Binocular coordination in response to stereoscopic stimuli

Simon P. Liversedge[1], Nicolas S. Holliman[2] & Hazel I. Blythe[1]

[1]University of Southampton, UK   [2]Durham University, UK

ABSTRACT

Humans actively explore their visual environment by moving their eyes. Precise coordination of the eyes during visual scanning underlies the experience of a unified perceptual representation and is important for the perception of depth. We report data from three psychological experiments investigating human binocular coordination during visual processing of stereoscopic stimuli. In the first experiment participants were required to read sentences that contained a stereoscopically presented target word. Half of the word was presented exclusively to one eye and half exclusively to the other eye. Eye movements were recorded and showed that saccadic targeting was uninfluenced by the stereoscopic presentation, strongly suggesting that complementary retinal stimuli are perceived as a single, unified input prior to saccade initiation. In a second eye movement experiment we presented words stereoscopically to measure Panum's Fusional Area for linguistic stimuli. In the final experiment we compared binocular coordination during saccades between simple dot stimuli under 2D, stereoscopic 3D and real 3D viewing conditions. Results showed that depth appropriate vergence movements were made during saccades and fixations to real 3D stimuli, but only during fixations on stereoscopic 3D stimuli. 2D stimuli did not induce depth vergence movements. Together, these experiments indicate that stereoscopic visual stimuli are fused when they fall within Panum's Fusional Area, and that saccade metrics are computed on the basis of a unified percept. Also, there is sensitivity to non-foveal retinal disparity in real 3D stimuli, but not in stereoscopic 3D stimuli, and the system responsible for binocular coordination responds to this during saccades as well as fixations.

KEYWORDS: BINOCULAR COORDINATION   STEREOSCOPIC STIMULI FUSION

1. INTRODUCTION

Humans use their eyes to process visual information, and when sampling the visual array, they make a series of saccadic eye movements, rapid ballistic rotations of the eyeballs. Between saccades the eyes remain comparatively still, and the still periods are termed fixations [1,2]. Fixations usually last anywhere between about 200 ms and a second and it is during fixations that the visual information is extracted and processed psychologically (see [3,4]). In order that we perceive a single, non-diplopic representation of the visual environment, it is necessary that the movements of the two eyes are precisely coordinated. A complex and sophisticated neurophysiological system is responsible for oculomotor control [5,6], and in recent years, there has been a significant degree of interest in how people coordinate their eyes binocularly, both in relation to simple 2D stimuli, as well as in relation to more complex stereoscopic and real 3D stimuli (see [7] for a recent comprehensive review of experimental studies investigating human binocular coordination). One reason for this is that binocular coordination is central to the perception of depth. Furthermore, most often oculomotor responses occur spontaneously in relation to visual stimuli that are presented either in real depth, or stereoscopically. To this extent, therefore, binocular responses to qualitatively different visual stimuli can be taken as an index of the extent to which those stimuli were effective in conveying depth information to the perceiver.

Briefly, let us first characterise binocular coordination (in relatively simple terms). When making a saccade, the eye moving away from the nose (temporally) typically makes a larger, faster movement during the initial portion of the saccade with more rapid acceleration than the eye moving towards the nose (nasally), thus leading to a transient divergence [8]. During the latter half of the saccade, the temporal eye has a relatively longer deceleration period, allowing

the eye moving nasally to reduce the magnitude of the disparity which had accumulated between the two eyes (i.e., to catch up). The disparity that has developed during the saccade, however, is not entirely eliminated by the time the saccadic movement ends and the eyes are fixating. Thus, at fixation onset the eyes are disparate by a small amount and during the fixation slow vergence movements that work to reduce any residual disparity are made. Note, however, that the residual disparity is frequently not eliminated by the time that the next saccade is initiated. Thus, it is the case that fixation disparity often occurs during visual processing, thereby raising the question of how the visual system produces a single, unified, "cyclopean" percept of the environment given retinal disparity between the two eyes.

In the first experiment that we report below [9], we directly investigated this question by measuring participants' binocular eye movements as they read sentences containing a stereoscopically presented target word, of which, part was presented exclusively to the left eye, and part exclusively to the right eye. For example, in the sentence "Everyone scattered as the infamous cowboy drew his gun", the target word was "cowboy", and this was presented under three viewing conditions. In the control condition the whole word was presented to both eyes; in the congruous condition "cowb" was presented to the left eye and "wboy" to the right; and in the incongruous condition "wboy" was presented to the left eye and "cowb" was presented to the right. In both the congruous and incongruous conditions the "w" and the "b" of both word portions overlapped to ensure that the word parts were appropriately spatially aligned, and therefore perfectly complementary. Note also that the rest of the words in the sentence were presented to both eyes in their entirety. We analysed movements of both eyes as participants made a saccade onto the target word to determine whether readers would target the whole word with both eyes (in which case this would indicate that the word portions had been perceived as a single, unified input), or instead target one or other word part with each eye (in which case this would indicate either that one of the two visual inputs had been suppressed, or that saccades for the two eyes were programmed independently).

In our second experiment, we investigated how much retinal disparity the visual system could tolerate before unified vision breaks down and we experience diplopia (double vision). To address this question we presented single words stereoscopically that were either correctly or incorrectly spelled. Binocular eye movements were recorded throughout testing. Participants were required to make a decision as to whether the words were spelled correctly, and to do this they had to fuse the stereoscopic stimuli. To assess the limits of Panum's Fusional Area, we also manipulated the degree to which the stereoscopically presented word stimuli were horizontally offset (i.e., by 0, 1 or 2 characters). Thus, by taking into account the degree of horizontal disparity in the manipulated experimental stimuli, along with precise measurement of the fixation positions of the two eyes, we were able to obtain a very accurate index of the degree of retinal disparity that existed when participants made their decision about the spelling of the word. In this way, we were able to quantify the degree of retinal disparity that participants could tolerate whilst being able to maintain a non-diplopic percept of the word. In this way, we computed the limits of Panum's Fusional Area.

The final experiment that we conducted is, perhaps, of most interest to researchers developing 3D displays. In this experiment we wished to directly compare human oculomotor responses to visual stimuli under three different viewing conditions. To do this, we used stimuli based on very simple LED point light sources arranged in depth. There were three viewing conditions: a 3D array of LED light sources; a stereoscopic 3D digitised photographic representation of this visual array (viewed through shutter goggles); and a 2D digitised photographic representation of the array. The experimental task was also very simple. Participants' eye movements were recorded as they made saccades in time to a regular beat provided by a metronome between each of the points of light in the display in a pre-determined order. Specifically, we were interested to know how responsive the binocular coordination system was to depth cues under the two 3D conditions (in comparison to the 2D baseline condition). To assess this, we quantified vergence movements during saccades between the light points at different depths under the LED and stereoscopic 3D conditions and compared this to that which occurred in the 2D control condition. In this way, we were able to assess the degree to which binocular eye movements to real 3D stimuli were similar to those for stereoscopic 3D stimuli. In addition, this manipulation allowed us to investigate whether participants were sensitive to the depth of a saccade target before they actually made an eye movement to that target. That is to say, we were interested to know exactly when participants first made depth appropriate vergence movements as they made saccades between LEDs at different depths. If participants made

vergence movements during a saccade to a target LED, then this would indicate that the depth of the target was perceived and impacted on saccade metric computation prior to saccade initiation. In contrast, if depth appropriate vergence movements did not occur until fixation onset, then this would suggest that the depth of the target did not influence saccade computations, and depth appropriate vergence movements were only initiated once the participant had directly fixated the target LED.

2. EXPERIMENT 1

2.1. Methods.

2.1.1. Participants. Participants were adults aged 18- to 21-years and were undergraduate volunteers from Durham University with normal, uncorrected vision.

2.1.2. Apparatus. Stereoscopic presentation of the target words was achieved through use of CRS shutter goggles, interfaced with a Pentium 4 computer and a Philips 21B582BH 24" monitor on which the sentences were displayed. Binocular eye movement recordings were taken with two Fourward Technologies Dual Purkinje Image eye trackers. The positions of both eyes were recorded every millisecond. The display monitor was set at a viewing distance of 100 cm. All sentences were presented as white text on a black background. At the specified viewing distance, each character subtended 0.29º of visual angle. All participants bit on a wax dental mould and used forehead rests during the experiment, to eliminate head movements.

2.1.3. Materials & design. We constructed 72 experimental sentences, each containing a target word that was a compound noun 6, 8 or 10 letters long. Three files of experimental stimuli were prepared with each presentation condition for each target word being rotated across files according to a Latin Square. All stimuli appeared in each presentation condition equally often, and no participant saw the same stimulus twice.
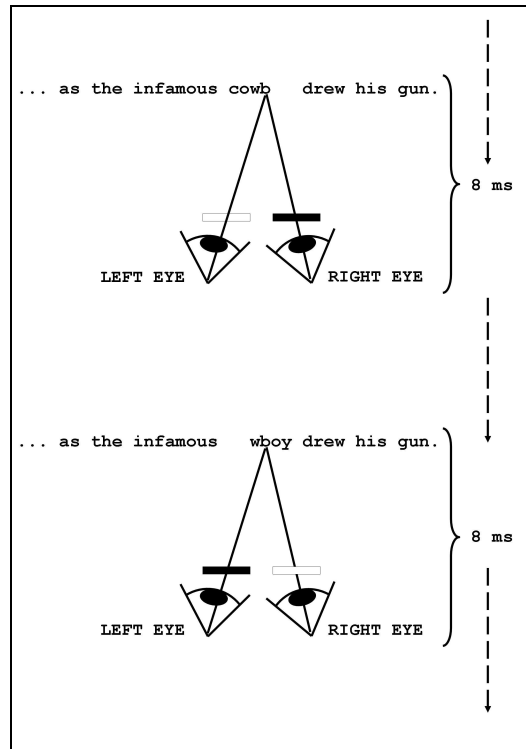
Figure 1. Example stimulus presentation, with the target word being "cowboy". There were three viewing conditions (the example here shows the incongruent condition). In the control condition the whole word was presented to both eyes; in the congruous condition "cowb" was presented to the left eye and "wboy" to the right; and in the incongruous condition "wboy" was presented to the left eye and "cowb" was presented to the right. In both the congruous and incongruous conditions the "w" and the "b" of both word portions overlapped to ensure that the word parts were appropriately spatially aligned, and therefore perfectly complementary. Note also that the rest of the words in the sentence were presented to both eyes in their entirety.

2.1.4. Procedure. Participants were presented with written instructions explaining the experimental procedure. They were told that they would be required to read the sentences normally and press a button to terminate the display when they had understood the sentences to the best of their ability. Prior to testing, the eye trackers were calibrated monocularly using a nine point display and the accuracy of the calibration was checked prior to every trial. Recalibrations were carried out whenever necessary. Participants were required to answer Yes/No comprehension questions using a button box that appeared on the screen after 24 of the experimental sentences. The experiment lasted about 45 minutes.

2.1.5. Analyses. Custom-designed software was used for the data analyses. Fixations were manually identified in order to avoid contamination by dynamic overshoots [10, 11].

2.2. Results. The results that were of primary interest in this experiment were the reading times for the critical word and the saccadic landing positions of each eye on the critical word. We computed first fixation duration (the duration of the first fixation on the target word), single fixation duration (the duration of the fixation on the target word when only one fixation was made before a different word was fixated) and gaze duration (the sum of all fixations on the target word prior to a different word being fixated) for the target word. These measures showed effects that were highly statistically reliable. For all three measures reading times were the same for the congruous and incongruous presentation conditions, and the incongruous and congruous conditions were longer than the control condition. Thus, the stereoscopic manipulation was influential in that it affected the time spent processing the word when it was first encountered. The

critical aspect of the data, however, concerned the saccadic landing positions on the target word. Again, the effects for this measure were highly reliable. The left eye initially fixated consistently, on average, one character to the left of the right eye. This reliable difference reflected the well documented finding that the eyes are disparate (with the lines of sight uncrossed) by a small amount at fixation onset. However, most importantly, the magnitude and direction of this disparity was uninfluenced by presentation condition; landing positions on the word were the same regardless of whether the target word was presented in its entirety (control condition), congruously or incongruously. Thus, participants computed the extent of their saccades to the target word based on the input of both eyes, rather than on the basis of the input from either eye alone.

2.3. Discussion. The results from our first experiment strongly suggest that participants fused the portions of the target word when it was presented stereoscopically. Furthermore, given the similarity of landing positions on the target word across presentation conditions, it appears that readers computed their saccade metrics on the basis of a single, unified percept. The data represent strong evidence to suggest that fusion underlies the perception of a "cyclopean" visual world during reading.

3. EXPERIMENT 2

3.1. Methods.

3.1.1. Participants. As in Experiment 1.

3.1.2. Apparatus. As in Experiment 1. All words were presented in block capitals, as white text on a black background. At the specified viewing distance, one character space subtended 0.37° of visual angle.

3.1.3. Materials & design. All participants viewed 100 trials, each trial consisting of a single 6-letter word. 10 counterbalanced files were created from the set of 100 words, so that every word appeared in each of the 5 experimental conditions, both correctly and incorrectly spelled. The misspellings were always a single-letter substitution, of either the 3rd or the 4th letter of the word (i.e. the misspelling was always in the centre of the word). The misspelling always created an obvious nonword (for example, changing HAMMER to HAQMER). The five experimental conditions were created through use of the shutter goggles, and refer to the horizontal displacement between the images sent to the two eyes. Words could either be presented as uncrossed by 1 or 2 character spaces, where the image presented to the left eye was shifted to the left on the display relative to the image presented to the right eye, aligned, where the two images were in the same location on the display, or crossed by 1 or 2 character spaces, where the image presented to the left eye was shifted to the right of the image presented to the right eye.

3.1.4. Procedure. Participants were instructed to look at the fixation cross which appeared on the left of the screen for one second at the beginning of reach trial, before looking at the word which was presented in the centre of the screen simultaneously with the offset of the fixation cross. This was to ensure that, on each trial, participants made a rightward saccade onto the target, rather than continuously fixating the centre of the display throughout the experiment. Participants were instructed to decide, for each trial, whether or not the word was correctly spelled, and to make their response using a yes/no button box. The left and right eye trackers were calibrated for each eye monocularly in turn (i.e., during calibration of the right eye, the left eye was occluded and vice-versa. Once both eyes had been calibrated accurately, the practice and experimental sentences were then presented. All participants had 5 practice trials before the experimental trials began. Following every four trials, the calibration was checked for accuracy, and the eye trackers were recalibrated if necessary. All participants were given a break half way through the experiment, and additional breaks were given as often as required.

3.1.5. Analyses. For every fixation, the disparity that was manipulated in the display was combined with that which was recorded to have occurred between the eyes, to form the categories presented in Table One.

Stereoscopic disparity		Binocular disparity					
		-3	-2	-1	0	1	2
	-2	U1	A	C1	C2	C3	C4
	-1	U2	U1	A	C1	C2	C3
	0	U3	U2	U1	A	C1	C2
	1	U4	U3	U2	U1	A	C1
	2	U5	U4	U3	U2	U1	A

Table One. The categories of total retinal disparity, formed by combining the disparity manipulated on the display with the disparity which occurred between the eyes. Negative numbers denote disparities which are uncrossed by the given number of character spaces, while positive numbers denote disparities which are crossed by the given number of character spaces. One character space = 0.37º.

3.2. Results. First, we examined vergence movements during both the initial saccade from the fixation cross to the target word, and during the initial fixation on the target word. There was no effect of the disparity of the parafoveal stereoscopic stimulus upon binocular coordination during the saccade, indicating that saccadic programming to the parafoveal stimulus was not modulated by the stereoscopic presentation. There was, however, a significant effect of disparity on vergence movements during fixations once the stimulus was directly fixated. During fixations on the target words, participants made direction-appropriate vergence movements in order to reduce the experienced retinal disparity (i.e. they converged their eyes when experiencing uncrossed retinal disparity and diverged their eyes when experiencing crossed retinal disparity). Furthermore, the magnitude of vergence movements was greater in response to larger magnitudes of experienced retinal disparity. Thus vergence movements during fixations on the target words were both direction- and magnitude-appropriate.

We then examined mean fixation duration, trial viewing time, and the number of fixations per trial as a function of retinal disparity. As retinal disparity increased (in either direction), fixation durations, trial viewing times, and the number of fixations per trial all increased, reflecting the greater difficulty experienced by the participants in completing the task. Finally, we examined response accuracy as our measure of whether participants had successfully fused the stimuli across the range of retinal disparities analysed.
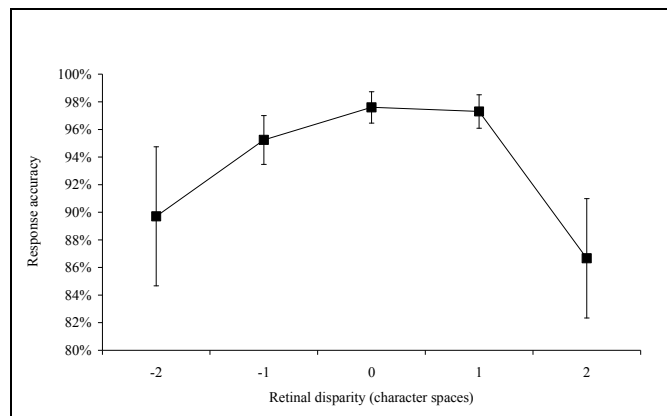

Figure 2. Response accuracy data from the lexical decision task. The x-axis shows retinal disparity; negative numbers denote disparities which are uncrossed by the given number of character spaces, while positive numbers denote disparities which are crossed by the given number of character spaces. One character space = 0.37º.

We found a significant decrease in response accuracy with greater retinal disparities (see Fig. 2). Within one character space of retinal disparity in either direction, response accuracy was between 95% and 98% correct. In contrast, this decreased by around 10% with retinal disparities of two character spaces, both crossed and uncrossed.

3.3. Discussion. These data clearly demonstrate the effects of retinal disparity upon both the vergence response, and then the subsequent ease and success of fusion. The lack of disparity-driven vergence movements during saccades onto the stereoscopic stimuli suggests a lack of sensitivity to disparity in the stereoscopic stimuli in the parafovea. In contrast, the strong effects of retinal disparity upon all measures of eye movement control during fixations on the stimuli provide strong evidence of participants' fusion. The decrease in response accuracy associated with retinal disparities of two character spaces in either direction suggests that the limits of Panum's fusional area for linguistic stimuli lie between 0.37 ° and 0.74º. The results from this study have implications for the presentation of text in stereoscopic displays, showing that there are clear limits for the magnitude of disparity that can be presented in order for viewers to be able to fuse and read text. These limits for fusion of linguistic stimuli are in line with those reported in experimental results from Jones et al. [12] for high contrast stimuli, and in line with earlier work surveyed in the same source.

4. EXPERIMENT 3

4.1. Methods.

4.1.1. Participants. The nine participants were adults aged 18- to 30-years, and were volunteers from the University of Southampton with normal, uncorrected vision.

4.1.2. Apparatus. Binocular eye movement recordings were taken with two Fourward Technologies Dual Purkinje image eye trackers. The position of each eye was recorded every millisecond. In the LEDs condition, two rows of coloured LEDs were suspended on rigid rods from the ceiling of a tunnel. The tunnel was built to fit up to the desk on which the eye trackers were mounted, with 3 walls and a ceiling all painted in black. The front of the tunnel was open to allow participants to view the stimuli within. The lid of the tunnel was hinged, allowing the LEDs to be lifted. There was a slot in the left side of the tunnel which allowed an LCD monitor to be slid into the tunnel for the calibration routine, when the lid was raised. During trials, the slot was covered with a black curtain.
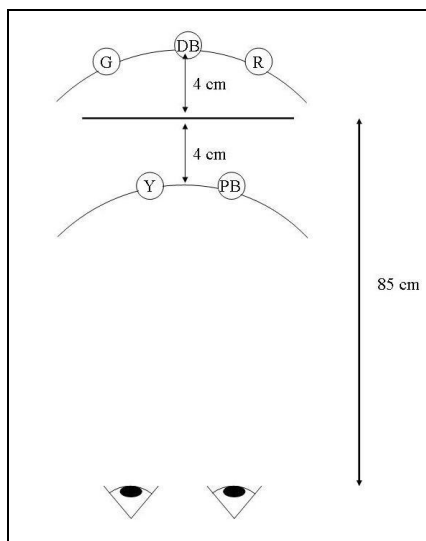

Figure 3. The viewing arrangement for the LEDs condition from a bird's eye perspective. The calibration monitor was placed at 85 cm, the viewing distance for the continuously present monitor on which the 2D and stereoscopic stimuli were presented in the other conditions. The two arcs represent the two (hypothetical) circles on which the rows of LEDs were placed; the front arc was at a distance of 89 cm and the back arc was at a distance of 81 cm. The letters show the colours of the LEDs – green, yellow, dark blue, pale blue, and red.

From left to right, the positions of the LEDs will be referred to as 1-5. The LEDs were green, yellow, dark blue, pale blue, and red, respectively. LEDs 1, 3, and 5 were presented on the back row, at a viewing distance of 89 cm. LEDs 2 and 4 were presented on the front row, at a viewing distance of 81 cm. The rows of LEDs were placed upon hypothetical circles, with the midpoint of each circle falling half-way between the participant's eyes and LEDs. This arrangement minimises the requirement for a change in the angle between the two eyes when looking between points on any given row. The outer circle, corresponding to the back row of LEDs, had a radius of 44.5 cm and the inner circle, corresponding to the front row of LEDs, had a radius of 40.5 cm. There was a distance of 12 cm between LEDs on the back row, and of 10.9 cm on the front row, corresponding to 7.68º in both cases. The calibration screen was positioned at a viewing distance of 85 cm when in place, i.e. at a mid-distance between the two rows of LEDs. In the stereoscopic and 2D conditions, the display monitor was set at a viewing distance of 85 cm, corresponding to the calibration distance in the LED condition. In all conditions, calibration routines and data collection were conducted using in-house software. The same software was used to present the images in the stereoscopic and 2D conditions. In the LEDs condition, the display screen was removed from the tunnel whilst the data were recorded. All participants bit on a wax dental mould and used forehead rests during the experiment, to eliminate head movements.

4.1.3. Materials and design. In the main experiment stimuli were viewed under all three conditions, and the order in which conditions were viewed was counterbalanced across participants. In the LEDs condition, participants viewed the LEDs within the tunnel described in the apparatus section above. For the stereoscopic condition, images of the LEDs were presented stereoscopically on a display screen at 85 cm viewing distance. The images were 1024 x 768 pixels. The images were captured using an orthoscopic camera arrangement so that the images when viewed had the same depth properties as the physical LED stimulus. A carefully aligned single camera (Olympus SP-500UZ) on a rail was used to photograph the LEDs; this avoided photometric errors and any remaining minor geometric errors were removed using the automatic stereo rectification tool in the Stereo Photomaker software package. Stereoscopic presentation of these images was achieved through use of CRS shutter goggles, interfaced with a Pentium 4 computer and a Philips 21B582BH 24" monitor on which the images were displayed. For the 2D condition, a single, central photograph of the LEDs was used. In each viewing condition (LEDs, 2D, and stereoscopic), there were three trials. In all trials, the specifically instructed eye movements (near version, far version, or combined movements) were preceded by an instruction to look around the array and ensure that they knew where each of the coloured LEDs was, before being asked to look straight ahead at the central LED whilst being given the trial-specific task. On trial one, participants were asked to move their eyes back and forth between LEDs 2 and 4 (the two on the front row), referred to as "near version" movements. On trial two, they were asked to move their eyes back and forth between LEDs 1, 3, and 5 (the three on the back row), referred to "far version" movements. On trial three, they were asked to move their eyes back and forth between all five LEDs (1, 2, 3, 4, and 5), from left to right and back again, looking at each in turn, referred to as "combined" movements. In all cases, the instructions were given to participants in terms of the colours of the LEDs, i.e., "Now look backwards and forwards between the yellow and the pale blue lights". Thus, we employed a 3 (viewing condition: LEDs vs. stereoscopic images vs. 2D image) x 3 (task: near version vs. far version vs. combined movements) design.

4.1.4. Procedure. All participants were first tested for stereopsis. Once this had been checked, the experimental trials were started. At the beginning of each viewing condition, the left and right eye trackers were calibrated for each eye monocularly in turn. Once both eyes had been calibrated accurately, the stimuli were presented. At the beginning of every trial within each viewing condition, each calibration was checked monocularly for accuracy and the eye trackers were recalibrated if necessary. All participants were given breaks between viewing conditions. The entire experiment lasted approximately one hour.
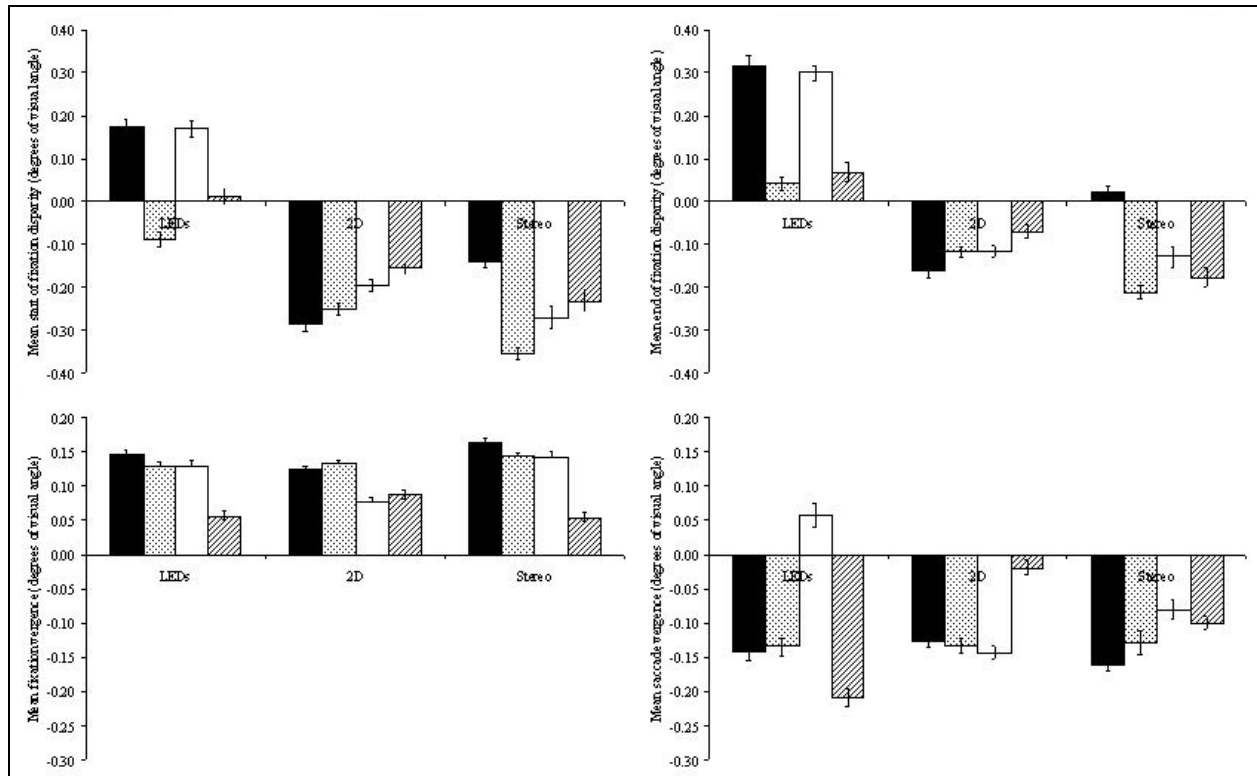
Figure 4. Black bars represent near version movements, dotted bars represent far version, white bars represent convergence, and striped bars represent divergence. The top left panel shows start of fixation disparity and the top right panel shows end of fixation disparity in degrees of visual angle, where positive values denote crossed fixations and negative values denote uncrossed fixations. The bottom left panel shows vergence during fixations and the bottom right panel shows vergence during saccades in degrees of visual angle, where positive values denote convergence and negative values denote divergence.

4.2. Results.

4.2.1. Saccade vergence. Here, we examined the change in vergence angle between the two eyes from the start to the end of the saccade. There were no overall differences between the three viewing conditions (LEDs, 2D, and stereoscopic viewing conditions). There were significant differences in saccade vergence between the four types of movement that were made (near version, far version, and convergent or divergent combined movements), as well as a significant interaction between the viewing condition and the movement type. The two eyes diverged during both near and far version movements in all viewing conditions, with this divergence being of similar magnitude in all conditions. The eyes converged during combined convergent movements when viewing the real LEDs, as might be expected. However, when viewing both the 2D images and the stereoscopic images the eyes diverged during the combined convergent movements. The magnitude of this divergence was greater in the 2D condition than in the stereoscopic condition. With respect to the combined divergent movements, the eyes diverged in all three conditions, although with varying magnitude. The magnitude of divergence was greatest for the real LEDs, and was smallest for the 2D condition. To summarise, vergence movements were made during saccades to correctly align the eyes at the depth of the stimulus when viewing the real LEDs, as might be expected. When viewing the 2D images, there was minimal modulation of the vergence movements by the stimulus, as a consequence of the fact that all the LEDs appeared to be at the same depth. The saccade vergence data from the stereoscopic condition were most similar to those from the 2D condition suggesting a lack of sensitivity to stereoscopic depth information in the parafovea.

4.2.2. Start of fixation alignment. For start of fixation alignment, there were significant differences between both the viewing conditions and the movement type, but there were no interactive effects. Fixations were primarily uncrossed for

both the 2D and the stereoscopic viewing conditions, but were primarily crossed for the LEDs. As can be seen in the top left panel of Figure 4, the effect of movement type was primarily driven by the LED data (although the interaction was not significant, the 2D and stereoscopic data did not show strong differences across conditions). Specifically, when viewing the LEDs, participants made more crossed fixations following near version and convergent movements, and more uncrossed fixations following far version movements. There was some suggestion in the stereoscopic data that the magnitude of uncrossed disparity was reduced during near version trials (i.e., the trials where participants were required to move their eyes back and forth between the two near positions).

4.2.3. Fixation vergence. For these analyses, the change in vergence angle between the two eyes from the start to the end of fixation was compared across conditions and preceding movement types. Again, there was a significant effect of movement type as well as a significant interaction between movement type and viewing condition. Convergence occurred during fixations following all types of movement, in all three viewing conditions. In fixations following near version movements, the magnitude of this convergence varied across viewing conditions. It was greatest for the stereoscopic condition, and smallest for the 2D condition. There was no difference between the viewing conditions for convergence during fixations following far version. For convergent combined movements, again, the magnitude of convergence varied across viewing conditions – it was smaller when viewing 2D images, while it was similar between the LEDs and the stereoscopic viewing conditions. Finally, for divergent combined movements, there was a similar pattern to convergent combined movements – the magnitude of convergence was similar for LED and stereoscopic viewing, but here was significantly greater in the 2D condition. This suggests that the alignment of the two eyes is being adjusted in a direction-appropriate manner in relation to the depth of the stimulus during fixations when viewing the real LEDs and the stereoscopic images, but not when viewing the 2D image.

4.2.4. End of fixation alignment. A similar pattern was observed here as for the start of fixation alignment data. There were effects of both viewing condition and movement type, but no interaction between the two. Here, fixations were uncrossed in both the 2D and the stereoscopic viewing conditions, but crossed for the LEDs. Again, the effect of movement type is clearest in the data from the LED and stereoscopic viewing conditions showing that there were more crossed fixations following near version movements (or that the magnitude of uncrossed disparity was reduced) compared to fixations following far version or combined movements. To be clear, by the end of fixation, the data show that alignment of the two eyes had become more appropriate for the depth of the fixated stimulus than was the case at the start of the fixation when stereoscopic images were viewed.

4.3. Discussion. These data show clear differences in binocular coordination between the three viewing conditions, with respect to the characteristics of binocular coordination that were elicited as a consequence of the different movement types being made. First, there were minimal differences between the types of movement when participants were viewing 2D images, as might be expected – in this case, the different movement types only affect saccade amplitude, as all five LEDs appear at the same depth. In contrast, there were differences between the movement types for both the real LEDs and for the stereoscopic images. Recall that for some movement types the participant is required to make horizontal movements between two points without the stimulus requiring any change in the depth of the point of fixation (version movements) while other movement types require both the horizontal position and the depth of the point of fixation to change on a fixation by fixation basis. The participants' binocular coordination ought, therefore, to reflect the demands of the task on different trials. The most interesting data here are those from the analysis of vergence movements during saccades and fixations. During saccades, real LED stimuli elicited direction-appropriate vergence in cases where it was necessary to change the depth at which the two eyes were aligned. In contrast, the eyes diverged during all movements when viewing stereoscopic and 2D images. To be clear, vergence movements during saccades were direction-appropriate only when viewing the real LEDs. Participants did not alter the alignment of the two eyes during saccades in order to move the point of fixation in depth as necessary when viewing the 2D image (as might be expected) but also when viewing the stereoscopic images. In contrast, during fixations the pattern was quite different. When the preceding movement had required the point of fixation to change in depth, the magnitude of vergence during fixations was determined by the depth of the fixated point for both the real LED and the stereoscopic images (there was little effect here for the 2D stimuli as might have been expected). In summary, vergence movements occur during both fixations and

saccades when viewing real stimuli in depth, in order to change the alignment of the eyes and, hence, the point of fixation, appropriately in depth. A different pattern of binocular coordination was elicited by the stereoscopic images, such that while vergence movements during fixations were direction-appropriate, there was little evidence to suggest that vergence movements during saccades were driven by the depth of the saccade target.

5. GENERAL DISCUSSION

In order for us to perceive a single unified visual representation of our environment, the movements of the eyes are precisely coordinated. Also, most often, oculomotor responses occur spontaneously in relation to different types of visual stimuli that are fixated. In this paper we reported three experiments in which we measured binocular coordination when different types of visual stimuli were viewed. In the first experiment we investigated saccadic targeting to parafoveal words that were presented stereoscopically and showed that part words that were presented uniquely to each eye were unified prior to fixation and that saccade metrics for both eyes were computed on the basis of this cyclopean percept. In the second experiment we used a similar technique to quantify Panum's Fusional Area for words and to demonstrate that the system responsible for the coordination of the eyes was very sensitive and responsive to stereoscopic disparity during fixations. Such disparity is frequently employed as the primary cue to depth in a variety of 3D displays. Finally, in the third experiment we directly compared binocular coordination during viewing of 2D, stereoscopic and real 3D visual stimuli. We showed that there was a parafoveal sensitivity to depth cues for real 3D stimuli only. For such stimuli participants made depth appropriate vergence movements during saccades to a target. For the stereoscopic stimuli, however, the oculomotor control system responded exclusively to depth cues in directly fixated stimuli. Our results show very clear qualitative differences in oculomotor behaviour in relation to stereoscopic stimuli compared to real 3D stimuli. Furthermore, if we take binocular eye movements to provide a measure of the extent to which visual stimuli effectively convey depth information, then our results suggest that stereoscopic representations of stimuli are limited in their efficacy relative to the natural stimuli themselves. One possible interpretation of these differences in binocular coordination between real 3D stimuli, 2D images, and stereoscopic representations is that the visual system must work harder in order to perceive depth from stereoscopic displays than from real 3D stimuli. For stereoscopic stimuli all vergence movements change the depth of the point of fixation are made during fixations; for such stimuli saccades do not serve the purpose of re-aligning the two eyes at the depth of the stimulus as is the case in real 3D scenes. Further research is necessary in order to determine whether it is the case that viewing stereoscopic images places the visual system under duress (in this and other ways), and whether the behavioural differences we have reported occur as a consequence of some form of visual fatigue.

6. REFERENCES

[1] Liversedge, S.P., & Findlay, J.M. "Saccadic eye movements and cognition." Trends in Cognitive Sciences, 4, 6-14 (2000).

[2] Rayner, K. "Eye movements in reading and information processing: 20 years of research." Psychological Bulletin, 124, 372-422 (1998).

[3] Rayner, K. and Pollatsek, A. [The psychology of reading]. Prentice-Hall (1989); reprinted by Erlbaum (1995).

[4] Findlay, J.M., & Gilchrist, I.D. [Active Vision: The Psychology of Looking and Seeing.] Oxford University Press (2003)

[5] Findlay, J.M. & Walker, R. "A model of saccade generation based on parallel processing and competitive inhibition." Behavioral and Brain Sciences 22, 661-674 (1999).

[6] Stone, L.S., Miles, A.M., & Banks, M.S. "Linking eye movements and perception". Journal of Vision, Special Issue (2003).

[7] Kirkby, J.A., Webster, L.A.D., Blythe, H.I., & Liversedge, S.P. "Binocular coordination during reading and non-reading tasks." Psychological Bulletin, 134, 742-763 (2008).

[8] Collewijn, H., Erkelens, C.J., & Steinman, R.M. "Binocular coordination of human horizontal saccadic eye movements." Journal of Physiology, 404, 157-182 (1988).

[9] Liversedge, S.P., Rayner, K., White, S.J., Findlay, J.M., & McSorley, E. "Binocular coordination of the eyes during reading." Current Biology, 16, 1726-1729 (2006).

[10] Deubel, H. & Bridgeman, B. "Fourth Purkinje image signals reveal lens deviations and retinal image distortions during saccadic eye movements." Vision Research, 35, 529-538 (1995).

[11] Liversedge, S.P., White, S.J., Findlay, J.M. & Rayner, K. "Binocular coordination of eye movements during reading." Vision Research, 46 (15), 2363-2374 (2006).

[12] Jones, G., Lee, D., Holliman, N., and D. Ezra, "Controlling perceived depth in stereoscopic images," in Stereoscopic Displays and Virtual Reality Systems VIII, Proceedings of SPIE 4297A, 2001.