UC Berkeley

UC Berkeley Electronic Theses and Dissertations

Title

Eye Movements in Amblyopia and the (un)natural Statistics of Eye Movements and Binocular Disparities in VR

Permalink

https://escholarship.org/uc/item/2vj1h9xn

Author

Aizenman, Avigael

Publication Date

2021

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Eye Movements in Amblyopia and the (un)natural Statistics of Eye Movements and Binocular Disparities in VR

by

Avigael M. Aizenman

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Vision Science

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Dennis M. Levi, Chair
Professor Michael Silver
Professor Dor Abrahamson

Fall 2021

Eye Movements in Amblyopia and the (un)natural Statistics of Eye Movements and Binocular Disparities in VR

Copyright 2021
by
Avigael M. Aizenman

Abstract

Eye Movements in Amblyopia and the (un)natural Statistics of Eye Movements and Binocular Disparities in VR

by

Avigael M. Aizenman

Doctor of Philosophy in Vision Science

University of California, Berkeley

Professor Dennis M. Levi, Chair

Advances in technology allow us to track gaze across multiple environments, from the laboratory to the natural world as well as in virtual reality. This has far reaching basic and clinical implications. This dissertation explores both. Our eyes are constantly in motion, helping us piece together a coherent visual perception of the natural world around us. Even during fixation, when the eyes are focused on an object of interest, tiny fixational eye movements shift the position of the image landing on the retina. Our oculomotor system is adapted to the regularities in our natural world, and as a result our visual system functions quickly, comfortably, and accurately in the natural environment. Previous work has shown oculomotor behavior is abnormal in visual disorders such as amblyopia. Chapter 2 of this dissertation describes a study to measure fixational stability in children with amblyopia currently undergoing treatment. This work demonstrates that fixational eye movements can be used as a nonverbal metric for the recovery of visual function in amblyopia. Chapter 3 measures the statistics of fixation and binocular disparity in the natural environment as well as virtual reality (VR) headsets. This work suggests that there are key differences in the statistics of fixation and disparity between the natural and VR environments. The vertical horopter and natural disparity statistics show a top back pitch in the natural environment. This top-back pattern is lacking in the disparity statistics of the virtual environment, and the differences are quite dramatic (in the lower visual field the natural environment has disparities that are much nearer than in VR, an effect of around 900 arcsec!). We additionally found that gaze is biased towards straight ahead viewing and farther fixation distances in VR, as compared to the natural environment. Chapter 4 follows up on the findings from the third chapter, and tests whether observers prefer content that is congruous with the statistics of the natural world in virtual reality. This work found that scene content that violates the top-back pitch pattern of the natural world leads to more discomfort. This suggests that the mismatch between the statistics of the natural and VR-environment documented in chapter 3, is leading to conflict and discomfort in VR headsets.

To my parents, Michal and Joshua Aizenman

My earliest memories are full of your encouragement for my curiosity. From helping me find snails in the backyard to examine, to driving me to the library to check out books on dinosaurs, your support bolstered my curiosity and intellectual development. Your love, encouragement, and support has been limitless at every step of this amazing journey. Thank you!

Contents

Contents **ii**

1 Introduction **1**
1.1 The importance of eye movements 1
1.2 Eye movements in the natural and virtual-reality environment 2
1.3 Deficits to oculomotor function in Amblyopia 4
1.4 Study rationale and dissertation approach 4

2 Fixational Stability as a Metric for the Recovery of Visual Function in Amblyopia **7**
2.1 Introduction 7
2.2 Methods 9
2.3 Data Analysis 11
2.4 Results 13
2.5 Discussion 20

3 The (Un)natural Statistics of Eye Movements and Binocular Disparities in VR Gaming Headsets **23**
3.1 Introduction 23
3.2 Methods 24
3.3 Results 31
3.4 Discussion 41
3.5 Conclusion 48

4 Discomfort Associated with the (Un)natural Statistics of VR Gaming Headsets **50**
4.1 Introduction 50
4.2 Methods 52
4.3 Results 53
4.4 Discussion 55

5 Conclusion and Future Directions **59**

Bibliography **62**
5.1 Appendix A: Pilot study I 72
5.2 Appendix B. Pilot Study II 77

Acknowledgments

It takes a village to train a scientist, and I feel that I am 'standing on the shoulders of giants.' My heartfelt thanks go to everyone who has supported me on this journey!

Foremost, I would like to thank my graduate advisor Professor Dennis Levi for his continued support during my PhD. Your steadfast guidance and wisdom have helped me every step of the way. It is easy to get lost in the nitty-gritty details of a project, and you have always helped me to see the bigger picture and the broader clinical impact of the work we do.

A special thank you to Professor Marty Banks for all of his support and wisdom during the last few years. From scientific knowledge to life knowledge, you have been endlessly generous with your mentorship and time. It has been a true honor to learn from you, thank you for being a pillar of support during every step of my PhD!

Great appreciation is due to the excellent mentors and scientist that made up my thesis committee. Thank you Professor Michael Silver and Professor Dor Abrahamson. You have helped guide my development as a scientist and improved my work at each stage of the PhD.

A special thank you is due to my previous research mentors, Professors Jeremy Wolfe and Robert Sekuler. Thank you both for introducing me to vision science and allowing me in all my naivete and excitement to pursue research in your labs.

There are many scientists and peers that I have had the pleasure of meeting and working with over the last few years. You have all contributed to my education in multi-faceted ways. Thank you to Professor George Koulieris and Agostino Gibaldi, it has been an absolute pleasure to work with you both! Thank you as well to my lab mates who have supported me, including Angie Godinez, Lauren Spano, and Sunwoo Kwon.

I would like to thank my family for their endless support. You have been such a positive source of vitality and light in my life, thank you for being my cheerleaders! To my husband Ryan, thank you for supporting me on this journey. You have moved across the country multiple times for my work, all the while allowing me to feel cared for and secure. You strengthen me every day, and constantly inspire me! To my sister Anbar, I have called at all hours of the day (and night), and you have been there to support me, no matter what. You inspire me to 'fight on,' and it is a privilege to have you in my life. To my parents, thank you for being my biggest cheerleaders! You have offered endless love and support, I couldn't have done this without you. Finally, my thanks to Ruggles (the dog) you have spent countless hours sitting by my side as I've written this dissertation. You are the hidden ghost author on all my work.

Chapter 1

Introduction

1.1 The importance of eye movements

Eye movements are critical for piecing together a coherent visual perception of the world that surrounds us. The fovea is a small depression in the retina where visual acuity is the highest. Acuity outside the foveal region degrades dramatically with increasing eccentricity. Saccades are rapid and ballistic eye movements that change our point of fixation, and which place the foveal region at the object of current interest. This ensures visual processing resources are maximized on the area of the visual field with the greatest interest. We make 2-3 saccadic eye movements a second, which are interspersed by fixations, where gaze is fixed on an object of interest. During fixation there is a subjective sense that the eye is motionless and fixed, but in reality the eyes are never truly still. Fixational eye movements are microscopic and largely unnoticed eye movements that are made during fixation. These eye movements include fixational saccades, slow drifts, and high frequency tremors.

Fixational saccades are the largest and fastest of the fixational eye movements, and are small saccades that happen during fixation. These small eye movements contribute to vision in profound and critical ways. Early studies examining the role of fixational saccades found that when the physiological motion of the eye was counteracted (effectively stabilizing the image on the retina), the perceived image would lose contrast and eventually fade away to a homogeneous field [25, 136, 70]. Fixational eye movements contribute to maintaining visibility during fixation by shifting the retinal image to avoid adaptation. This facilitates neural responses in visual neurons that would otherwise adapt to a stationary image. Recent work has shown compelling evidence that fixational eye movements additionally represent a critical stage of information processing, enabling the retina to represent space in a temporal fashion, and to begin feature processing [96].

Another important aspect of vision involves the coordination of binocular eye movements, known as vergence eye movements. Vergence eye movements function to align the fovea of each eye with targets located at difference distances/depth from the observer. Other types of eye movements (such as saccades) are conjugate - the two eyes are moving in the same

direction. Vergence eye movements, in contrast, are considered to be disjunctive as the eyes are rotating in opposite directions. If an observer is tracking an object that is approaching, the eyes will rotate inwards (convergence), while tracking an object that is receding will cause the eyes to rotate outward (divergence). Vergence eye movements facilitate and maintain our perception of the world as a single image. Vergence movements can be triggered by differences between the left and right eye's retinal image [90]. These differences are known as binocular disparity, and can be used to compute very precise depth information about the 3d layout of the scene. This compelling sense of depth from disparity is known as stereopsis [113]. The vergence system is also neurally linked to the focusing response of the eye to ensure single vision is accompanied by the clearest image quality possible. The focusing response of the eye, known as accommodation, changes the conformation of the crystalline lens to bring objects in best focus for the clearest image possible. Accommodation and vergence are neurally linked, and changes in one can drive the other. This cross-talk between the vergence and accommodation systems provides clear single vision in the natural environment.

1.2 Eye movements in the natural and virtual-reality environment

Humans have evolved for many years in natural environments. Optimizing our function in the natural world has shaped the design of our perceptual systems. These adaptations can be hardwired and present from birth, or may be learned and fine-tuned over the lifetime. Previous work has shown the visual system distributes resources effectively by constantly and actively exploiting regularities in the natural environment [31, 36]. For example, the shape of the binocular horopter is consistent with the distribution of naturally occurring binocular disparities [113, 38]. The horopter is a curved surface in 3d space with special properties - points on this surface fall on corresponding points in the left and right retina. The horopter represents the distance where binocular fusion is guaranteed and stereoacuity is optimized. The horopter has a top back pitch. The finding that the distribution of naturally occurring binocular disparities show the same bias is an effective adaptation, as determining which point in each eye's image arose from the same point in the scene is computationally complex. Solving this computation, known as the correspondence problem, would be particularly daunting if the natural environment was comprised of uniformly distributed small objects. Our environment is not random in this fashion, but instead contains regularities. For example, the natural world contains many opaque objects, and near objects occlude those that are farther away. Additionally, as our world is structured by gravity, we tend to be surrounded by many surfaces that are earth-horizontal (like the ground and tabletops) or earth vertical (walls and trees). The constraints of the natural visual environment are manifested in the brain's search for solutions to the binocular correspondence problem, allowing for an efficient and restrictive search. The binocular visual system has evolved to function best in the natural environment as evidenced by its effective adaptation to environmental

regularities.

Alongside evidence that the binocular visual system has finely adapted to the statistics of the natural environment, much is known about oculomotor behavior in the natural world. People don't fixate the world randomly, but tend to fixate objects they're planning on interacting with, or surfaces they're walking on [64, 63, 81]. In contrast, very little is known about oculomotor behavior in simulated environments like virtual-reality (VR) gaming headsets. This is due in part to the complexities and challenges involved with eye tracking in these environments. As a result, basic questions about oculomotor behavior in VR have gone unanswered. As a result, we don't know where people tend to look in headsets, and we also don't know what average fixation distance looks like. The binocular disparity users are experiencing is also unknown, and crucially it's unclear whether the binocular disparity in VR headsets aligns with the regularities in the natural environment. There is evidence that the binocular disparity observers experience in the natural environment is aligned with the statistics of the natural world [113]. If there is a mismatch in the disparity statistics between the natural and VR environment, this may be a source of conflict that leads to discomfort in these headsets.

As previously described in Section 1.1, the focusing response of the eye (accommodation) is neurally linked with the binocular coordination of eye movements (vergence). This means that in the natural environment, the eyes are accommodating and converging to the same distance. This ensures that when fixating an object at some distance, the ultimate perception is a clear and single image. In VR headsets, the eye is accommodated to the optical screen, which is at a fixed distance. At the same time, the eye is constantly changing rotation and alignment in order to best match the simulated distance of the object of interest, which may not be at the optical distance of the screen [55, 62, 122]. This set up in VR headsets is creating a difference in accommodation and vergence distance, which is breaking the typical neural coupling of these responses. This difference in accommodation and vergence is known as the vergence-accommodation conflict. Previous work has shown that the greater the mismatch between vergence and accommodation distance in stereoscopic displays (such as VR headsets) the longer it takes to fuse an image, and there are deficits to disparity scaling and stereopsis [51, 132, 3]. Importantly, the vergence-accommodation conflict is a cause of discomfort in users of stereoscopic displays - the larger the conflict the more discomfort users experience [51, 57, 108]. When efforts are made to reduce the mismatch between vergence and accommodation distance, performance on visual tasks improve, and users report a more comfortable viewing experience. To avoid the vergence-accommodation conflict in stereoscopic displays, fixation distance should generally match the optical distance of the screen. It would be useful to know where users are converging relative to the optical distance of the screen to help inform whether the screen is at an appropriate distance to minimize the conflict. Eye tracking users in VR headsets, and quantifying the statistics of fixation and binocular disparity helps inform VR headset design to improve users' comfort and experience with this technology.

1.3 Deficits to oculomotor function in Amblyopia

Research has shown that distinctive features of eye movements are altered in visual disorders such as amblyopia. Amblyopia is a developmental abnormality that arises due to abnormal visual experience early in life. Most commonly, amblyopia arises when the two eyes are misaligned (strabismic amblyopia) or when the two eyes have unequal refractive power (anisometropia). Amblyopia is clinically important as it is the second most common cause of vision loss in infants and young children (aside from refractive error) impacting 3-5% of the population. Amblyopia leads to a constellation of perceptual deficits in the (weaker) amblyopic eye including impaired visual acuity, contrast sensitivity, stereopsis, and form vision. Amblyopia is additionally associated with oculomotor abnormalities including eccentric and unsteady fixation [19, 116] and abnormal saccadic eye movements (increased saccadic latency and number of corrective saccades [35]).

In people with normal vision and healthy oculomotor control, the eyes are constantly in motion, even during fixation. These small fixational saccades that arise play a critical role in vision [70, 96]. When people with amblyopia attempt to maintain steady fixation on a target, the instability of the amblyopic eye leads to drifts and subsequent saccades to correct for the eye's position. These corrective eye movements however, are error prone, and move the eye farther away from the intended target [19]. This leads to decreased fixational stability in the amblyopic eye, and is particularly prevalent in the case of strabismic amblyopia. Importantly, there is a significant correlation between the visual acuity and fixational stability of the amblyopic eye [19, 106, 116]. This means that worse visual acuity is associated with more fixational instability in the amblyopic eye. Strikingly, the fixational eye movements of the amblyopic eye account for more than half of the variance in the visual acuity of observers with strabismic and anisometropic amblyopia [19]. Reduced visual acuity is the 'sine qua non' of amblyopia, and the degree to which visual acuity is reduced in the amblyopic eye directly relates to the depth and severity of a patient's amblyopia. As such, it would be useful to establish whether changes in visual acuity with treatment in amblyopia are accompanied by changes to fixational stability. Previous work has relied on a cross sectional approach of measuring the visual acuity and fixational stability of patients at one point in time. There has been little work to understand longitudinal changes in these measures, and how treatment for amblyopia may impact the relationship between fixational stability and clinical metrics. It would be useful to establish whether functional improvements with treatment in amblyopia are accompanied by changes to fixational stability. This relationship would establish whether fixational stability can be used as an objective measure for monitoring treatment in amblyopia and possibly other disorders.

1.4 Study rationale and dissertation approach

The work described in this dissertation raises important considerations and concerns regarding eye movements in environments such as virtual reality, which don't necessarily align with

the statistics of the natural world we've adapted to. Additional work explores the possibility of using eye movement metrics to track disease progression in amblyopia. Future interventions for amblyopia that rely on virtual reality are being developed and tested in academia as well as the private sector [40]. It is important to understand the effect of virtual reality environments on perception before advancing interventions for visual disorders which rely on this technology.

Chapter 2 presents the results of pilot study which tracked changes to eye movements and visual function with treatment for amblyopia. The goal of this study was to establish whether fixational stability can be used as a metric for the recovery of visual function in amblyopia. Previous work has documented a robust correlation between visual acuity in the amblyopic eye and fixational stability. This correlation indicates that an amblyopic eye with worse visual acuity will also tend to have worse fixational stability. Reduced visual acuity is the 'sine qua non' of amblyopia, and the degree to which visual acuity is reduced in the amblyopic eye directly relates to the depth and severity of a patient's amblyopia. As such, it would be useful to establish whether changes in visual acuity with treatment in amblyopia are accompanied by changes to fixational stability. This relationship would establish whether fixational stability can be used as an objective measure for monitoring treatment in amblyopia and possibly other disorders. We tracked children's fixational stability during patching treatment over time and found fixational stability changes alongside improvements in visual acuity. Specifically we found that with treatment, patients' visual acuity improved, and their fixational saccades showed reduced amplitude. The improvements in visual acuity were proportional to the change in fixational saccade amplitude. This suggests fixational stability can be used as an objective measure for monitoring treatment in amblyopia and other disorders. Ongoing research both in academic institutions and industry are exploring the use of specially designed dichoptic video games as interventions for amblyopia. As these gamified treatments become accessible, and eye tracking technology becomes more prevalent, the possibility of using quantitative eye movement based metrics of a given treatment's efficacy may prove to be very useful, and may allow for remote monitoring of patients.

Chapter 3 presents two studies that set out quantify and compare the statistics of eye movements and binocular disparity in both natural and VR head mounted display environments. The human visual system evolved in the natural environment which contains statistical regularities. Binocular vision has adapted to these regularities such that depth perception and binocular eye movements are more precise, faster, and performed with greater comfort in environments that are consistent with these regularities. We measured the statistics of eye movements and binocular disparities in VR-gaming environments and found that they are quite different from those in the natural environment. Fixation direction and distance are more restricted in VR. In addition, fixation distance is farther in VR. The pattern of binocular disparity across the visual field is less regular in VR and does not conform to a prominent property of naturally occurring disparities. The disparity pattern increases the likelihood of experiencing double vision in VR-gaming environments. We determined from our fixation statistics the optimal screen distance to minimize discomfort due to the vergence-accommodation conflict and the optimal nasal-temporal positioning of HMD screens to max-

imize the binocular field of view. Previous work has documented the discomfort associated with VR headsets, and our findings inform specific improvements to VR headset design that will improve user comfort.

Chapter 4 follows up on the finding from chapter 3, that the pattern of binocular disparity across the visual field is less regular in VR and violates regularities seen in the disparity distribution of the natural environment. Incompatibility with the statistics of the natural environment may cause viewer discomfort and reduced visual performance in the VR environment. We designed a user experiment to test whether scene consistent content that is congruous with the statistics of the natural environment leads to greater viewer comfort and better performance in VR headsets. We found observers prefer scene content that is congruous with the statistics of the natural environment. This work suggests practical guidelines for VR developers to produce content that is faithful to the statistics of the natural environment.

Chapter 2

Fixational Stability as a Metric for the Recovery of Visual Function in Amblyopia

2.1 Introduction

Amblyopia is a developmental visual disorder commonly associated with strabismus or anisometropia. Amblyopia is clinically important as it is the second most frequent cause of vision loss in infants (aside from refractive error) impacting 3-5% of the population [74, 98, 134]. Amblyopia leads to impaired visual acuity, contrast sensitivity, stereopsis, and form vision as well as oculomotor deficits. The oculomotor abnormalities seen in amblyopia are often associated with strabismus, and include eccentric and unsteady fixation [49, 20] and abnormal saccadic eye movements (increased saccadic latency and number of corrective saccades: [35, 99]).

It has long been known that for people with normal oculomotor control and healthy vision that the eyes are never perfectly still, and are constantly in motion [130]. Work over the years has shed a light on the importance of eye movements during fixation. Our visual system can detect stationary objects only because small eye movements prevent the retinal image from staying still for too long, and adaptation from taking effect. If eye movements are eliminated, visual perception fades to a homogeneous field [92]. Recent work has shown that even at the level of small fixational saccades, these eye movements represent a critical stage of information processing, allowing the retina to begin extracting features [96]. Additionally, these microscopic eye movements directly improve acuity by precisely aligning the image on the retina and fine-tuning the retinal image motion [52].

When people with amblyopia attempt to maintain steady fixation on a target, the instability of the amblyopic eye leads to drifts and subsequent saccades to correct for the eye's instability, however these corrective eye movements are less accurate than in observers with healthy vision, and move the eye further away from the intended landing position [19]. As

a result, fixational and oculomotor stability is reduced and abnormal in the amblyopic eye, particularly in strabismic amblyopia [19, 106, 116].

Importantly, there is a significant correlation between the visual acuity and the fixational stability of the amblyopic eye [19, 106, 116]. This means that worse visual acuity is associated with more fixational instability in the amblyopic eye. Strikingly, the fixational eye movements of the amblyopic eye account for more than half of the variance in the visual acuity of observers with strabismic and anisometropic amblyopia [19]. In addition, previous work has shown that multifocal evoked potentials and electroretinogram measurements in the amblyopic eye are reduced, and show a robust correlation with the degree of fixational instability [138]. These findings raise the possibility that unsteady fixation may play a role in the perceptual deficits seen in amblyopia. However, it is currently unclear whether it is the unsteadiness of fixation that limits visual acuity, or visual acuity limiting fixational stability in amblyopic vision (a chicken and egg problem).

Previous work to understand the directionality of the causality between visual acuity and fixational stability have shown contradictory findings. When visual acuity impairments were introduced with retinal defocus, fixational stability remained consistent, even when the fellow eye visual acuity was matched to that of the amblyopic eye in observers with amblyopia [91]. Additional work has shown that random jittering of acuity targets leads to reduced visual acuity [18]. Both of these findings suggest that fixational stability places limits on visual acuity. However, previous works found that in strabismic amblyopia, when gratings were stabilized (using afterimages), grating acuity was not significantly improved as compared to an unstabilized condition [48]. These contradictory findings make it difficult to tease apart whether fixational stability is the limiting factor on visual acuity, or vice versa.

Reduced visual acuity is the 'sine qua non' of amblyopia, and the degree to which visual acuity is reduced in the amblyopic eye directly relates to the depth and severity of a patients' amblyopia. As such, it would be useful to establish whether changes in visual acuity with treatment in amblyopia are accompanied by changes to fixational stability. This relationship would establish whether fixational stability can be used as an objective measure for monitoring treatment in amblyopia and possibly other disorders. This is especially important because treatment of amblyopia is most effective in children, who may be too young to respond to subjective tests, such as reading a visual acuity chart. An objective measure that aligns with treatment efficacy would provide a more effective monitoring for these young patients. Additionally, previous work has proposed an intervention that uses biofeedback to directly improve fixational stability in strabismic amblyopia [19]. Recent work that has followed up on this recommendation does suggest that biofeedback fixation training leads to more stable fixation in strabismic amblyopia [67]. However, more work is needed to inform the usefulness of training fixational stability directly, and how these changes to fixational stability following training may relate to clinical improvements.

Ongoing research is exploring the use of specially designed dichoptic video games to facilitate treatment for amblyopia [10]. Recent work has shown interactive video game interventions in virtual reality are effective in rehabilitating stereopsis in amblyopia [40]. As these gamified treatments become more accessible, and eye tracking technology becomes less ex-

pensive and more prevalent, the possibility of using quantitative eye movement based metrics of a given treatment's efficacy may prove to be very useful. Eye tracking may additionally allow for remote monitoring of patients in real time as they carry out their prescribed vision therapy. Additionally, using a non-verbal metric like fixational stability as a proxy for visual acuity is particularly useful for young children and toddlers with visual disorders that are unable to read an acuity chart.

Our aim in this pilot experiment is to track observers with amblyopia as they are treated, and to monitor improvements in clinical tests (such as visual acuity and stereopsis) alongside any changes to fixational stability. To compare with these observers, we additionally tested children with healthy vision at the same intervals as our clinical population. Our goal is to establish whether improvements in clinical measures (such as visual acuity) with treatment are accompanied by changes to fixational stability in amblyopia. This pilot study investigates the possibility of using eye movements as a metric of a treatment efficacy in amblyopia.

2.2 Methods

Participants

The study included children with anisometropic or strabismic amblyopia, and a comparison group of children with normal vision, recruited from the Meredith Morgan Eye Center at UC Berkeley. Children with amblyopia were undergoing treatment, primarily patching the dominant eye, and the experimenters were blinded as to the treatment details. Amblyopia was defined as a difference in the best-corrected visual acuity between the two eyes of 0.2 logMAR (the logarithm of the minimum angle of resolution, where 0.0 logMAR = 20/20 Snellen acuity). Amongst the 5 observers with amblyopia, 3 had amblyopia due to anisometropia (defined as a difference in refractive corrections of $> 0.75D$ spherical equivalent) with no strabismus; and 3 had amblyopia due to strabismus, as revealed using the standard cover test procedure. The comparison group of visually normal children had a VA of 20/25 (0.1 logMAR) or better in each eye, an inter-eye acuity difference of 1 line or less (0.1 logMAR), normal stereopsis (20 seconds of arc with the Randot Stereoacuity Test) and no history of treatment for amblyopia or the presence of any amblyogenic condition. Details of the observers' characteristics are given in Table 2.1. All observers and their parents gave written consent and assent before the commencement of data collection. This research followed the tenets of the Declaration of Helsinki and was approved by the Committee for Protection of Human Subjects at the University of California, Berkeley.

Apparatus

Participants were seated 57 cm from a 21" ViewSonic G225f monitor, and positioned in a chin rest such that the viewing eye was parallel to and gazing at the center of the screen. The non-dominant eye was occluded by an eye patch. Eye movements were recorded with an Eyelink II tracking at 500 Hz.

Table 2.1: Visual characteristics of observers tested

Type	Age	Acuity Session 1	Acuity Session 3	Intervention
Strabismic	7	20/200	20/125	Patching (3-4 hrs/day)
Anisometropic	9	20/25-1	20/25+2	Dichoptic training (1/week)
Strabismic	8	20/63+1	20/40	Patching (10 hrs/week)
Anisometropic	11	20/100	20/50-1	Patching (2 hrs/day)
Strabismic	6	20/80	20/32	Patching (10 hrs/week)
Non-Amblyopic	9	20/20-1	20/20-1	N/A
Non-Amblyopic	10	20/16-1	20/16	N/A
Non-Amblyopic	9	20/25-1	20/25	N/A
Non-Amblyopic	10	20/16	20/20	N/A
Non-Amblyopic	9	20/16-2	20/16	N/A

Measuring Fixational Stability

To measure fixational stability, the dominant eye of each observer was patched and observers were asked to look at a 1° colorful smiley face (as shown in Fig. 2.1) on an otherwise black screen for a 20-second intervals The smiley face changed color halfway through the trial (10 sec ±3, the exact timing was randomized), and observers were asked to press the spacebar on a keyboard as soon as they detected the color change. Observers were instructed to continue fixating on the new colored smiley face. The color change detection task was added in order to increase engagement during the experiment. There were ten 20-second trials in total in a session.

Clinical Measures

Visual acuity was measured using the Bailey Lovey acuity chart [8]. Measures for stereopsis were taken using the Randot Circles Stereotest as well as the Asteroid Test [125]. The Asteroid test is shown on an autostereoscopic tablet and presents observers with four dynamic random-dot stereograms, one of which contains a target with disparity. The Asteroid test evaluates a wider range of stereoacuity than standard clinical tests which is useful when testing patients who may have coarse residual stereopsis [125].

Experiment design

Each observer participated in 3 sessions over the course of 3 months. This duration was chosen as over 3 months of patching treatment for amblyopia one would expect to see measurable improvements in clinical vision tests. During each of the 3 sessions, the clinical tests were performed and fixational stability was measured.

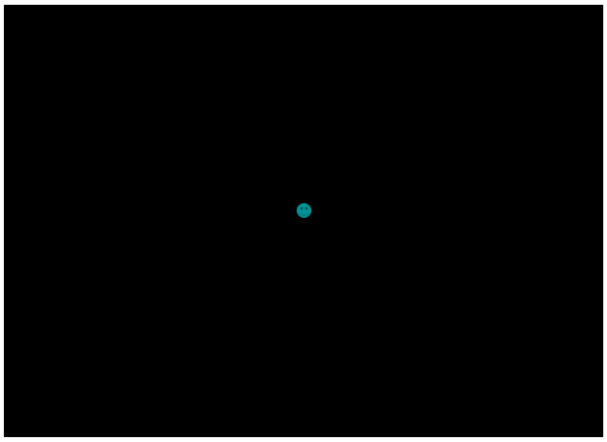

Figure 2.1: Example of a trial. A one degree smiley face was shown for 20 seconds against an otherwise black background, and observers were instructed to maintain fixation on the center of the smiley face.

2.3 Data Analysis

Quantifying Fixational Stability

Fixational stability is typically quantified based on the area that encloses a certain percentage (most conventionally, 68% or 1 S.D.) of eye positions. The most commonly used fixational stability metric, Bivareate Contour Ellipse Area (BCEA), makes the assumption that fixational position follows a normal distribution [115, 118], an assumption that is typically violated [17, 60]. The Isoline Area method (ISOA) does not assume that fixation position follows a normal distribution [16]. In order to quantify ISOA, we estimated the probability density function of eye position using kernel density estimation [133]. We next calculated the area corresponding to 68% (1 S.D.) of fixation in degrees squared. This measure corresponds to fixational stability, as shown in Fig. 2.2A.

Identifying Fixational Saccades

Fixational saccades were identified using a saccade detector algorithm. Eye position samples with a velocity $> 7°$/sec were labelled as candidate saccades. An acceleration threshold of $350°$/sec was used to refine the start (acceleration $> 350°$/sec) and end (acceleration $< 350°$/sec) of fixational saccade eye movements.We manually inspected all trials to remove any falsely-identified or missed fixational saccades. Fig. 2.2B presents the horizontal and vertical eye position traces for a 5-second epoch from four different observers, two with amblyopia (top) and two observers with normal vision (bottom). Fixational saccade eye movements are labelled. The eye movement identification process used captures both microsaccades

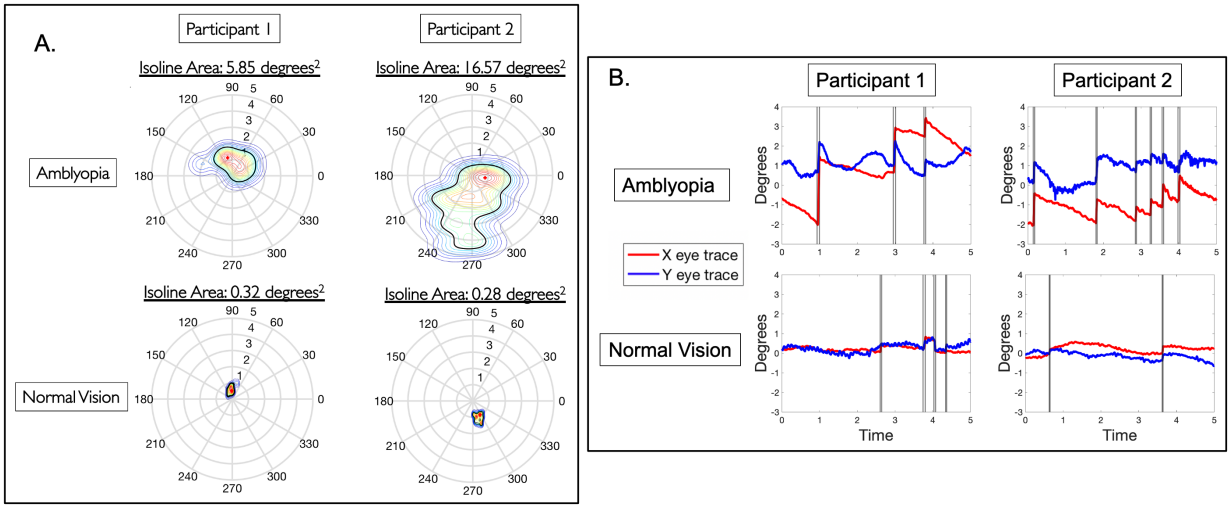

Figure 2.2: Examples of measures used to quantify fixational stability and fixational eye movements. A) ISOA for 4 different observers for 4 separate 20-second fixation trials. Eccentricity from 1-5° is shown, and the black contour corresponds to 68% of highest density fixations. ISOA corresponds to the area (in degrees^2) of the black contours. B) 5-sec of eye position traces are shown for observers with amblyopia and those with normal vision. Fixational saccades (which include both microsaccades as well as larger eye movements) are highlighted in grey.

(fixational eye movements less than 1°) as well as larger saccadic eye movements. Growing evidence points towards common shared characteristics between microsaccades and saccades, suggesting a unified neural generator for these eye movements [93, 68, 69]. Additionally, as amblyopic eyes show increased fixational instability, microsaccades may have a greater amplitude than 1°. Using a somewhat arbitrary threshold of less than 1° for microsaccade identification may remove useful eye movements. Taking this into consideration, we include both microsaccades and saccades in our fixational saccade analysis.

In order to evaluate characteristics of fixational eye movements in patients being treated for amblyopia, and observers with normal vision, we assessed three characteristics of fixational eye movements: frequency, error magnitude, and amplitude. We compare these fixational eye movement characteristics to clinical tests and the ISOA metric. We additionally confirm whether changes in visual acuity with treatment in amblyopia are accompanied by changes in fixational stability.

2.4 Results

Clinically measurable improvements with treatment

Patching treatment, where the dominant eye is covered for a duration, typically leads to improvements in visual acuity in amblyopia. To determine whether observers with amblyopia showed improvements to their visual acuity with patching, Fig. 2.3 shows visual acuity in the third session plotted versus visual acuity in the first session. Observers with healthy vision all fall near or on the 1:1 unity line, while observers being treated for amblyopia all fall below the 1:1 unity line. To compare the change in visual acuity between the patients being treated for amblyopia and the observers with normal vision, we computed the pre-post ratio, or the ratio of visual acuity in session 1 and session 3 for each observer. A pre-post ratio greater than 1 corresponds to an improvement in visual acuity, while a pre-post ratio of exactly 1 corresponds to no change over sessions. A one sample t-test found that the pre-post visual acuity ratio for the observers with amblyopia was significantly greater than 1 ($t(4) = 3.12$, $p = .03$), while the pre-post ratios for the observers with normal vision was not significantly different from 1 ($t(4) = .86$, $p = .44$). This means the visual acuity of healthy observers with no interventions showed little change, while the observers with amblyopia all undergoing treatment (predominantly patching), showed significant improvements to visual acuity over the course of three months.

Relationship between visual acuity and stereopsis

To visualize how stereoacuity tracks visual acuity, Fig 2.4A and Fig 2.4B show visual acuity (in MAR) by stereopsis as measured by the Asteroid test and the Randot Circles Stereotest, respectively. Visual acuity is positively correlated with stereopsis as measured by the Asteroid test ($p< 0.00001$, Slope $= 0.79 \pm 0.22$ (95% CI: 0.35,1.21), $r^2 = 0.50$) and the Randot Circles test ($p< 0.00001$, Slope $= 0.39 \pm 0.08$ (95% CI: 0.23,0.55), $r^2 = 0.51$).

Relationship between visual acuity and fixational stability

In section 2.4 we established that visual acuity improved in the treatment group. Next we can ask what the relationship is between fixational stability measures and clinical tests. Figure 2.5 shows the significant positive correlation between visual acuity and fixational stability as measured by Isoline Area ($p= 0.0003$, Slope $= 0.55 \pm 0.16$ (95% CI: 0.24,0.86), $r^2 = 0.41$). The finding that worse visual acuity is associated with worse fixational stability is a replication of previous research [19, 116]. We additionally found that visual acuity shows a positive correlation with fixational saccade amplitude ($p< 0.00001$, Slope $= 1.59 \pm 0.22$ (95% CI: 1.16,2.02), $r^2 = 0.65$), a finding that replicates previous work [19]. This suggests that worse visual acuity is associated with larger fixational saccade amplitudes during attempted fixation.

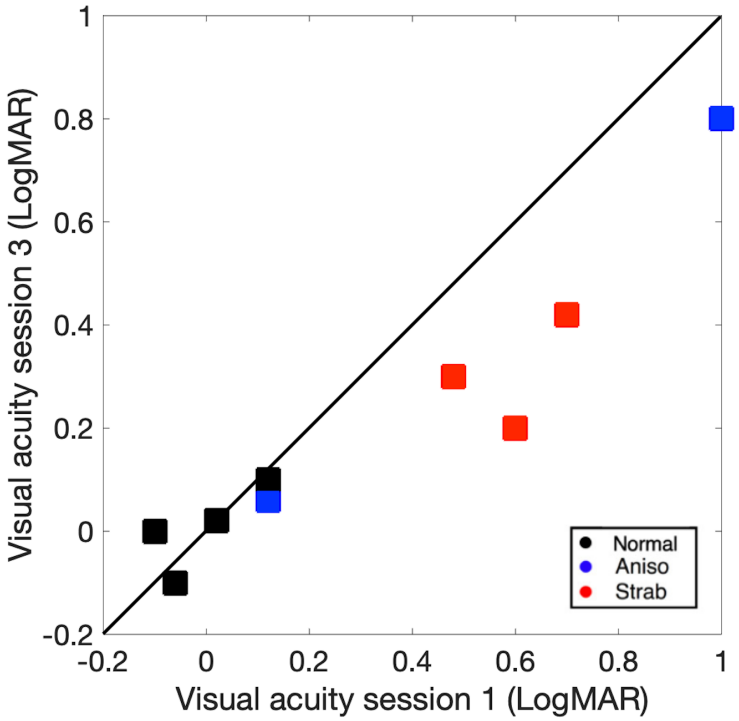

Figure 2.3: Visual acuity as measured during the third (and final) session plotted versus visual acuity in the first session. Visual acuity is shown in LogMAR units for strabismic (red) and anisometropic (blue) amblyopia, and observers with normal and healthy vision (black). The black line represents the 1:1 unity line, data that fall on this line show no change over sessions.

Relationship between stereopsis and fixational stability

Figure 2.6A shows the positive correlation between stereopsis as measured by the Asteroid test and fixational stability ($p< 0.0001$, Slope $= 0.56 \pm 0.14$ (95% CI: 0.29,0.83), $r^2 = 0.46$) while Figure 2.6B show the correlation between stereopsis as measured by the Randot stereotest and fixational stability ($p= 0.0001$, Slope $= 0.89 \pm 0.30$ (95% CI: 0.30,1.48), $r^2 = 0.42$). Both these findings suggest that worse stereopsis is associated with reduced fixational stability, a replication of previous work [116]. The Asteroid test is an interactive stereotest that can be played on a tablet. It assesses a wider range of stereothresholds than the standard clinical Randot stereotest (2,000 arcsec versus 400 arcsec). Figure 2.6C shows the relationship between performance on the Randot circles stereotest and on the Asteroid tests, and despite the differences between these tests, they are positively correlated ($p< 0.00001$, Slope $= 0.94 \pm 0.23$ (95% CI: 0.49,0.1.39), $r^2 = 0.73$). Taking into consideration the limited range of disparity in the Randot circles test, the Asteroid test is particularly useful when

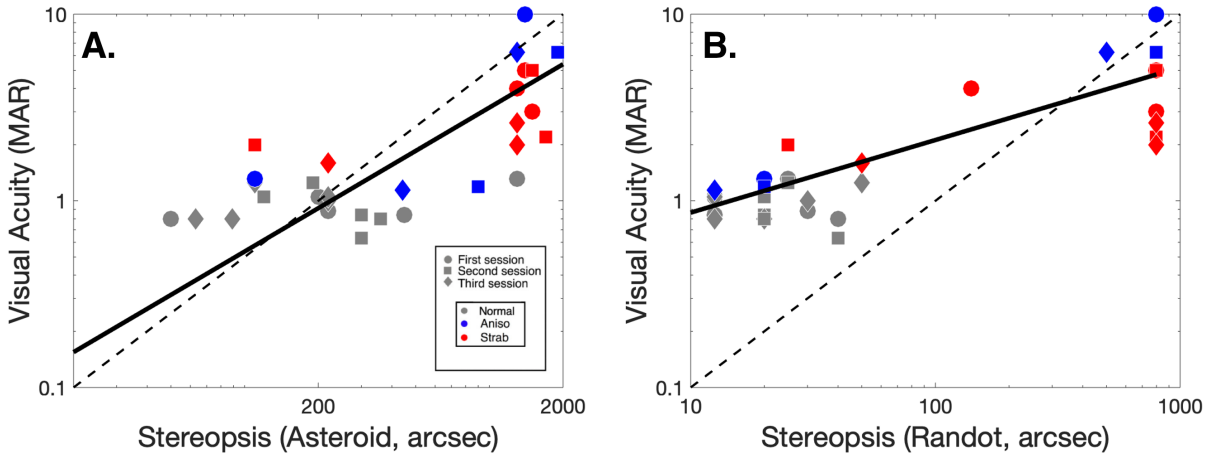

Figure 2.4: Visual acuity (in MAR) plotted versus stereoacuity. Each plot shows 3 data points per observer, one point per session. Circles correspond to data from session 1, squares to data from session 2, and diamonds to data from session 3. This is shown for strabismic (red) and anisometropic (blue) amblyopia, and observers with normal and healthy vision (grey). If the data were perfectly correlated, data points would fall on the dotted black line, while the solid black line is a power function fit to the data, which is linear on a log scale. A: Visual acuity versus stereopsis as measured by the Asteroid tests. B: Visual acuity versus stereopsis as measured by the Randot Circles Stereotest. The greatest disparity tested is 400 arcsec. Observers who fail the test are given a score of 800 (2 times the largest disparity tested).

considering patient populations (such as amblyopia) with limited residual stereopsis.

Relationship between fixational stability measures

Figure 2.7A shows the positive correlation between fixational saccade error and Isoline Area ($p< 0.00001$, Slope $= 0.59 \pm 0.06$ (95% CI: 0.47,0.71), $r^2 = 0.84$), and Figure 2.7B shows the positive correlation between fixational saccade amplitude and Isoline Area ($p< 0.00001$, Slope $= 0.37 \pm 0.07$ (95% CI: 0.23,0.51), $r^2 = 0.59$). These are both findings that replicate Chung et al. (2015). This means people with reduced fixational stability (as measured by Isoline Area) are additionally likely to have increased fixational saccade error and fixational saccade amplitude. Figure 2.7C shows the positive correlation between fixational saccade frequency and fixational saccade amplitude ($p< 0.0001$, Slope $= 0.52 \pm 0.11$ (95% CI: 0.30,0.74),$r^2 = 0.46$) . More frequent fixational saccades are associated with larger fixational saccades.

Table 2.2 shows all other comparisons between eye movement and clinical metrics. No other pairwise comparisons between Isoline Area, fixational saccade metrics and clinical

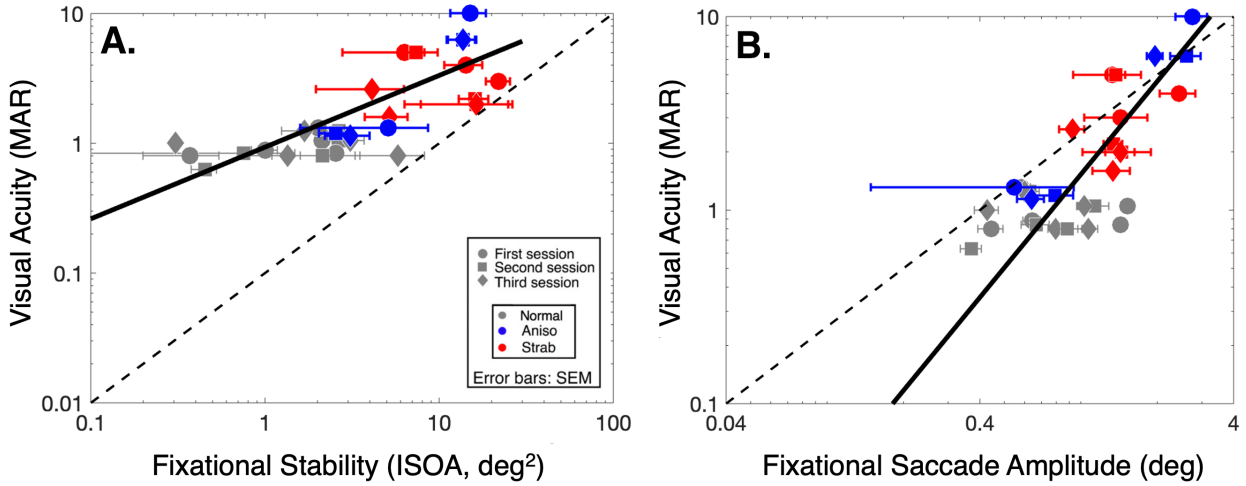

Figure 2.5: Each plot shows 3 data points per observer, one point per session. The solid black line represents the best fit line, if the data were perfectly correlated, data points would fall on the dotted black line. Error bars are SEM. A: Visual acuity (in MAR) versus fixational stability as quantified by the ISOA metric. B: Visual acuity versus fixational saccade amplitude (in degrees of visual angle).

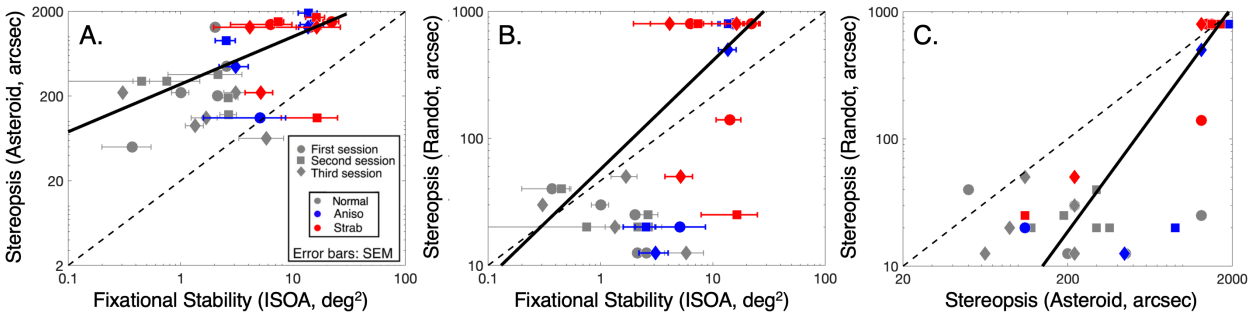

Figure 2.6: Each plot shows 3 data points per observer, one point per session. The solid black line represents the best fit line, if the data were perfectly correlated, data points would fall on the dotted black line. Error bars are SEM. A: Stereopsis (as measured by the Asteroid test) versus fixational stability. B: Stereopsis as measured by the Randot Circles Stereotest versus fixational stability. The greatest disparity tested is 400 arcsec. Observers who fail the test are given a score of 800 (2 times the largest disparity tested). C: Stereopsis (as measured by the Randot Circles Stereotest) versus Stereopsis (as measured by the Asteroid test). The results from these two tests are positively correlated.

Table 2.2: All other comparisons between Isoline Area, fixational saccade metrics and clinical measures. None showed a significant correlation.

Comparison	r^2	Slope
Fixational saccade error x Fixational saccade frequency	0.23	0.31
Fixational saccade error x Fixational saccade amplitude	0.34	0.64
Isoline area x Fixational saccade frequency	0.32	0.20
Stereopsis (Asteroid) x Fixational saccade amplitude	0.38	1.00
Stereopsis (Asteroid) x Fixational saccade error	0.30	0.67
Stereopsis (Asteroid) x Fixational saccade frequency	0.13	0.74
Stereopsis (Randot) x Fixational saccade amplitude	0.25	1.20
Stereopsis (Randot) x Fixational saccade error	0.27	1.00
Stereopsis (Randot) x Fixational saccade frequency	0.07	0.8
Visual acuity x Fixational saccade error	0.2	0.58
Visual acuity x Fixational saccade frequency	0.15	0.87

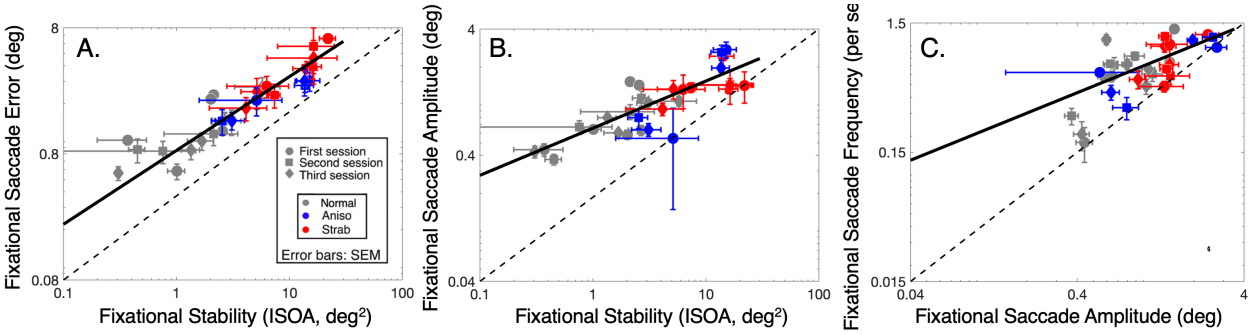

Figure 2.7: Each plot shows 3 data points per observer, one point per session. The solid black line represents the best fit line, if the data were perfectly correlated, data points would fall on the dotted black line. Error bars are SEM. A: Fixational Saccade Error (in degrees) versus fixational stability as measured by the Isoline Area metric. B: Fixational Saccade Error (in degrees) versus fixational stability as measured by the Isoline Area metric). C: Fixational Saccade Frequency (per second) versus Fixational Saccade Amplitude (in degrees).

measures showed a correlation.

Relationship between improvements to clinical measures and fixational stability with treatment

The most important results of this study are whether the clinical improvements shown are accompanied by improvements to fixational stability in the patient group. These are the

clinical measures that fixational stability improvements may track. In order to quantify changes between sessions, we calculated the pre/post ratio between session 1 and session 3 for both the clinical metrics and fixational stability measurements. A ratio of 1 means there was no change in that measure between session 1 and session 3 (similar to the analysis in section 2.4.

Fixational saccade amplitude tracks changes in visual acuity

We first calculated the ratio between the median performance in the first and third session. The correlation for this analysis is shown in Figure 2.8A. We found a statistically significantly correlation between the change (pre/post) in visual acuity and the change (pre/post) in fixational saccade amplitude (p= .04, Slope = 1.07 ± 0.44 (95% CI: 0.21,1.93),r^2 = 0.43). Importantly, the slope $\approx$ 1, implying that the improvement in visual acuity is proportional to the change in fixational saccade amplitude. Splitting the data into subgroups reveals a significant correlation in observers with amblyopia currently patching (p= .02, Slope = 1.25,r^2 =.88). Amongst observers with normal vision there is no significant correlation (p= .83, Slope = $-.06$,r^2 = .04) as neither their saccade amplitude nor visual acuity change significantly.

As visual acuity improves with treatment, the amplitude of fixational saccades during attempted fixation decreases in the amblyopic eye. This relationship becomes apparent when considering patients currently undergoing treatment. This means that the patients recovering visual acuity with treatment are also those with a reduction in fixational saccade amplitude with treatment. Although the correlation is apparent, it's unclear whether fixational saccade amplitude limits visual acuity or vice-versa.

Isoline Area tracks changes in visual acuity

The change (pre/ post) in isoline area showed a correlation with changes to visual acuity (p=0.01, Slope = 0.64 ± 0.19(95% CI: 0.26,1.00), r= 0.59), as shown in Fig 2.8B. Splitting the data into subgroups does not reveal a significant correlation in observers with amblyopia currently patching (p= .18, Slope = .6,r^2 = .51). Amongst observers with normal vision there is no significant correlation (p= .40, Slope = $-.02$,r^2 = .25). When restricting the data to observers with normal vision, or observers with amblyopia, the range of changes reported is constrained and the number of subjects is reduced. This explains the insignificant correlation when restricting the subject groups by amblyopia status.

Fixational stability does not track changes in stereopsis

We also found that changes to fixational stability (both Isoline Area and various fixational saccade metrics) and changes to stereopsis are not correlated (all $r^2 < 0.38$). In order to have fine stereopsis, good acuity is critical. However, even if both eyes have good acuity, stereopsis is not guaranteed. Patients with constant strabismus and good visual acuity in both eyes will typically be stereoblind [66]. Even with the recovery of visual function seen in

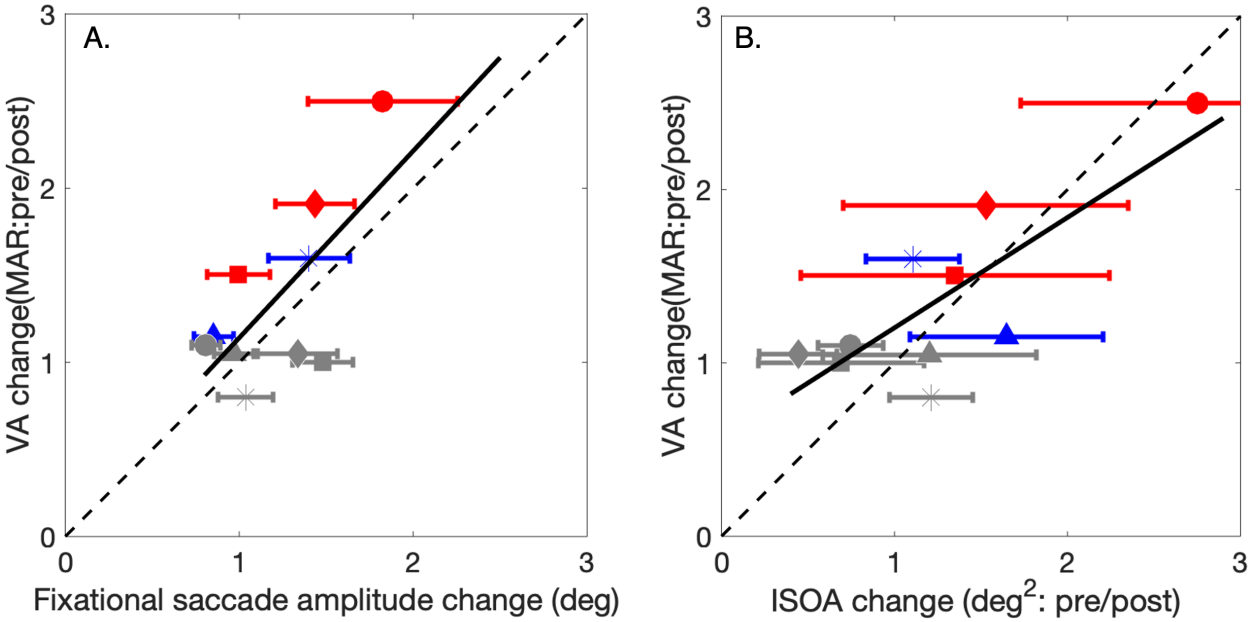

Figure 2.8: Each point represents a pre-post ratio. The solid black line represents the best fit line, if the data were perfectly correlated, data points would fall on the dotted black line. Error bars are RMS. A: Pre-post visual acuity versus pre-post fixational saccade amplitude B: Pre-post visual acuity versus pre-post fixational stability (as measured by the Isoline Area metric).

observers with amblyopia, stereopsis is not recovering at a commensurate rate with fixational stability (by proxy visual acuity), which explains our findings.

No other comparisons between changes to fixational saccade metrics and changes in clinical measures showed a correlation (as shown in Table 2.3). When quantifying how fixational stability tracks clinical tests (as in Figure 2.4 - Figure 2.7), some comparisons that showed robust correlations when comparing performance by session, showed no relationship when comparing the pre-post ratios of the same measures. One such example is the correlation between fixational saccade error and fixational stability (Fig 2.7A, $r^2 = .84$), which was not reflected in the comparison of the pre-post ratios ($r^2 = .15$). Computing the pre-post ratio as a measure of change leads to a narrow range (values between ∼1-3) as interventions over the course of three months of treatment can only lead to so much improvement. More observers may be needed to reveal a significant relationship amongst such a small range of possible improvements. Additionally, it is possible the changes to these measurements are not commensurate, and there is a lag between improvements that isn't captured over a short 3-month period.

Table 2.3: All other comparisons between changes to Isoline Area, fixational saccade metrics and clinical measures. None showed a significant correlation.

Comparison (the pre-post ratio was calculated for all measures)	**r^2**	**Slope**
Fixational saccade amplitude x Fixational saccade frequency	0.001	-0.008
Fixational saccade amplitude x Fixational saccade error	0.15	0.14
Fixational saccade amplitude x Isoline area	0.12	0.20
Fixational saccade error x Fixational saccade frequency	0.08	0.07
Fixational saccade error x Isoline area	0.15	0.33
Fixational saccade frequency x Isoline area	0.21	0.16
Stereopsis (Asteroid) x Fixational saccade amplitude	0.19	3.2
Stereopsis (Asteroid) x Fixational saccade error	0.38	-3.21
Stereopsis (Asteroid) x Fixational saccade frequency	0.07	-1.35
Stereopsis (Asteroid) x Isoline area	0.029	0.25
Stereopsis (Randot) x Fixational saccade amplitude	0.15	1.18
Stereopsis (Randot) x Fixational saccade error	0.22	-0.19
Stereopsis (Randot) x Fixational saccade frequency	0.20	003
Stereopsis (Randot) x Isoline area	0.30	0.67
Stereopsis (Randot) x Stereopsis (Asteroid)	0.37	-0.06
Visual acuity x Fixational saccade error	0.02	-0.03
Visual acuity x Fixational saccade frequency	0.02	0.80
Visual acuity x Stereopsis (Asteroid)	0.03	-0.01
Visual acuity x Stereopsis (Randot)	0.18	0.22

2.5 Discussion

In current best clinical practice, reduced visual acuity characterizes amblyopia. Considering that increased fixational unsteadiness in the amblyopic eye is correlated with poor visual acuity, it would be useful to establish whether improvements in visual acuity with treatment are accompanied by improved fixational stability in the amblyopic eye. Our preliminary results suggest that that visual acuity in the non-dominant eye and stereopsis are both correlated with fixational stability (Isoline Area). We additionally found that fixational saccade error and fixational saccade amplitude are correlated with fixational stability. Fixational saccade frequency and fixational saccade amplitude were also correlated. These results all replicate previous findings [19, 106, 116].

Importantly, we additionally showed that with treatment, visual acuity in the amblyopic eye improves, and these improvements are correlated with increased fixational stability and smaller amplitudes during attempted fixation. Presumably, the larger fixational saccades take the eye farther away from the fixated target. Interestingly, changes to fixational stability (and by proxy visual acuity) are not matched by changes to stereopsis, reflecting the

inconsistent relationship between visual acuity and stereopsis [66].

In young children early diagnosis is associated with the best outcomes for recovering visual function in amblyopia as neural plasticity is highest [131, 65]. As amblyopia does not develop after 6-8 years of age [77], a nonverbal and objective measure of the deficits associated with amblyopia are particularly useful. Our findings provide preliminary evidence that there are fixational saccade metrics that track recovery of visual function in amblyopia that may be useful as non-verbal metrics of treatment efficacy.

Novel interventions for amblyopia that utilize virtual reality technology are becoming more prevalent [7, 40, 140]. Additionally, eye tracking technology is becoming cheaper to implement, and is now included in some commercially available head mounted displays [127]. Alongside these technological advances, the recent COVID-19 pandemic has highlighted the importance of remote monitoring and consulting in optometric care [75]. As virtual reality interventions for amblyopia become more widespread, developers of these interventions may want to consider monitoring fixational stability. This is a quick and easy way to collect an objective measure that tracks treatment efficacy, and these measurements can be collected and monitored remotely.

Our small sample size of participants with anisometropia or strabismus does not allow us to make statistical inferences based on the etiology of our observers' amblyopia. Previous work has suggested fixation is particularly unstable in the case of strabismic amblyopia [19, 116]. Future work should strive to establish whether there are differences in recovery and fixational stability dynamics between observers with strabismic or anisometropic amblyopia.

Healthy and normal fixation requires contradictory functionality - we must keep our eyes still enough to examine the tiniest details in our world, but if we were to fixate perfectly, minimizing all eye movements, the entire world would fade to a homogeneous field [92, 68]. The smallest fixational eye movements are implicated in a variety of functional roles from precisely adjusting the position of objects of interest on the retina to processing fine spatial details, and encoding spatial information both temporally and spatially [95, 96]. Reduced fixational stability in amblyopia may have far reaching implications, beyond low level deficits. Previous work has found abnormalities in attentional control by the amblyopic eye [107, 85, 87] and Verghese (2019) has suggested the reduced fixational stability leads to attentional deficits[126]. Future research should work to understand the role of fixational stability in higher level cognitive function, such as attention. In the meantime, preliminary evidence suggests that the rehabilitation of fixational stability may have wide reaching implications for amblyopic vision.

The significant correlation between improvements to fixational stability and visual acuity with treatment suggests that visual acuity may be aided by reduced fixational stability. Although the causality remains an ambiguous chicken and egg problem, training stable fixation in amblyopia may prove useful for improving visual acuity, particularly for those with strabismus. In the past, biofeedback training was used to improve fixational stability in observers with amblyopia [34, 100]. Recent work has shown that biofeedback fixation training may improve fixational stability in the amblyopic eye of patients with strabismic amblyopia [67]. However, this type of training, where auditory feedback is given based on eye position,

was never widely prescribed, likely due to the expensive technology needed and the time consuming nature of the intervention. As biofeedback testing/training has taken place in controlled lab environments, it is additionally unclear whether the benefits of such training transfer outside of the lab, once biofeedback is removed. Ideally, such biofeedback interventions should lead to long lasting improvements once biofeedback training has concluded. Future work should establish the transfer and persistence of improved fixational stability with such training methods. If these interventions showed promise, the implications could be far-reaching. With advances in virtual reality and eye tracking technology, prescribing such biofeedback interventions for patients to work on at home may be possible in the future. Our data as well as others [19] suggest that such interventions should focus on reducing the amplitude of fixational saccades.

Chapter 3

The (Un)natural Statistics of Eye Movements and Binocular Disparities in VR Gaming Headsets

3.1 Introduction

Having two eyes to view the world is both advantageous and challenging. The advantage is that the differences in the two views—binocular disparities—can be used to precisely compute the 3D layout of the scene. The challenge is the difficulty in determining binocular correspondence: Which points in the two eyes' images came from the same feature in the scene? Consider trying to solve the correspondence problem in an environment consisting of sparse small objects randomly distributed in three dimensions. In every direction, all distances would be equally probable, so disparities would be very broadly distributed. Accordingly, the search for solutions to binocular correspondence would have to occur over a very large range of disparities.

But the natural environment is very different from this. It contains many occluding surfaces and many earth-horizontal and earth-vertical surfaces. Furthermore, viewers do not fixate randomly, but rather fixate behaviorally significant points. These environmental and oculomotor constraints are manifested in the brain's search for solutions to binocular correspondence, allowing a much more restricted and efficient search than would otherwise be needed. In fact, the human visual system has adapted to these constraints such that it functions best (faster, more accurately, and with greater comfort) in environments that are like the natural environment [22, 113, 51].

Another important aspect of visual function is the coordination of binocular eye movements and the focusing response of the eyes: *i.e.,*, vergence (converging or diverging the eyes to be aligned on the object of interest) and accommodation (changing the power of the eye lens to focus the object of interest). These responses are neurally coupled [73, 101]. As a consequence, converging (or diverging) the eyes causes the eye lens to increase (or decrease)

power. And accommodating by increasing (or decreasing) the lens power causes the eyes to converge (or diverge) [32, 101]. Stereoscopic displays, including head-mounted displays (HMDs), require the visual system to uncouple these responses because the viewer may have to converge or diverge to fuse an object in front of or behind the screen while maintaining accommodation at the screen distance [55]. This is known to cause discomfort and even nausea [45, 51, 62, 108, 122, 57], impairment in visual performance [3, 51], and distortions of 3D percepts [132, 72]. An important goal of the work reported here is to determine the statistics of vergence-accommodation conflicts in VR gaming in order to provide guidelines for minimizing the conflict.

When people make upward, leftward, and rightward saccades, they tend to diverge the eyes. When they make downward saccades, they tend to converge [30, 21, 37]. These biases in saccadic-related vergence are consistent with the statistics of the natural environment and thereby enable the oculomotor system to make accurate movements in the real world. A goal of our work is to determine to what extent the statistics of virtual scenes conform to natural statistics and to make recommendations on how to modify the statistics to aid oculomotor behavior.

The screens in HMDs have wider temporal fields (toward the ears) than nasal fields (toward the nose). This increases the total field of view (the regions seen by one or the other eye) but decreases the binocular field of view (the regions that are imaged on corresponding regions in the two eyes). A goal of the work presented here is to measure the tendency of fixation distances to determine the screen placements that would maximize the binocular field of view.

3.2 Methods

Disparities & Eye Movements in the Natural Environment

We measured the distributions of gaze directions and distances, and the distribution of retinal disparities across the central visual field while people engaged in everyday activities in the natural environment using techniques that have been previously developed [113, 37, 56, 71, 1].

Apparatus

Agostino Gibaldi and Vasha Dutell used a custom head-mounted device to measure the 3D structure of the scene and where subjects fixated in those scenes [29, 39]. The apparatus, shown in the left panel of Fig. 3.1, consists of two outward-facing cameras (Sony XCD-MV6) that capture stereoscopic images of the scene in front of the subject, and a binocular eye tracker (SR Research EyeLink II Eye Tracker - RRID: SCR_009602) that measures gaze direction for both eyes. The cameras allowed us to determine the distances of points in the visual scene over a 75×58° field of view at 30Hz with a resolution of 640×480.

Figure 3.1: Apparatus for the natural and VR-gaming environments. Left: Apparatus for measuring disparities and eye movements in the natural environment. The custom head-mounted device has two cameras and a binocular eye tracker. A backpack contains the host computer and power supply. Right: HMD used to measure eye movements and disparities in the VR-gaming environment. Handheld controller enables game interaction.

Participants

Four people with normal or corrected-to-normal vision participated. They were 24-38 years of age. The experimental protocol was approved by the Institutional Review Board at our university in accordance with the Declaration of Helsinki. Participants signed informed consent forms before participating.

Everyday Tasks

Subjects each performed six tasks that were chosen because they are representative of everyday activities. The tasks were:

- *Make Sandwich*: Subjects assembled a peanut-butter-and-jelly sandwich.
- *Edit Text*: Subjects revised a text document on a desktop computer from a printed document with pencilled edits.
- *Play Video Game*: Subjects played a shooting game on a desktop display using the mouse as the controller.
- *Order Coffee*: Subjects ordered a coffee at a local café while socializing.
- *Indoor Walk*: Subjects walked through a campus building looking for an office.

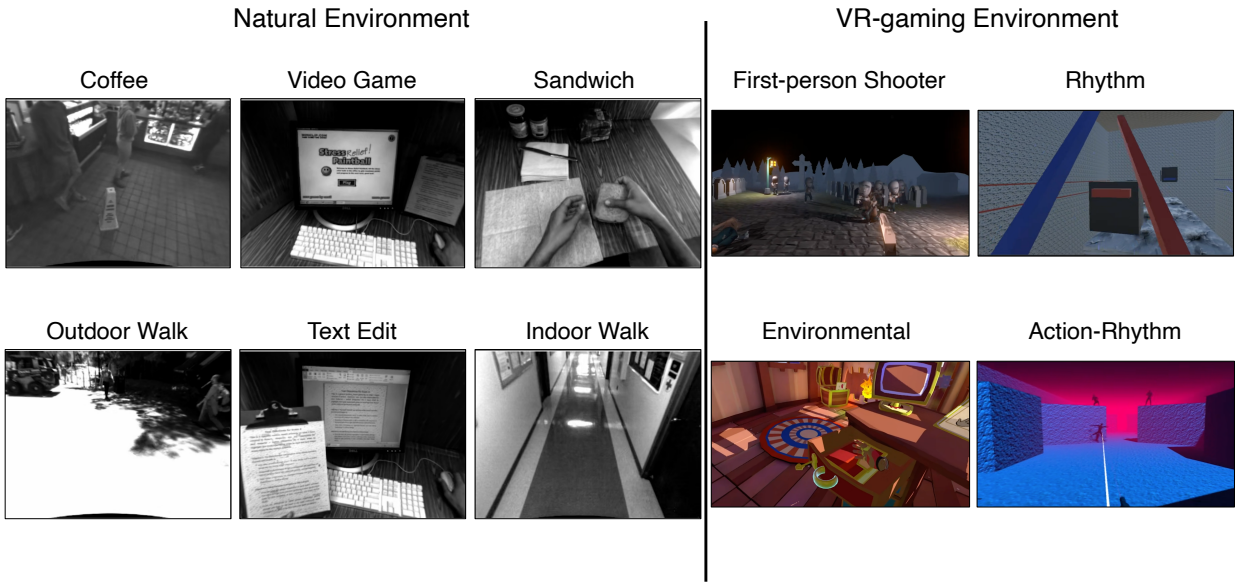

Figure 3.2: Examples of images from natural and VR-gaming environments. Left) Single frames from each of the six natural tasks. Right) Single frames from each of the four VR video games.

- *Outdoor Walk*: Subjects walked in natural and urban outdoor areas.

Examples are provided in the left panel of Fig. 3.2.

Calibration

A custom calibration procedure was designed to define an objective reference for binocular gaze data and map the binocular fixation point into the 3D scene reconstructed by the stereo scene camera.

The procedure was performed at the beginning of each experimental session. The subject's head was carefully aligned to the calibration apparatus using a custom bite bar that allowed three directions of translation (X, Y, & Z) and three rotations (pitch, yaw,& roll) to accurately position the eyes with respect to the calibration display screen. We also had an adjustment for the inter-ocular distance of each subject. When the subject was aligned, the surface normal from the center of the screen intersected the midpoint of the subject's inter-ocular axis (*i.e.*,, the cyclopean eye) at 1m from the screen. The inter-ocular axis was parallel to the horizontal midline of the screen, setting head yaw and roll. The reference orientation for head pitch was the Camper Plane (the plane passing through the ears' canals

and the tip of the anterior nasal spine) which was parallel to the floor. To optimize the eye-tracker calibration and the working range of the device, we adjusted the pitch of the cameras relative to the subject's head for each task. Pitch was 0° for *Indoor Walk* and *Outdoor Walk*, -4° (head pitched upward) for *Order Coffee*, and -8° for *Make Sandwich*, *Edit Test*, and *Play Video Game* during calibration. Those pitches, of course, affected the orientation of the cameras relative to the head when the subjects performed the various tasks with their head in natural orientations. The pose of the stereo scene camera in the head-referenced coordinate system was estimated by displaying a checkerboard pattern on the calibration display screen and then acquiring images of that pattern.

Post-processing

Binocular gaze direction and retinal disparity were computed in post-processing.

The eye tracker returned the pixel position of the fixation point for each eye on the calibration display screen. The binocular fixation point was then computed as the point where the two gaze-direction vectors intersected or came closest to intersecting. The head pose relative to the screen was used to transform the screen-referenced gaze data into real-world, cyclopean-eye–referenced coordinates using the Camper Plane to set the reference azimuth and elevation for the estimated binocular gaze directions.

The eye tracker does not measure eye torsion (rotation about the line of sight). To incorporate torsion, we modeled the torsion of each eye with the binocular extension of Listing's Law (L2) [124, 120, 112]. We used an L2 gain of 0.8, which is the average gain across people [112]. That assumption is not critical because using gains from 0.6–1 causes essentially no change in the estimated retinal disparities.

To reconstruct the 3D structure of the visual scene, we used common computer vision techniques available in OpenCV [80]. The stereo pair was first undistorted [12, 139] and stereoscopic disparity was computed from the rectified images [50]. The extrinsic parameters of the stereo camera, estimated at calibration [83], were used to transform stereoscopic disparity into a 3D point cloud of the scene. The stereo camera pose with respect to the calibration screen was used to transform the 3D scene into the head-referenced coordinate system, consistent with the subject's gaze direction. Finally, the 3D scene was projected onto two virtual cameras, simulating the pose of the left and right eyes, to compute the retinal disparities experienced by the subject given where he/she was fixating.

To combine data across tasks, we used the American Time Usage Survey (`https://www.bls.gov/tus/`). The Survey describes the amount of time adults spend on average on a large number of everyday activities. We had five people indicate which of our six tasks was most representative of each of the Survey tasks. We averaged their responses. From the averages, we computed a weight for each of our tasks in order to derive the best estimate of how our tasks represent the time people spend during waking hours. This procedure yielded the weighted statistics for gaze distance and direction (Fig. 3.4) and retinal disparity (Fig. 3.7).

Database

The data were ported into a database consisting of ∼880,000 stereo pairs, disparity data, and fixation data. The database is publicly available. The URL has been redacted to maintain anonymity.

Disparities & Eye Movements in Video Games

We measured the distributions of gaze direction and distance, and the distribution of retinal disparity across the visual field during video-game play in an HMD. Unfortunately, video-game companies did not allow access to the 3D structure of virtual scenes during game play. To circumvent this issue, we developed four video games in Unity (version 2019.3.8f1), and saved gaze data and depth buffers during game play. The four games were designed to be representative of popular VR video games (Sec. 3.2).

Depth Buffer Acquisition

To save the 3D geometry of the environment during game play, we acquired the scene depth using Render Textures in Unity. At runtime, a depth render texture is created where each pixel value of the texture contains a high-precision depth value. The depth value represents Unity view-space depth ranging non-linearly between [0,1] with a precision of 16 bits, depending on the platform and game configuration. We converted from buffer values to distances in meters:

$$scale = \frac{\left(\frac{1}{Z_{min}} - \frac{1}{Z_{max}}\right)}{levels}$$
$$Z = Z_b scale + \frac{1}{Z_{max}}$$
$$z = \frac{1}{Z}$$

where $levels$ is 2^{16}, Z_{min} and Z_{max} are the nearest and farthest values, Z_b is a buffer value, Z is distance in diopters, and z is distance in meters. Textures were acquired for each game for the left eye at a minimum of 40 depth frames per second. Saving these textures to disk during runtime can affect game play by reducing frame rate. To ensure the best user experience, we down-sampled the textures by a factor of 4, encoding them to 363×403 PNG images before saving to disk. We found that this resolution was more than adequate for measuring fixation and disparity statistics. Examples of the depth buffers for each game are shown in Fig. 3.3.

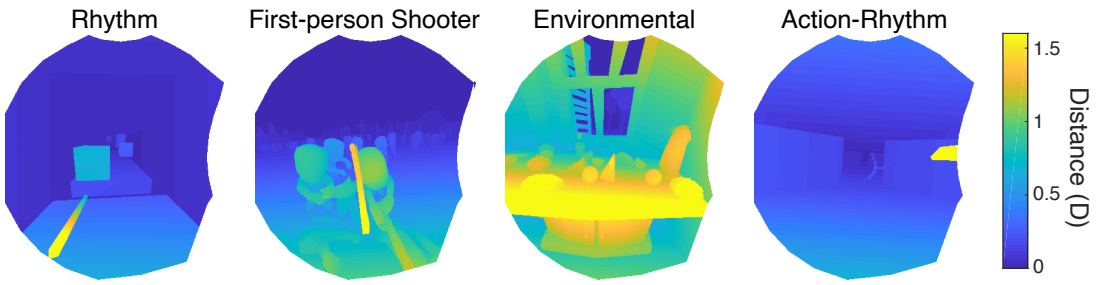

Figure 3.3: Example depth-buffer values from the video games. One frame seen by the left eye is shown from left to right for the Rhythm, First-person Shooter, Environmental, and Action-Rhythm games. Colors represent distance in diopters as indicated by the color bar.

Apparatus

Video games were presented using the HTC Vive Pro Eye headset, shown in the right panel of Fig. 3.1, which includes a built-in eye tracker (Tobii XR). The Tobii XR SDK V1.8.3 [119] and Vive SRanipal SDK V1.1.0.1 [128] were used to access tracking data at 90Hz. According to the manufacturer, tracking accuracy is 0.5–1.1° [127]. The HMD includes two OLED screens, one for each eye, with a resolution of 1400×1600 pixels per eye.

We measured the monocular and binocular fields of view in the Vive Pro Eye. To do this, we generated a row and column of colored cubes each 2cm wide and high in the virtual scene at a distance of 100cm. Participants wore the headset and viewed the cubes with just the left eye or just the right. To assess the horizontal field of view, they indicated the leftmost and rightmost cubes that were visible to the left and right eyes. They did the same for the highest and lowest visible cubes. The results differed slightly from one subject to another because the distance from their eyes to the screen differed. From the average measurements, we determined that the monocular fields extend ∼47° from straight ahead temporally (*i.e.*, left limit for left eye and right limit for right eye) and ∼36° nasally (right and left limits for left and right eyes, respectively). Thus the monocular fields are each ∼73° horizontally and ∼93° vertically. Consequently, with the eyes in forward and parallel gaze (*i.e.*, vergence = 0°), the binocular field is ∼72° wide and ∼93° high. These values agree reasonably well with previous reports [127].

According to the manufacturer, the optical distance from the viewer's eye to the screen is 65cm (1.54 diopters). We confirmed this using a camera with short depth of field positioned where the eye is meant to be. We focused the camera on the displayed content and then, without changing focus, moved the camera to an optical bench where we translated it relative to an eye chart to find best focus distance.

The games were run on a PC with Windows 10 64-bit operating system, an Intel(R) core(TM) i7-8700k processor with 3.7GHz, 48GB RAM, and two NVIDIA TITAN V graphics

cards. Video-game frame rate reached ∼80Hz.

Participants

Ten people with normal or corrected-to-normal vision participated. They were 23–37 years of age. The experimental protocol was approved by the Institutional Review Board at our university in accordance with the Declaration of Helsinki. Participants signed informed consent forms before participating.

Video Games

Participants each played four video games for 3min each. The order of game presentation was counterbalanced using a Latin Square design. Our games were designed to be representative of the most popular VR games. We used data from from Steam [114], the video-game distribution platform, to guide our game designs. The selected games have a representative range of depths (far, middle, and near/reach space) and tasks (first-person shooter, rhythm game, environment simulation). The games were:

- *Rhythm Game (mid/near depth task)*: Cubes representing the beats of background music move toward the player. The player swipes at the cubes with a saber. This game is similar to *Beat Saber*®, the 3rd most popular VR game [114].
- *First-Person Shooter Game: (near/mid/far depth task)*: Zombies in a haunted graveyard approach the player. Players kill them using a gun and axe. This game is similar to *Arizona Sunshine*®, the 4th most popular VR game [114].
- *Environmental Simulation Game (near depth task)*: To escape a cabin, players must complete tasks that are revealed as they explore the cabin. This game is most similar to *Job Simulator*®, the 21st most popular VR game [114].
- *Action-Rhythm First-Person Shooter Game (mid/fear depth task)*: Players are transported forward along a path. Enemies appear randomly and shoot at the player who must shoot the enemies or dodge the bullets to avoid being hit. This game is most similar to *Pistol Whip*®, the 17th most popular VR game [114].

Examples are provided in the right panel of Fig. 3.2.

Calibration & Validation

At the beginning of each session, the participant placed and adjusted the HMD on the head to a comfortable position that enabled a full field of view. They also adjusted the separation between the left and right screens to match the inter-ocular distance.

We then calibrated the eye tracker. The Vive Pro Eye provides a five-point calibration procedure, but it did not provide sufficient accuracy for our measurements and slippage of

the HMD on the participant's head can invalidate the calibration. To attain more accurate and consistent tracking, we developed our own procedure. A small target was displayed at different positions in the central visual field, and the participant was instructed to fixate its center and press a button once he/she thought fixation was accurate. The targets were displayed at distances of 1.5 and 10m. They were shown in random order at both distances in 10 positions: straight ahead and at eccentric points in a 2×2 matrix with corners at ±10°. The procedure was performed before and after each game play. To assess tracking accuracy, we computed the gaze direction required to fixate the target center and the root-mean-square error (RMS) with respect to the measured direction. We computed RMS before and after the game play. 3 minute videogame sessions in which RMS exceeded 0.8° were discarded (which occurred about 1/3 of the time), and the participant repeated the pre-calibration, game play, and post-calibration. We chose 0.8° as our criterion because that value is similar to the repeatability of the eye tracker.

Post-processing

Binocular gaze direction and retinal disparity were computed in post-processing.

The data from the eye tracker were used to compute the pixel position of the fixation point for each eye in the left depth buffer image, and their binocular combination. The depth buffer and eye position, returned by the eye tracker, were used to transform the screen-referenced gaze data into real-world, cyclopean-eye–referenced coordinates, using the screen center to set the reference azimuth and elevation for the estimated binocular gaze directions.

The depth buffers were used to reconstruct the 3D scene [15]. The gaze data were then mapped into the reconstructed scene, and the 3D scene was projected into the left and right eyes to compute the retinal disparities experienced by the subject given where they were fixating [38]. To incorporate eye torsion, we used the same approach used for the natural-disparity experiment.

For summary statistics, we wanted to combine the data across the four games. Unlike what we did with the tasks in the natural-environment study, we weighted the data from the games equally because each type of game is popular in the gaming community. This yielded average statistics for gaze direction and distance (Fig. 3.4) and retinal disparity (Fig. 3.7).

3.3 Results

Fixation Directions & Distances

The left and right panels of Fig. 3.4 show the statistics of binocular eye movements in natural and VR-gaming environments, respectively.

The upper left panel shows the distribution of gaze directions relative to the head for the natural environment. We find that the direction of gaze is most commonly straight ahead and slightly down relative to primary position. Secondary directions—leftward, rightward,

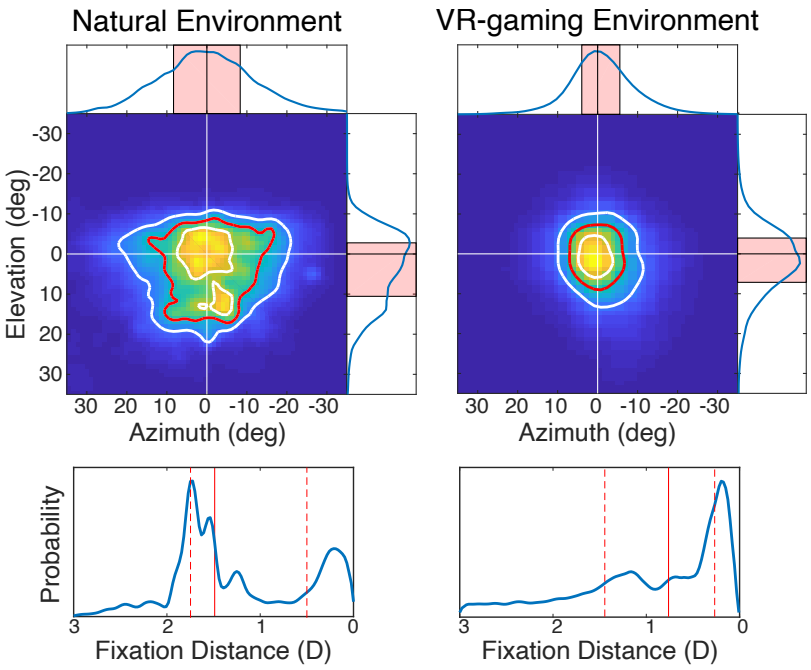

Figure 3.4: Probabilities of fixation directions and distances in natural and VR-gaming environments. Upper panels: The probability of fixation directions. Horizontal gaze direction is on the horizontal axis and vertical on the vertical axis. The left panel plots the probability of different gaze directions in the natural environment and the right panel the probability in the VR-gaming environment. Red contours represent 50% of fixations. White contours are 25th and 75th percentiles. Marginal probabilities are shown on the right and above. Pink areas represent 50% of the fixation directions. The natural environment data are based on the weighted combination across the six tasks; the VR data on the average for the four video games. Lower panels: The probability of fixation distances in diopters in the natural and VR-gaming environments. Near distances are on the left and far on the right. Median fixation distances are represented by red lines and 25th and 75th percentiles by red dashed lines.

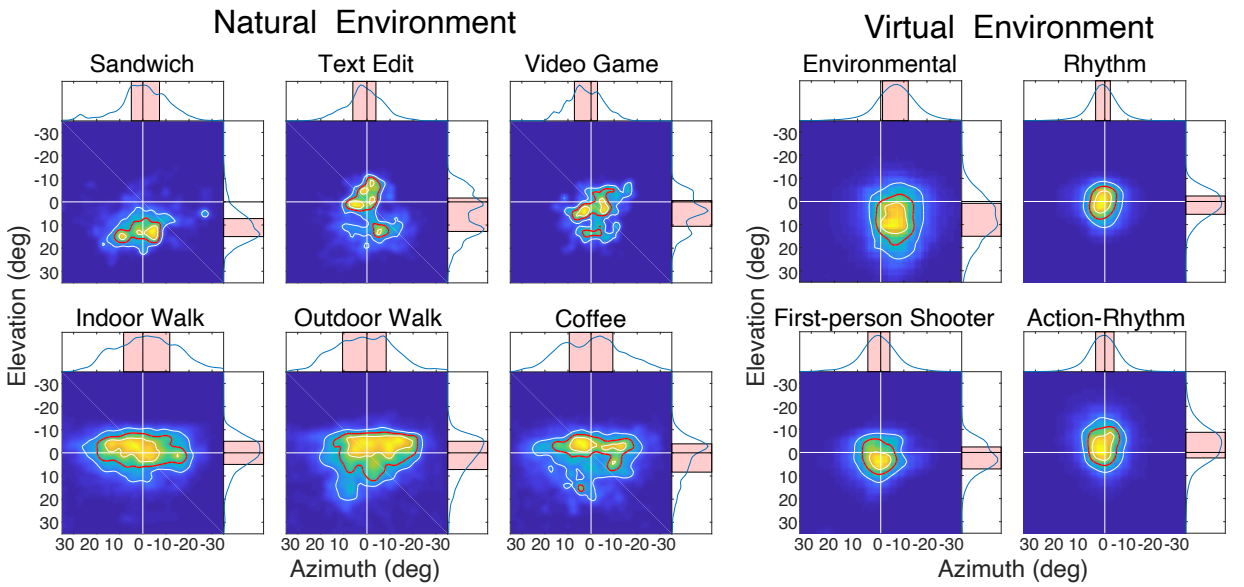

Figure 3.5: Probability of fixation direction for each natural task and video game. Horizontal gaze direction is plotted on the horizontal axes and vertical on the vertical axes. Red contours encircle 50% of fixations. White contours are 25th and 75th percentiles. Marginal probabilities are shown on the right and top. Pink areas contain 50% of fixations.

upward, and downward—are the next most common [113, 37, 56, 117]. There are few gaze directions more than 15° from straight ahead because when people attempt to look at more eccentric points they usually execute a combined eye and head rotation [9, 44, 84].

The upper right panel of Fig. 3.4 shows the distribution of gaze directions for the VR-gaming environment. These data also show a strong tendency to look straight ahead. But the distribution of gaze directions is clearly narrower than that for natural environments. In fact the VR distribution is essentially isotropic because it does not contain as many fixations in secondary directions (especially left, right, and down) as occur in the natural environment. The narrowing of the distribution of fixation directions in an HMD has been reported by others who have hypothesized, as we do, that people tend to make smaller eye movements and larger head movements due to restricted field of view in HMDs compared to natural viewing [109, 111, 84] (see Sec. 3.4).

The left and right panels of 3.5 show the distributions of gaze direction relative to the head for each task in the natural environment (left panel) and VR-gaming environments (right panel). We observe a broad range of distributions, and that fixation direction is task dependent in the natural environment. We also find that the narrow distribution of gaze directions around straight ahead viewing in VR is remarkably consistent across tasks and visual environments.

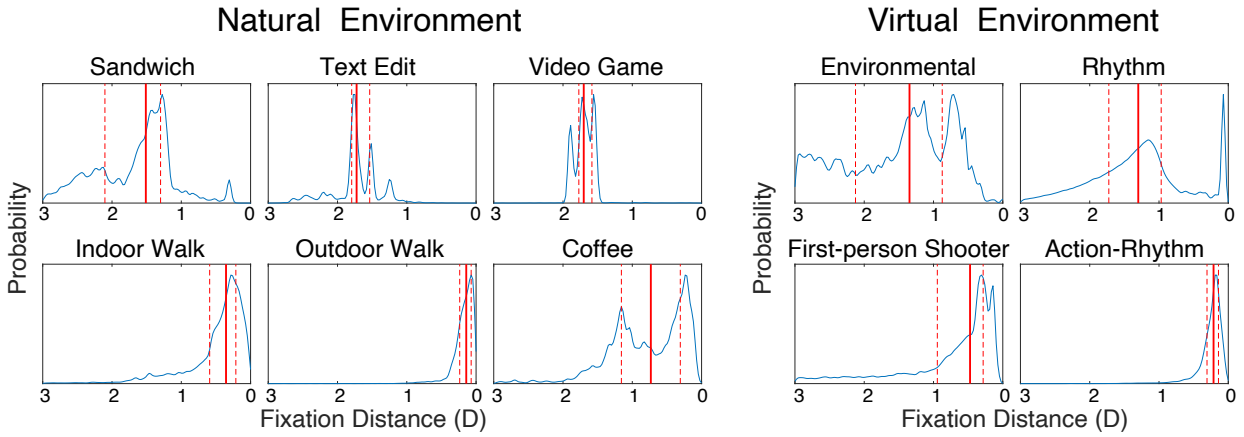

Figure 3.6: Probability of fixation distances for each natural task and video game. Fixation distance is plotted in diopters: Near distances on the left and far on the right. Median fixation distances are represented by red lines and 25th and 75th percentiles by red dashed lines.

The fact that directions are concentrated near straight ahead in the VR-gaming environment, and that this pattern is consistent across tasks and visual environments, is useful information for foveated rendering applied to video games [43, 82, 4]. Specifically, one might achieve essentially the same compute-time benefit as achieved with rendering coupled with eye tracking by not doing eye tracking and simply expanding the sharply rendered region to cover the great majority of fixation directions: roughly the central 10–15°.

The lower left panel of Fig. 3.4 shows the distribution of fixation distances in the natural environment. We observe a broad distribution with a median value of ∼70cm (1.5D); that distance is indicated by the solid red line. We also find that the distance of gaze varies significantly from one task to another as shown in the left panel of Fig 3.6. When walking outdoors the most common distance is ∼500cm (0.2D). When making a sandwich the most likely distance is ∼62cm (1.6D).

The lower right panel of Fig. 3.4 shows the distribution of fixation distances in the VR environment. We observe a much narrower and farther distribution than in the natural environment. The bulk of fixation distances is from 80–350cm, although the distances vary somewhat from on game to another (as shown in the right panel of Fig 3.6). The median distance is ∼125cm (0.8D), indicated by the solid red line. This is nearly twice the distance observed in the natural environment. Thus, on average, players of video games in HMDs tend to look farther and over a narrower range of distances than they do when performing everyday tasks in the natural environment. We examine the consequences in Sec. 3.4.

When a person looks at a near object off to the left or right, the object is closer to one eye than the other creating a larger retinal image in the closer eye. When the object is also up or down, the person must make a vertical vergence movement to fixate the object

accurately [102] and this can produce discomfort [53]. The left panels of Fig. 3.4 show that this combination of near gaze in an oblique direction is relatively rare in the natural environment. The right panels show that this combination is even rarer in the VR-gaming environment. Thus, the vertical disparities experienced in the VR environment are quite small and probably not problematic.

Disparity Statistics

The left and right panels of Fig. 3.7 show the statistics of horizontal disparity in natural and VR-gaming environments, respectively. It is important to note that the disparities are in retinal coordinates, and that to determine disparities in those coordinates, we needed to know both the 3D scene geometry and where participants fixated in those scenes.

The left panels reveal clear regularities in naturally occurring disparities. The upper panel shows median horizontal disparities across the visual field. There is a striking change from the lower to the upper field. The median disparity in the lower visual field (*i.e.*, below fixation) is positive (crossed) while the median disparity in the upper field is negative (uncrossed). These are large tendencies. For example, 10° above fixation, 70% of the disparities are negative. Thus, given where people tend to fixate, the natural environment creates a pattern of disparities that is slanted top back. The natural data also exhibit a systematic change from the left to the right field. Median disparity changes from negative (uncrossed) on the left to zero near the fovea to negative again on the right. The top-back pitch of the data is highlighted in the lower panel, which shows the median and range of disparity from the lower to upper field.

The left panel of Fig. 3.8 shows the statistics of horizontal disparity for each of the natural tasks. The majority of tasks in the natural environment show this pattern of far disparity in the upper visual field and near disparity in the lower visual field, although the magnitude of the top-back pitch of the data varies by task. The exception to this pattern is the *Text Edit* and *Video Game tasks*. In both cases, the computer, a near object in the scene, appears in the upper visual field leading to near disparity above fixation. Despite these exceptions, when the data are weighted by the ATUS, median disparity shows a clear far above and near below fixation pattern.

For humans to perceive depth from disparity, the visual system must determine which points in the left-eye's image correspond to points in the right-eye's image. The visual system utilizes the environmental regularities mentioned above to solve this binocular correspondence problem. Specifically, the search for disparity in a given location in the visual field is centered on corresponding retinal points. For every retinal location in one eye there is a location in the other eye that forms a pairing with special status in binocular vision. Rays projected from those corresponding-point pairs intersect in the world on a surface called the *binocular horopter* [78, 47]. The horopter is pitched top back [76, 110, 22]. So for objects above current fixation to fall on the horopter they must be farther than fixation while objects below fixation must be nearer. The horopter is also farther on the left and right (relative to the zero disparity surface) than in the center.

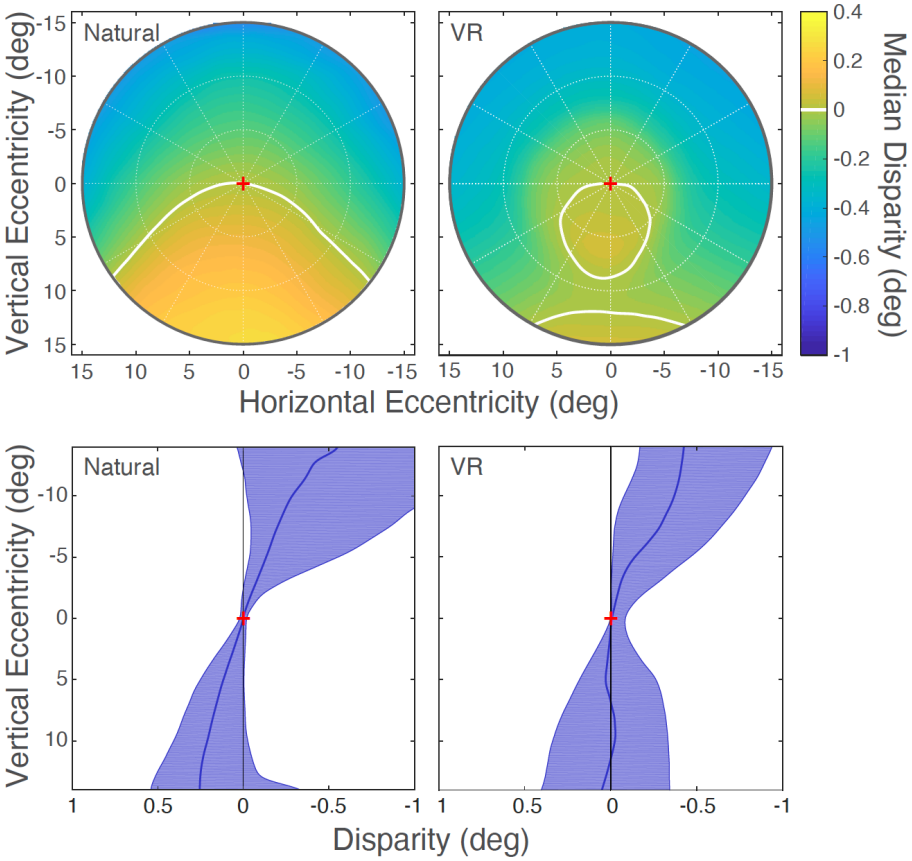

Figure 3.7: Median disparity as a function of field position for natural and VR-gaming environments. Upper panels: Median horizontal disparity for each field position. Fovea is in the middle. Upper field is up and left field is left. The white contours represent zero disparity. The left panel shows the data from the natural environment. The data have been generated by the weighted combination across the six tasks. The right panel shows the data from the VR-gaming environment. The data have been generated from the average across the four games. Lower panels: Cross sections along the vertical meridian. Disparity near the vertical meridian is plotted as a function of vertical eccentricity. Data for the natural and VR environments are in the left and right panels, respectively. The thick blue curves are the medians. Shaded areas indicate disparities between the 25th and 75th percentiles.

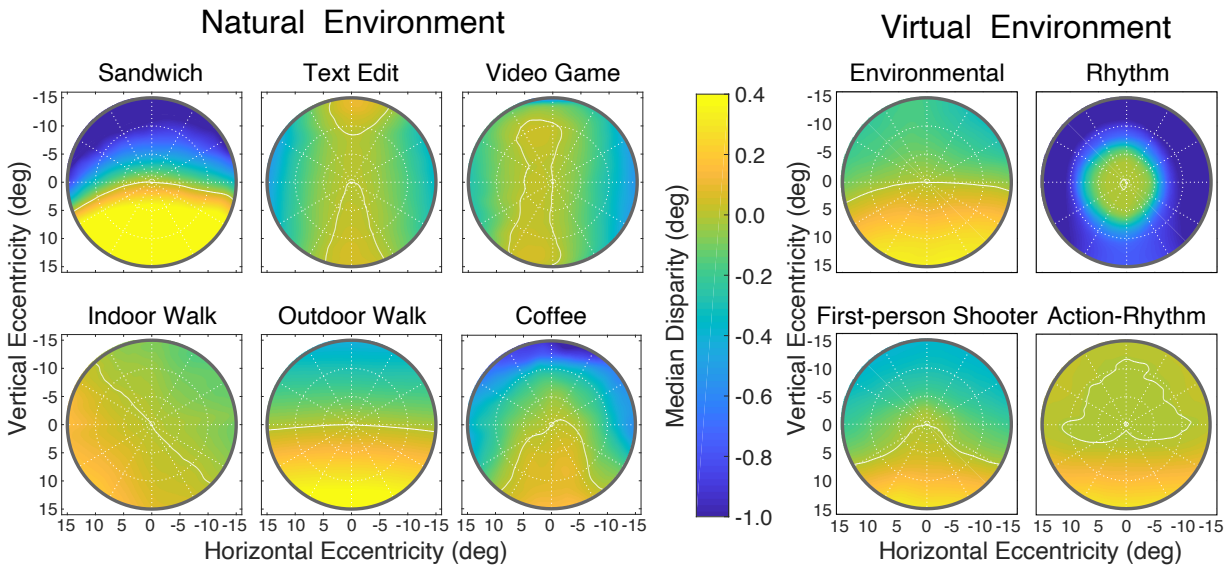

Figure 3.8: Median disparity across the visual field for each natural task and video game. Median horizontal disparity is plotted in each panel. Fovea is in the middle of each heatmap. White curves show where median disparity is zero. The data for the six natural tasks are shown on the left and the data for the four video games on the right.

Why is the horopter important? Binocular vision is best for objects on or near the horopter: fusion is guaranteed and depth discrimination is most precise [13, 88, 129, 11, 105, 33, 78]. Importantly, the shape of the horopter is quite similar to the central tendency of the natural-disparity statistics (Fig. 3.7). Therefore, fusion and accurate stereopsis are guaranteed for the most likely natural scenes.

The disparity statistics are also relevant to oculomotor behavior. When people make upward saccadic eye movements to a stimulus whose distance is ambiguous, their eyes diverge and when they make downward saccades their eyes converge [137, 37, 30, 21]. These vergence biases are consistent with natural-disparity statistics. Consequently, the biases ensure that when the eyes land at the end of a saccade in the real world they will be fixating the most likely distance of the new target. This speeds up visual processing because it minimizes the likelihood of having to make another vergence movement to accurately fixate the new target.

Thus, it is very important that the horopter and oculomotor biases are compatible with the statistics of the natural environment. Otherwise, these biases would be counter-productive.

We next turn to the disparities in the VR-gaming environment. The upper right panel of Fig. 3.7 shows median disparities across the visual field in that environment. The data were obtained by averaging across all subjects and the four video games. The median disparities are qualitatively similar to those from the natural environment. The VR statistics exhibit a

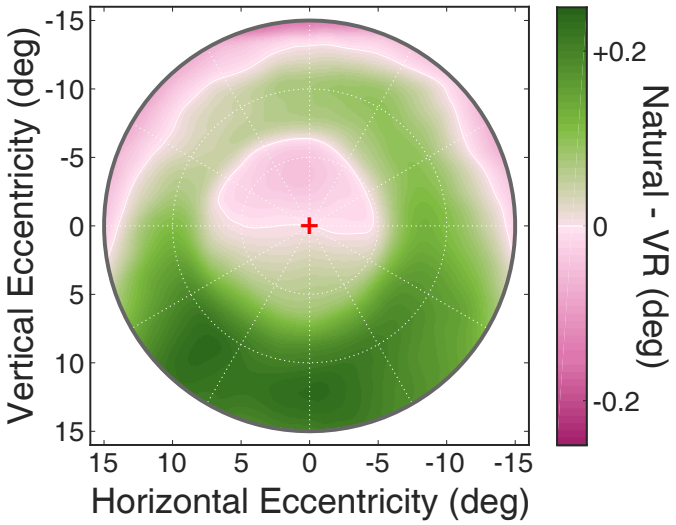

Figure 3.9: Differences between median disparity in natural and VR-gaming environments. The difference—natural minus VR—is plotted for all field positions. The color bar indicates disparity difference in degrees. Green regions are where the natural disparities are more positive (uncrossed) than the VR disparities. Purple is where they are more negative (crossed).

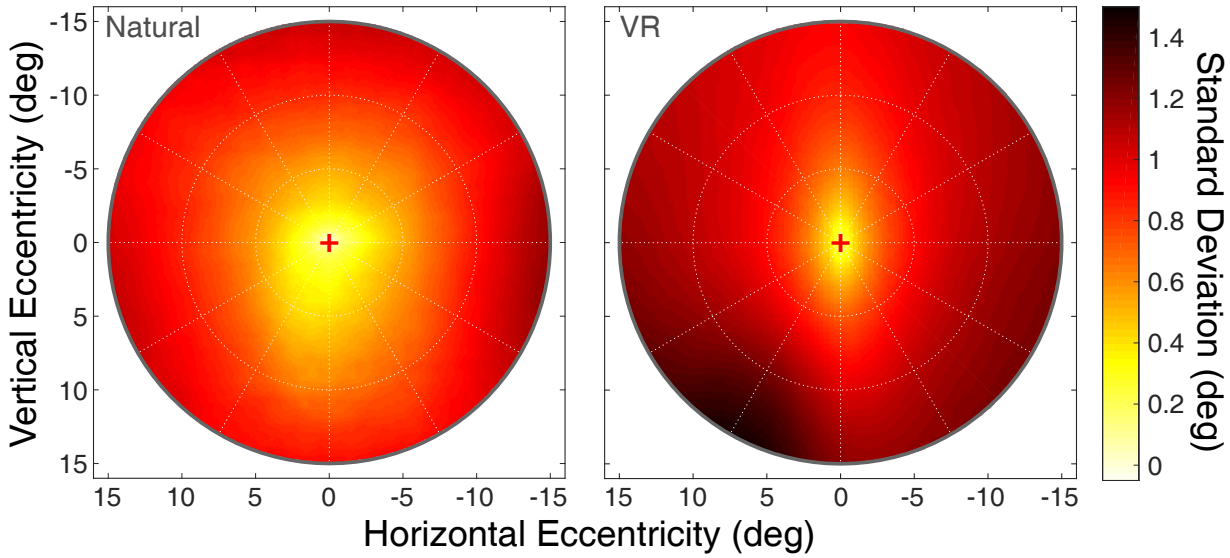

Figure 3.10: Standard deviation of horizontal disparity. The left panel shows the data from the natural environment and the right panel the data from the VR-gaming environment.

bottom-to-top change from positive to negative disparity (near to far) and the left-to-right change from negative to zero and back to negative. But these changes are much smaller in the VR environment than in the natural. We highlight this in Fig. 3.9, which plots the difference between the medians (natural–VR) for each position in the visual field. There is a prominent difference in the lower field where disparity is decidedly more positive in the natural than in the VR environment. The right panel of Fig. 3.8 shows the statistics of horizontal disparity for each of the VR environments. The statistics of the *Environmental* and *First-Person Shooter* game show a qualitatively similar disparity pattern to tasks in the natural environment, such as the *Outdoor Walk*. In contrast, the *Rhythm* and *Action-Rhythm* games are inconsistent with the disparity pattern of far above and near below fixation seen in the natural environment. Unlike the natural-environment data, the bottom-to-top change in the aggregate VR data is not large enough to match the horopter's pitch. And the left-right change is not large enough to match the horopter's horizontal curvature. We hypothesize that solving the binocular correspondence problem, obtaining fusion, achieving precise stereo vision, and making accurate vergence during saccadic eye movements is compromised in the VR-gaming environment for these tasks.

We next examined the variability of disparity in the two environments (Fig. 3.10). In the natural environment (left panel), the standard deviation increases roughly in proportion to eccentricity from a value close to 0° at the fovea to 2–3° at an eccentricity of 10°. This systematic change in disparity variation is reflected in the functional structure of the binocular visual system. The range of disparities that produce a fused image (*i.e.*, not a double image) grows in proportion to retinal eccentricity [78, 46]. The standard deviation in the VR environment increases more with eccentricity than in the natural environment, particularly in the left and right visual fields. We explored an implication of this finding by calculating from the disparity statistics, the probability of experiencing double vision across the visual field. We centered the range of fusable disparities on the horopter [5, 78]. We used data from Cooper et al. [22], Grove et al. [42], Schreiber et al. [104], and Nakayama [76] to fit a surface representing the horopter over the visual field. We used data from Ames et al. [5], Hampton & Kertesz [46], and Ogle [78] to construct a range of fusable disparities around the horopter. We then calculated for each field position the proportion of observed disparities that would fall outside of the fusable range. The results for the natural and VR-gaming environments are plotted in the left and right panels of Fig. 3.11, respectively. Clearly, the proportion of disparities that could produce double vision or suppression is greater in the VR-environment, particularly in the left and right visual fields.

We also observe that the spread of horizontal disparity in the natural environment is much greater than the spread of vertical disparity. Specifically, the aspect ratio of the joint distribution of horizontal and vertical disparity is ∼20:1. This statistical property is manifest in the binocular visual system. For example, cortical neurons in primates have much more variation in their preferred horizontal disparity than in their preferred vertical disparity [23, 28]. Furthermore, when presented stereoscopic stimuli in which the direction of disparity (*e.g.*, horizontal, vertical, or oblique) is ambiguous, humans exhibit a strong bias to assume that the direction is horizontal [123, 89]. The spread of horizontal disparity relative to that

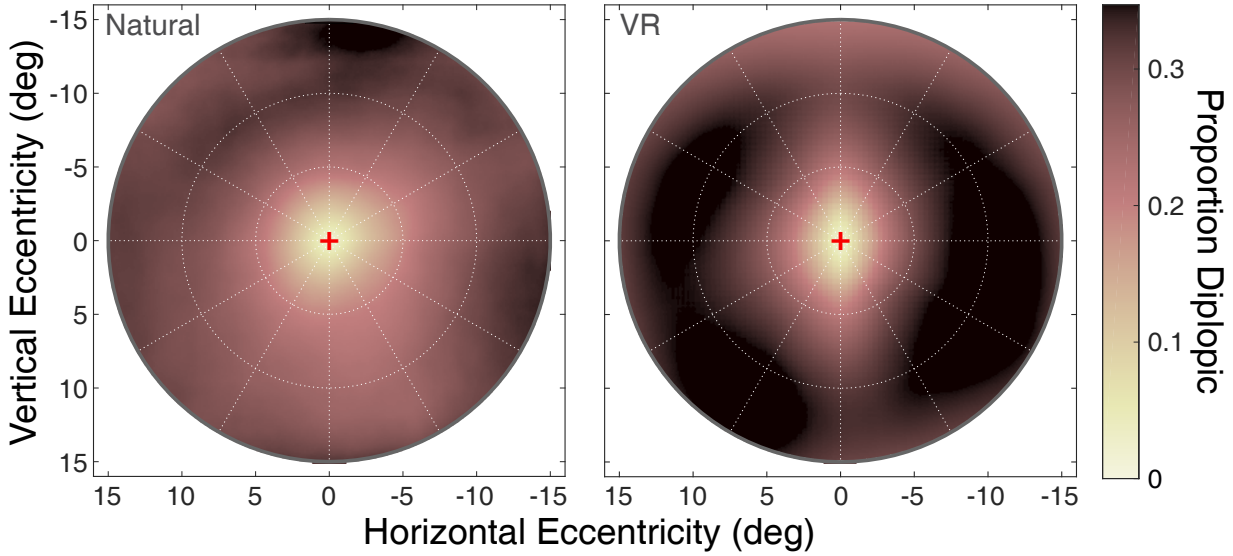

Figure 3.11: Proportion of disparities that may produce double vision in natural and VR-gaming environments. Left: Proportion for the natural environment. Right: Proportion for the VR environment.

of vertical disparity in the VR-gaming environment is ∼16:1, which is quite similar to the natural ratio. Thus, this aspect of disparity in the virtual environment is consistent with natural statistics.

Comparing Natural and VR-gaming Environments

Fig. 3.12 shows the correlations between tasks for the natural and VR-gaming environments, left panel shows fixation direction, right panel shows median disparity, and middle panel shows fixation distance. For fixation direction (Fig. 3.12, left panel) we observe correlations between the *Text Edit* task in the natural environment and almost all the video games in the VR environment. We also observe correlations between the *Video Game* task and *Outdoor Walk* task in the natural environment with the *First-Person Shooter* game. For fixation distance (Fig. 3.12, middle panel) we observe correlations between both walking tasks and the *First-Person Shooter* game in the VR environment, as well as between the *Action-Rhythm* game and the *Outdoor Walk* and *Coffee* tasks. For median disparity, (Fig. 3.12, right panel) we observe correlations between the *Sandwich* and *Coffee* tasks and the *Environmental* and *First-Person Shooter* video games. *Indoor walk* additionally correlated with the *First-Person Shooter* game.

The comparisons of aggregate data for fixation direction, distance, and binocular disparity generally found differences between the natural and VR-gaming environments. More

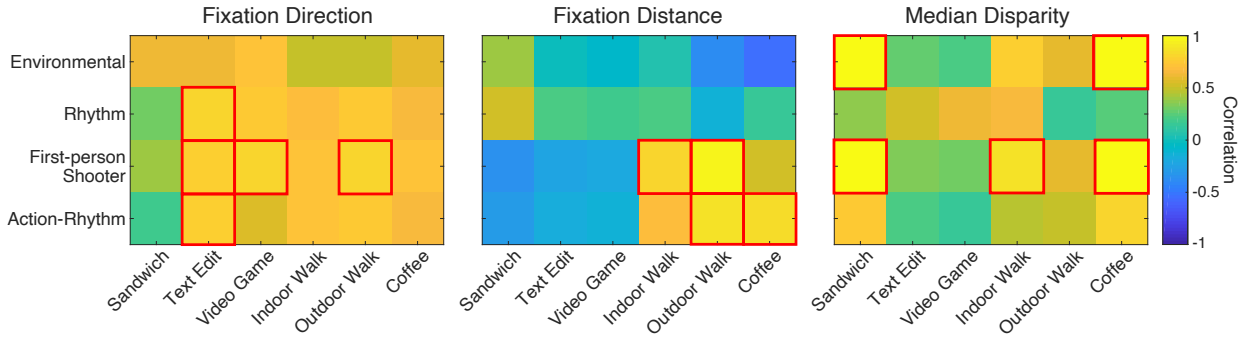

Figure 3.12: Correlations between data from the natural and VR-gaming environments. Pearson's correlation was computed between the two environments for fixation direction (left), fixation distance (center), and median disparity (right). Each square represents the correlation between the data from a natural task (horizontal axis) and the data from a VR game (vertical axis). Red squares indicate correlations > 0.75.

focused comparisons between the tasks used in these environments suggest the statistics for specific VR games better match the natural environment than others. For example, the *First-Person Shooter* and *Environmental* games feature environments (outdoor graveyard and indoor building respectively) consistent with the natural world, and this is reflected in the correlations with the disparity statistics of the natural environment. In contrast, the *Rhythm* and *Action-Rhythm* games, although representative of popular VR environments, are inconsistent with the disparity statistics of the natural environment. Clearly the statistics of fixation and binocular disparity are not only affected by the hardware constraints of the HMD, but are also influenced by the content of the video game environment.

3.4 Discussion

Vergence-Accommodation Conflict

When a viewer fixates an object in the natural environment, two oculomotor responses—vergence and accommodation—occur such that the resulting percept is single and sharp. Vergence and accommodation are negative-feedback control systems [32, 24, 101]. The vergence part takes disparity as input and generates converging or diverging eye movements to null the disparity at the fovea. The accommodation part takes retinal blur as input and adjusts focus to minimize the blur.

The vergence and accommodation parts of the control system work to drive their respective outputs to the same distance in the environment, so it makes sense that they communicate with one another through cross-links. Because of the cross-links, the act of converging

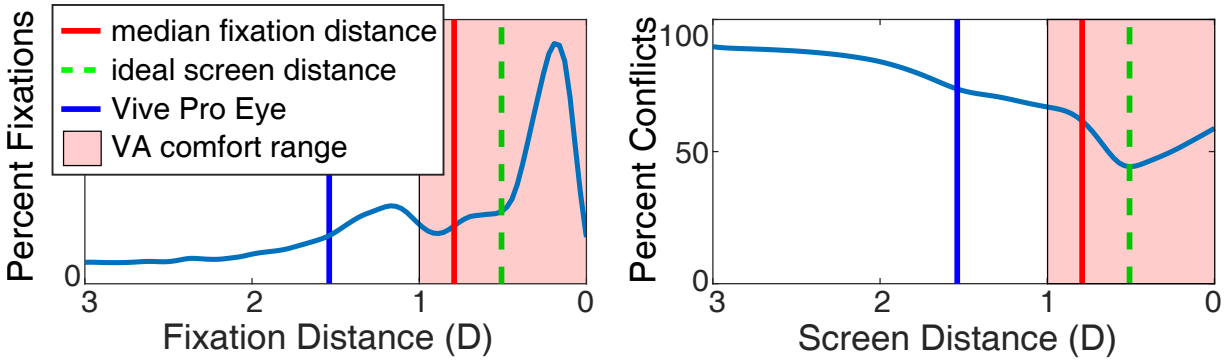

Figure 3.13: Probability of vergence-accommodation conflicts in video-game play. Left) Percentage of fixations of various distances. The horizontal axis is fixation distance in diopters and the vertical axis is the percentage of fixation distances averaged across the four games. Median fixation distance during game play is indicated by the red solid line at 0.8D (125cm). Optical distance of the screen in the HTC Vive Pro Eye is indicated by the solid blue line at 1.54D (65cm). The pink patch represents a ±0.5D comfort range for the vergence-accommodation conflict, centered on the median fixation distance. Right) Percentage of fixations generating uncomfortable vergence-accommodation conflict. The horizontal axis is screen distance in diopters and the vertical axis is the percentage of conflicts that exceed ±0.5D. The screen distance that minimizes the percentage of bothersome conflicts is indicated by the green dashed line at 0.51D (196cm).

or diverging causes the lens to change power (vergence accommodation) and the act of accommodating nearer or farther causes vergence movements (accommodative vergence). The cross-coupling increases speed and accuracy in the natural environment because the distance to which the eyes must converge and accommodate are the same [24].

The vergence-accommodation cross-coupling is counter-productive for viewing stereoscopic displays such as HMDs. In such displays, vergence must be to the distance of the virtual object of interest for a single, fused image to be seen. But the light comes from the display screen so accommodation must be to the screen distance for a sharp image to be seen. Thus, the distances for appropriate vergence and appropriate accommodation are often quite different. The difference is the *vergence-accommodation conflict*. When the conflict is non-zero, the visual system must work against the cross-coupling in order to fuse and sharpen the images.

The vergence-accommodation conflict is a cause of discomfort in users of sterescopic displays: the larger the conflict, the more the discomfort [51, 108, 57]. The discomfort risk is associated with age. Specifically, adolescents and young adults report more discomfort than middle-age adults [135] due to presbyopia in the older people [27]. The conflict also impairs visual performance [3, 51] and causes distortions of 3D percepts [132, 72].

The vergence-accommodation conflict is to first approximation always zero in the natural environment [61], but it can be quite large in VR environments. We used our measurements of content and fixations while playing video games to determine the distribution of such conflicts. Specifically, we used the distribution of fixation distances during game play (Fig. 3.4) to determine how frequently those vergence distances would be nearer or farther than the screen distance by ±0.5D, thereby creating a conflict large enough to cause discomfort [108]. Fig. 3.13 shows the results. The left panel shows the percentage of fixations at various distances, averaged across the four games. The median fixation distance is represented by the vertical red line. The right panel shows the percentage of conflicts greater than ±0.5D, given the distribution of fixation distances, as a function of screen distance. The screen distance in the Vive Pro Eye is indicated by the vertical blue line. The dashed green line represents the screen distance that would minimize conflicts. Obviously, it is much farther than the actual screen distance. Thus, discomfort due to vergence-accommodation conflicts would be minimized by nearly tripling the screen distance to 196cm (0.51D). (The screen distances of other commercial devices (*e.g.*, Oculus DK1, DK2 & CV1; HoloLens 1 & 2) are greater.) Of course, the degree of mismatch will depend strongly on the specific demands of the virtual environment and task (Fig. 3.8). Designers of HMDs and video games can use our data to better match screen and fixation distance to improve viewer comfort and performance [57].

Field of View in HMDs vs Natural Viewing

We observed (Fig. 3.4), as others have, that the direction of gaze is concentrated more straight ahead in HMDs than in natural viewing [84, 109, 54, 111]. We hypothesize that this is due to: 1) the smaller field of view in HMDs, and 2) how eye movements affect field of view in HMDs compared to natural viewing.

With respect to the first item, the horizontal and vertical fields of view in natural viewing are respectively ∼200° and ∼150°. The horizontal and vertical fields in HMDs are much smaller. In the Vive Pro Eye they are 73° and 93°. Because of the limited field, HMD users must rotate their heads more frequently to see objects of potential interest than they do in natural viewing.

With respect to the second item, eye movements affect the field of view differently in HMDs and natural viewing. In HMDs, the part of the virtual world an eye can see is fixed to the head because the display device is fixed to the head. As a consequence, making leftward and rightward eye movements does not expand the field seen by an eye; they simply shift the visible field across the retina. This is more complicated in natural viewing. The nasal field limit is imposed by the nose and bony orbit. The temporal limit is imposed by the *ora serrata*: the position in the retina where photoreceptors terminate. Thus the nasal limit is fixed to the head and the temporal limit is fixed to the retina. As a result, leftward and rightward eye movements expand the field seen by an eye. If one makes a leftward (or rightward) eye movement in natural viewing, the visible field expands leftward

(or rightward). We hypothesize therefore that viewers make larger eye movements in natural viewing than in HMDs because this allows them to expand the effective visible field.

Screen Displacement

The screens in most HMDs have a wider temporal field than nasal field, which increases the total field of view (the regions seen by one or the other eye). But this temporal bias decreases the binocular field of view (the regions that are imaged on corresponding regions in the two eyes). Fig. 3.14 explains this. It shows how screen size and positioning and fixation distance affect the binocular field of view. The left and right panels show respectively the situations with the eyes fixating at infinity (parallel lines of sight) and at a near distance. The upper and lower halves of the figure show the situations when the screens are symmetric about the line of sight (*i.e.*, eyes fixating ahead at infinity) and when the screens are shifted nasally. The width of the field seen by both eyes on corresponding retinal regions is indicated by *fov*. With symmetric screens (upper panels) the binocular field of view is widest (and identical to the two monocular fields) when the eyes are converged at infinity. As the eyes converge, the lines of sight intersect the screens at successively more nasal points, and the binocular field narrows. The ellipses at the bottom of the upper panels represent the fused binocular images. The red grid is the part of the field seen by the left eye and the green grid is the part seen by the right eye. The binocular field of view is the intersection of the two monocular fields. The total field of view is the union of the monocular fields. With nasally shifted screens (lower panels), the binocular field is wider when the eyes are converged.

Fig. 3.15 shows how the width of the binocular field of view depends on fixation distance and whether the screens are shifted nasally or temporally relative to straight ahead. The screens in the simulation are both 117cm wide at an optical distance of 65cm (as in the Vive Pro Eye). The widest binocular field for symmetric screens (*i.e.*, shift = 0cm) is 84° and is achieved when the eyes are converged at infinity. The Vive Pro Eye has temporal shifts of ∼10cm so the binocular field (yellow dotted line) is only 72° in that device when the eyes are fixated at infinity. Temporal shifts decrease the binocular field and nasal shifts increase it, especially at nearer fixation distances. Our data on the statistics of fixations in VR video games (Fig. 3.4) revealed a median fixation distance of ∼150cm (0.7D), which is indicated by the red arrow. For this fixation distance, symmetric screens (shift = 0cm) yield a binocular field of ∼81° while asymmetric screens like the Vive Pro Eye (shift = -10cm) yield a binocular field of just 70°. A wider binocular field of view is achieved for the median fixation distance by shifting the screens nasally by 5cm. Furthermore, the binocular field is wider for nearly all fixation distances with 5cm nasal shifts than with no shift or temporal shifts. This expansion of the binocular field is maintained when subjects make leftward or rightward movements while keeping the same fixation distance. Thus, HMDs would be more effective in presenting stereo information for the most likely fixation distances if the screens were shifted nasally.

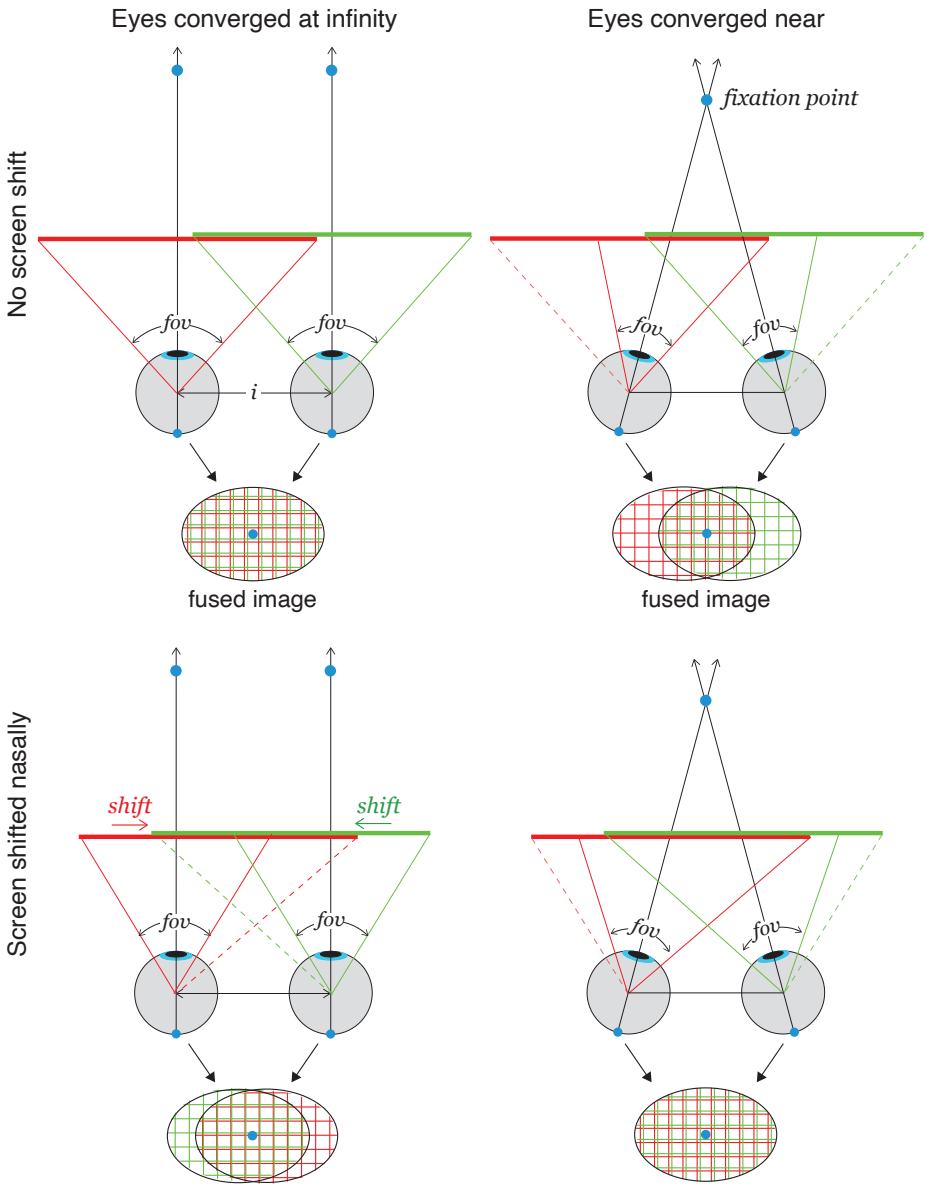

Figure 3.14: Geometry of binocular field of view. In the upper panels, the screens are symmetric about the lines of sight. In the lower panels, the screens are shifted nasally. The eyes are converged at infinity and at a near distance in the left and right panels, respectively. The display screens are represented by the thick red and green lines. The foveas are indicated by blue dots at the back of the eyes. The binocular field of view is represented by *fov*. The binocularly fused images are indicated by ellipses below the eyes. The red grid represents the part of the screen seen by the left eye and the green grid the part seen by the right eye. The foveas are indicated again by blue dots. In the upper left panel, the screen parts seen by the two eyes are superimposed, so the binocular field is the same width as the monocular fields. In the upper right panel, the fused images are displaced temporally because the eyes are converged. The binocular field is the part where the red and green grids are superimposed. It is narrower than in the left panel. In the lower right panel, the eyes are converged so the nasal shifts of the screens creates a wider binocular field of view.

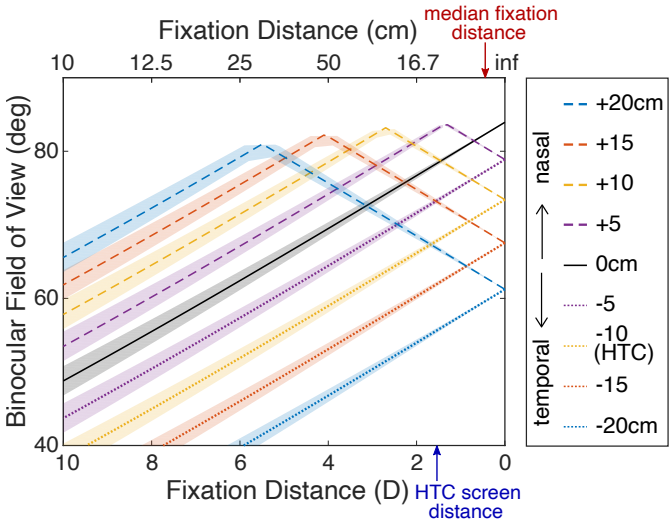

Figure 3.15: Binocular field of view, fixation distance, and screen position. The width of the binocular field is plotted as a function of the distance to which the eyes are converged and the horizontal shifts of the two display screens. Fixation distance is plotted in diopters on the lower axis and centimeters on the upper. Curves of different colors represent field size for different screen shifts. Black is no displacement (screens symmetric with lines of sight with forward gaze and eyes converged at infinity). Dashed lines represent displacements of both screens nasalward (toward the nose). Dotted lines represent displacement temporalward. An inter-ocular distance of 6.33cm is assumed; shaded areas represent ±1 standard deviation of inter-ocular distance [26]. The yellow dotted line represents field size for the HTC Vive Pro Eye which has a temporalward shift of ∼10cm. The blue arrow indicates screen distance in the Vive Pro Eye and the red arrow the median fixation distance in the VR-gaming statistics.

Adverse Effects Due to Deviations from Natural Environment

There are a variety of negative consequences and implications of presenting environments that do not conform to the regularities we observed for the natural environment.

1) Binocular fusion is determined by the 3D location of an object relative to the horopter and Panum's fusion area. As mentioned earlier (Sec. 3.3), the horopter is pitched top-back. This means that surfaces that are also slanted top-back are more likely to create a fused impression than surfaces that are pitched top-forward. A compelling example of this is the *Venetian-blind effect* [86, 121]. (A demonstration is provided in Fig. 3.16) A pattern of vertical stripes on a planar surface is viewed binocularly. The surface is then rotated about the horizontal axis. When the slant is top-forward, the pattern is not properly fused and a series of steps in depth is seen: a Venetian blind. When the slant is top-back, the pattern can be properly fused and the illusory depth steps are not seen. Thus surfaces that are consistent with the top-back pitch of the horopter are more fusable than surfaces that are

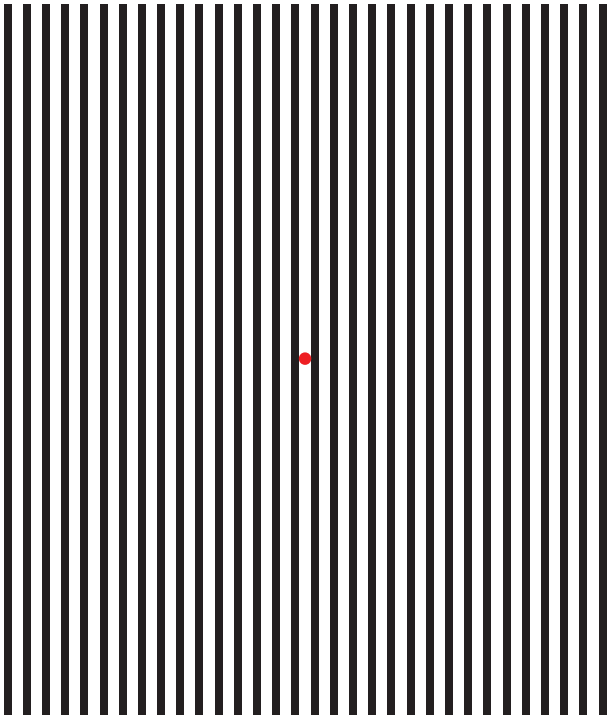

Figure 3.16: Vertical stripes that when viewed binocularly at an angle creates the impression of Venetian blinds. Place the figure on a flat surface. Rotate the surface about a horizontal axis so the pattern is slanted top-forward from your vantage point. Fixate the red dot with both eyes, being careful not to move the eyes from that dot. As you steadily fixate, you will see the apparent surface break up into horizontal steps like a Venetian blind. Find the slant that creates the effect. Now do the same with the pattern slanted top-back. A much greater slant is required to see the "blinds" in the top-back case than in the top-forward case. Adapted from [86, 121].

inconsistent.

2) Ergonomic researchers advise computer users to pitch desktop displays slightly top back to minimize viewing discomfort [6, 41]. The top-back pitch is consistent with the pitch of the horopter and with natural disparity statistics (Sec. 3.3). Environments that do not conform to the horopter produce more discomfort.

3) Panum's fusional area is centered on the horopter and increases in proportion to retinal eccentricity [46, 78], which means that the objects in the parafovea and periphery can have larger disparities before they produce a double (non-fused) percept. In the natural environment, the range of disparities is proportional to retinal eccentricity (Fig. 3.10), so the probability of experiencing non-fused, double vision is roughly constant across the visual field (Fig. 3.11). Our observations for the VR-gaming environment show that the range of

disparities in that environment is not proportional to eccentricity (Fig. 3.10). In particular, the range in the left and right visual fields is quite large, so double vision may be experienced more often in that environment than in the real world (Fig. 3.11). Furthermore, video games do not generally incorporate depth-of-field blur as it is experienced in the real world, which increases the likelihood of diplopia because Panum's fusion area is smaller for sharp than for blurred objects [103].

4) Oculomotor behavior should be consistent with natural statistics. When people make upward saccades, they tend to diverge the eyes. This is the same but to a lesser degree for leftward and rightward saccades. When people make downward saccades, they tend to converge [37, 30, 21]. These biases are useful because they ensure that the eyes at the end of a saccade are most likely to be aligned with the new fixation target. Because the statistics in the VR-gaming environment are not congruent with those in the natural environment, the relationship between saccades in different directions and the appropriate vergence is disrupted and should cause delays in the acquisition of new targets in the VR environment.

5) As discussed in Sec. 3.4, the vergence-accommodation conflict causes discomfort, poorer performance, and distortions of 3D perception [62, 122, 59, 58]. We found that such conflicts are common in the VR-gaming environment because players tend to fixate consistently farther in the virtual scene than the distance of the screen (Fig. 3.4). Thus, it is commonplace for significant vergence-accommodation conflicts to occur in that environment.

3.5 Conclusion

In summary, we found that fixation directions and distances are more restricted in VR-gaming environments than in the natural environment and that fixation distances are considerably farther in virtual environments. We used our data to calculate the screen distance and positioning that minimize discomfort and maximize the binocular field of view, respectively. We also found that the patterns of retinal disparity encountered in VR-gaming and natural environments are quite different from one another. The pattern is more variable in the virtual environment and does not exhibit the top-back pitch to the same degree as we observed in the natural environment. It is likely that the human visual system is well adapted to the average disparity statistics of the natural environment as we perform many different tasks across visual environments. In contrast, HMDs are typically used for very specific types of tasks, and the visual system is maladapted for some of these. Our data suggest concrete steps that developers of this technology can take in order to improve viewing conditions in HMDs. For example, in the future, it may be helpful if HMD designs allowed for the viewing condition (e.g., virtual screen distance) to be adjusted to suit the task. This would allow users to adjust the screen to be far for first-person shooter style games, and to be near for games that contain visual environments close to the observer.

This is the first set of experiments to compare and contrast the statistics of fixation and binocular disparity between the natural and VR-gaming environment. We chose different

tasks in each environment, as it was unclear if the statistics between these environments would differ. The aggregate data (as discussed in section 3.3) suggest clear differences between the natural and virtual environments, but statistics from specific video games suggest that some virtual environments are more consistent with the natural world than others. Future experiments should consider equating the hardware (real vs HMD) and content (natural viewing vs VR games) to be as similar as possible. For example, one could limit FOV in natural viewing to be consistent with the restrictions in HMDs, or one could create content for VR that is as consistent as possible with the natural environment. By matching hardware or content, one can determine how each influence the statistics of fixation and disparity in different environments.

Our study was limited to one type of headset and just four video games. It would be useful to expand this type of analysis to other headsets and other types of VR experience. It would be interesting as well to measure head movements as people experience natural and virtual environments in order to compare the combined eye and head movements made in these environments. We showed how the binocular field of view can be widened for common fixations, but this comes with a narrowing of the total field of view. It would be useful to determine the best trade-off between expanding the binocular field versus expanding the total field.

Chapter 4

Discomfort Associated with the (Un)natural Statistics of VR Gaming Headsets

4.1 Introduction

The natural environment is structured in ways that have significant impact on visual experience. As our environment is structured by gravity, many surfaces that surround us are earth vertical (such as walls and trees) or earth horizontal (the ground, floors, and tabletops). Additionally, our own behavior within the natural environment is constrained. People tend to view the environment with their head upright and fixate objects they are interacting with, or surfaces on which they are walking. These environmental and behavioral properties lead to statistical regularities in the images formed on the retinas and the types of eye movements people make.

The differences in each eye's image (known as binocular disparity) can be used to compute very precise depth information about the 3d layout of the scene. The challenge is the difficulty in determining binocular correspondence: Which points in the two eyes' images came from the same feature in the scene? Consider trying to solve the correspondence problem in an environment consisting of sparse small objects randomly distributed in three dimensions. In every direction, all distances would be equally probable, so disparities would be very broadly distributed. Accordingly, the search for solutions to binocular correspondence would have to occur over a very large range of disparities. The environmental and oculomotor constraints in the natural environment and in our own behavior are manifested in the brain's search for solutions to binocular correspondence, allowing a much more restricted and efficient search than would otherwise be. The human binocular visual system is adapted to these naturally occurring regularities. As a result, depth perception and eye movements in the real world are generally fast, precise, and performed with comfort [22, 113, 51].

For humans to perceive depth from disparity, the visual system must determine which

points in the left-eye's image correspond to points in the right-eye's image. The visual system utilizes the environmental regularities mentioned above to solve this binocular correspondence problem. Specifically, the search for disparity in a given location in the visual field is centered on corresponding retinal points. For every retinal location in one eye there is a location in the other eye that forms a pairing with special status in binocular vision. Rays projected from those corresponding-point pairs intersect in the world on a surface called the *binocular horopter* [78, 47]. The horopter is pitched top back [76, 110, 22]. So for objects above current fixation to fall on the horopter they must be farther than fixation while objects below fixation must be nearer. The horopter is also farther on the left and right (relative to the zero disparity surface) than in the center. Why is the horopter important? Binocular vision is best for objects on or near the horopter: fusion is guaranteed and depth discrimination is most precise [13, 88, 129, 11, 105, 33, 78]. Importantly, the shape of the horopter is quite similar to the central tendency of the natural-disparity statistics (Fig. 3.7). Therefore, fusion and accurate stereopsis are guaranteed for the most likely natural scenes.

In Section 3.3 we showed that virtual environments, such as virtual-reality (VR) games in HMDs, are not compatible with the regularities to which the visual system has become adapted. Specifically, in the natural environment, there is a dramatic and striking change in disparity between the upper and lower visual fields. The median disparity in the lower visual field (i.e., below fixation) is positive (crossed) while the median disparity in the upper field is negative (uncrossed). The VR disparity statistics show a qualitatively similar pattern of far disparity in the upper visual field and near disparity in the lower visual field, but these changes are much smaller in the VR-environment than in the natural environment. As Figure 3.7 shows, the median disparity in the lower visual field of the natural environment is more positive in the VR-environment (a difference of 900 arcseconds, which is quite notable). Additionally, the disparity distribution in the upper visual field did not show a consistent decrease with vertical eccentricity in VR, instead there are fluctuations between the VR-environment having more near or far disparity than the natural environment. This work additionally showed the proportion of disparities that could produce double vision or suppression is greater in the VR-environment, particularly in the left and right visual fields. These results suggest a mismatch between the disparity statistics of the natural and virtual reality environment.

As a result, it is probable that solving the binocular correspondence problem, obtaining fusion, achieving precise stereovision, and making accurate vergence during saccadic eye movements is compromised in VR-gaming environment. Taking this into account, incompatibility with the statistics of the natural environment may well cause viewer discomfort and lower visual performance in the VR environment. The main purpose of the work presented here is to test whether scene consistent content that is congruous with the statistics of the natural environment leads to greater viewer comfort and better performance in VR headsets.

4.2 Methods

We designed an experiment in a virtual reality head mounted display (HMD) to test whether having scene content consistent with the statistics of the natural environment matters with respect to viewer comfort and performance. To our knowledge, this is the first time the hypothesis has been tested in virtual environments. The following methodology was developed following two pilot experiments (described in Appendix A and Appendix B, in Sections 5.1 and 5.2). This pilot testing led to the current experimental design which includes controls to eliminate other confounding sources of discomfort such as the vergence-accommodation conflict. This conflict is known to cause discomfort and even nausea [45, 51, 62, 108, 122, 57], impairment in visual performance [3, 51], and distortions of 3D percepts [132, 72].

Apparatus

The methods and apparatus were identical to those described in 3.2 except for the material presented by the HMD and the participant's task. Participants were shown black text on a white page and asked to read for comprehension. The planar page was slanted top back (consistent with the horopter) or top forward (inconsistent with the horopter).

Participants

16 observers participated in the study. They were 20-61 years of age, and all had better than 20/32 visual acuity as measured by the Bailey Lovey chart, and stereothresholds of less than 30 arcsec on the Randot stereopsis test. All observers could read the content presented in the HMD clearly, and accurately.

Procedure

The experiment involved one session on a single day. Participants were shown black text on a white page and asked to read out loud. The text was taken from *Harry Potter and the Sorcerers Stone* [94]. Each trial had two intervals, in between the presentation of each interval was a 1sec interstimulus interval. There were two different types of trials: Slant trials (2/3 of the trials) and Magnitude trials (1/3 of trials). For the Slant trials, in one interval, the planar page was slanted top back (consistent with the horopter) and in the other interval the page was slanted top forward (inconsistent with the horopter). The order of which interval contained a consistent or inconsistent page was randomized to avoid order effects. The magnitude of the slant was the same in both intervals. Following both pages, the participant was asked to respond which page was more comfortable, the first (numerical keypad 1) or the second (numerical keypad 2). We additionally assessed performance by measuring how long it took for each observer to read the page in each condition. For the Magnitude trials, observers were shown two intervals where the sign of the slant was the same (both intervals contained a top back or top forward slant) although the magnitudes of the slants were

different. These Magnitude trials acted as catch trials and reveal any biases in observers' response, or strategies (for example only responding '1'). If observers performed the task correctly, we anticipate a preference should emerge for the interval containing the smallest slant. The smallest slant tested (20°) is similar to the top back pitch of the vertical horopter (16.6° at a distance of 66cm) [22, 79]. All other slants tested deviate quite dramatically (especially in the case of the 50° slant) from the top back pitch of the vertical horopter.

There were two experimental conditions: *consistent* and *inconsistent*. In the *consistent* condition, the textured plane was slanted top back, which is consistent with the vertical horopter [22, 79]. In the *inconsistent* condition the plane was slanted top forward, which is inconsistent with the horopter. The distance to the center of the textured plane was 66cm, the optical distance of the screen for the headset used. This ensured the vergence-accommodation conflict was minimized, and the greatest comfort for the observer [51]. The slants of the plane were 20, 30, 40, and 50° at the distance off 66cm: top back in the *consistent* condition and top forward in the *inconsistent* condition. Fig. 4.1 provides examples of *consistent* and *inconsistent* stimuli for Slant trials, as well as an example of a Magnitude trial. There were 30 trials in total, each trial lasted around 1min, depending on the observers' reading speed.

4.3 Results

Figure 4.2A shows the percent of trials the *consistent* interval was preferred. Higher values indicate greater preference for the *consistent* interval, while the dashed lined indicates no difference in preference between *consistent* and *inconsistent* content. Observers preferred the *consistent* page more often than the *inconsistent* page (one sided t-test against a population with a central tendency of 50: $t(15) = 3.06$, $p = 0.004$). This means that in virtual reality, observers prefer scene content that is congruous with the statistics of the natural environment, as compared to content that is inconsistent with environmental regularities. Figure 4.2B shows the percent of trials the *consistent* interval was preferred versus slant magnitude. A one-way repeated measures ANOVA over page slant showed no effect of slant on the preference for consistent scene content ($F(3) = 1.26$, $p = 0.29$). This means that regardless of the magnitude of the slant, observers consistently preferred scene content congruous with the natural environment.

Figure 4.3A shows the time to read a page on average in the *consistent* and *inconsistent* sessions. The central tendency is slightly lower in the *consistent* condition (*consistent* 24.6 seconds per page, *inconsistent* 25.1 seconds per page). A paired t-test revealed that observers read significantly faster in the *consistent* intervals as compared to the *inconsistent* intervals ($t(15) = 2.33$, $p = 0.03$). Figure 4.3B shows the average words read per minute for *consistent* and *inconsistent* sessions by the slant tested. A four (page slant) by two (scene consistent content) repeated measures ANOVA revealed a borderline main effect of scene consistent content ($F(1) = 4.33$, $p = .05$), and no effect of slant ($F(3) = 1.38$, $p = 0.26$) or interaction ($F(3) = 1.18$, $p = 0.33$). These results show no effect of slant on reading speed. One might

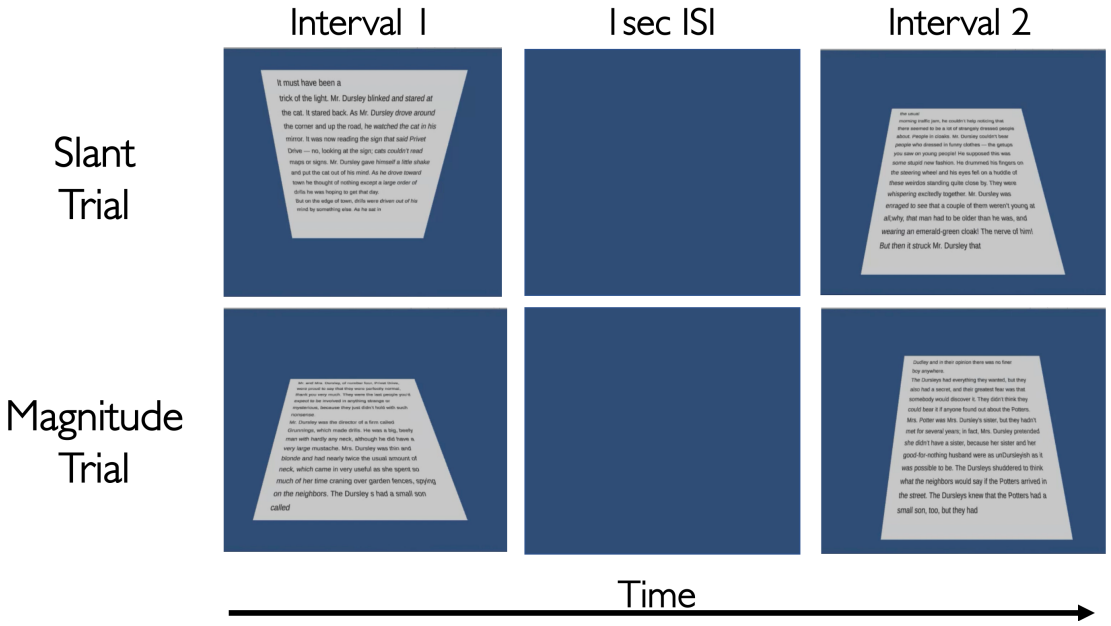

Figure 4.1: Examples of the two different types of trials and the *consistent* and *inconsistent* conditions. The top row shows an example of an Slant trial. Here the distance is 66cm and the slants are -30 (top forward, *inconsistent*) in the first interval and +30° (top back, *consistent*) in the second interval. The bottom row shows an example of a magnitude trial. Here the distance is 66cm and the slants are +50 (top back) in the first interval and +20° (top back) in the second interval.

expect that in the case of the more extreme slants (40° and 50°), reading speed might naturally be slower than for smaller slants. In the experiment, observers were instructed to read the text out loud to ensure compliance with the reading task. Previous work has shown reading out loud takes longer than reading silently [14]. Observers may have hit a ceiling for rate of speech during the task. This may have led to no effect of slant on reading speed, as readers are limited by their rate of speech. If testing involved silent reading, it is possible observers would have been able to read faster, which may reveal an effect of slant magnitude and a more robust difference between reading speed for *consistent* and *inconsistent* scene content.

Figure 4.4A shows the results from the Magnitude trials, which comprised two intervals where the sign of the slant was the same (both intervals contained a top back or top forward slant), but the magnitude of the slant was different. Higher values indicate a greater preference for the interval which contained the smaller slant. On average, observers preferred the smaller slant and this difference was significant (one sided t-test against a population with a central tendency of 50: $t(15) = 5.48$, $p < .0001$). These results confirm observers were performing the task correctly, and weren't adopting biased strategies (such as only press-

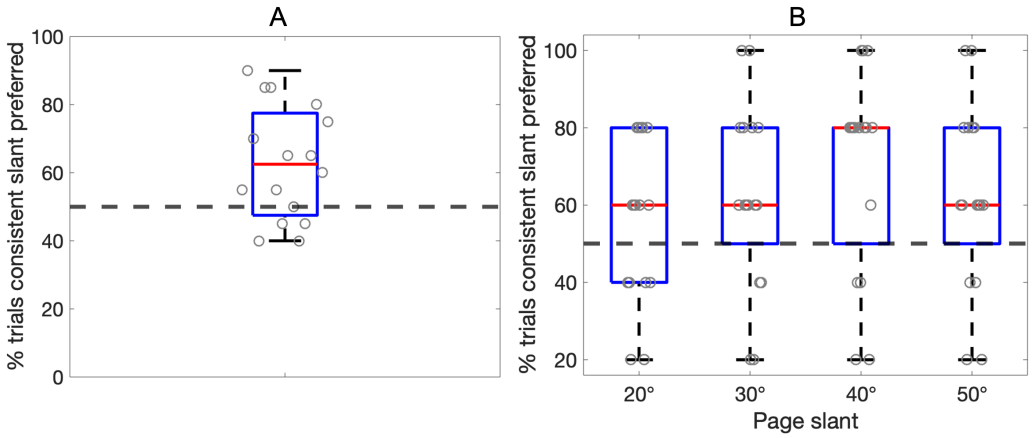

Figure 4.2: A) The percent of trials the scene consistent page slant was preferred. B) The percent of trials the scene consistent page slant was preferred versus page slant. For both plots, the median is shown as the bar inside of each box. The bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers. Outliers are defined as values that are more than 1.5 times the interquartile range away from the bottom or top of the box. The dashed line indicates no difference in preference between consistent and inconsistent content.

ing one of the response keys). Figure 4.4B shows the percent of trials a smaller slant was preferred over the difference in slants between intervals (Δ slant). Δ slant was computed as the absolute value of the difference between the slants presented in the first and second interval. A one-way ANOVA showed no main effect of Δ slant ($F(2) = 2.49$,$p=0.10$). These results show that the interval containing the smaller slant is consistently preferred, regardless of the magnitude of the difference between the slants presented. These findings align with environmental regularities, as the smallest slant tested (20°) is similar to the top back pitch of the vertical horopter (16.6° at a distance of 66cm), where comfort is maximized [22, 113].

4.4 Discussion

In this experiment we investigated whether scene consistent content that is congruous with the statistics of the natural environment leads to greater viewer comfort, and better performance in VR headsets. We isolated the variable of interest - the pitch of the visual content - while keeping all other aspects of the visual scene and procedure the same between both intervals of a trial. To our knowledge this is the first time this hypothesis has been directly tested in the virtual reality environment. Our results show that in virtual reality, observers prefer scene content that is consistent with regularities from the natural environment. We additionally found a small but consistent effect of scene consistent content on reading speed.

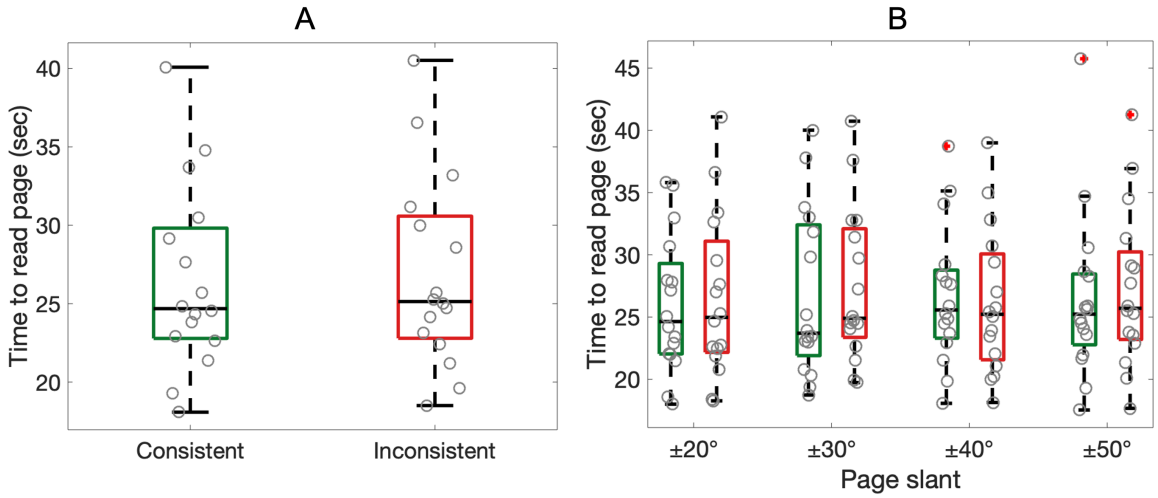

Figure 4.3: A. Average time to read a page split by condition (*consistent* and *inconsistent*). B. The average time to read a page split by condition (*consistent* and *inconsistent*) versus page slant. The median is shown as the black bar inside of each box. The color of the box indicates condition (green *consistent*, red *inconsistent*). The bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers.Outliers are defined as values that are more than 1.5 times the interquartile range away from the bottom or top of the box.

Observers read faster on average in scenes that were congruent with regularities in the natural environment.

There are a variety of negative consequences for presenting visual content in VR headsets that does not conform to environmental regularities in the natural world. Binocular fusion is determined by the 3D location of an object relative to the horopter and Panum's fusion area. The top back pitch of the horopter means that surfaces that are also slanted top-back are more likely to create a fused impression than surfaces that are pitched top-forward. A compelling example of this is the *Venetian-blind effect* [86, 121]. (A demonstration is provided in Fig. 3.16). Additionally, ergonomics researchers advise computer users to pitch desktop displays slightly top back to minimize viewing discomfort [6, 41]. The top-back pitch is consistent with the pitch of the horopter and with natural disparity statistics (Sec. 3.3). Environments that do not conform to the horopter produce more discomfort.

Oculomotor behavior is additionally consistent with natural statistics. When people make upward saccades, they tend to diverge the eyes. This is the same but to a lesser degree for leftward and rightward saccades. When people make downward saccades, they tend to converge [37, 30, 21]. These biases are useful because they ensure that the eyes at the end of a saccade are most likely to be aligned with the new fixation target. Because the statistics in the *inconsistent* condition are not congruent with those in the natural environment, the re-

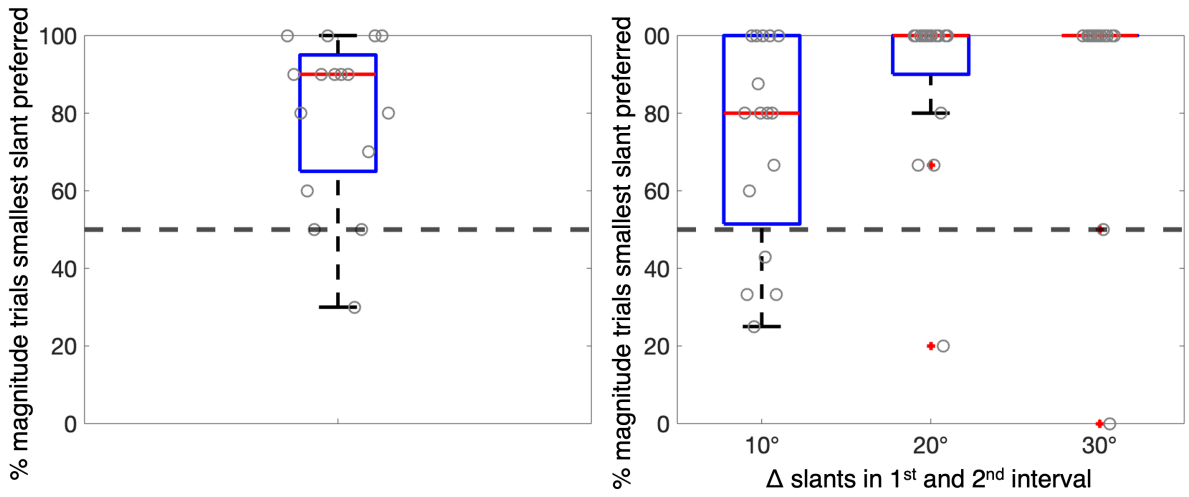

Figure 4.4: A) The percent of trials the page with the smaller slant was preferred for the magnitude trials. B) The percent of trials the page with the smaller slant was preferred versus Δ slant. For both plots, the median is shown as the bar inside of each box. The bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually as a cross. The dashed line indicates no difference in preference between consistent and inconsistent content. Note the unique distribution of responses to each question may alter the appearance of the box plot. For example, in the Δ 30° condition, the data are tightly clustered. There is no way to visualize the 25th and 75th percentiles in this case, and only the median is shown.

lationship between saccades in different directions and the appropriate vergence is disrupted and should cause delays in the acquisition of new targets in the VR environment.

Although observers showed a preference for visual content consistent with environmental regularities, we only found a small effect for reading speed between *consistent* and *inconsistent* scene content. Although observers read faster with *consistent*, content, it is possible the binocular visual system is able to adapt to the incongruous content, and high performance in the reading task we tested is mostly preserved. All participants tested were students or academics, who spend much of their time reading. Expertise with reading may have led to adaptations in virtual reality which facilitated similar performance across conditions. Additionally, as observers were required to read out loud, rate of speech may have led to a ceiling effect on reading speed. Future work may consider testing observers as they read silently to themselves. Alternatively, Taking expertise into account, It may be useful to test observers with an unfamiliar task that forces eye movements to random parts of the visual scene. This may reveal an increase in vergence errors and performance deficits in behavioral tasks. Future directions may include eye tracking during the task, and a more vigorous assessment

of vergence errors when comparing performance in visual environments with congruous or incongruous content.

These findings suggest practical guidelines for the engineers and developers designing VR headsets and applications. In most cases, their goal is to create visual content that is comfortable and non-fatiguing and which produces precepts that are true to the natural environment. Our results suggest that the developers of visual content for virtual reality should review the depth information across the visible field, and to consider where users are likely to fixate. An ideal visual scene, that is consistent with natural scene statistics, will feature visual information that is farther away from the observer above and near information below in the visual field. The top-back pitch of the horopter (and ideally, visual content in the scene) doesn't require a large difference in distance between near and far content. The pitch of the horopter at 66cm (the optical distance of the screen in the Vive Pro Eye HMD, and an appropriate estimate for depth in a work-space environment) is 16.6°. Such modifications to visual content in virtual reality may be subtle, but can have dramatic consequences to improve viewer comfort and performance.

Chapter 5

Conclusion and Future Directions

The advances in technology that allow us to track gaze across multiple environments, from the laboratory to the natural world as well as virtual reality, have far reaching implications. The work presented in this dissertation highlights how this technology can be used for diagnosing and monitoring the progression of visual disorders and how gaze tracking can inform the development of new technologies.

The second chapter in this dissertation presents a pilot study showing that fixational stability tracks changes in visual acuity in patients currently undergoing treatment for amblyopia. Reduced visual acuity is the 'sine qua non' of amblyopia, and the degree to which visual acuity is reduced in the amblyopic eye directly relates to the depth and severity of a patient's amblyopia. This work shows changes in visual acuity with treatment in amblyopia are accompanied by changes to fixational stability. This suggests fixational stability can be used as an objective measure for monitoring treatment in amblyopia. As early diagnosis in children is associated with the best recovery outcomes, a nonverbal and objective measure of the deficits associated with amblyopia are particularly useful. New interventions for amblyopia that utilize virtual reality technology are becoming more prevalent [7, 40, 140]. Simultaneously, eye tracking technology is becoming cheaper to implement, and is included in some commercially available head mounted displays [127]. As virtual reality interventions for amblyopia become more widespread, it may be useful to additionally monitor fixational stability. This is a quick and easy way to collect an objective measure that tracks treatment efficacy, and these measurements can be collected and monitored remotely.

The third chapter in this dissertation presents two studies that quantified and compared the statistics of eye movements and binocular disparity in both natural and VR head mounted display environments. The human visual system has evolved in the natural environment which has consistently contained environmental regularities. Binocular vision has adapted to these regularities such that depth perception and binocular eye movements are more precise, faster, and performed with greater comfort in environments that are consistent with these regularities. We measured the statistics of fixation and binocular disparities in VR-gaming environments and found that they are quite different from those in the natural environment. Fixation direction and distance are more restricted in VR. In addition, fixation distance is

farther in VR. The pattern of binocular disparity across the visual field is less regular in VR and does not conform to properties of naturally occurring disparities. Specifically, the vertical horopter and natural disparity statistics show a top back pitch in the natural environment. This top-back pattern is lacking in the disparity statistics of the virtual environment, and the differences are quite dramatic (in the lower visual field the natural environment has disparities that are nearer than in VR, an effect of around 900 arcsec). The variability of disparity is greater in VR, and the pattern increases the likelihood of experiencing double vision in VR-gaming environments. Additionally, we found observers tend to fixate twice as far away as the optical distance of the screen in VR, which leads to a significant vergence accommodation conflict. We determined from our fixation statistics the optimal screen distance to minimize discomfort due to the vergence-accommodation conflict. These findings inform specific improvements to VR headset design that can improve user comfort.

The fourth chapter in this dissertation follows up on the finding from chapter 3, that the pattern of binocular disparity across the visual field is less regular in VR and violates regularities seen in the disparity distribution of the natural environment. Incompatibility with the statistics of the natural environment may cause viewer discomfort and reduced visual performance in the VR environment. We designed a user experiment to test whether scene consistent content that is congruous with the statistics of the natural environment leads to greater viewer comfort and better performance in VR headsets. We found observers prefer scene content that is congruous with the statistics of the natural environment. We also found observers perform faster on a reading task in environments that are congruous with the statistics of the natural environment. This work suggests guidelines for VR developers to produce content that is faithful to the statistics of the natural environment.

Virtual reality represents a new and exciting technology. VR has far-reaching application from medicine, and business, to the classroom. There are many applications across disciplines for VR technology. However, the findings that there is a mismatch between the natural and VR environment raise concerns for the widespread adoption of this technology. The consequences of spending significant time in environments that violate the statistical regularities of the natural world is unclear. Previous work has documented the discomfort associated with the vergence accommodation conflict in VR environments, and work in rats has shown neurons associated with spatial learning show reduced activity in VR as compared to the natural world [51, 2]. This dissertation has shown that environments which violate the statistics of the natural world lead to greater discomfort and worse performance in VR. In the future, it would be useful to document the ways VR differs from the natural environment, whether these differences lead to discomfort, and how these different sources of discomfort may interact with each other. These assessments help inform ways to improve VR technology and user experience. Ultimately, more research is needed to understand the effects of VR environments on the visual system. This is particularly true when we consider children, and the effects of VR on the developing visual system. There is concern that exposure could harm and damage the developing visual system [97]. As neural plasticity is highest in childhood, it is unethical to subject children to potential changes in their visual system that might be long lasting. Therefore VR is not approved for young children (under

13 years of age) as effects on development are unknown.

Interventions for amblyopia that rely on virtual reality technology are currently being developed and tested in both industry and the private sector [7, 40, 140]. The findings that fixational stability tracks visual function in amblyopia suggest that it may be useful for these interventions to additionally monitor fixational stability as an objective measure of treatment efficacy. These virtual reality interventions promise an engaging VR experience to facilitate the recovery of visual function, which would (ideally) transfer to the natural environment. Although these interventions appear a great improvement on patching treatment, providing clinical improvements with novel and engaging technology, considerations should be made as to the differences between the VR training environment and transfer to the natural world. The mismatch between vergence and accommodation, as well as the disruption of environmental regularities in virtual reality raise concerns. These disruptions, between the natural environment that the visual system has adapted to, and VR, suggest that depth perception and binocular eye movements are less precise, slower, and less comfortable in VR environments [51, 22]. Future interventions for amblyopia which rely on VR and hope to rehabilitate binocular function, should consider the safety and feasibility of training in virtual environments which violate the statistics of the natural world. These interventions should additionally consider using content that is faithful to the statistics of the natural environment for optimal transfer of visual function to the natural world.

Bibliography

[1] Isayas B Adhanom et al. "GazeMetrics: An open-source tool for measuring the data quality of HMD-based eye trackers". In: *ACM Symposium on Eye Tracking Research & Applications*. 2020, pp. 1–5.

[2] Zahra M Aghajan et al. "Impaired spatial selectivity and intact phase precession in two-dimensional virtual reality". In: *Nature neuroscience* 18.1 (2015), pp. 121–128.

[3] Kurt Akeley et al. "A stereo display prototype with multiple focal distances". In: *ACM Transactions on Graphics (TOG)* 23.3 (2004), pp. 804–813.

[4] Rachel Albert et al. "Latency requirements for foveated rendering in virtual reality". In: *ACM Transactions on Applied Perception (TAP)* 14.4 (2017), pp. 1–13.

[5] Adelbert Ames, Kenneth N Ogle, and Gordon H Gliddon. "Corresponding retinal points, the horopter and size and shape of ocular images". In: *JOSA* 22.11 (1932), pp. 575–631.

[6] Dennis R Ankrum, EE Hansen, and Kristie J. Nemeth. "The vertical horopter and the angle of view". In: *Work with Display Units* 94 (1995), pp. 655–665.

[7] Benjamin T Backus et al. "Use of virtual reality to assess and treat weakness in human stereoscopic vision". In: *Electronic Imaging* 2018.4 (2018), pp. 109–1.

[8] Ian L Bailey and Jan E Lovie. "New design principles for visual acuity letter charts." In: *American journal of optometry and physiological optics* 53.11 (1976), pp. 740–745.

[9] Graham R Barnes. "Vestibulo-ocular function during co-ordinated head and eye movements to acquire visual targets." In: *The Journal of Physiology* 287.1 (1979), pp. 127–147.

[10] Daphne Bavelier et al. "Removing brakes on adult brain plasticity: from molecular to behavioral interventions". In: *Journal of Neuroscience* 30.45 (2010), pp. 14964–14971.

[11] Colin Blakemore. "The range and scope of binocular depth discrimination in man". In: *The Journal of Physiology* 211.3 (1970), pp. 599–622.

[12] Jean-Yves Bouguet. "Camera calibration toolbox for matlab". In: *http://www.vision.caltech.edu/bou html* (2004).

[13] David Brewster. *On the knowledge of distance given by binocular vision*. Neill, 1844.

[14] Marc Brysbaert. "How many words do we read per minute? A review and meta-analysis of reading rate". In: *Journal of Memory and Language* 109 (2019), p. 104047.

[15] Andrea Canessa et al. "A dataset of stereoscopic images and ground-truth disparity mimicking human fixations in peripersonal space". In: *Scientific Data* 4.1 (2017), pp. 1–16.

[16] Eric Castet and Michael Crossland. "Quantifying eye stability during a fixation task: a review of definitions and methods". In: *Seeing and Perceiving* 25.5 (2012), pp. 449–469.

[17] Claudia Cherici et al. "Precision of sustained fixation in trained and untrained observers". In: *Journal of vision* 12.6 (2012), pp. 31–31.

[18] Susana TL Chung and Harold E Bedell. "Effect of retinal image motion on visual acuity and contour interaction in congenital nystagmus". In: *Vision Research* 35.21 (1995), pp. 3071–3082.

[19] Susana TL Chung et al. "Characteristics of fixational eye movements in amblyopia: Limitations on fixation stability and acuity?" In: *Vision research* 114 (2015), pp. 87–99.

[20] KJ Ciuffreda, RV Kenyon, and L Stark. "Fixational eye movements in amblyopia and strabismus." In: *Journal of the American Optometric Association* 50.11 (1979), pp. 1251–1258.

[21] Han Collewijn, Casper J Erkelens, and Robert M Steinman. "Binocular co-ordination of human vertical saccadic eye movements." In: *The Journal of Physiology* 404.1 (1988), pp. 183–197.

[22] Emily A Cooper, Johannes Burge, and Martin S Banks. "The vertical horopter is not adaptable, but it may be adaptive". In: *Journal of Vision* 11.3 (2011), pp. 20–20.

[23] Bruce G Cumming. "An unexpected specialization for horizontal disparity in primate primary visual cortex". In: *Nature* 418.6898 (2002), pp. 633–636.

[24] Bruce G Cumming and Stuart J Judge. "Disparity-induced and blur-induced convergence eye movement and accommodation in the monkey". In: *Journal of Neurophysiology* 55.5 (1986), pp. 896–914.

[25] Robert William Ditchburn. "Eye-movements and visual perception." In: (1973).

[26] Neil A Dodgson. "Variation and extrema of human interpupillary distance". In: *Stereoscopic Displays & Virtual Reality Systems XI*. Vol. 5291. International Society for Optics and Photonics. 2004, pp. 36–46.

[27] Alexander Duane. "Normal values of the accommodation at all ages". In: *Journal of the American Medical Association* 59.12 (1912), pp. 1010–1013.

[28] Jean-Baptiste Durand, Simona Celebrini, and Yves Trotter. "Neural bases of stereopsis across visual field of the alert macaque monkey". In: *Cerebral Cortex* 17.6 (2007), pp. 1260–1273.

[29] Vasha DuTell et al. "Integrating High Fidelity Eye, Head and World Tracking in a Wearable Device". In: *ACM Symposium on Eye Tracking Research and Applications*. 2021, pp. 1–4.

[30] James T Enright. "Changes in vergence mediated by saccades." In: *The Journal of Physiology* 350.1 (1984), pp. 9–31.

[31] David J Field. "Relations between the statistics of natural images and the response properties of cortical cells". In: *Josa a* 4.12 (1987), pp. 2379–2394.

[32] Edgar F Fincham and John Walton. "The reciprocal actions of accommodation and convergence". In: *The Journal of Physiology* 137.3 (1957), p. 488.

[33] Franz Peter Fischer. "III. Experimentelle Beiträge zum Begriff der Sehrichtungsgemeinschaft der Netzhaute auf Grund der binokularen Noniusmethode". In: *Fortgesetzte Studien über Binokularsehen. Pflugers Archiv für die Gesamte Physiologie des Menschen und der Tiere* 204 (1924), pp. 234–246.

[34] Merton C Flom, DG Kirschen, and Harold E Bedell. "Control of unsteady, eccentric fixation in amblyopic eyes by auditory feedback of eye position." In: *Investigative ophthalmology & visual science* 19.11 (1980), pp. 1371–1381.

[35] Christina Gambacorta et al. "Both saccadic and manual responses in the amblyopic eye of strabismics are irreducibly delayed". In: *Journal of vision* 18.3 (2018), pp. 20–20.

[36] Wilson S Geisler. "Visual perception and the statistical properties of natural scenes". In: *Annu. Rev. Psychol.* 59 (2008), pp. 167–192.

[37] Agostino Gibaldi and Martin S Banks. "Binocular eye movements are adapted to the natural environment". In: *Journal of Neuroscience* 39.15 (2019), pp. 2877–2888.

[38] Agostino Gibaldi, Andrea Canessa, and Silvio P Sabatini. "The active side of stereopsis: Fixation strategy and adaptation to natural environments". In: *Scientific Reports* 7 (2017), p. 44800.

[39] Agostino Gibaldi, Vasha DuTell, and Martin S Banks. "Solving Parallax Error for 3D Eye Tracking". In: *ACM Symposium on Eye Tracking Research and Applications*. 2021, pp. 1–4.

[40] Angelica Godinez et al. "Scaffolding depth cues and perceptual learning in VR to train stereovision: a proof of concept pilot study". In: *Scientific Reports* 11.1 (2021), pp. 1–16.

[41] Etienne Grandjean, W Hünting, and M Pidermann. "VDT workstation design: Preferred settings and their effects". In: *Human Factors* 25.2 (1983), pp. 161–175.

[42] Philip M Grove, Hirohiko Kaneko, and Hiroshi Ono. "The backward inclination of a surface defined by empirical corresponding points". In: *Perception* 30.4 (2001), pp. 411–429.

[43] Brian Guenter et al. "Foveated 3D graphics". In: *ACM Transactions on Graphics (TOG)* 31.6 (2012), pp. 1–10.

[44] Daniel Guitton and Michel Volle. "Gaze control in humans: Eye-head coordination during orienting movements to targets within and beyond the oculomotor range". In: *Journal of Neurophysiology* 58.3 (1987), pp. 427–459.

[45] Jukka Häkkinen et al. "Simulator sickness in virtual display gaming: A comparison of stereoscopic and non-stereoscopic situations". In: *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices & Services*. 2006, pp. 227–230.

[46] David R Hampton and Andrew E Kertesz. "The extent of Panum's area and the human cortical magnification factor". In: *Perception* 12.2 (1983), pp. 161–165.

[47] Hermann von Helmholtz. *Treatise on physiological optics*. Vol. 3. Courier Corporation, 2013.

[48] RF Hess. "Eye movements and grating acuity in strabismic amblyopia". In: *Ophthalmic Research* 9.4 (1977), pp. 225–237.

[49] RF Hess. "On the relationship between strabismic amblyopia and eccentric fixation." In: *British Journal of Ophthalmology* 61.12 (1977), pp. 767–773.

[50] Heiko Hirschmüller. "Stereo processing by semiglobal matching and mutual information". In: *IEEE Transactions on Pattern Analysis & Machine Intelligence* 30.2 (2007), pp. 328–341.

[51] David M Hoffman et al. "Vergence–accommodation conflicts hinder visual performance and cause visual fatigue". In: *Journal of Vision* 8.3 (2008), pp. 33–33.

[52] Janis Intoy and Michele Rucci. "Finely tuned eye movements enhance visual acuity". In: *Nature communications* 11.1 (2020), pp. 1–11.

[53] David Kane, Robert T Held, and Martin S Banks. "Visual discomfort with stereo 3D displays when the head is not upright". In: *Stereoscopic Displays & Applications XXIII*. Vol. 8288. International Society for Optics and Photonics. 2012, p. 828814.

[54] Tobit Kollenberg et al. "Visual search in the (un) real world: How head-mounted displays affect eye movements, head movements and target detection". In: *Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications*. 2010, pp. 121–124.

[55] Frank L Kooi and Alexander Toet. "Visual comfort of binocular and 3D displays". In: *Displays* 25.2-3 (2004), pp. 99–108.

[56] Rakshit Kothari et al. "Gaze-in-wild: A dataset for studying eye and head coordination in everyday activities". In: *Scientific Reports* 10.1 (2020), pp. 1–18.

[57] George Alex Koulieris et al. "Accommodation and comfort in head-mounted displays". In: *ACM Transactions on Graphics (TOG)* 36.4 (2017), pp. 1–11.

[58] George Alex Koulieris et al. "Near-eye display and tracking technologies for virtual and augmented reality". In: *Computer Graphics Forum*. Vol. 38. 2. Wiley Online Library. 2019, pp. 493–519.

[59] Gregory Kramida. "Resolving the vergence-accommodation conflict in head-mounted displays". In: *IEEE Transactions on Visualization & Computer Graphics* 22.7 (2015), pp. 1912–1931.

[60] Girish Kumar and Susana TL Chung. "Characteristics of fixational eye movements in people with macular disease". In: *Investigative ophthalmology & visual science* 55.8 (2014), pp. 5125–5133.

[61] Vivek Labhishetty et al. "Lags and leads of accommodation: Fact or fiction?" In: *Journal of Vision* 8.3 (2020), pp. 33–33.

[62] Marc Lambooij et al. "Visual discomfort and visual fatigue of stereoscopic displays: A review". In: *Journal of Imaging Science & Technology* 53.3 (2009), pp. 30201–1.

[63] Michael F Land and Mary Hayhoe. "In what ways do eye movements contribute to everyday activities?" In: *Vision Research* 41.25-26 (2001), pp. 3559–3565.

[64] Michael Land, Neil Mennie, and Jennifer Rusted. "The roles of vision and eye movements in the control of activities of daily living". In: *Perception* 28.11 (1999), pp. 1311–1328.

[65] Dennis M Levi. "Prentice award lecture 2011: removing the brakes on plasticity in the amblyopic brain". In: *Optometry and vision science* 89.6 (2012), p. 827.

[66] Dennis M Levi, David C Knill, and Daphne Bavelier. "Stereopsis and amblyopia: A mini-review". In: *Vision research* 114 (2015), pp. 17–30.

[67] Otto Alexander Maneschg et al. "Fixation stability after surgical treatment of strabismus and biofeedback fixation training in amblyopic eyes". In: *BMC ophthalmology* 21.1 (2021), pp. 1–9.

[68] Susana Martinez-Conde, Stephen L Macknik, and David H Hubel. "The role of fixational eye movements in visual perception". In: *Nature reviews neuroscience* 5.3 (2004), pp. 229–240.

[69] Susana Martinez-Conde, Jorge Otero-Millan, and Stephen L Macknik. "The impact of microsaccades on vision: towards a unified theory of saccadic function". In: *Nature Reviews Neuroscience* 14.2 (2013), pp. 83–96.

[70] Susana Martinez-Conde et al. "Microsaccades counteract visual fading during fixation". In: *Neuron* 49.2 (2006), pp. 297–305.

[71] Jonathan S Matthis, Jacob L Yates, and Mary M Hayhoe. "Gaze and the control of foot placement when walking in natural terrain". In: *Current Biology* 28.8 (2018), pp. 1224–1233.

[72] Michael Mauderer et al. "Depth perception with gaze-contingent depth of field". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2014, pp. 217–226.

[73] LAWRENCE E Mays. "Neural control of vergence eye movements: convergence and divergence neurons in midbrain". In: *Journal of Neurophysiology* 51.5 (1984), pp. 1091–1108.

[74] Roberta McKean-Cowdin et al. "Prevalence of amblyopia or strabismus in asian and non-Hispanic white preschool children: multi-ethnic pediatric eye disease study". In: *Ophthalmology* 120.10 (2013), pp. 2117–2124.

[75] Manbir Nagra et al. "The effect of the COVID-19 pandemic on working practices of UK primary care optometrists". In: *Ophthalmic and Physiological Optics* 41.2 (2021), pp. 378–392.

[76] Ken Nakayama. "Geometric and physiological aspects of depth perception". In: *Three-dimensional imaging*. Vol. 120. International Society for Optics and Photonics. 1977, pp. 2–9.

[77] Gunter K von Noorden. "New clinical aspects of stimulus deprivation amblyopia". In: *American journal of ophthalmology* 92.3 (1981), pp. 416–421.

[78] Kenneth N Ogle. "Researches in binocular vision." In: (1950).

[79] Kennneth N Ogle and Vincent J Ellerbrock. "Cyclofusional movements". In: *Archives of Ophthalmology* 36.6 (1946), pp. 700–735.

[80] OpenCV. *Open Source Computer Vision Library*. 2015.

[81] Aftab E Patla and Joan N Vickers. "How far ahead do we look when required to step on specific locations in the travel path during locomotion?" In: *Experimental brain research* 148.1 (2003), pp. 133–138.

[82] Anjul Patney et al. "Towards foveated rendering for gaze-tracked virtual reality". In: *ACM Transactions on Graphics (TOG)* 35.6 (2016), p. 179.

[83] Adrian Penate-Sanchez, Juan Andrade-Cetto, and Francesc Moreno-Noguer. "Exhaustive linearization for robust camera pose and focal length estimation". In: *IEEE Transactions on Pattern Analysis & Machine Intelligence* 35.10 (2013), pp. 2387–2400.

[84] Kevin Pfeil et al. "A comparison of eye-head coordination between virtual and physical realities". In: *Proceedings of the 15th ACM Symposium on Applied Perception*. 2018, pp. 1–7.

[85] Amelie Pham, Marisa Carrasco, and Lynne Kiorpes. "Endogenous attention improves perception in amblyopic macaques". In: *Journal of vision* 18.3 (2018), pp. 11–11.

[86] David Piggins. "Moirés maintained internally by binocular vision". In: *Perception* 7.6 (1978), pp. 679–681.

[87] Ariella V Popple and Dennis M Levi. "The attentional blink in amblyopia". In: *Journal of Vision* 8.13 (2008), pp. 12–12.

[88] Simon JD Prince and Richard A Eagle. "Stereo correspondence in one-dimensional Gabor stimuli". In: *Vision Research* 40.8 (2000), pp. 913–924.

[89] Holger A Rambold and Frederick A Miles. "Human vergence eye movements to oblique disparity stimuli: Evidence for an anisotropy favoring horizontal disparities". In: *Vision Research* 48.19 (2008), pp. 2006–2019.

[90] C Rashbass and Gv Westheimer. "Disjunctive eye movements". In: *The Journal of Physiology* 159.2 (1961), pp. 339–360.

[91] Rajkumar Nallour Raveendran, William Bobier, and Benjamin Thompson. "Reduced amblyopic eye fixation stability cannot be simulated using retinal-defocus-induced reductions in visual acuity". In: *Vision research* 154 (2019), pp. 14–20.

[92] LA Riggs and F Ratliff. "The effects of counteracting the normal movements of the eye". In: *Journal of the Optical Society of America*. Vol. 42. 11. AMER INST PHYSICS CIRCULATION FULFILLMENT DIV, 500 SUNNYSIDE BLVD, WOODBURY ... 1952, pp. 872–873.

[93] Martin Rolfs, Reinhold Kliegl, and Ralf Engbert. "Toward a model of microsaccade generation: The case of microsaccadic inhibition". In: *Journal of vision* 8.11 (2008), pp. 5–5.

[94] JK Rowling. *Harry Potter and the sorcerer's stone (book 1)*. Arthur A. Levine Books, 1997.

[95] Michele Rucci and Martina Poletti. "Control and functions of fixational eye movements". In: *Annual review of vision science* 1 (2015), pp. 499–518.

[96] Michele Rucci and Jonathan D Victor. "The unsteady eye: an information-processing stage, not a bug". In: *Trends in neurosciences* 38.4 (2015), pp. 195–206.

[97] Simon K Rushton and Patricia M Riddell. "Developing visual systems and exposure to virtual reality and stereo displays: some concerns and speculations about the demands on accommodation and vergence". In: *Applied Ergonomics* 30.1 (1999), pp. 69–78.

[98] R Sachsenweger. "Problems of organic lesions in functional amblyopia". In: *International strabismus symposium*. Karger Publishers. 1968, pp. 63–66.

[99] C Schor. "A directional impairment of eye movement control in strabismus amblyopia." In: *Investigative ophthalmology & visual science* 14.9 (1975), pp. 692–697.

[100] C Schor and W Hallmark. "Slow control of eye position in strabismic amblyopia." In: *Investigative ophthalmology & visual science* 17.6 (1978), pp. 577–581.

[101] Clifton M Schor. "A dynamic model of cross-coupling between accommodation and convergence: Simulations of step and frequency responses." In: *Optometry & Vision Science* 69.4 (1992), pp. 258–269.

[102] Clifton M Schor, James S Maxwell, and Scott B Stevenson. "Isovergence surfaces: The conjugacy of vertical eye movements in tertiary positions of gaze". In: *Ophthalmic & Physiological Optics* 14.3 (1994), pp. 279–286.

[103] Clifton M Schor, Ivan Wood, and Jane Ogawa. "Binocular sensory fusion is limited by spatial resolution". In: *Vision Research* 24.7 (1984), pp. 661–665.

[104] Kai M Schreiber et al. "The surface of the empirical horopter". In: *Journal of Vision* 8.3 (2008), pp. 7–7.

[105] Robert A Schumer and Bela Julesz. "Binocular disparity modulation sensitivity to disparities offset from the plane of fixation". In: *Vision Research* 24.6 (1984), pp. 533–542.

[106] Aasef G Shaikh et al. "Abnormal fixational eye movements in amblyopia". In: *PLoS One* 11.3 (2016), e0149953.

[107] Vineeta Sharma, Dennis M Levi, and Stanley A Klein. "Undercounting features and missing features: evidence for a high-level deficit in strabismic amblyopia". In: *Nature neuroscience* 3.5 (2000), pp. 496–501.

[108] Takashi Shibata et al. "The zone of comfort: Predicting visual discomfort with stereo displays". In: *Journal of Vision* 11.8 (2011), pp. 11–11.

[109] Ludwig Sidenmark and Hans Gellersen. "Eye, head and torso coordination during gaze shifts in virtual reality". In: *ACM Transactions on Computer-Human Interaction (TOCHI)* 27.1 (2019), pp. 1–40.

[110] John Siderov, Ronald S Harwerth, and Harold E Bedell. "Stereopsis, cyclovergence and the backwards tilt of the vertical horopter". In: *Vision Research* 39.7 (1999), pp. 1347–1357.

[111] Vincent Sitzmann et al. "Saliency in VR: How do people explore virtual environments?" In: *IEEE Transactions on Visualization & Computer Graphics* 24.4 (2018), pp. 1633–1642.

[112] Rizwan AB Somani et al. "Visual test of Listing's law during vergence". In: *Vision Research* 38.6 (1998), pp. 911–923.

[113] William W Sprague et al. "Stereopsis is adaptive for the natural environment". In: *Science Advances* 1.4 (2015), e1400254.

[114] Steam. *Virtual Reality on Steam (top sellers)*. 2020. URL: `https://store.steampowered.com/vr/#p=0&tab=TopSellers`. (accessed: 10.22.2020).

[115] Robert M Steinman. "Effect of target size, luminance, and color on monocular fixation". In: *JOSA* 55.9 (1965), pp. 1158–1164.

[116] Vidhya Subramanian, Reed M Jost, and Eileen E Birch. "A quantitative study of fixation stability in amblyopia". In: *Investigative ophthalmology & visual science* 54.3 (2013), pp. 1998–2003.

[117] Benjamin W Tatler and Benjamin T Vincent. "Systematic tendencies in scene viewing". In: *Journal of Eye Movement Research* 2 (2008), p. 2263.

[118] George T Timberlake et al. "Retinal location of the preferred retinal locus relative to the fovea in scanning laser ophthalmoscope images". In: *Optometry and vision science* 82.3 (2005), E177–E187.

[119] Tobii. *Downloads*. 2020. URL: `https://vr.tobii.com/sdk/downloads/`. (accessed: 09.21.2020).

[120] Douglas Tweed. "Visual-motor optimization in binocular control". In: *Vision Research* 37.14 (1997), pp. 1939–1951.

[121] Christopher W Tyler. "Binocular moiré fringes and the vertical horopter". In: *Perception* 9.4 (1980), pp. 475–478.

[122] Matthieu Urvoy, Marcus Barkowsky, and Patrick Le Callet. "How visual fatigue and discomfort impact 3D-TV quality of experience: a comprehensive review of technological, psychophysical, and psychological factors". In: *Annals of Telecommunications-Annales des Télécommunications* 68.11-12 (2013), pp. 641–655.

[123] Raymond Van Ee and Clifton M Schor. "Unconstrained stereoscopic matching of lines". In: *Vision Research* 40.2 (2000), pp. 151–162.

[124] LJ Van Run and Albert V Van den Berg. "Binocular eye orientation during fixations: Listing's law extended to include eye vergence". In: *Vision Research* 33.5-6 (1993), pp. 691–708.

[125] Kathleen Vancleef et al. "ASTEROID: a new clinical stereotest on an autostereo 3D tablet". In: *Translational vision science & technology* 8.1 (2019), pp. 25–25.

[126] Preeti Verghese, Suzanne P McKee, and Dennis M Levi. "Attention deficits in amblyopia". In: *Current opinion in psychology* 29 (2019), pp. 199–204.

[127] HTC Vive. *Specs and Details*. 2020. URL: `https://enterprise.vive.com/us/product/vive-pro-eye-office/`. (accessed: 09.21.2020).

[128] HTC Vive. *VIVE Eye Tracking SDK (SRanipal))*. 2020. URL: `https://developer.vive.com/resources/vive-sense/sdk/vive-eye-tracking-sdk-sranipal/`. (accessed: 09.21.2020).

[129] Björn NS Vlaskamp, Phillip Guan, and Martin S Banks. "The venetian-blind effect: A preference for zero disparity or zero slant?" In: *Frontiers in Psychology* 4 (2013), p. 836.

[130] Hermann Von Helmholtz. *Helmholtz's treatise on physiological optics*. Vol. 3. Optical Society of America, 1925.

[131] David K Wallace et al. "Time course and predictors of amblyopia improvement with 2 hours of daily patching". In: *JAMA ophthalmology* 133.5 (2015), pp. 606–609.

[132] Simon J Watt et al. "Focus cues affect perceived depth". In: *Journal of Vision* 5.10 (2005), pp. 7–7.

[133] Stanisław Weglarczyk. "Kernel density estimation and its application". In: *ITM Web of Conferences*. Vol. 23. EDP Sciences. 2018.

[134] Cathy Williams et al. "Prevalence and risk factors for common vision problems in children: data from the ALSPAC study". In: *British Journal of Ophthalmology* 92.7 (2008), pp. 959–964.

[135] Shun-nan Yang et al. "Stereoscopic viewing and reported perceived immersion and symptoms". In: *Optometry & Vision Science* 89.7 (2012), pp. 1068–1080.

[136] Alfred L Yarbus. *Eye movements and vision*. Plenum Press, 1967.

[137] David S Zee, Edmond J Fitzgibbon, and Lance M Optican. "Saccade-vergence interactions in humans". In: *Journal of Neurophysiology* 68.5 (1992), pp. 1624–1641.

[138] Bin Zhang et al. "Effects of fixation instability on multifocal VEP (mfVEP) responses in amblyopes". In: *Journal of vision* 8.3 (2008), pp. 16–16.

[139] Zhengyou Zhang. "A flexible new technique for camera calibration". In: *IEEE Transactions on Pattern Analysis & Machine Intelligence* 22.11 (2000), pp. 1330–1334.

[140] Peter Žiak et al. "Amblyopia treatment of adults with dichoptic training using the virtual reality oculus rift head mounted display: preliminary results". In: *BMC ophthalmology* 17.1 (2017), pp. 1–8.

5.1 Appendix A: Pilot study I

Methods

We designed an experiment in the HMD to test whether having scene content consistent with the statistics of the natural environment matters with respect to viewer comfort and performance. To our knowledge, this is the first time the hypothesis has been tested in virtual environments.

Apparatus

The methods and apparatus were identical to those described in 3.2 except for the material presented by the HMD and the participant's task. Participants were shown black text on a white page and asked to read for comprehension. The planar page was slanted top back (consistent with the horopter) or top forward (inconsistent with the horopter).

Participants

10 people with normal or corrected-to-normal vision participated. The were 21-37 years of age.

Procedure

Participants read 30min of text in two sessions conducted on consecutive days. In the *consistent session* the page was slanted at an angle consistent with the top-back pitch of the vertical horopter [22, 79]. In the *inconsistent session* the page was slanted top forward with the same magnitude as in the cues-consistent session but the opposite sign; the 3D orientation of the page in this session was therefore inconsistent with the vertical horopter and natural-environment statistics. Fig. 5.1 shows examples of a *consistent* and *inconsistent* page. In each session, participants read for 10min each three engaging news articles one page at a time. Participants pressed a button to advance to the next page. The distance of the page was randomized after each button press to be 28.6, 40, or 66cm. By using different distances we made sure that participants had to make vergence eye movements from time to time. The slants at those respective distances were 7.3, 10.1, and 16.6° in the consistent session. They were the opposite in the inconsistent session. The slants in the consistent sessions co-varied with distance to maintain consistency with the horopter whose top-back slant also co-varies with distance [22]. We also presented slants of twice those values in both sessions.

After each session, participants filled out a discomfort questionnaire in which they indicated on a 5-point scale how tired their eyes feel, how clear their vision is, how tired and sore their neck and back are, how their eyes feel, and how their head feels. Participants were instructed to rate each symptom relative to how they felt at the end of the session relative to the beginning. After completing the second session, participants also completed

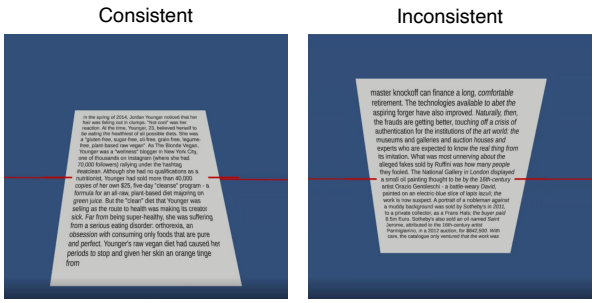

Figure 5.1: Examples of the images in the *consistent* and *inconsistent* sessions. In these examples, slants are 33.2 and -32.2°, respectively. The red tick marks on the pages enabled viewers to align them with the horizon indicated by a red line.

a questionnaire indicating on a 9-point scale which session they found most fatiguing, irritated their eyes the most, felt a worse headache in, and generally preferred least. Following the questionnaires on both days, participants completed a short reading quiz that contained five questions of varying difficulty for each article. This motivated participants to read the articles carefully, and provided a measure of comprehension.

Results

Figure 5.2 shows the median reported symptoms on the symptom questionnaire: The more severe the reported symptom, the higher the black bar, which represents the median. For all questions, the median severity of each symptom was remarkably similar for both *consistent* and *inconsistent* symptoms, and the distributions clearly overlap to such a great extent that we did not pursue significance testing. Figure 5.3 shows the results for the display evaluation questionnaire, where participants were asked to directly compare their discomfort symptoms in the *consistent* and *inconsistent* sessions. Higher positive values are consistent with more favorable ratings for the *consistent* session, while a value of 0 indicates no difference in preference between sessions. The fatigue, eye irritation, and headache questions all have a median of 0, meaning on average there was no difference in preference between *consistent* and *inconsistent* session for these symptoms. The fourth question, directly asking observers to compare the two sessions and to report which session they preferred actually shows a slight preference for the *inconsistent* session, but the broad spread of the responses to this question (that clearly overlaps with 0) suggest very inconsistent responses.

Figure 5.4A shows the results from the comprehension test which was meant to motivate participants to read closely, and to gauge differences in comprehension between the *consistent* and *inconsistent* sessions. Although the *inconsistent* session showed worse comprehension on average than the *consistent* session, a t-test did not reveal a significant difference in comprehension between sessions ($t(9) = .76$, $p = .47$). Figure 5.4B shows the number of

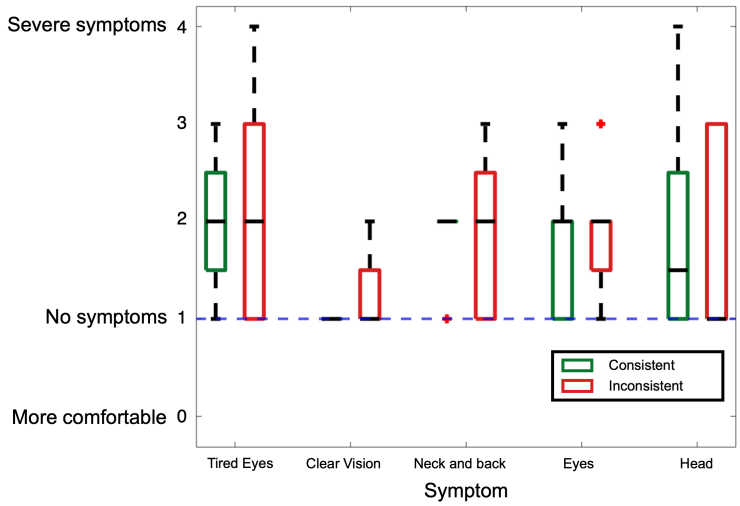

Figure 5.2: Results from the symptom questionnaire shown with a box plot. The median severity of the reported symptom is plotted for each of the five questionnaire items, and is shown by the black bar in the box. The color of the box indicates condition (green consistent, red inconsistent). The bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually as a cross. The blue dotted line indicates a response of 'no symptoms'. Note the unique distribution of responses to each question may alter the appearance of the box plot. For example, when asked about the clearness of their vision in the consistent condition, all observers responded '1'. There is no way to visualize the 25th and 75th percentiles in this case, and only the median is shown.

words per minute read on average in the *consistent* and *inconsistent* sessions. The central tendency and spread of the plots are quite similar in both conditions, showing no effect of scene consistent content on reading speed

Discussion

The findings from this pilot study did not show an effect of scene consistent content on viewer comfort and performance. These findings were initially surprising, as previous work has shown the importance of the horopter for fusion and binocular vision [46, 78]. Surfaces that are slanted top-back (consistent with the vertical horopter) are more likely to create a fused impression than surfaces that are pitched top-forward. Additional work has shown oculomotor behavior to be consistent with natural statistics - when people make upward saccades, they tend to diverge the eyes [37, 30, 21]. Section 3.3 showed that the disparity statistics show a much smaller change over the visual field, a pattern which is inconsistent with the far-above and near-below disparity pattern seen in the natural environment. As the

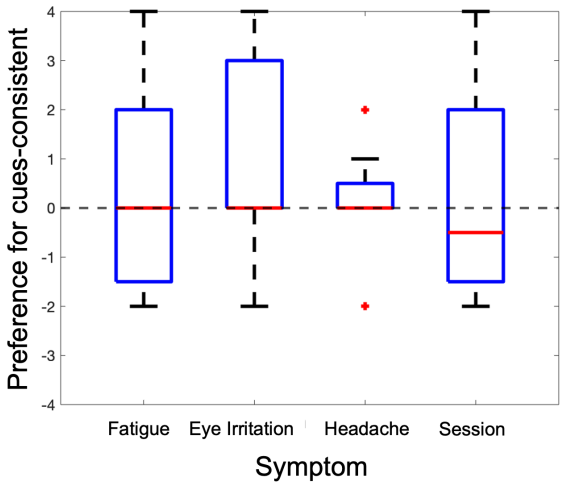

Figure 5.3: Results from the display evaluation questionnaire. The median comparative rating of a pair of sessions (consistent and inconsistent) is plotted for each of the four questionnaire items as the bar inside each of the boxes.The bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually as a cross. The dashed line indicates a response of "no difference"

disparity statistics in the VR-gaming environment are not congruent with those in the natural environment, the relationship between saccades in different directions and the appropriate vergence is disrupted and should cause delays in the acquisition of new targets in the VR environment. We anticipated that these violations to the regularities in the VR-gaming environment would lead to discomfort and reduced perceptual performance. However, this is not what we found. Our results show no effect of scene consistent content on comfort, performance, or speed in a reading task.

Five of the observers that initially participated in this pilot study agreed to be interviewed following the experiment. These participants raised a few concerns which may have impacted our findings. Based on these interviews, we designed a second pilot study. For the duration of the experiment observers were asked to stand. The participants interviewed felt that standing with the headset on led to notable discomfort, specifically pain in the neck and back. Observers felt this discomfort overpowered any of the more subtle discomfort symptoms we were measuring. In order to address these concerns, in the next experiment, observers remained seated while wearing the headset. Observers also felt that their interest between the 6 articles varied, and that they performed better and felt less discomfort while reading articles of interest. In the next experiment we used a visual search task that was equally engaging over *consistent* and *inconsistent* sessions. Finally, observers felt that they were unfamiliar with the HMD and how to set it up properly. As a result, participants felt they

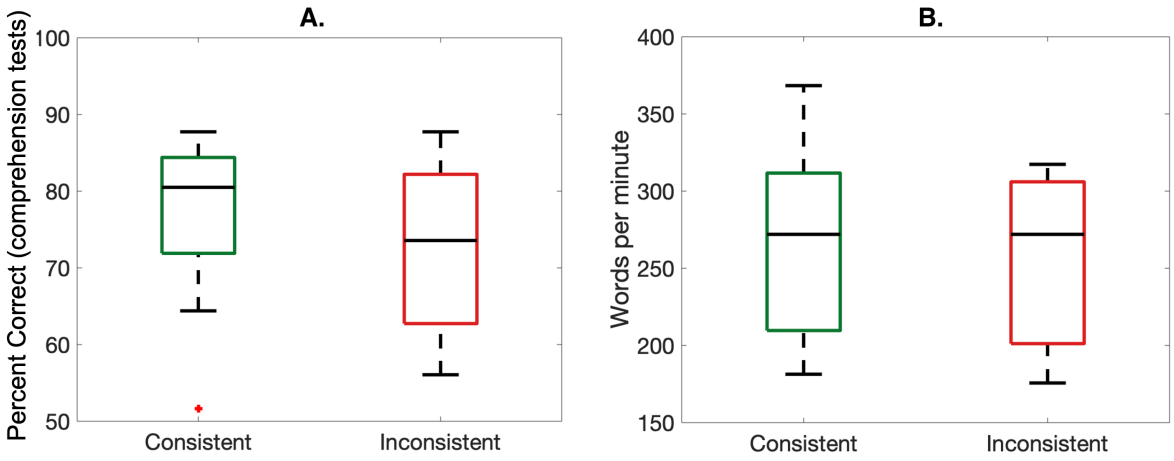

Figure 5.4: A) Results from the comprehension tests shown as the percent of correct responses split by session (consistent and inconsistent). Questions for the comprehension tests were taken from the text all observers read. B) The average number of words read per minute split by session (consistent and inconsistent). For both plots, the median is shown as the bar inside of each box. The color of the box indicates condition (green consistent, red inconsistent).The bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually as a cross.

did not know how to properly put on the headset on the first day of the study, which led to blurry images with more optical aberrations. This led to greater discomfort regardless of the viewing condition tested on the first day. By the second day of the experiment observers felt they had sufficient expertise with the headset to properly adjust the HMD to achieve the clearest image possible. With our counterbalancing procedure, if observers all experienced a less comfortable viewing experience on the first day, this would average out and eliminate any differences between the tested conditions. In the following pilot study, there was an extra practice session to ensure participant familiarity with the headset.

5.2 Appendix B. Pilot Study II

Methods

Based on the results from our first pilot study, we designed an experiment in the HMD to test whether having scene content consistent with the statistics of the natural environment affects viewer comfort and performance. We changed the overall paradigm used to better control for the difficulty of the task observers performed in the headset as well as to ensure observers were familiar with how to optimally set up the head sets.

Apparatus

The HMD and controllers were the same as in the previous experiment.

Participants

10 people with normal or corrected-to-normal vision participated. They were 22-32 years of age.

Procedure

The experiment involved three sessions conducted on three different days. The first session involved training in which participants familiarized themselves with the HMD and how to position it in order to see the whole field of view clearly. The second and third sessions were experimental sessions in which participants performed a number-search task for three blocks of trials each 10min long. The stimuli were texture planes on which black numbers (1–10) were positioned randomly. There was a circular fixation target below each number. Participants were told to fixate the targets in sequence from the smallest to the largest number. When they successfully fixated the correct target, it turned red and disappeared. Participants then searched for the next number. After fixating the last number in a given stimulus, participants pressed a button to advance to the next stimulus.

There were two experimental conditions: *consistent* and *inconsistent*. In the *consistent* condition, the textured plane was slanted top back, which is consistent with the vertical horopter [22, 79]. In the *inconsistent* condition it was slanted top forward, which is inconsistent with the horopter. The distance to the center of the textured plane was 29, 40, or 66cm. The slants of the plane were 7.3, 10.1, and 16.6° at those respective distances: top back in the *consistent* condition and top forward in the *inconsistent*. The top-back slants are the slant values of the vertical horopter at the three distances. Fig. 5.5 provides examples of *consistent* and *inconsistent* stimuli. We also presented conditions in which the slant values were twice as large: *i.e.*, 14.6, 20.2, and 33.2°. Again in half the sessions the slants were top back and in the other half they were top forward.

After each session, participants filled out a discomfort questionnaire in which they indicated on a 9-point scale how tired their eyes feel, how clear their vision is, how tired and sore

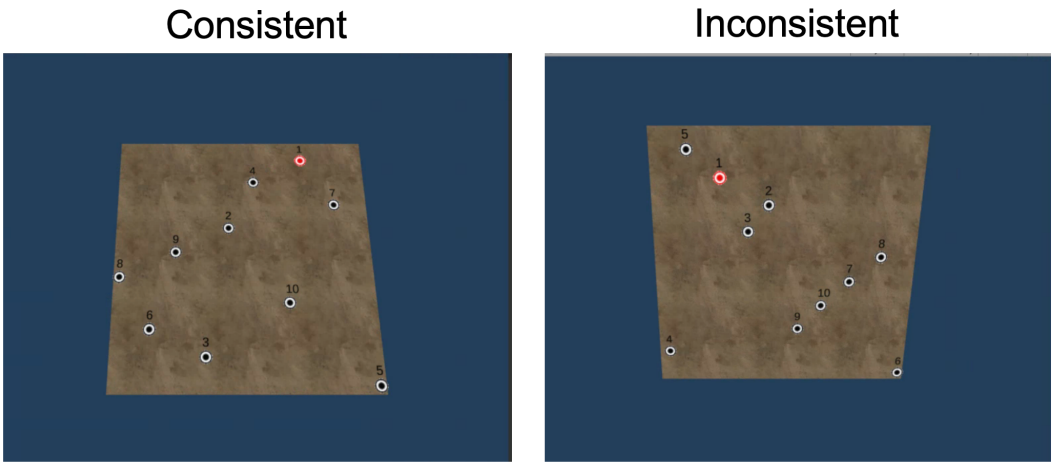

Figure 5.5: Examples of the stimuli in the *consistent* and *inconsistent* conditions. Here the distance is 66cm and the slants are +33.2 (top back) and -33.2° (top forward). The numbers and fixation targets next to them are visible. The currently fixated target is highlighted in red.

their neck and back are, how their eyes feel, and how their head feels. Participants were told to rate each symptom in terms of how they felt at the end of the session compared to the beginning. After completing the second session, participants also completed a questionnaire indicating on a 9-point scale which session they found most fatiguing, irritated their eyes the most, felt a worse headache in, and generally preferred least.

We also assessed performance by measuring how long it took for each observer to complete the task in each condition.

Results

Figure 5.6 shows the median reported symptoms on the symptom questionnaire: The more severe the reported symptom, the higher the black bar, which represents the median. For all questions, the median severity of each symptom was remarkably similar for both *consistent* and *inconsistent* symptoms, and the distributions clearly overlap to such a great extent that we did not pursue significance testing. Figure 5.7 shows the results for the display evaluation questionnaire, where participants were asked to directly compare their discomfort symptoms in the *consistent* and *inconsistent* sessions. Higher positive values are consistent with more favorable ratings for the *consistent* session, while a value of 0 indicates no difference in preference between sessions. The fatigue, eye irritation, and headache questions all have a median of 0, meaning on average there was no difference in preference between *consistent* and *inconsistent* session for these symptoms. The fourth question, directly asking observers to

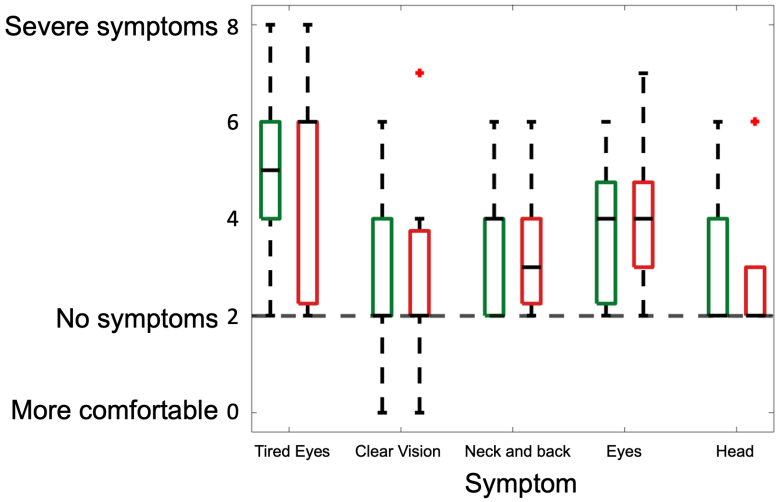

Figure 5.6: Results from the symptom questionnaire shown with a box plot. The median severity of the reported symptom is plotted for each of the nine questionnaire items, and is shown by the black bar in the box. The color of the box indicates condition (green consistent, red inconsistent). The bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually as a cross. The black dotted line indicates a response of 'no symptoms'.

compare the two sessions and to report which session they preferred shows a slight preference for the *consistent* session, but the broad spread of the responses (that clearly overlaps with 0) to this question suggest very inconsistent responses.

Figure 5.8A shows the median reaction time (RT) from the visual search task for *consistent* and *inconsistent* sessions. RT was calculated as the time between when the previous target was destroyed, and time at which the current search target was first fixated. The difference in RT between sessions is negligible, and there was not a significant difference between slant conditions ($t(9)$ = -.25, p =.81). Figure 5.8B plots RT by the distance the page was presented at. As there was not a significant difference in average RT between conditions, the median RTs for the *consistent* and *inconsistent* sessions were averaged for each distance. A one-way repeated measures ANOVA found a main effect of distance ($F(2)$ = 30.34, $p < .001$). Bonferroni-corrected posthoc t-tests shows the average RT decreases significantly as the page distance increases, and that the RTs at each page distance are significantly different from the other RTs (all p <.004). The lowest RT is found at the 66 cm page distance, which matches the optical distance of the screen in the HTC vive pro eye as is described in 3.2. At this distance, the vergence-accommodation conflict (VAC) is minimized to the greatest extent, and users have the most comfortable viewing experience [51]. The magnitude of the mismatch between the vergence and accommodation distance

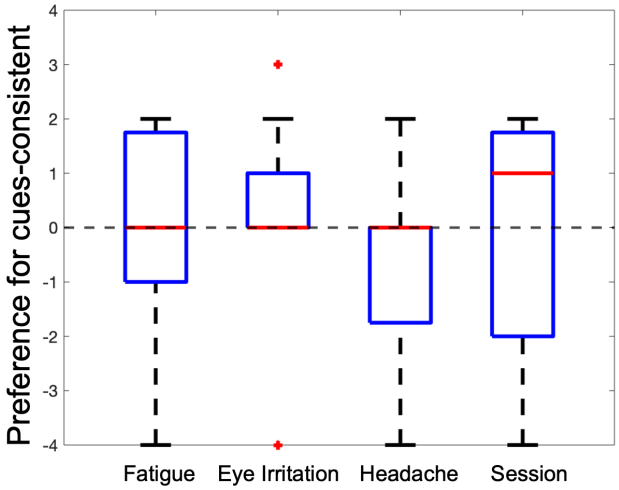

Figure 5.7: Results from the display evaluation questionnaire. The median comparative rating of a pair of sessions (consistent and inconsistent) is plotted for each of the four questionnaire items as the red bar inside each of the boxes. The bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually as a cross. The dashed line indicates a response of "no difference"

may be impacting observers' speed on the visual search task.

Discussion

The findings from this second pilot study did not show an effect of scene consistent content on viewer comfort and performance. There was no difference in average discomfort and speed between the *consistent* and *inconsistent* sessions for a visual search task. Figure 5.8B suggests a possible explanation as to why there is no effect of scene consistent content. Observers were consistently faster at the search task when the page distance matched the optical distance of the screen. In these trials, the vergence and accommodation distance were matched. Previous work has documented the effect of the VAC in stereoscopic displays. A greater conflict is associated with discomfort and even nausea [45, 51, 62, 108, 122, 57], impairment in visual performance [3, 51], and distortions of 3D percepts [132, 72]. Our results show that a greater VAC is associated with longer RTs. The majority of trials in this pilot study (around 2/3) contained a significant VAC. The discomfort from the VAC experienced in all experimental sessions may override discomfort from scene inconsistent content. In a final study (see 4.2) scene content was presented at the optical distance of the screen to minimize the VAC.

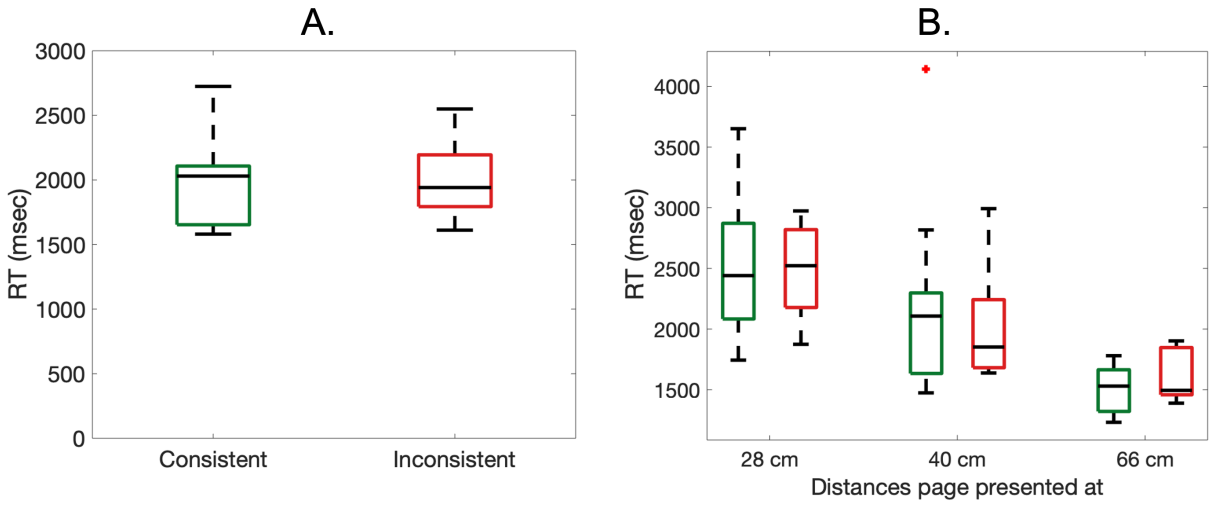

Figure 5.8: Reaction time (RT) from the visual search task. RT was calculated as the time from when the previous target was destroyed to the first time point the observer fixated the next target. A) Median RT for *consistent* and *inconsistent* sessions, averaged over all viewing distances and slants presented. B) Median RT for *consistent* and *inconsistent* sessions for each of the viewing distances tested. For both plots, the median is shown as the bar inside of each box. The color of the box indicates condition (green consistent, red inconsistent).The bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. The whiskers extend to the most extreme data points not considered outliers, and the outliers are plotted individually as a cross.