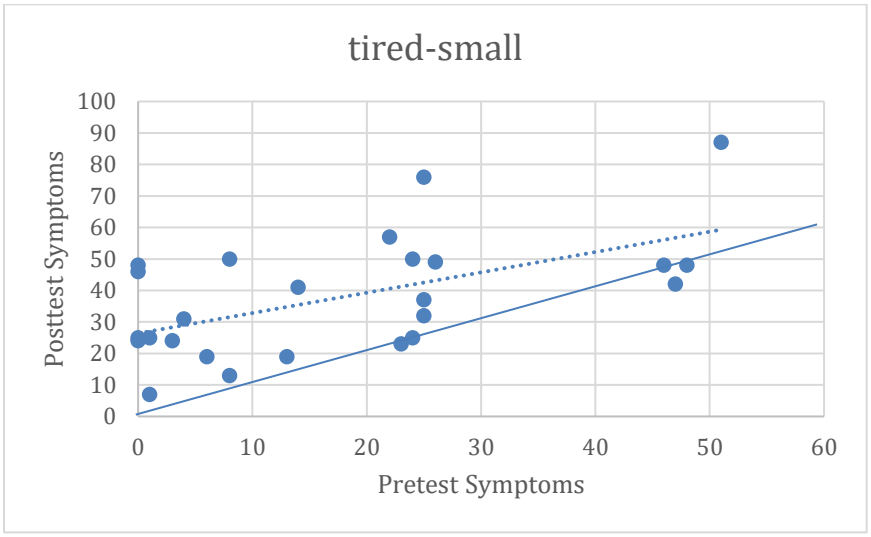
Target Size	pretest symptoms	posttest symptoms	difference symptoms
small	16.16	32.36	15.12
medium	23.17	34.64	10.4
large	24.54	35.8	10.6
	F=2.66; P=0.08	F=0.4723, P=0.626	F=1.04, P=0.35

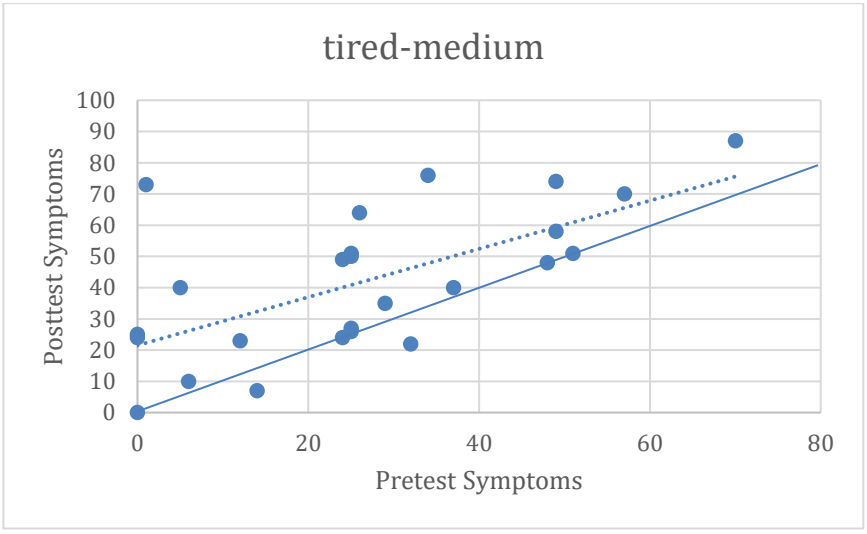
There is no significant difference between posttest backache symptoms as a function of target size, as in the eyestrain posttest data. The backache data refute the first hypothesis for explaining the lack of difference between the posttest values because there is no rationale that the backache stimulus conditions in our study could differ by target size to cancel out pretest biases. The backache data argue that we must accept the second hypothesis that subjects drop their pretest biases in the face of equal discomfort experiences across target size conditions.

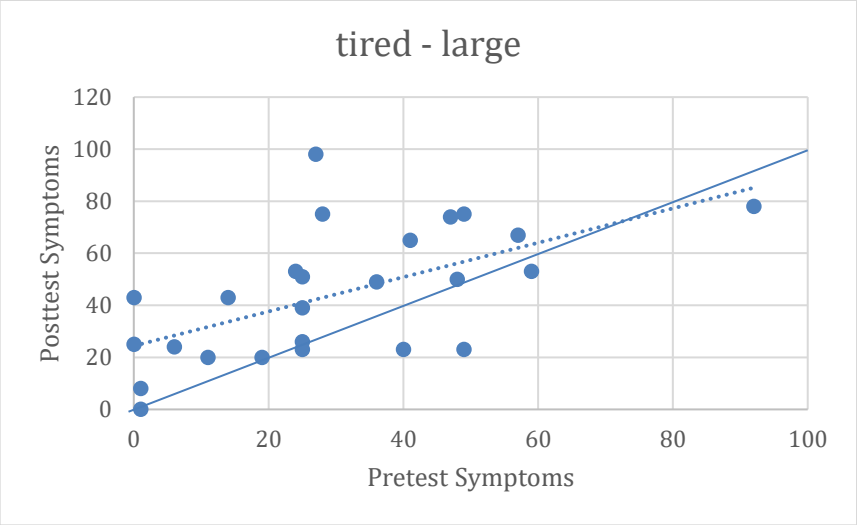
The validity of the difference calculation of symptoms also depends on the assumption that the pretest symptoms are stable throughout a session and add to symptoms

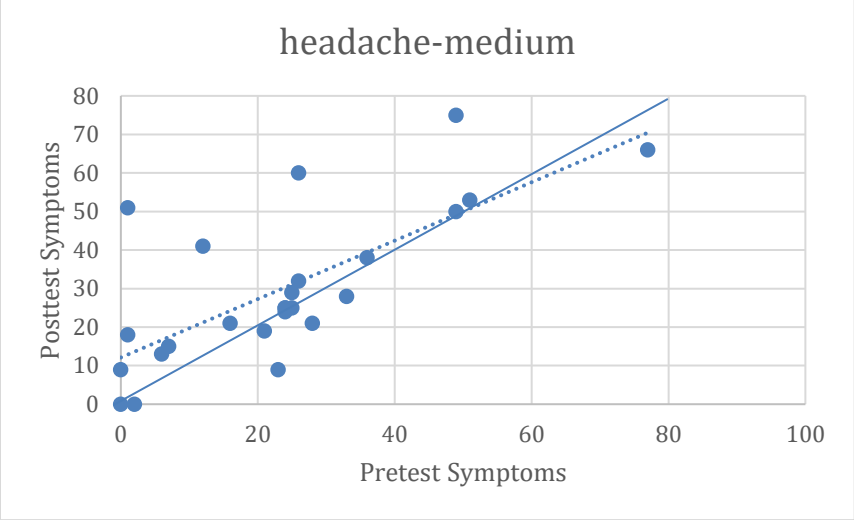
generated by that session. The following data show that this assumption is not necessarily true.

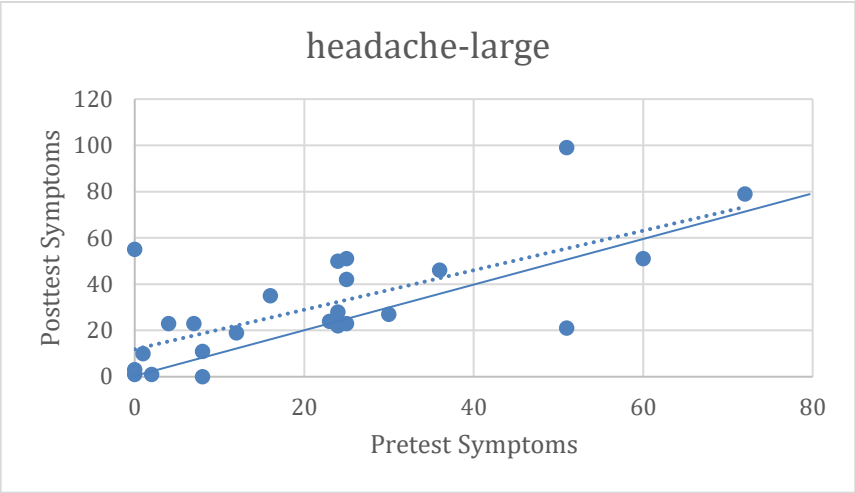
Figure 5. Graphs of posttest symptoms versus pretest symptoms for the tired and headache symptoms for 3 target sizes: small, medium, and large.






tired - large
Posttest Symptoms
Pretest Symptoms
0
20
40
60
80
100
120


headache-medium
Posttest Symptoms
Pretest Symptoms
0
10
20
30
40
50
60
70
80
100



The solid line in each graph is the 1-to-1 line. Data points on the 1-to1 line indicate that posttest symptoms equaled pretest symptoms. That is, the experimental condition induced no symptoms. Most subjects in most conditions displayed posttest symptoms greater than pretest symptoms, as one would expect for pretest-posttest additivity. We show here only the five graphs for those symptoms and target size where data points are below the 1-to-1 line. The dotted line is the linear regression for all 25 subjects. Some subjects' data fall well below the 1-to-1 line suggesting the unlikely scenario that our experimental condition reduced the severity of the symptom relative to its severity at the start of the test (e.g., like taking aspirin). A more likely explanation of such a result is that the subject's pretest symptom dissipated during the experimental session, and that the stimulus itself did not induce that particular symptom. Dissipation of a pretest symptom throughout a session violates the difference method assumption that pretest symptoms are stable throughout a session and add to symptoms generated in the session. These results show that sometimes, the additivity assumption about pretest and stimulus-induced symptoms fails for some subjects.

Given the presence of bias in pretest surveys, skew in difference values, and non-additivity of pretest and posttest values, we reran our statistics on the posttest data, which were mostly not skewed and contained no clear bias and were not subject to symptom additivity problems. Table G shows the mean symptom ratings from posttest symptom surveys, including the results of parametric and Friedman ANOVA analysis, for each symptom and target size. It is abundantly clear from these data that there is no effect of target size on any of these symptoms by either analysis. The visually connected tired and eyestrain symptoms were relatively more severe than the others. The headache symptom was relatively milder than the other three symptoms, and its stronger skew suggests that most subjects were probably not experiencing headache as a result of the tasks. This behavior is similar to Shibata et als.' finding that headache is only weakly related to 3D discomfort. The backache symptom does not change with target size condition and is almost certainly caused by sitting 20 minutes on a stool with head in a chinrest.

Table G. Mean posttest symptoms for all symptom types and target sizes, including statistical analysis.

	Tired	Backache	Eyestrain	Headache
Small	37.84	32.36	42.56	26.68
Medium	42.16	34.64	42.2	29.88
Large	44.2	35.8	42.48	30.72
Parametric	F=0.84, p=0.84	F=0.47, p=0.63	F=0.00256, p=0.997	F=0.496, p=0.61
Friedman	chiSq=.27, p=0.87	chiSq=0.53, p=0.77	chiSq=1.37, p=0.50	chiSq=1.31, p=0.52

We do not yet know the source of the small target bias. There are two common sources of bias that might be related to our small target bias: test time-of-day, and experimental session order. We first looked at the time-of-day a test was run as a factor in causing differences of comfort in the pretest surveys. Our protocol had called for running each session as close to the same time of day as possible, preferably in the morning. This was not always achievable given the limited availability of subjects. The question is: did the small target sessions occur mainly in the morning before subjects experience visual fatigue, while medium and large target sessions occurred later in the day when subjects might be expected to have more visual fatigue? If so, we might expect subjects to be less fatigued when filling out the pretest survey before testing with the small target. We collected the PC date and time-of-day data for each symptom session file, which were generated when each PC symptom session file closed. We then ran a parametric one-way ANOVA of session time by target size. Mean session times were 1:11PM (small target), 1:29PM (medium target), and 1:40PM (large target), and ranged from 10:59AM to 7:57PM. The differences between mean session time were not significant ($F=.021$, $P=0.81$). We conclude that session time-of-day did not bias our pretest symptom surveys.

We then investigated experimental session order as a potential cause of bias. For example, if subjects were to become progressively more bored as they passed through our three experimental sessions, they might rate the later sessions as more fatiguing. If the small target appeared mostly in the first experimental session, then we might then expect lower pretest symptoms for the small target. Note that we randomly assigned target sizes to

experimental sessions for each subject to prevent a session order effect. If our sessions prove reasonably randomized, there should be no order effect. Table H shows how many times each target size was presented in experimental sessions 3, 4, and 5. Overall, the three target sizes were well distributed across sessions, though not identically. It appears to us that the session order was sufficiently randomized.

Table H. The distribution of target size presentations presented to 25 subjects during experimental sessions 3,4, and 5.

Session 3	Session 4	Session 5
small = 8 times	small = 11 times	small = 6 times
medium = 10 times	medium = 6 times	medium = 9 times
large = 7 times	large = 8 times	large = 10 times

One final question is whether subjects experienced any symptoms at all because they viewed our experimental targets. This question is highly relevant because if none of our stimuli cause symptoms, the target size condition would not matter. This question can be answered by using the nonparametric sign test to compare all posttest symptoms (collapsed over symptom type) to all pretest symptoms (collapsed over symptom type). The bias in the pretest values would only affect the magnitude of pretest to posttest differences but is unlikely to affect the sign of the difference between posttest and pretest surveys. The average symptom score difference between posttest and pretest symptoms was 14.1 (out of 100), the sign test z score was -10.45465, and p was not different from zero (highly significant). We conclude that our experimental conditions did cause significant symptoms, but they did not differ by target size.

Discussion

This study showed that target size did not differentially affect 3D visual comfort even though our vergence stimuli did cause 3D visual discomfort. Shibata et al. reported that for their posttest surveys of the tired and eyestrain symptoms, subjects experienced symptom severities of approximately 50% of full scale. In comparison, our subjects reported approximately 40% of full-scale symptoms for posttest symptoms. We conclude that our conditions produced strong enough symptoms to reveal a significant target size effect if target size was a significant factor.

We conclude that target size is not a concern for causing symptoms for those dealing with 3D content, at least within the constraints of normally binocular subjects tracking smooth stereoscopic 3D motion in depth in clear and single vision. Most 3D viewing is unlikely to be as rigorous as our 20 minutes of continuous viewing of motion in depth within the limits of clear and single vision, so we are confident in our suggestion that the target size factor is not a serious concern for causing oculomotor discomfort symptoms.

Why didn't target size affect symptoms in this study? There were three possible sources of visual discomfort in this study, each of which could have caused symptoms that were equal across target size: (1) accommodation/convergence mismatch (i.e., how much accommodation differed from convergence), (2) convergence angle (i.e., convergence without regard to accommodation), and (3) visual motion-in-depth.

First, accommodation/convergence mismatch was the target of this study. Like heterophoria, it can place uncomfortable loads on fusional vergence which Sheard's criterion

can assess. Target size can affect fusional vergence reserves and, therefore, should affect resistance to discomfort. It appears that our Sheard's criterion standard for constructing our stimuli failed to cause target size-induced differences of accommodation/convergence mismatch discomfort. An interesting Convergence Insufficiency Treatment Trial study group finding might be instructive at this point[18]. They observed that once convergence insufficiency patients demonstrate all three basic signs of convergence insufficiency - high near exophoria, low near fusional vergence amplitude, and receded near point of convergence - the severity of the patient's symptoms were no longer proportional to the magnitudes of those three signs when they exceeded the level required for basic diagnosis. They claimed that this behavior implied that symptom severity saturates once the clinical signs reach a certain level of defect. This is not to say that our subjects have a binocular anomaly, but rather that under our conditions, normals may have experienced symptoms of discomfort because of accommodation/convergence mismatch that saturated at the same level for all target sizes and hence neutralized target size-related discomfort differences.

Second, it is also conceivable that subjects in our study experienced discomfort just because they were using repeated dynamic convergence over 20 minutes. Shibata et al.[2] and Kim et al.[15] both reported such convergence angle discomfort in conditions where subjects viewed targets containing congruently changing accommodation and convergence. This type of discomfort is analogous to the discomfort experienced when performing prolonged pencil-pushup vision training exercises, which stimulate accommodation and convergence congruently. This source of vergence discomfort would not be expected to correlate with Sheard's Criterion and would not be responsive to target size. We did not test for discomfort

related to congruent changes of accommodation and convergence because we did not have the equipment to do so.

The third factor that might induce discomfort in our study is motion-related discomfort caused by the combination of the looming depth cue and the changing convergence angle depth cue connected with our experimental targets. The looming depth cue is information about changing depth caused by changing target size. The changing convergence angle cue is information about motion-in-depth caused by changing convergence angle. Additionally, motion parallax created by head movement during 3D viewing creates illusory motion-in-depth that conflicts with the other depth cues. 3D motion discomfort can happen when motion cues conflict. Therefore, our subjects may have been experiencing some degree of visual motion-in-depth discomfort. This source of discomfort in our study would not be surprising because visual motion discomfort, known as visually induced motion sickness (VIMS), is a known component of 3D viewing discomfort, especially for motion-in-depth[19]. We would not expect this proposed motion-in-depth discomfort to vary with target size, which would be in accordance with our results. In suggesting that our subjects were experiencing visual motion-in-depth discomfort it is assumed that subjects reported their discomfort indirectly through the tired eyes and eyestrain symptoms. There are survey questions that more directly assess visual motion discomfort, but they were not used because we were focused on accommodation/convergence mismatch discomfort.

It is possible that all three of these sources of discomfort contributed to our subjects' responses. For instance, the convergence angle and looming causes of discomfort, by

themselves, may have been sufficient to cause symptoms. Any accommodation/convergence mismatch discomfort and other causes of discomfort could have pushed our subjects' responses into discomfort saturation. The predicted result would be symptoms unrelated to the target size.

We noted in raw data samples, even in this carefully selected normally binocular subject population, that while some subjects were more symptomatic with smaller target size (the accommodation/convergence mismatch hypothesis), others' discomfort did not change with target size (the convergence angle hypothesis), while a third group experienced higher symptoms with larger target size (the visual motion discomfort hypothesis). If these individual behaviors were repeatable in others, it would imply that different subjects might not respond the same way to these three types of 3D discomfort stimuli. That is a question for future research to answer.

Finally, it may be asked how we failed to find target size effects on symptoms – which are logically related to accommodation/convergence mismatch, while Shibata et al. did find symptoms related to accommodation/convergence mismatch. We believe that the reason for this is methodological differences. Their targets were 4° diameter random dot stereograms, which maintained a constant visual angle through all depth conditions. This constant visual angle would suppress motion-in-depth perception that might otherwise arise from stereopsis and convergence corollary discharge. Moreover, their random dot stimulus would be a more demanding binocular target than our Maltese cross, requiring precise ocular alignment, and therefore penalizing poor oculomotor performance with strong symptoms in the face of accommodation/convergence mismatch. Our targets, on the other hand, subtended

visual angles, which covaried with disparity, as in naturalistic imagery and most 3D content. Our images generated strong motion-in-depth perception and, therefore created the potential for visual motion sickness symptoms. Our formed outlined images would be arguably more robust in maintaining clarity and fusion in the face of small occasional ocular misalignments and, therefore, not as perturbed by accommodation/convergence mismatch. Putting these differences together, we concluded that our targets were less likely to reveal symptoms from accommodation/convergence mismatch and more likely to reveal symptoms from visual motion sickness, which may partially explain the differences between our studies.

The source of the bias observed in our pretest symptom surveys is unknown, though we have hypotheses about how it might have developed. First, it must be recognized that the bias in our pretest data that small targets were rated less symptomatic than large was the opposite to our experimental hypothesis, which held that small targets should be more symptomatic than large. The “opposite” nature of the bias indicates that subjects probably did not inherit the bias from discussions with the experimenters.

Subjects needed two types of information to reveal a bias in the pretest surveys: (1) knowledge of which target size should be, in their minds, more symptomatic, and (2) knowledge of which target size was going to be used in the experimental session immediately following a pretest symptom survey.

We believe that the vergences session (the second session) caused the bias of which target should be rated as more symptomatic. During the vergence sessions, subjects underwent repeated convergence and divergence episodes that ranged from blur to diplopia. Most everyone who experiences these types of tests regards them as uncomfortable, whether

in a clinical testing environment or in our experimental apparatus. Because subjects performed vergences with all three experimental target sizes, they gained an experience in the relationship of target size with discomfort during vergence testing. The type of bias learned - that large targets were more uncomfortable than small – may have developed because more time and effort were spent converging than diverging because convergence has more range than divergence, and 3D targets become very large during 3D convergence. Hence, subjects spent most of their time during the vergences feeling uncomfortable while watching large targets. We suggest that they carried this experience with them to the experimental sessions. Experimental practice sessions were also run before experimental sessions. Still, they could not have caused an association of discomfort with target size because the vergence task during practice was too brief (one oscillation) to cause discomfort. A bias from knowledge gained in the vergences session is inherent in our experimental design because we had to measure the vergences before the experimental sessions to calculate the appropriate vergence demands for the experimental sessions. Moreover, we had to use the same targets in vergences and experimental sessions because vergence stimulus spatial characteristics significantly affect vergence behavior.

We believe that subjects learned which target size they were to see after a pretest survey from our practice sessions. For the most part, practice sessions were intended to be done just once before the experimental sessions so as to acquaint subjects with the stereoscopic vigilance task and how to enter responses to it. MN recalls that approximately 60% of subjects only needed one showing of the practice session to learn the task. Still, approximately 40% of subjects needed a redo of the practice session before each

experimental session because their sessions were separated by multiple days. The practice software, which allowed the selection of target size for demonstration purposes, was used to select the same practice target size as in the following experimental session. For those 40% of subjects who saw the practice session before each experimental session, the practice would have yielded a strong cue to the target size to appear in the experimental session following the pretest survey. This practice information would have allowed subjects to apply the bias developed in the vergences session to the pretest surveys.

Limitations

Our study population might not be as sensitive to binocular discomfort as the general population or subjects in other studies. We scrupulously removed subjects from our study population who showed evidence of being outside of the clinically normal binocular population. We note that Shibata et al's study, on the basis of phorias and vergences published in their study, included subjects who would not have been admitted to our study. Eleven of their 24 subjects appeared to have either convergence insufficiency (1), fusional insufficiency (5), or basic esophoria (5). Those eleven subjects would have been more prone to convergence-related symptoms than our subjects. Another limit of our conclusions is that they apply to dynamic convergence to motion-in-depth and might not correlate to comfort during convergence to static targets at various distances.

Conclusions:

Based on the comprehensive analysis of the results and discussions presented, several critical insights emerge regarding the relationship between target size, dynamic 3D vergences and 3D discomfort. The findings reveal that while dynamic 3D vergences exhibit a proportional increase with the size of dynamic fixation targets, discomfort induced by sustained fusion of 3D targets in motion remains largely unaffected by target size, contrary to our initial hypotheses. Despite marginal significance observed in “difference” eyestrain symptoms through parametric analysis, subsequent scrutiny revealed issues of skewness in the data, potentially influencing interpretation. Notably, the presence of biases in pretest ratings, wherein subjects reported lower symptoms for small targets, highlights the importance of considering some adjustments for continuing studies. Ultimately, the examination of posttest symptom surveys confirmed the lack of significant effects of target size on discomfort symptoms.

References:

1. Hoffman DM, Girshick AR, Akeley K, and Banks MS. Vergence-accommodation conflicts hinder visual performance and cause visual fatigue. J of Vision (2008) 8(3):33, 1-30, doi:https://doi.org/10.1167/8.3.33

2. Shibata T, Kim J, Hoffman DM, Banks MS. The zone of comfort: Predicting visual discomfort with stereo displays. Journal of Vision; 2011 11(8):11.doi:https://doi.org/10.1167/11.8.11

3. Sugita, N., Yamaga, T., Yoshizawa, M., Abe, M., Tanaka, A., Yambe, T., & Nitta, S.-I. (2014). Visual fatigue induced by accommodation convergence mismatch while viewing three-dimensional television. 2014 IEEE 3rd Global Conference on Consumer Electronics (GCCE), 250–251.

4. McCormack, G. L., & Hogan, K. M. (2019). Clear and single binocular vision in near 3D displays. Optometry and Vision Science: Official Publication of the American Academy of Optometry, 96(11), 850–859. https://doi.org/10.1097/OPX.0000000000001439

5. Kim, Y., Kim, J., Hong, K., Yang, H. K., Jung, J.-H., Choi, H., Min, S.-W., Seo, J.-M., Hwang, J.-M., & Lee, B. (2012). Accommodative response of integral imaging in near distance. Journal of Display Technology, 8(2), 70–78. https://doi.org/10.1109/jdt.2011.2163701

6. Soomro, S. R., & Urey, H. Visual acuity response when using the 3D head-up display in the presence of an accommodation-convergence conflict. Journal of Information Display,2020;21(2), 93–101. https://doi.org/10.1080/15980316.2019.1697766

7. McCormack, G. L., & Kulowski, K. A. (2021). Image size and the range of clear and single binocular vision in 3D displays. Optometry and Vision Science: Official Publication of the American Academy of Optometry, 98(8), 947–958. https://doi.org/10.1097/OPX.0000000000001745

8. Sheard C. The Prescription of Prisms. Am J Optom, 1934;11:364–78.

9. Percival AS. The Relation of Convergence to Accommodation and its Practical Bearing. Ophthalmological Review 1892;11:313–28.

10. Feldman JM, Cooper J, Eichler R. The Effect of Stimulus Parameters (Size, Complexity, Depth, and Line Thickness) on Horizontal Fusional Amplitudes in Normal Humans. Binocular Vision and Eye Muscle Surgery Quarterly 1993;8:23-30.

11. Jones R, Stephens GL. Horizontal Fusional Amplitudes. Evidence for Disparity Tuning. Invest Ophthalmol Vis Sci 1989;30:1638-42.

12. Hofstetter HW. The Zone of Clear Single Binocular Vision. Arch Am Acad Optom 1945;22:301-61.

13. Labishetty V, Cholewiak SA, and Banks MS. Contributions of the fovea and non-foveal retina to the eye's focusing response. Journal of Vision (2019) 19(12):18, 1–

14. Sheedy J and Bergstrom N. Performance and Comfort on Near-Eye Computer Displays. Optometry and Vision Science (2002); 79(5): 306-312.

15. Kim, J., Kane, D., & Banks, M. S. (2014). The rate of change of vergence-accommodation conflict affects visual discomfort. Vision Research, 105, 159–165. https://doi.org/10.1016/j.visres.2014.10.021

16. Kuze J. and Ukai K. Subjective evaluation of visual fatigue caused by motion images. Displays (2008); 29: 159-166.

17. Keim, Robert. Understanding Parametric Tests, Skewness, and Kurtosis. All About Circuits, August 24, 2020. https://www.allaboutcircuits.com/technical-articles/understanding-the-normal-distribution-parametric-tests-skewness-and-kurtosis/

18. Bade A, Boas M, Gallaway M, Mitchell GL, Scheiman M, Kulp MT, Cotter SA, and Rouse M. Relationship Between Clinical Signs and Symptoms of Convergence Insufficiency (2013);9099):988-995.

19. Howarth PA. Potential Hazards of Viewing 3-D Stereoscopic Television, Cinema and Computer Games: A Review. OphthalPhysiolOpt2011;31:111-22.