Investigating eye movement acquisition and analysis technologies as a causal factor in differential prevalence of crossed and uncrossed fixation disparity during reading and dot scanning.

Kirkby J. A.[1]*, Blythe H. I.[2], Drieghe D.[2], Benson V.[2], & Liversedge S. P.[2]

1. Department of Psychology, Bournemouth University, UK

2. Psychology, University of Southampton, UK

*Corresponding author

Department of Psychology

Bournemouth University

Talbot Campus

Fern Barrow

Poole

BH12 5BB

Tel: +44 (0) 1202 965251

Email: jkirkby@bournemouth.ac.uk

Abstract

Previous studies examining binocular coordination during reading have reported conflicting results in terms of the nature of disparity (e.g., Liversedge, White, Findlay, & Rayner, 2006; Kliegl, Nuthmann, & Engbert, 2006). One potential cause of this inconsistency is differences in acquisition devices and associated analysis technologies. We tested this by directly comparing binocular eye movement recordings made using SR Research EyeLink 1000 and the Fourward Technologies Inc. DPI binocular eye tracking systems. Participants read sentences or scanned horizontal rows of dot strings; for each participant half the data were recorded with the EyeLink and the other half with the DPIs. The viewing conditions in both testing laboratories were set to be very similar. Monocular calibrations were used. The majority of fixations recorded using either system were aligned, although data from the EyeLink system showed greater disparity magnitudes. Critically, for unaligned fixations, the data from both systems showed a majority of uncrossed fixations. These results suggest that variability in previous reports of binocular fixation alignment is attributable to the specific viewing conditions associated with a particular experiment (variables such as luminance and viewing distance), rather than acquisition and analysis software and hardware.

Key words: Binocular coordination, eye movements, reading and non-reading tasks

Until recently an implicit assumption was held amongst researchers that during reading the two eyes are precisely coupled so that both eyes fixate the same letter within a word, ensuring that the visual system is supplied with matching visual inputs. A body of work investigating binocular coordination during reading has accumulated, however, which has demonstrated that some degree of disparity is present during a substantial proportion of fixations (Blythe et al., 2006; Jainta, Hoorman, Kloke, & Jaschinski, 2010; Juhasz, Liversedge, White, & Rayner, 2006; Kliegl, Nuthmann, & Engbert, 2006; Kirkby, Blythe, Benson, & Liversedge, 2010; Liversedge, White, Findlay, & Rayner, 2006; Nuthmann & Kliegl, 2009; Shillcock, Roberts, Kreiner, & Obregón, 2010; see, Kirkby, Webster, Blythe, & Liversedge, 2008, for a review of binocular coordination during reading and non-reading tasks). Despite this disparity, readers do not typically experience diplopia (double vision). A single visual representation is primarily achieved by the coordination of the two eyes; however, the evidence thus far indicates that the visual system is frequently required to construct a fused perceptual representation from two disparate retinal inputs.

Fixation disparity has been characterised both in terms of its magnitude, typically measured in character spaces, and in terms of its direction as aligned, crossed, or uncrossed (Liversedge, White et al., 2006; see Figure 1). Several studies have now reported the proportions of aligned, crossed, and uncrossed fixations to be relatively constant during both reading and non-reading tasks (Blythe et al., 2006; Juhasz et al., 2006; Kirkby et al., 2010; Liversedge, White et al., 2006). These studies have found the majority of fixations to be aligned but, within those that were unaligned, the majority were in an uncrossed direction; relatively few crossed fixations were observed. During fixations, corrective vergence movements were found to be predominantly convergent, such that the magnitude of disparity was reduced by the end of fixation.

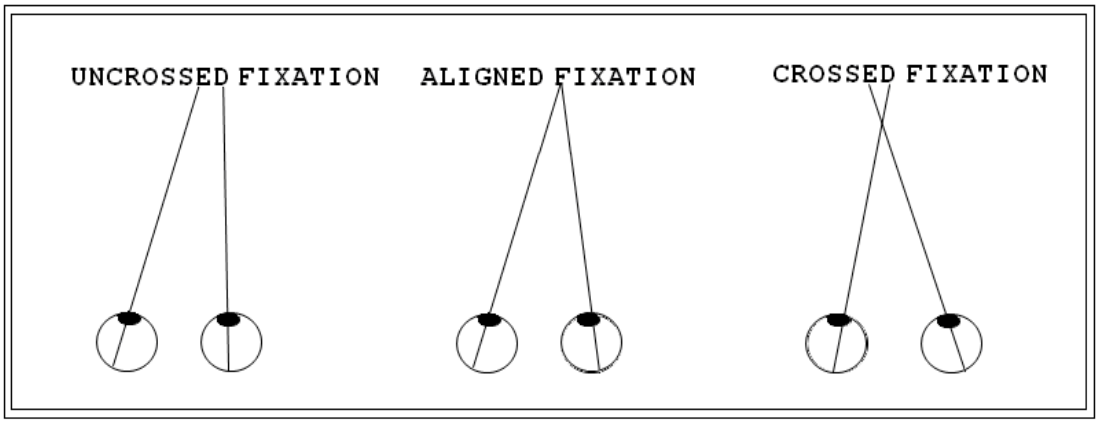


Figure 1. *Categories of fixation disparity. Aligned fixations are those where both eyes' positions are within one character space at the plane of text. Crossed fixations are those where the two eyes' lines of sight are aligned in front of the plane of text such that the lines of sight at the plane of text are literally crossed, and fixations are disparate by at least one character space. Uncrossed fixations are those where the two eyes' lines of sight are aligned behind the plane of text such that the fixation positions at the plane of text are uncrossed by at least one character space.*

In contrast, however, other researchers have reported the majority of unaligned fixations during reading to be crossed (Nuthmann & Kliegl, 2009; Shillcock, Roberts, Kreiner & Obregón, 2010). In a study by Nuthmann and Kliegl (2009), analyses of binocular data based on the Potsdam-Sentence-Corpus were reported. Their findings were consistent with the data reported from other studies, in that fixation disparity occurred during approximately half of all fixations. On average, they found the absolute magnitude of disparity at the start of fixations to be 1.22 character spaces, which was then reduced to 1.03 character spaces by the end of fixations. Intriguingly, however, Nuthmann and Kliegl found that unaligned fixations were predominantly in a crossed direction – the opposite pattern to that observed in previous binocular coordination studies (e.g., Blythe et al., 2006; Juhasz et al., 2006; Kirkby et al., 2010; Liversedge, Rayner, et al., 2006; Liversedge, White et al., 2006).

A number of possible factors have been proposed in order to account for these different findings: (1) the particular eye tracking systems used to acquire binocular data (i.e., Dual Purkinje Image trackers or the EyeLink 1000[1]); (2) the software associated with analysing binocular data; (3) the luminance of the room during data collection; (4) viewing distance; (5) font size; (6) colour combination of text stimuli (black text on white background or vice versa); (7) individual differences in readers; (8) the calibration procedure employed (monocular vs. binocular viewing during calibration); (9) the language of the stimuli (in reading experiments); (10) whether the stimuli were formatted as sentences or paragraphs (Kirkby et al., 2008; Nuthmann & Kliegl, 2009; Shillcock et al., 2010). The primary aim of this experiment was to investigate the first of these possible explanatory factors by making a direct comparison of the binocular eye movement data recorded by two Dual Purkinje Image eye trackers and an EyeLink 1000 eye tracker, whilst keeping all other factors constant.

We included two manipulations in this experiment in order to make the comparison of the two eye tracking systems as broad as possible. First, we included both sentences and dot strings in our stimuli. Previous work has found that when adults scan along rows of dot strings, eliciting comparable patterns of fixations and saccades to those typically observed during reading, fixation disparity is highly similar to that observed during reading (Kirkby et al., 2010). In the present experiment we made a direct comparison of binocular coordination on the two tasks, as measured by the two different eye tracking systems. We predicted that if different eye trackers and associated analysis software caused differences in disparity alignments, then we would obtain crossed disparities in the EyeLink eye tracker data

[1] We used the latest generation of the EyeLink 1000 system, which is able to take samples at 2000 Hz monocularly and 1000 Hz binocularly. Because of the 2000 Hz maximum sample rate, this eye-tracker is sometimes referred to as EyeLink 2000, however, the manufacturer's name is EyeLink 1000.

and uncrossed disparities in the DPI eye tracker data. We also examined landing positions on the dot strings. Landing positions are monocular measures relating to saccadic targeting that are commonly reported in both reading and non-reading eye movement studies, and again, we wished to examine whether landing positions were similar in data from the two eye tracking systems. Second, we included a target word in every sentence that was manipulated for frequency. A hallmark effect of cognitive control of eye movements during reading is that reading times are longer on low frequency words than on high frequency words (see Rayner, 1998; 2009); again, we were keen to establish unambiguously that the magnitude of any such effect would be similar in the data from the two eye tracking systems.

Methods

Participants

Twelve adult participants took part in the experiment; all were students at the University of Southampton. All had English as their first language and had normal, uncorrected vision. Participants either earned course credits as partial fulfilment of course requirements or were paid £6 per hour in cash for volunteering to take part.

Apparatus

Dual Purkinje Image laboratory. Two Dual Purkinje Image (DPI) eye trackers were used, recording the positions of the two eyes simultaneously every millisecond. A Pentium 4 computer was interfaced with the eye trackers and all experimental stimuli were presented on a Philips 21B582BH 20" monitor set at a viewing distance of 100 cm. Participants were required to bite on a sterilised bite bar covered in dental wax, and to lean forward onto two forehead rests to minimise head movements. Low stimuli luminance (white text on a black background) and a dark experimental room were necessary for accurate eye movement recordings, as is

standard for DPI eye tracking experiments. DPI eye trackers have an extremely high spatial resolution (<0.1°).

EyeLink 1000 laboratory. An EyeLink 1000 eye tracker (SR Research) was used to monitor participants' binocular eye movements. A Dell Precision computer was interfaced with the eye tracker, and all experimental stimuli were presented on a ViewSonic P227F 20" monitor set at a viewing distance of 100 cm. The movements of each eye were monitored every millisecond. Participants were required to place their chin on a chin rest and lean forward onto a forehead rest to minimise head movements. The EyeLink 1000 has a spatial resolution of <0.5°. To make testing conditions as comparable as possible with the DPI laboratory, the EyeLink laboratory was also kept dark during testing sessions although this is not a requirement of the system. We chose to match viewing conditions by mimicking those typically applied in the DPI laboratory for the trials run in the EyeLink laboratory. An alternative setup in which we would have mimicked typical viewing conditions of the EyeLink laboratory in the DPI laboratory would have been close to impossible. For example, black text presented on a white background (most often used in EyeLink experiments) would have been accompanied by increased brightness and resulted in pupil shrinkage, thereby making it much more difficult to track the purkinje reflections of each eye.

Materials.

Two eye tracking tasks were employed during the study; a reading and a dot scanning task.

Reading task. In the reading task, 40 experimental sentences were constructed, each of which contained a target word. The sentence frames were constructed such that the target word could either be a high or a low frequency six-letter word. The stimuli were all single line sentences with simple syntactic structures designed to ensure comprehension. The sentences were presented in white, Courier

New font size 14, on a black background. Each character space extended 0.17˚ of visual angle. The low-frequency target words had a mean frequency of 3.5 counts per million (range: 1 to 10 per million) and the mean frequency for high frequency target words was 147 counts per million (range: 71 to 492 per million); this difference in frequency was highly significant (t (39) = 10.27, $p < 0.001$). All target word frequencies were taken from Francis and Kuĉera (1982) but the differences in frequency were also significant according to the norms collected in the HAL corpus (Burgess & Livesay, 1998) and in the SUBTL corpus (Brysbaert & New, 2009). Examples of the experimental sentence stimuli with the frequency manipulation are given in Table 1.

Table 1. Example of an experimental sentence frame with a high and a low frequency target word embedded in the sentence; target words are presented in italics (although they did not appear in italics during the experiment).

Condition	Sentence
Low-frequency target word	He didn't master the *jargon* until well into his last year.
High-frequency target word	He didn't master the *theory* until well into his last year.

In addition to the experimental sentences 10 practice sentences were constructed, five of which included a high-frequency target word and five a low-frequency target word; the practise sentences were constructed in a similar way to the experimental stimuli. Five practice sentences were presented at the start of each session (DPI and EyeLink). After 15% of the sentences, a comprehension question requiring a yes/no response was presented; these were distributed randomly throughout the experimental session.

Dot scanning task. Five horizontal arrays of white dot strings on a black background were presented (see Figure 2). Each array consisted of six target strings; each target string consisted of six dots. Each dot had a diameter of 0.29˚ of visual angle. All targets in the row were presented simultaneously and remained visible throughout the trial.

Figure 2. *Non-linguistic stimuli for the dot string scanning task; each target string (six dots) covered 1.74˚ of visual angle.*

Calibration

Calibration procedures were similar during both the eye tracking sessions (DPI and EyeLink). Left and right eye calibrations were performed monocularly (*e.g.*, when calibrating the left eye the right was manually occluded and vice versa; Liversedge, White et al., 2006). During calibration the participant was instructed to look at each of three fixation points presented horizontally to the left, centre and right of the screen. Monocular eye positions were recorded for each of these fixation points and then checked for accuracy. This was then repeated for the other eye and again checked for accuracy.

In the Eyelink system during calibration the initial fixation position was accepted by the experimenter when the pupil appeared stable; the remaining fixation positions were automatically recorded by the calibration system when a stable fixation was detected. The calibration procedure for the DPI system requires that calibration fixations are accepted manually by the experimenter when the eye is considered

stable. Both systems require validation following the initial calibration during which calibration data are used to provide an indication of the participant's gaze position, relative to the calibration matrix. In the Eyelink system the validation procedure was essentially identical to the initial calibration procedure. Based on the two data sets (initial calibration and validation), the discrepancy between the two is computed for each point. An error of < 0.2° was accepted as an accurate calibration for each target and recalibration was performed if the validation error was > 0.2°. For the DPI system validation was evaluated visually by the experimenter. Based on the initial calibration the system provided a dot at the point of fixation and the participant was required to refixate and maintain fixation on each of the calibration points in turn. In this way, the experimenter was able to evaluate the discrepancy between where the participant was fixating and where the system recorded that they were looking. The calibration fixation points extended .29° and the amount of fixation error accepted was estimated as .14°. For both systems, if the fixation error was greater than the limits described, the calibration procedure was repeated. These calibration and validation procedures are standard, and were the same as those used in the experiments reported in the papers that are central to the current research questions. It was for this reason that we strictly adhered to these procedures in the current study. When a successful calibration was completed the experimental stimuli were presented. Following every trial during the experiment the calibration accuracy was verified and at that point recalibration was carried out if necessary.

Design and Procedure

Eye movement data from all participants were collected in two separate laboratories, the DPI and EyeLink laboratory, within the School of Psychology at the University of Southampton. The eye movement data were acquired in two testing sessions that took place consecutively (i.e., on the same day, with one testing session straight after the other with only a 10 min break). The experimental procedures were

identical during the eye tracking sessions using the EyeLink 1000 and the DPI eye trackers.

We adopted a within-subjects, repeated-measures Latin square design with two independent variables, word frequency (high vs. low) and laboratory (EyeLink vs. DPI). Participants read 20 sentences (plus practice sentences) in the EyeLink laboratory, and 20 sentences (plus practice sentences) in the DPI laboratory. Half the participants read half of the sentences containing the high frequency word with the remaining sentences containing the low frequency word. The remaining participants read the sentences containing the counterpart target word.

During the reading task, participants were instructed to read the sentences normally for comprehension and answer the comprehension questions that were presented periodically as accurately as possible. In the dot scanning task the participants were required to scan the dot strings from left to right until they reached the last string. Participants were instructed to fixate each dot string as a whole, and to move their eyes from one string to the next in time with the beat of a metronome (set at 60 beats per minute). All trials (reading and dot scanning tasks) were self-terminated by a button press.

Analysis

All data were analysed using in-house software (we took the raw, horizontal position output from the EyeLink and converted it from pixels to degrees using custom-built software). Using the streams of raw data, fixations and saccades were manually identified (see Figure 3) in order to avoid contamination by dynamic overshoots (Liversedge, White et al., 2006; Deubel & Bridgeman, 1995).

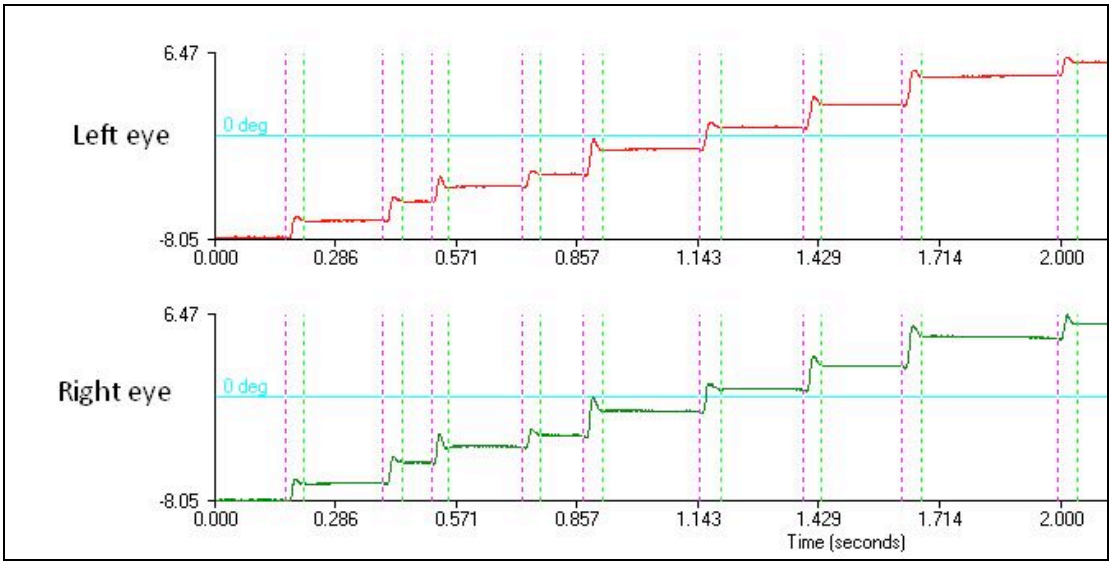


Figure 3. *Manual demarcation of saccades and fixations in the binocular data stream, where all dynamic overshoot is excluded from the fixation period (vertical axis represents horizontal eye position in degrees of visual angle, horizontal axis represents time in seconds).*

To calculate fixation disparity, the horizontal position of the right eye was subtracted from that of the left eye at both the start and the end of fixations. As per Liversedge, White et al. (2006) fixations were categorized as aligned or unaligned; aligned fixations were all those fixations where the points of the two eyes were within one character space (or dot space) of each other (0.17° during reading and 0.29° during dot scanning). Unaligned fixations, where the eyes were more than one character space apart, were further categorised as being uncrossed or crossed. Crossed fixations were those where the left eye's point of fixation was more than one character space to the right of the right eye's point of fixation. Conversely, uncrossed fixations were those where the left eye's point of fixation was more than one character space to the left of the right eye's point of fixation.

Data removed prior to analyses included fixations less than 80 ms or more than 1200 ms, and fixations where the magnitude of absolute disparity measured more than 2 standard deviations from the mean for the individual participant (5.1% of the

data; see Blythe et al., 2006; Kirkby et al., 2010; Liversedge, White et al., 2006). The final data set consisted of 4811 fixations.

Results

Throughout the Results section we have conducted Analyses of Variance and *t*-tests considering participants (F_1, t_1) and items (F_2, t_2) as random variables (Clark, 1973). On the comprehension questions, the participants' mean accuracy was 88% correct.

Sentence reading

Monocular - global measures. First we report the mean fixation durations, saccade lengths, and regression frequencies observed during sentence reading (presented in Table 2). These data were analysed using paired-samples *t*-tests. Significant differences were found between the data from the DPI eye trackers and the data from the EyeLink tracker: when reading sentences in the EyeLink laboratory, participants had longer fixation durations, made larger saccades, had longer sentence reading times, and made more fixations and regressions per sentence than when reading in the DPI laboratory. As these differences were not of primary interest in this paper, we defer discussion of them until the General Discussion.

Table 2. *Mean fixation duration, saccade amplitude, total sentence reading time, number of fixations and regression frequency for the data collected during the DPI and EyeLink eye tracking sessions.*

	DPI	EyeLink	t_1 (df)	t_2 (df)
Fixation duration				
Mean	217 ms	245 ms	3.10 (11)**	6.44 (39)***

SD	89 ms	129 ms		
Progressive saccade amplitude				
Mean	1.63°	1.95°	5.64 (11)***	8.87 (39)***
SD	0.83°	1.11°		
Total sentence reading time				
Mean	2880 ms	3687 ms	4.04 (11)**	4.34 (39)***
SD	1068 ms	1576 ms		
Number of fixations per sentence				
Mean	10.03	11.75	2.62 (11)*	3.12 (39)**
SD	3.30	4.55		
Regression frequency				
Mean	19%	23%	3.09 (11)**	2.72 (39)**
SD	12%	13%		

Note. *$p<0.05$. **$p\leq0.01$. ***$p\leq0.001$

Binocular - global measures. The primary question under investigation was whether measures of binocular disparity were similar for data collected with the DPI eye trackers and data collected with the EyeLink eye tracker when all other experimental conditions were held constant between the two laboratories. The distributions of disparities that we observed from each of the two eye tracking systems at the starts and ends of fixations are shown in Figure 4.

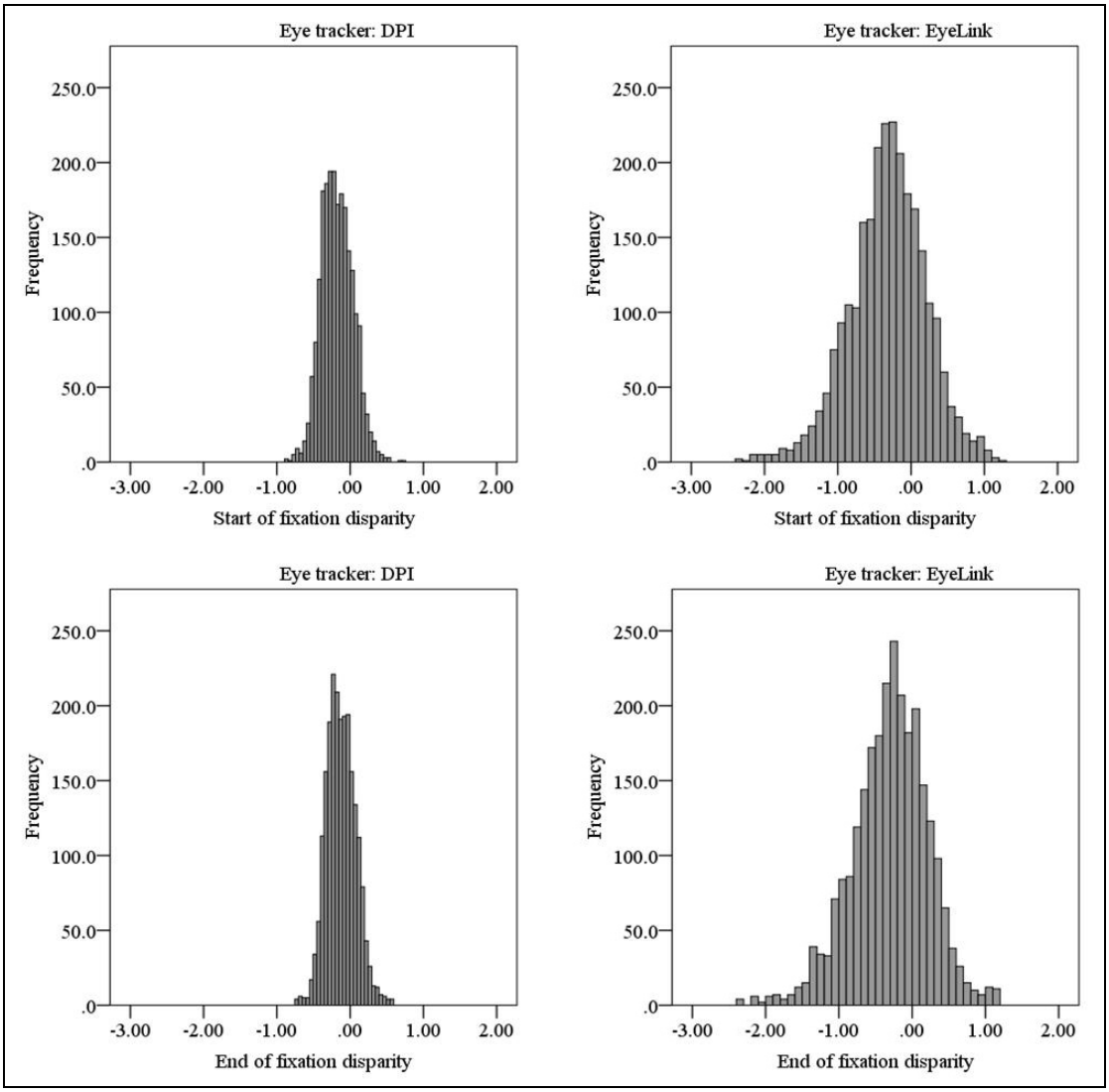


Figure 4. *Distributions of disparities observed at the start (top panels) and end (bottom panels) of fixations. Fixation disparities are given in degrees of visual angle, and were calculated as the difference between the left and right eye positions. Positive values correspond to crossed disparities, where the left eye was fixating to the right of the right eye. Negative values correspond to uncrossed disparities, where the left eye was fixating to the left of the right eye. The left panels show data collected using the DPI eye trackers, and the right panels show data collected using the EyeLink eye tracker.*

For statistical analysis of these data, we examined both the absolute magnitude and the direction (crossed or uncrossed) of disparity measured at the starts and ends of

fixations in the two laboratories; these data were analysed using 2 (eye tracker: DPI vs. EyeLink) x 2 (sample point: start vs. end of fixation) repeated-measures ANOVAs. The magnitudes of absolute fixation disparity data are shown in Figure 5. We found that, at fixation onset, the mean magnitude of fixation disparity in data from the DPI trackers was 0.23° (SD 0.16°). Since each character space extended 0.17° this meant that, on average, the two eyes were more than one character space apart. This finding is entirely consistent with the literature on binocular eye movements during reading and non-reading tasks from DPI eye trackers (Blythe et al., 2006, Juhasz et al., 2006; Kirkby et al., 2010; Liversedge, White et al., 2006; see Kirkby et al., 2008, for a review).

When comparing the two eye tracking systems, we found the mean magnitude of disparity in data from the EyeLink tracker to be significantly greater than that in data from the DPI trackers (0.48°, SD 0.38°; F_1 (1, 11) = 13.55, $p < 0.01$; F_2 (1, 39) = 268.66, $p < 0.001$). Additionally, as can be seen in both Figures 4 and 5, the fixation disparity data from the EyeLink were more broadly distributed around the mean than the data from the DPIs (standard deviations of 0.38° and 0.16°, respectively for the start of fixation data, and 0.38° and 0.13°, respectively for the end of fixation data). The main effect of sample point was also significant – the absolute magnitude of fixation disparity decreased from the start to the end of fixations (F_1 (1, 11) = 6.46, $p = 0.03$; F_2 (1, 39) = 133.36, $p < 0.001$). This effect is also consistent with the literature showing that, during fixation, vergence movements occur that reduce the magnitude of fixation disparity (Blythe et al., 2006; Jainta et al., 2010; Kirkby et al., 2010; Liversedge, White et al., 2006; Nuthmann & Kliegl, 2009; see Kirkby et al., 2008, for a review).

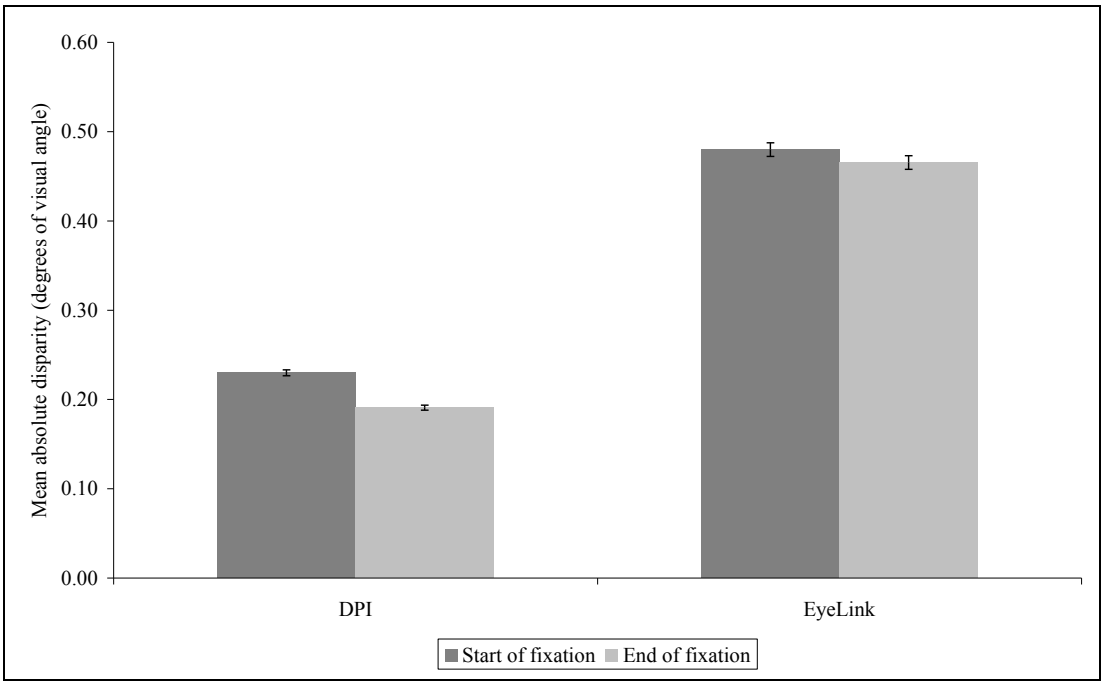


Figure 5. *Mean absolute disparity magnitude at the start and end of fixation for the data collected using the DPI and EyeLink systems. Error bars show the standard error of the mean.*

The interaction between eye tracker and sample point was also significant (F_1 (1, 11) = 11.06, p = 0.007; F_2 (1, 39) = 17.66, p < 0.001). The reduction in the absolute magnitude of disparity from the start to the end of fixation was only significant in the data collected with the DPI trackers (t_1 (11) = 3.66, p = 0.004; t_2 (39) = 11.19, p < 0.001). While there was a numerical reduction in the magnitude of disparity by the ends of fixations in data collected with the EyeLink tracker, this was not significant (t_1 (11) = 0.76, p = 0.46; t_2 (39) = 2.78, p = 0.01). In the Appendix we provide a simulation and discussion of why this difference between the EyeLink and the DPI is not necessarily due to differences in vergence movements made by the participants (and instead due to differences in the accuracy of the eye trackers). In summary, the analyses showed that data collected with the DPI eye trackers contain smaller overall magnitudes of fixation disparity compared to the EyeLink, as well as

smaller standard deviations. We also found a greater difference between disparity at the starts and ends of fixations in the DPI data set compared to the EyeLink data set.

Next we examined the alignment characteristics of the two eyes during fixation. This was of primary interest as contrasting results have been reported in the literature with respect to whether the majority of unaligned fixations are crossed or uncrossed. The mean proportions of aligned, crossed and uncrossed fixations at the start and end of fixations are presented in Figure 6. As can be seen in Figure 6, the overall pattern of alignment was highly similar in the data obtained from both the DPI and the EyeLink eye trackers; furthermore, although the intra-individual variability was larger in the Eyelink compared to the DPI data, the pattern of alignment was highly similar in the data for each participant, (see Figure 7). We compared the proportion of aligned fixations using a 2 (eye tracker: DPI vs. EyeLink) x 2 (sample point: start vs. end of fixation) repeated-measures ANOVA. Congruent with our analysis of the magnitude of fixation disparity, we found that the proportion of aligned fixations was greater in the data set from the DPI trackers than that found in the data set from the EyeLink tracker (F_1 (1, 11) = 10.47, p = 0.01; F_2 (1, 39) = 86.90, $p < 0.001$). There was a significant increase in the proportion of aligned fixations from the start to the end of fixations (F_1 (1, 11) = 23.21, p = 0.001; F_2 (1, 39) = 38.29, $p < 0.001$). There was also a significant interaction between the eye tracker and the sample point (F_1 (1, 11) = 14.19, p = 0.003; F_2 (1, 39) = 18.34, $p < 0$ .001). The proportion of aligned fixations increased significantly between the start (43%) and the end (51%) of the fixation in the data set collected with the DPI trackers (t_1 (11) = 4.99, $p < 0.001$; t_2 (39) = 7.46, $p < 0.001$). However, there was no significant increase in the proportion of aligned fixations by the end of the fixation (25%) compared to the start of fixation (24%) in the data set collected with the EyeLink tracker (t_1 (11) = 0.90, p = 0.39; t_2 (39) = 0.73, p = 0.47; again see the Appendix for a simulation relating to this effect).

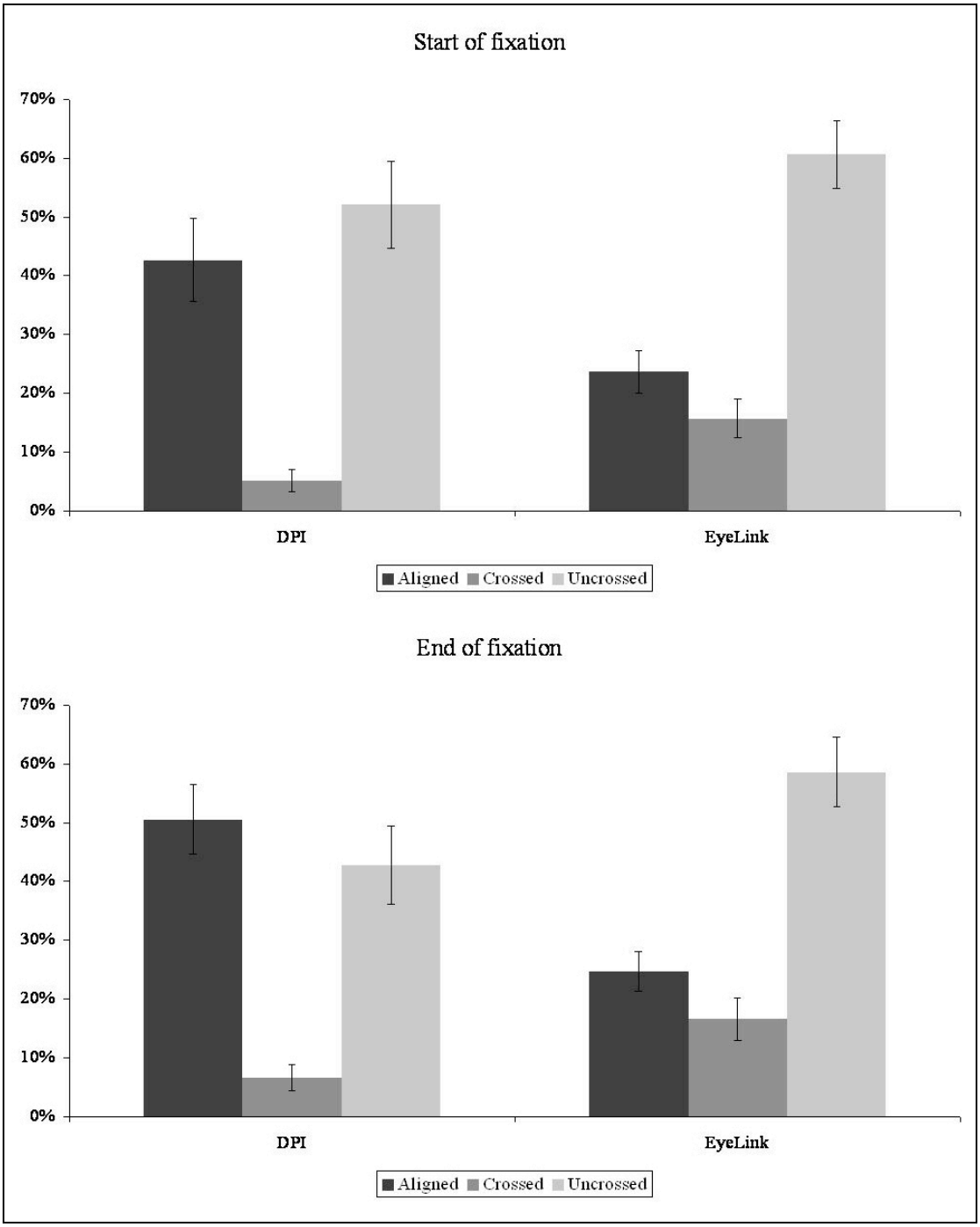


Figure 6. *Mean fixation alignment proportions. The top panel shows data from the start of fixations, while the bottom panel shows data from the end of fixations. Aligned fixations are those where the two eyes' points of fixation were within one character space of each other. Uncrossed fixations are those where the left eye's point of fixation was more than one character space to the left of the right eye's point of fixation. Crossed fixations are those where the left eye's point of fixation was more*

than one character space to the right of the right eye's point of fixation. Error bars show the standard error of the mean.

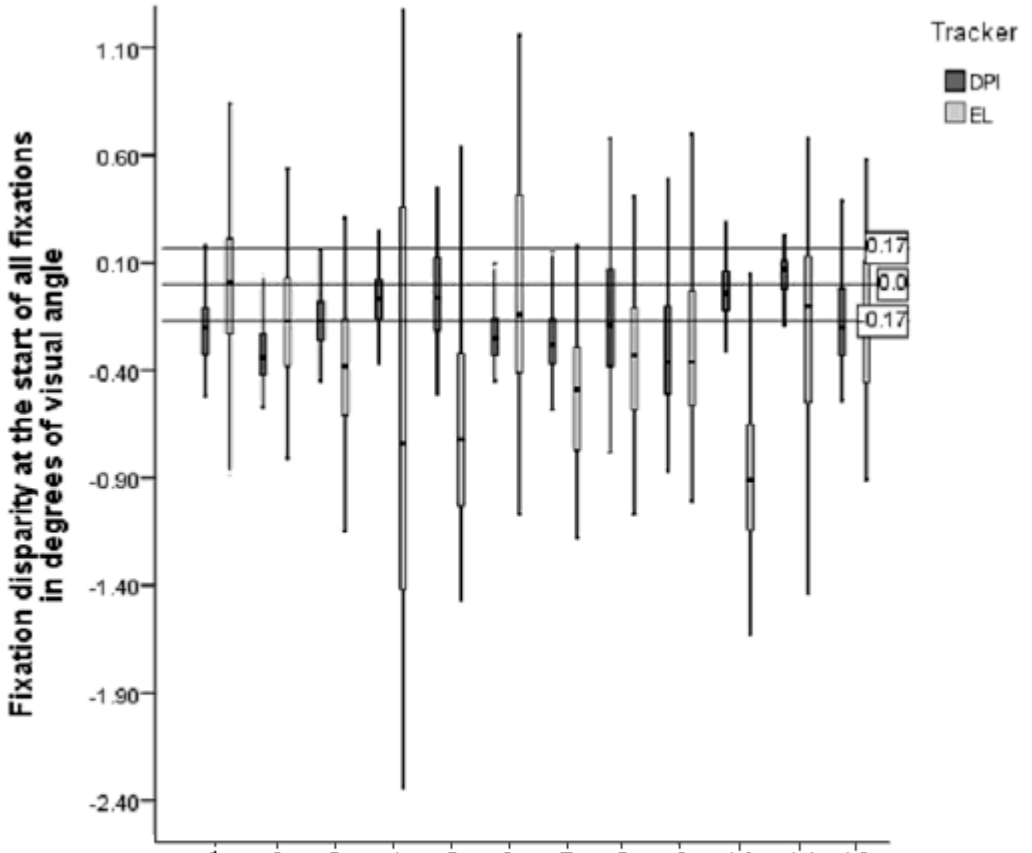


Figure 7. *Boxplots of fixation disparities observed at the start of fixation for each participant in both the DPI and Eyelink systems; disparity is reported in degrees of visual angle, positive values correspond to crossed disparities, and negative values correspond to uncrossed disparities. There are two different categorizations shown:. In the figure, the two lines at +/- 0.17° represent a maximum disparity of half a character in either direction and fixations were considered aligned when the observed disparity fell within that range. In the table below the figure each participant is categorized for each eye tracking system in relation to 0° of disparity. "A"*

represents a perfectly aligned fixation, "U" represents an uncrossed fixation, and "C" represents a crossed fixation.

Of principal concern in the present experiment was the direction of disparity within the unaligned fixations (crossed vs. uncrossed). As can clearly be seen in Figure 6, more uncrossed fixations were observed than crossed fixations in both the DPI and the EyeLink data sets. To our knowledge, this is the first observation of a higher proportion of uncrossed than crossed unaligned fixations in a binocular data set from an EyeLink eye tracker. While the higher proportion of uncrossed fixations is typical for a binocular data set from DPI eye trackers (Blythe et al., 2006; Juhasz et al., 2006; Kirkby et al., 2010; Liversedge, White et al., 2006), to our knowledge, all previous studies that have collected binocular eye movement data using an EyeLink eye tracker have observed more crossed than uncrossed fixations.

Within the fixations categorized as unaligned, we compared the proportion of crossed fixations using a 2 (eye tracker: DPI vs. EyeLink) x 2 (sample point: start vs. end of fixation) repeated-measures ANOVA. The difference in the proportion of crossed fixations between the data sets collected from the two different eye trackers was significant by items but not by participants (F_1 (1, 11) < 0.001, p = 0.99; F_2 (1, 39) = 5.86, p = 0.02). The important point to note is that while there is a trend for a higher proportion of crossed fixations in the EyeLink data set compared to the DPI data set, both data sets clearly contain a much higher proportion of uncrossed than crossed fixations. The proportion of crossed fixations at fixation offset (DPI: 13%; EyeLink: 22%) was larger than at fixation onset (DPI: 9%; EyeLink 21%; F_1 (1, 11) = 16.67, p = 0.001; F_2 (1, 39) = 15.91, p < 0.001). Given that this analysis was based only on those fixations categorized as unaligned, these data indicate, therefore, that the proportion of uncrossed fixations must have decreased from the start to the end of fixations (because if the proportion of crossed fixations within those categorised as

unaligned increased, then the proportion of uncrossed fixations in that category must necessarily have decreased). This result suggests that there is a stronger tendency to converge the eyes during a fixation than to make a divergent eye movement, consistent with previously reported binocular data sets (Blythe et al., 2006; Kirkby et al., 2010; Liversedge, White et al., 2006). The interaction between eye tracker and sample point for the proportion of crossed fixations was significant across items and approached significance in the participants analysis (F_1 (1, 11) = 3.43, p = 0.09; F_2 (1, 39) = 7.92, p = 0.01). Most importantly for the current study, patterns of binocular alignment were found to be highly similar in the data recorded with both the DPI and the EyeLink eye trackers under very similar experimental conditions in the two laboratories.

We also examined whether the software used for analysis might lead to differences in the reported fixation disparities. We processed the data from the EyeLink again, this time using the standard, commercially-available software (DataViewer; SR Research Ltd.). We compared the same data set when processing the samples containing eye location for each millisecond using DataViewer to when it had been processed using our "home developed" custom software. For each fixation, as described in Section 2.6, our custom software allowed us to examine disparity at sample points at both the start and at the end of the fixation. In contrast, DataViewer provides a single disparity value for each fixation, which is the average of all samples during that fixation. When processed using DataViewer, the data set was found to contain a mean absolute fixation disparity of 0.65°, significantly more than the mean start of fixation disparity (0.48°; t_1 (11) = 2.36, p = 0.04; t_2 (39) = 14.88, $p < 0.001$) or end of fixation disparity (0.47°; t_1 (11) = 2.31, p = 0.04; t_2 (39) = 16.97, $p < 0.001$) values that were found from the data being processed through our custom software.

We also examined the direction of fixation disparity. When processed using DataViewer, we observed that 19% of fixations were aligned, 18% were crossed, and

62% were uncrossed. As can be seen from comparisons with the EyeLink data in Figure 6, the overall pattern is highly similar. Thus, these data clearly show that, within fixations classed as unaligned, the majority were uncrossed in our EyeLink data set and this was not a consequence of the software that was used to process the data.

There are at least two possible causes of the discrepancy in absolute magnitude of fixation disparity as calculated by the two sets of software. First, as described above, DataViewer generates an average disparity value based on all the pairs of sample values during each fixation, whereas the custom software delivers individual pairs of start and end of fixation sample points. If it were the case that the disparity between the eyes varied substantially during a fixation, and that variability resulted in an average increase in disparity over the duration of the entire fixation, then the resulting average based on all pairs of samples through the fixation could be larger than that based on pairs of values at the start and end of fixation. Here we have in mind either the possibility that the eyes might make gradual divergent and corresponding convergent movements during a fixation, or alternatively, there is the possibility that brief but quite large microsaccades could occur in one or both eyes during a fixation (Engbert & Kliegl, 2004).

In order to test whether the discrepancy in absolute disparity magnitudes could have occurred due to differences in the manner of computation, we took the continuous data stream from our custom software and, for a subset of fixations, calculated the average disparity between the designated start and end of fixation points based on all the pairs of sample values in between[2]. We then conducted a

[2] Accessing individual samples to compute average disparity per fixation was labour intensive, and for this reason we made these computations based on the data from 138 fixations (rather than the whole data set). Despite this, this subset of data provided more than adequate statistical power for our analyses (far more than the other analyses we report based on subjects and items).

within-fixation analysis, to determine whether the averaging process generated a higher estimate of fixation disparity than that using only fixation start and end sample points. We found no significant difference between the averaged value and either the value computed on the basis of start sample points (t (345) = 0.70, p = 0.48) or that computed on the basis of end sample points (t (345) = 1.16, p = 0.25).

A second possible cause of the difference in disparity from the two sets of software is that DataViewer uses an algorithm to determine the locations of saccade onsets and offsets while in our custom software this is done manually. An example of the differences that arise between these two procedures is shown in Figure 8.

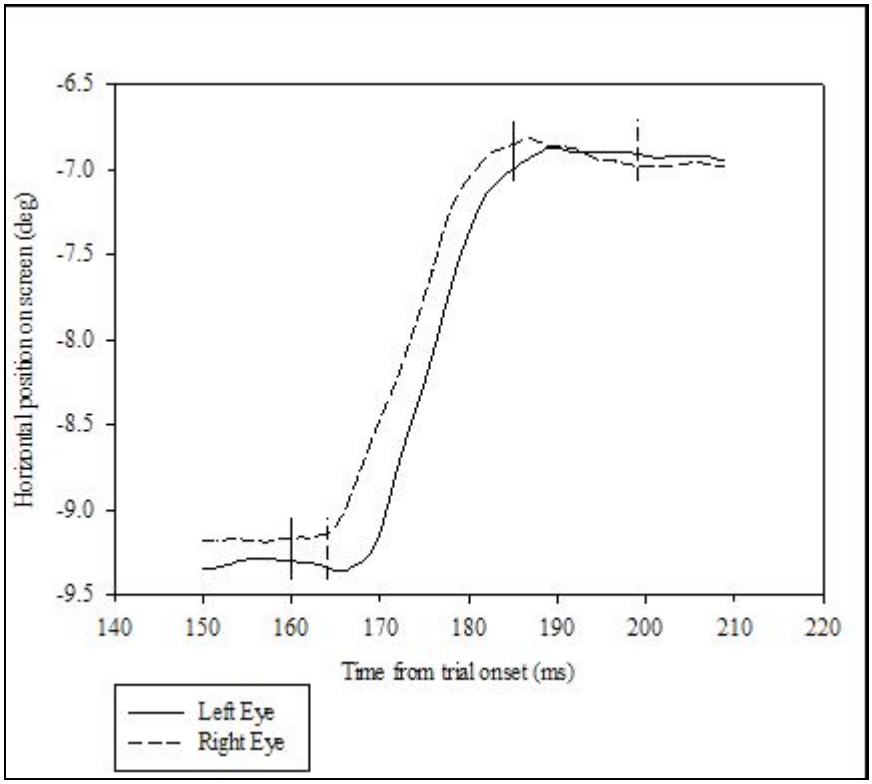


Figure 8. *An example saccade, showing the relative positions of saccade onsets and offsets as selected (a) by the algorithm in DataViewer (solid vertical lines) and (b) by hand, in custom-designed software (dashed vertical lines).*

Once again, mapping the hand-selected and algorithm-selected fixation start and end points onto the same data sets is an extremely time consuming process that must be carried out by hand. However, we did formally examine a small proportion

of fixations from our data set, as well as visually inspecting numerous eye movement records segmented according to the two systems, and it is clearly the case that while the two methods generally produced similar temporal locations for saccade onsets, there was a tendency for manual selection to be more conservative during segmentation, such that the end of the saccade was marked as occurring later than was the case when the point is selected by the algorithm. Thus, this seems to be the likely cause of the difference that we observed. Note, also, that the disparity reduces to its minimum point within the initial 100 ms or so of a fixation onset, on average (Jainta et al., 2010). The later the saccade offset is marked, the further into this initial portion of the fixation it will be, during which time disparity is being reduced. It seems plausible, therefore, that more conservative, manual selection of saccade offsets might result in smaller calculations of fixation disparity.

These analyses show that: (1) larger fixation disparities were calculated from the same data set when it is processed through DataViewer than through our custom software; (2) this difference is unlikely to be due to DataViewer's averaging process for calculating disparity during fixations; and (3) this difference is likely to result from earlier demarcation of the saccade offset when determined by the DataViewer algorithm as opposed to being determined by manual selection.

Binocular - local measures. Recall that we included a critical word in the sentences that was either high or low frequency. Previous research has found no effects of word frequency on binocular coordination during reading (Juhasz et al., 2006). However, the inclusion of these target words allowed us to compare the magnitude of word frequency effects – an extremely common manipulation in psycholinguistic experiments – as measured by DPI and EyeLink eye trackers. These data are summarised in Figure 9.

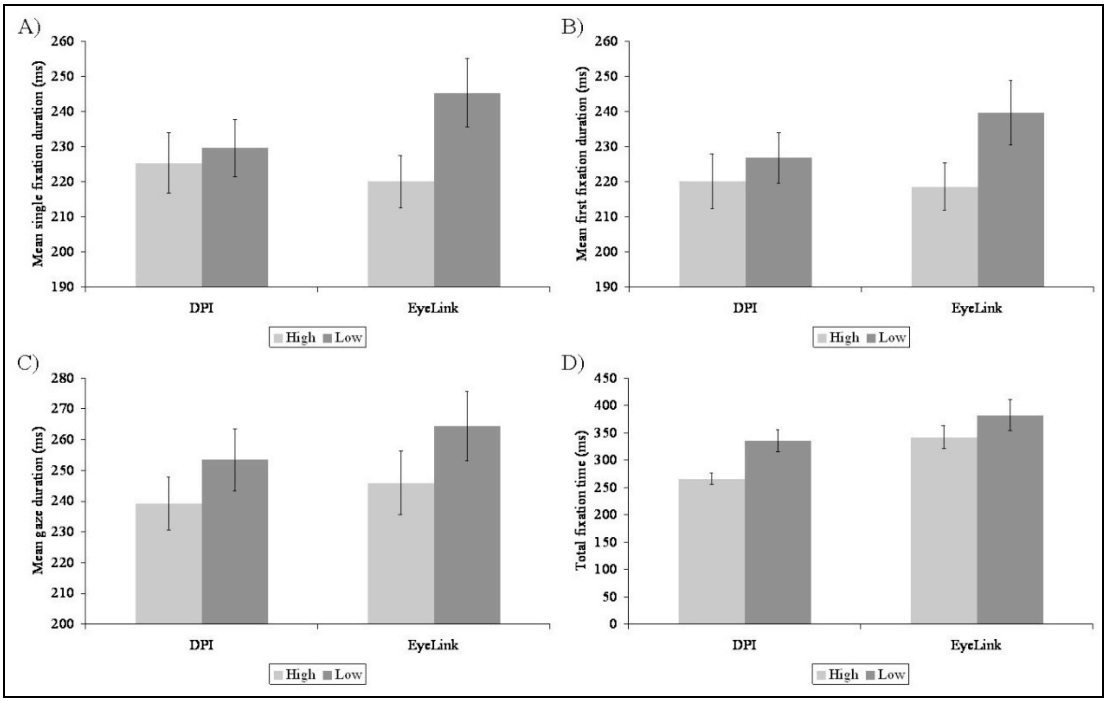


Figure 9. *Means and standard deviations for the analyses of target word frequency. Panel A shows single fixation duration data, Panel B shows first fixation duration data, Panel C shows gaze duration data, and Panel D shows total fixation time data. Error bars show standard errors of the mean.*

Consistent with prior research (see Rayner, 1998, 2009), reading times were longer on low frequency words than on high frequency words. This effect was significant by participants but not by items for single fixation durations (F_1 (1, 11) = 11.29, p = 0.01; F_2 (1, 39) = 1.20, p = 0.28) and gaze durations (F_1 (1, 11) = 5.40, p = 0.04; (F_2 (1, 39) = 2.79, p = 0.10), and was significant across both participants and items for first fixation durations (F_1 (1, 11) = 5.37, p = 0.04; F_2 (1, 39) = 4.70, p = 0.04) and for total fixation times (F_1 (1, 11) = 11.80, p = 0.006; F_2 (1, 39) = 5.03, p = 0.03). There was no overall difference in reading times between the data sets from the EyeLink and the DPI eye trackers (all Fs < 4, all ps > 0.05). Furthermore, there were no significant interactions between eye tracker and word frequency (all Fs < 3, all ps > 0.1). These data show very clearly that word frequency effects are highly similar in data sets from both EyeLink and DPI eye trackers.

Dot scanning

Binocular coordination. The inclusion of dot strings in this experiment allowed us to make a direct comparison of binocular coordination during reading and dot scanning as measured by both DPI and EyeLink eye trackers. These data are summarised in Table 3.

Table 3. Means and standard deviations for absolute disparity magnitudes. Data are presented from both the starts and ends of fixations, from both the DPI and the EyeLink data set, comparing data from the sentence reading and dot scanning tasks. Standard deviations are given in parentheses.

	Sentence reading	Dot scanning
Start of fixation		
DPI	0.23 (0.16)	0.16 (0.14)
EyeLink	0.48 (0.38)	0.54 (0.42)
End of fixation		
DPI	0.19 (0.13)	0.17 (0.14)
EyeLink	0.47 (0.38)	0.54 (0.43)

These data were analysed with 2 (eye tracker: DPI vs. EyeLink) x 2 (task: sentence reading vs. dot scanning) repeated-measures ANOVAs. For both start and end of fixation data, there was a significant effect of eye tracker (F_1 (1, 11) = 32.67, $p < 0.001$ and F_1 (1, 11) = 34.73, $p < 0.001$, respectively). Absolute disparity magnitudes were found to be larger in the EyeLink data set than in the DPI data set. There were no significant effects of task (F_1 (1, 11) = 0.33, $p = 0.58$ and F_1 (1, 11) = 1.11, $p = 0.31$, for start and end of fixation respectively), nor were the interactions between eye tracker and task significant (F_1 (1, 11) = 2.24, $p = 0.16$ and F_1 (1, 11) =

1.22, $p = 0.29$, for start and end of fixation respectively). Thus, these data show binocular coordination, at least in terms of disparity measures, to be highly similar when reading and when scanning along rows of dots, a finding consistent with Kirkby et al. (2010).

Landing position distributions. We also examined landing position distributions on the dot strings, to examine whether both DPI and EyeLink eye trackers generated similar data sets in terms of another commonly reported measure. These data are shown in Figure 10.

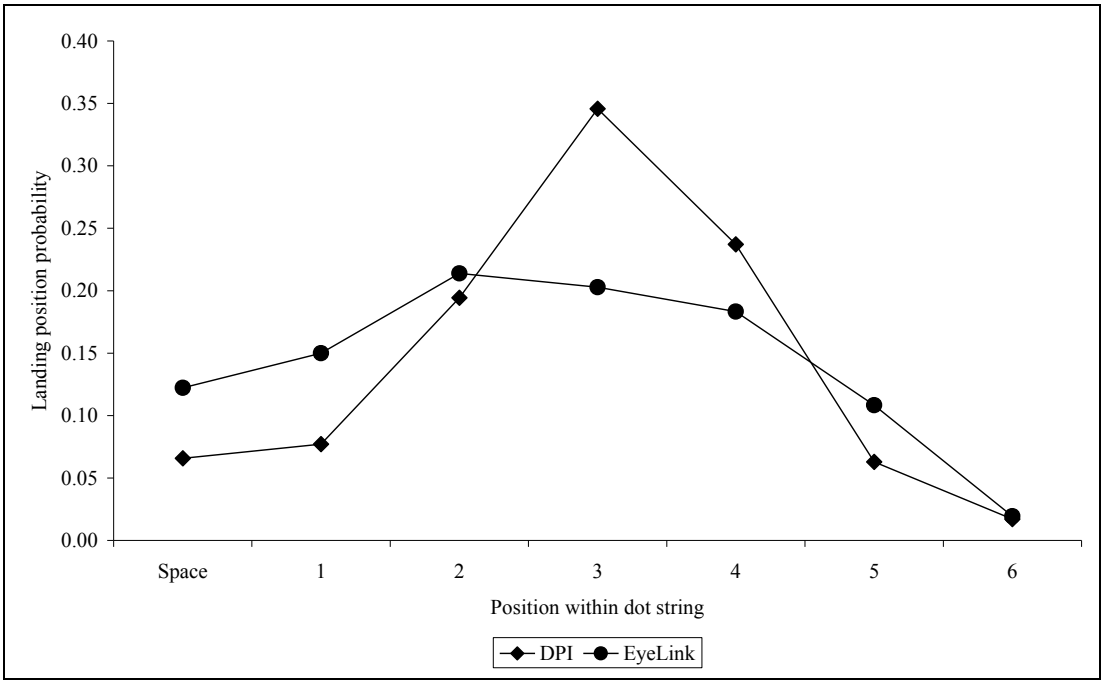


Figure 10. *Landing position distributions on the dot strings from the DPI and EyeLink data sets. Values on the x-axis refer to individual dots within the strings.*

Landing positions were measured in terms of the dot within the string of six dots that the right eye landed on following the initial saccade onto the string. The mean landing position in the DPI data set was 2.87, and the mean landing position in the EyeLink data set was 2.57 – participants were measured to be landing slightly further to the right within the dot strings in the DPI data set (t_1 (11) = 2.30, $p = 0.04$).

Discussion

The aim of this investigation was to compare data collected using two different eye tracking devices in order to inform the ongoing debate concerning variability in the proportion of crossed and uncrossed fixations that are observed in binocular eye movement experiments investigating reading. To summarise, the following pattern of effects was observed: (1) within the fixations classed as unaligned, both DPI and EyeLink data sets contained a higher proportion of uncrossed than crossed fixations; (2) the pattern of alignment was not affected by the software used to process the data; (3) greater magnitudes of fixation disparity were found in the EyeLink data set compared to the DPI data set; (4) greater reductions in disparity through vergence movements were observed during fixations in the DPI data set compared to the EyeLink data set; (5) a broader range of fixation disparity was observed in the EyeLink data set compared to that observed in the DPI data set; (6) no influence of task demands (reading vs. dot scanning) was found on the basic characteristics of binocular coordination; (7) equal word frequency effects were observed in both the DPI and EyeLink data sets; (8) landing positions on the dot strings were found to be slightly further to the right in the DPI data set compared to the EyeLink data set.

We were primarily interested in whether the apparent inconsistency in the direction of fixation disparity reported in previous research might be attributable to the particular eye tracking system used to acquire the data. In both data sets, DPI and EyeLink, we found the majority of unaligned fixations to be uncrossed, with a fairly small minority of crossed fixations. We also found that corrective vergence movements were more often convergent than divergent during fixation (reducing fixation disparity). These data clearly demonstrate that, within identical experimental set-ups, both the DPI and the EyeLink eye trackers provide similar measures of

fixation disparity, and that in the current experimental study they showed disparity to be predominantly uncrossed. Thus, the conflicting results within the literature concerning the direction of disparity cannot be attributed to the different eye tracking systems or the software used in the different laboratories. Rather, it seems likely that binocular alignment is influenced by the specific experimental conditions, for example, the luminance of the room during data collection, viewing distance, font size, etc. (Kirkby et al., 2008; Nuthmann & Kliegl, 2009; Shillcock et al., 2010). Certainly, these data demonstrate that at 1 m viewing distance, with font size 14 white text presented on a black background, the majority of unaligned fixations were uncrossed.

Despite the similarity of the overall pattern of effects, we did find subtle differences between the two data sets in terms of the magnitude of disparity. The absolute magnitude of fixation disparity was significantly greater in the EyeLink data set than in the DPI data set and, therefore, fewer fixations were classed as aligned. Wyatt (2010) demonstrated that changes in pupil size can affect the accuracy of camera-based eye trackers. Drewes, Masson, and Montagnine (2012) demonstrated a method to compensate for this artefact by calibrating during constricted and dilated pupil conditions, using pupil size as an index to weight the calibrations. This could, potentially, be the source of increased fixation disparity in the Eyelink system. There was also a greater standard deviation in the fixation disparity data from the EyeLink compared to the DPI. These differences between the DPI and the EyeLink suggest that the EyeLink system is somewhat less sensitive to small changes in ocular alignment during fixation compared to the DPI system. We observed that the reduction in disparity magnitude during fixations as a consequence of vergence movements was only significant in the DPI data set; the lack of this effect in the EyeLink data set is related to its increased standard deviation (see Appendix). It is important to note that the increased standard deviations found in the EyeLink data,

which are possibly due to the comparatively higher noise level found in the Eyelink system compared to the DPI system (see Appendix Panel C), also have the potential to lead to more misclassifications of fixation alignment. This is because the calculation of the absolute value of fixation disparity can be influenced by the standard deviations of the distributions of the fixation position data for the left and the right eye from which it is calculated (as illustrated in the Appendix). Therefore, the noisier the initial fixation location data, the higher the chances that the deviations between the actual and the reported fixation locations will result in a misclassification of fixation alignment.

In monocular measures from the reading task, we found unexpected global differences between the DPI and the EyeLink data sets: when reading in the EyeLink laboratory, participants had longer fixation durations and sentence reading times, made larger amplitude saccades, and made more fixations and regressions per sentence than when reading in the DPI laboratory. Ordinarily, such effects might be taken to indicate a difference in processing difficulty between two experimental conditions – longer fixation durations and reading times generally indicate greater processing difficulty. This may not be the case here, however. What we can be sure of is that these differences did not occur due to differences in the experimental stimuli used in the two laboratories. Recall that identical stimuli were used in both laboratories (split between two counterbalanced lists) and that the same participants were tested in both laboratories. Furthermore, similar viewing and lighting conditions were used in both laboratories. In our view, the most plausible explanation of these effects is that they arose due to the somewhat less comfortable experimental conditions experienced in the DPI laboratory compared to the EyeLink laboratory. In the DPI laboratory, participants were required to bite on a sterilised bite bar covered in dental wax, to lean forward onto two forehead rests and to have Velcro straps secured behind their head.

In contrast, in the EyeLink laboratory participants simply placed their chin in a chin rest and their forehead against a restraint. Given the increased discomfort associated with testing in the DPI laboratory, the participants may have read the sentences more quickly in an effort to finish the testing session as quickly as possible. Importantly, the global measures in both data sets were well within typical ranges that are reported in the literature (see Rayner, 1998, 2009 for reviews), and, in both data sets, we found standard word frequency effects that did not vary as a result of the eye tracker with which the data had been collected. There was one other small unexpected difference between the DPI and the EyeLink data sets. We found a difference in the average landing site in the dot scanning task such that participants landed very slightly closer to the beginning of a string of dots when tested in the EyeLink than in the DPI tracker. We consider that this effect may well be spurious. In the dot scanning task there was no evidence of speeded sampling in the DPI laboratory relative to that in the EyeLink laboratory, presumably because a metronome was used as a guide as to when to make saccades in the dot scanning task, resulting in similar performance in the two laboratories, in terms of oculomotor timings, for this task.

Our comparison of sentence reading and dot scanning under identical experimental conditions allowed us to directly examine whether this task difference impacted on binocular coordination. Our results supported the conclusions of Juhasz et al. (2006), that processing difficulty (as indexed by the relatively high processing demands of reading compared to a non-linguistic task such as dot scanning) does not influence binocular coordination. This finding is also consistent with the data reported by Kirkby et al. (2010) in which the pattern of binocular coordination during a dot scanning task was found to be very similar to data from reading experiments.

Perhaps the most important theoretical conclusion that we can draw on the basis of this methodological paper is that the differences in the direction of binocular

disparity effects in reading that have been reported in the literature did not arise as a consequence of the device that was used to acquire those data, nor the software used to analyse them. This finding is somewhat reassuring for those working in the field of eye movement research. Furthermore, these results strongly suggest that the conflicting effects reported in the literature must have arisen due to reasons that are much more theoretically interesting. That is to say, the current study strongly suggests that factors such as the particular viewing conditions under which binocular eye movements are recorded and the nature of the visual stimuli (lighting conditions, font and background colour, viewing distance etc.) are all potential candidate causes for differences in the direction of disparity. Clearly, further research is needed to determine which of such experimental factors affect the direction of fixation disparity that occurs during reading.

References

Blythe, H. I., Liversedge, S. P., Joseph, H. S. S. L., White, S. J., Findlay, J. M., & Rayner, K. (2006). The binocular coordination of eye movements during reading in children and adults. *Vision Research, 46,* 3898-3908.

Brysbaert, M. & New, B. (2009). Moving beyond Kuĉera and Francis: A critical evaluation of current word frequency norms and the introduction of a new and improved word frequency measure for American English. *Behavior Research Methods, 41*, 977 – 990.

Clark, H. H. (1973). The language as fixed effect fallacy: a critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior, 12*, 335-359.

Coltheart, M. (1981). The MRC Psycholinguistic Database. *Quarterly Journal of Experimental Psychology, 33A,* 497-505.

Deubel, H. & Bridgeman, B. (1995). Perceptual consequences of ocular lens overshoot during saccadic eye movements. *Vision Research, 35,* 2897-2902.

Drewes J., Masson G.S., & Montagnine A. (2012). Shifts in reported gaze position due to changes in pupil size: ground truth and compensation. *Proceedings of the Symposium on Eye Tracking Research and Applications*. 209-212.

Engbert, R. & Kliegl, R. (2004). Microsaccades keep the eyes' balance during fixation. *Psychological Science, 15,* 431-436.

Francis, W. & Kuĉera, H. (1982). Frequency Analysis of English Usage: Lexicon and Grammar. Boston: Houghton Mifflin.

Howard, I. P. (1999). The Helmholtz - Hering debate in retrospect. *Perception, 28,* 543–549.

Jainta, S., Hoorman, J., Kloke, W.B., & Jaschinski, W. (2010). Binocularity during reading fixations: Properties of the minimum fixation disparity. *Vision Research, 50,* 1775-1785.

Juhasz, B. J., Liversedge, S. P., White, S. J., & Rayner, K. (2006). Binocular coordination of the eyes during reading: Word frequency and case alternation affect fixation duration but not fixation disparity. *Quarterly Journal of Experimental Psychology, 59,* 1614-1625.

King, W. M. & Zhou, W. (2000). New ideas about binocular coordination of eye movements: is there a chameleon in the primate tree? *The Anatomical Record, 261,* 153-161.

Kliegl, R., Nuthman, A., & Engbert, R. (2006). Tracking the mind during reading: The influence of past, present, and future words on fixation durations. *Journal of Experimental Psychology: General*, 135, 12-35.

Kirkby, J. A., Blythe, H. I., Benson, V., & Liversedge, S. P. (2010). Binocular coordination during scanning of simple dot stimuli. *Vision Research, 50*. 171-180.

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

Liversedge, S. P., Rayner, K., White, S. J., Findlay, J. M., & McSorley, E. (2006). Binocular Coordination of the Eyes during Reading. *Current Biology, 16,* 1726-1729.

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

Nuthmann, A. & Kliegl, R. (2009). An examination of binocular reading fixations based on sentence corpus data. *Journal of Vision, 9,* 1-28.

Paterson, K. B., Jordan, T. R., & Kurtev, S. (2009). Binocular fixation disparity in single word displays. *Journal of Experimental Psychology: Human Perception and Performance, 35,* 1961-1968.

Rayner, K. (1998). Eye Movements in Reading and Information Processing: 20 Years of Research. *Psychological Bulletin, 124,* 372-422.

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62,* 1457-1506.

Shillcock, R., Ellison, M.T., & Monaghan, P. (2000). Eye-fixation behaviour, lexical storage and visual word recognition in a split processing model. *Psychological Review*, 107, 824-851.

Shillcock, R., Roberts, M., Kreiner, H., & Obregon M. (2010) Binocular foveation in reading. *Attention, Perception and Psychophysics, 72*, 2184 - 2203.

Wyatt H.J. (2010). The human pupil and the use of video-based eyetrackers. *Vision Research, 50,* 1982-1988.

Yang, Q. & Kapoula, Z. (2003). Binocular coordination of saccades at far and at near in children and in adults. *Journal of Vision, 3,* 554-561.

Appendix

Our analysis showed vergence movements in the direction of reduced fixation disparity at the end of the fixation compared to at the beginning of the fixation. We were surprised to observe an interaction between vergence movements and the type of tracker such that the DPI showed larger and statistically significant vergence movements, whereas the numerically smaller vergence movements in the EyeLink system failed to reach statistical significance (see Figure 5).

However, careful scrutiny of the data and mathematical simulations showed that this interaction should not be interpreted to indicate that there were smaller vergence movements in the EyeLink data set, but instead likely arises due to the greater standard deviations in the distributions of fixation locations that occurred in the EyeLink data set compared to the DPI data set, and the mathematical operations carried out to calculate vergence movements during fixations.

As a reminder, vergence movements are calculated on the basis of 4 values: the fixation location of the left eye at the beginning and end of a fixation and the same two values for the right eye. Three operations were carried out on these data to obtain vergence movements. First of all, the difference was calculated between the fixation location of the left and the right eye. This was done both for the values at the beginning and end of the fixation. Secondly, the absolute value of these differences was then taken. Finally, the difference was then calculated between the absolute fixation disparity at the beginning and the end of the fixation to obtain the value for vergence movement.

Below we report the results of our simulations showing that with nearly identical mean values for different distributions of fixation locations, a larger spread

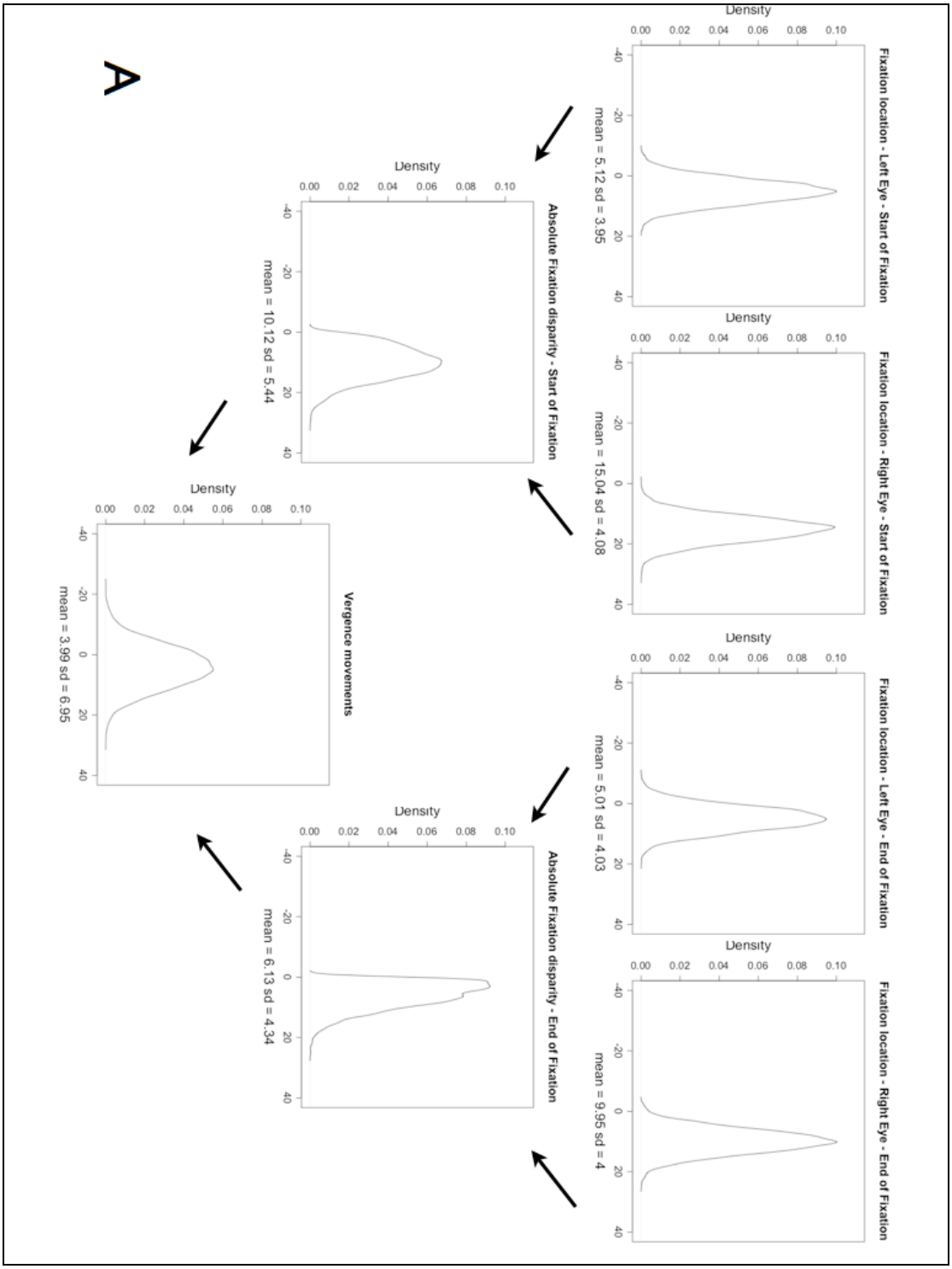
in the distribution of the fixation location data will result in reduced vergence movements even though the other properties of the distribution are equal.

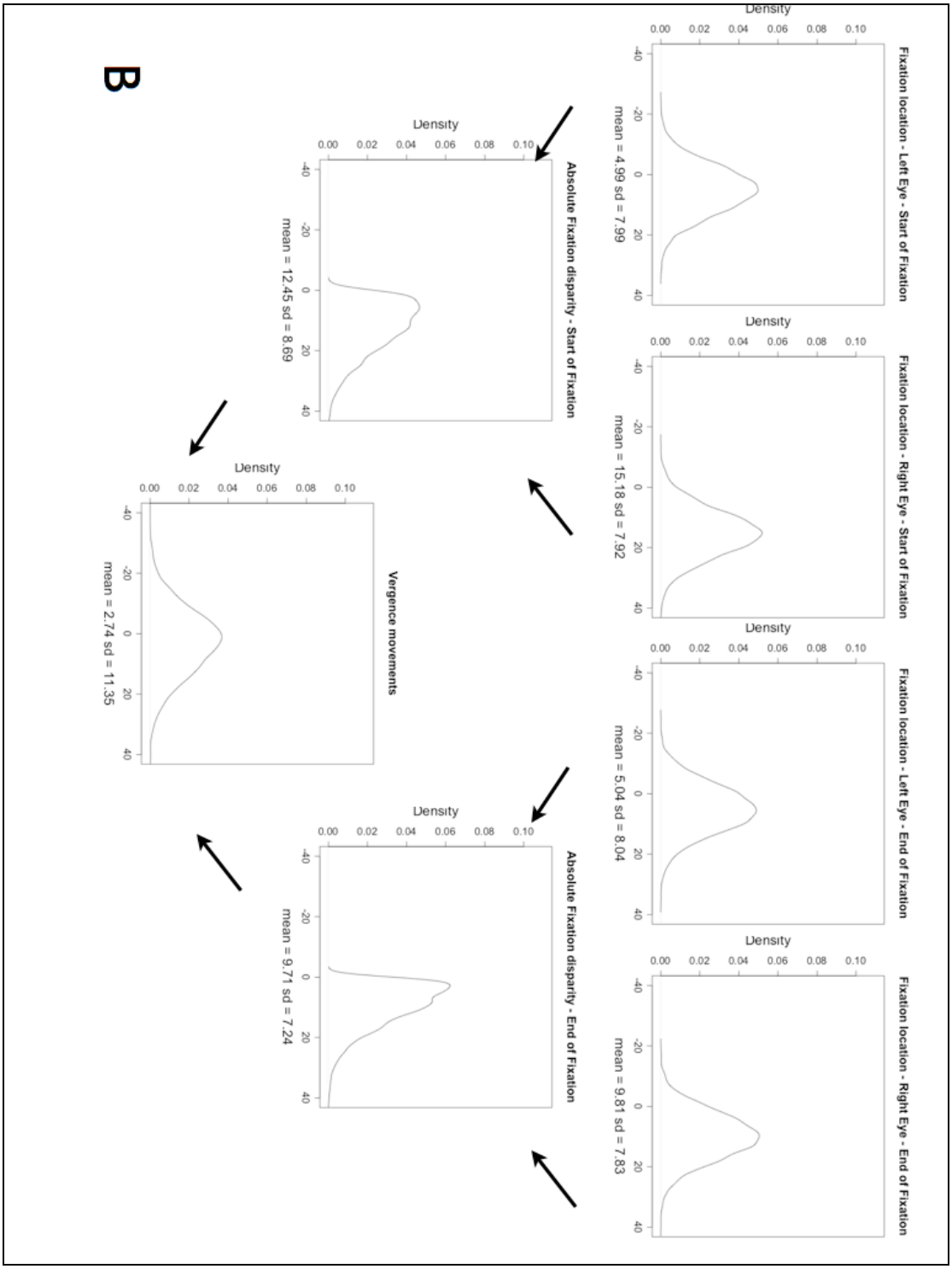
For these simulations we used the rnorm command in the R library (2010) to generate 5000 values approximately following a normal distribution with pre-specified mean and standard deviation. Our initial values for the means of the fixation location distributions were 5 and 15 at the beginning of fixation for respectively the left and right eye, and resp. 5 and 10 for at the end of fixation. These values were only chosen for didactic purposes. The assumption of independent noise (i.e. standard deviation) in the fixation position for each eye is clearly appropriate for the DPI eye trackers where a separate eye tracker is determining the fixation position of each eye. For the EyeLink, the noise in the fixation position for each eye might be correlated, however it is unclear if this is the case and if so, which form and magnitude this correlation would take. For reasons of simplicity, we also assumed independent noise in the fixations positions of the eyes as reported by the EyeLink. Panel A shows the distribution for a standard deviation with value 4. In Panel B we doubled the standard deviation to 8 but kept the means the same. Vergence movements were reduced with a bigger standard deviation as was the case when similar simulations were run with different sets of means for the initial fixation location distributions.

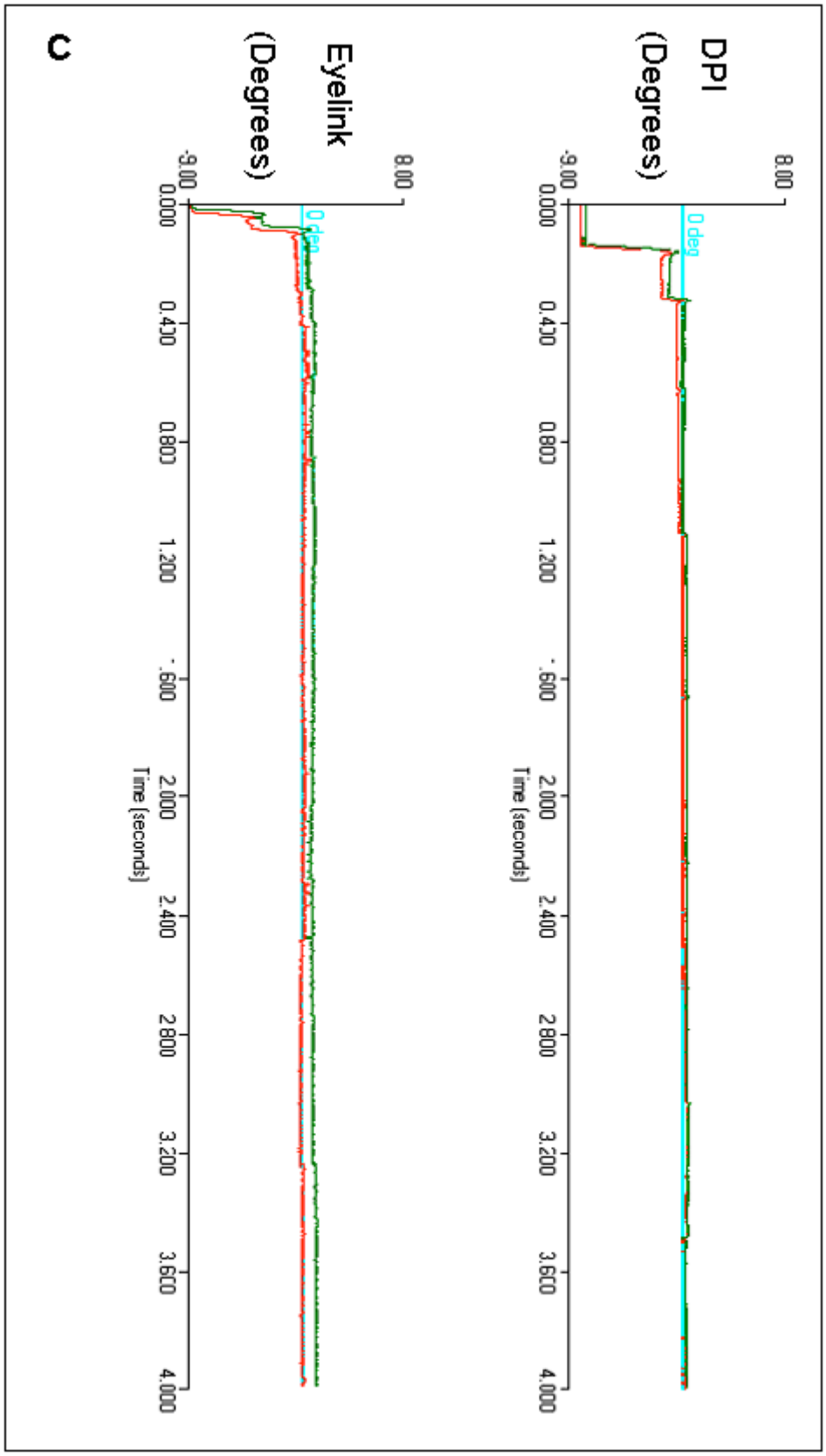
What these simulations indicate is that the smaller vergence movements that we report for the EyeLink system compared to the DPI system arose as a result of the mathematical operations carried out on the fixation location data. They did not arise due to theoretically interesting psychological processes. Specifically, they occurred due to the comparison of vergence movements based on data with a larger than a smaller spread, in all likelihood due to the comparatively higher accuracy of the DPI system compared to the EyeLink system (see Panel C for plots of raw data from both

the DPI system (top) and Eyelink system (bottom) during a prolonged fixation period of 4 seconds; this panel demonstrates the reduced noise level in the DPI eye tracking system compared to the Eyelink system; however, what remains clear is that the nature of alignment during this fixation was similar in both data sets).

.


A
Fixation location - Left Eye - Start of Fixation
mean = 5.12 sd = 3.95
Fixation location - Right Eye - Start of Fixation
mean = 15.04 sd = 4.08
Fixation location - Left Eye - End of Fixation
mean = 5.01 sd = 4.03
Fixation location - Right Eye - End of Fixation
mean = 9.95 sd = 4
Absolute Fixation disparity - Start of Fixation
mean = 10.12 sd = 5.44
Absolute Fixation disparity - End of Fixation
mean = 6.13 sd = 4.34
Vergence movements
mean = 3.99 sd = 6.95
Density


B
Density
Fixation location - Left Eye - Start of Fixation
mean = 4.99 sd = 7.99
Fixation location - Right Eye - Start of Fixation
mean = 15.18 sd = 7.92
Fixation location - Left Eye - End of Fixation
mean = 5.04 sd = 8.04
Fixation location - Right Eye - End of Fixation
mean = 9.81 sd = 7.83
Absolute Fixation disparity - Start of Fixation
mean = 12.45 sd = 8.69
Vergence movements
mean = 2.74 sd = 11.35
Absolute Fixation disparity - End of Fixation
mean = 9.71 sd = 7.24


C
DPI
(Degrees)
8.00
0 deg
-9.00
0.000
0.400
0.800
1.200
1.600
2.000
2.400
2.800
3.200
3.600
4.000
Time (seconds)
Eyelink
(Degrees)
8.00
0 deg
-9.00
0.000
0.400
0.800
1.200
1.600
2.000
2.400
2.800
3.200
3.600
4.000
Time (seconds)

Acknowledgements

J.A.K. was supported by the Leverhulme Trust. H.I.B. was supported by an Economic and Social Research Council Postdoctoral Fellowship. S.P.L. was supported by Economic and Social Research Council grants (RES-000-22-4128 & ES/I032398/1) and Leverhulme Trust grant (F/00 180/AN).