# Extra-retinal signals support the estimation of 3D motion

Andrew E. Welchman [a,*], Julie M. Harris [b], Eli Brenner [c]

[a] School of Psychology, University of Birmingham, Edgbaston, Birmingham B15 2TT, UK
[b] School of Psychology, University of St. Andrews, St. Mary's College, South Street, St. Andrews KY16 9JP, UK
[c] Faculty of Human Movement Sciences, Van der Boechorststraat 9, 1081 BT Amsterdam, The Netherlands

ABSTRACT

In natural settings, our eyes tend to track approaching objects. To estimate motion, the brain should thus take account of eye movements, perhaps using retinal cues (retinal slip of static objects) or extra-retinal signals (motor commands). Previous work suggests that extra-retinal ocular vergence signals do not support the perceptual judgments. Here, we re-evaluate this conclusion, studying motion judgments based on retinal slip and extra-retinal signals. We find that (1) each cue can be sufficient, and, (2) retinal and extra-retinal signals are combined, when estimating motion-in-depth. This challenges the accepted view that observers are essentially blind to eye vergence changes.

© 2009 Elsevier Ltd. All rights reserved.

## 1. Introduction

How do we estimate the motion of an approaching object? If the eyes are stationary (Fig. 1A) the retinal position of the object will change differentially on the two retinae, and the brain could use this change in absolute retinal disparity to detect the motion. However, an observer will normally try to track a moving object, thus minimizing absolute retinal disparities. If so, the observer could use extra-retinal cues about pursuit eye movements to judge the object's motion (Fig. 1B). This idea has a long history (Helmholtz, 1910; Sherrington, 1918), but whilst extra-retinal cues to horizontal version (eyes moving in the same direction) are accepted to play a role in lateral motion perception (Brenner, 1991; Freeman & Banks, 1998; Turano & Massof, 2001), previous studies have suggested that extra-retinal cues to horizontal vergence (eyes moving in opposite directions) provide very poor information about objects moving in depth (Erkelens & Collewijn, 1985b; Regan, Erkelens, & Collewijn, 1986). The latter studies reported that tracking a large moving object does not lead to a perception of motion-in-depth unless other static structures are visible. This suggests that changing vergence information is not provided by extra-retinal cues, but rather by the retinal slip of static structures (Fig. 1D).

To isolate binocular information, Erkelens and Collewijn (1985b) employed stimuli that changed in disparity but had constant retinal size. This is potentially problematic because it can result in a strong conflict between binocular cues and retinal size (looming) cues to motion-in-depth – specifically, binocular signals

specify motion whilst retinal size cues specify no movement of the object. An indication that this might be critical is that some motion-in-depth can be perceived without static references if large targets are replaced by small ones (Regan et al., 1986; Brenner, Van Den Berg, & Van Damme, 1996; Harris, 2006). This suggests that extra-retinal cues can be useful when cue conflict is less evident. Thus, the extra-retinal contribution to the estimation of motion-in-depth may need to be re-evaluated.

Here we test whether observers can discriminate motion-in-depth based on extra-retinal cues. In Experiment 1, we measured this ability by comparing motion direction discrimination thresholds when the eyes were already moving at target motion onset (extra-retinal cues available) and when they were stationary at motion onset (retinal cues available). Our results suggest that extra-retinal cues can be sufficient to support estimates of motion-in-depth. In Experiment 2 we examined the trial-by-trial relationship between psychophysical judgments and the amount of retinal and extra-retinal information available. We found that psychophysical judgments were best accounted for on the basis that observers combine retinal and extra-retinal signals to judge direction of motion. We conclude that, in contrast to the accepted view, extra-retinal signals can be used to support perceptual judgments of motion-in-depth.

## 2. Experiment 1

### 2.1. Method

Observers (the authors and 3 naïve participants) sat 80 cm from a 39.2 by 29.3 cm (27.5 × 20.75 deg) computer screen with their
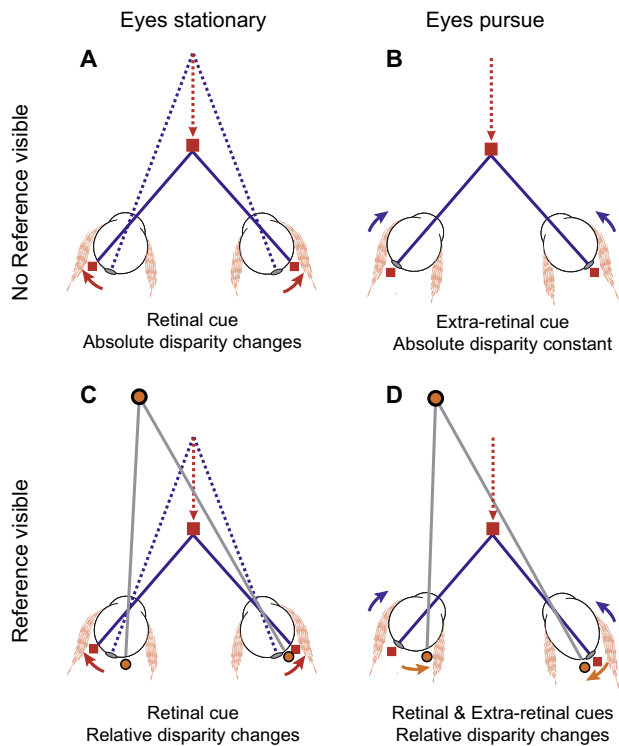
**Fig. 1.** An illustration of the information available for motion-in-depth estimation. (A) If the eyes are still, target motion produces a retinal motion signal (retinal slip of the square target). (B) This retinal slip signal is reduced as the eyes pursue the target, so that with perfect pursuit (and hence no retinal slip) in the dark, information about eye rotation must derive from extra-retinal sources. (C) The presence of a static reference (circle) provides information about motion-in-depth from the changing relative disparity, irrespective of any eye movements. (D) Moving the eyes in the presence of a static marker also produces retinal slip of the marker that provides information about eye rotation.

chin and head supported. The screen's spatial resolution (815 by 611 pixels) was improved using anti-aliasing techniques. Stereo-images were presented sequentially using shutter spectacles (60 Hz per eye). Experiments were conducted in a totally dark room. Periodic room illumination prevented dark adaptation. Stimuli were drawn with the CRT's red gun as the shutter spectacles have little cross-talk for red images. A red filter in front of the CRT removed residual light from the black screen. Observers pressed one of two buttons on the computer's mouse to indicate whether the target moved towards them or away from them. They were instructed to indicate the direction in which the target moved just before it disappeared. An audible beep provided feedback after correct responses.

Observers were instructed to fixate a small square target (6.2 × 6.2 arcmin) surrounded by 200 equilateral triangles (with side lengths 0.72 deg). The positions of the triangles in this background stimulus were randomly assigned at the start of each trial, avoiding a 1.43 deg radius around the target. To mask any visible changes in the triangles as they shifted across the screen (such as aliasing effects from the interaction between the triangles and the pixels), each triangle rotated at 120 deg/s about its centre. Half the triangles rotated clockwise, the other half anti-clockwise. Experimental conditions were interleaved and each observer performed 1000 trials. To ensure that observers' judgments were based on binocular information, the retinal size of the target and background stimulus remained constant irrespective of the simulated depth position. Thus there was a conflict between looming and binocular cues to motion-in-depth. This conflict was large for the background, however, since the target was small, the 'missing'

changing size following the removal of the background was not compelling. Had looming been provided after the background was removed, the maximum change in target size would have been an expansion of the target's edges by 1.4 arcmin. We measured thresholds in three conditions:

### 2.1.1. Eyes moving: large-field

To measure performance based on extra-retinal cues, the observers' eyes had to be smoothly pursuing the target (through symmetrical changes in vergence) before performance was measured. Otherwise subjects may base their judgments on the retinal signals that initiate the eye movements. To achieve this, we initially surrounded the target by a large, structured background and varied the simulated depth position of the background and target together over time, without changing retinal size (Fig. 2A: solid line). This large stimulus (22 × 17 deg) promotes accurate pursuit (Erkelens & Collewijn, 1985a) whilst motion-in-depth should be imperceptible (Erkelens & Collewijn, 1985b; Regan et al., 1986). To make sure that motion-in-depth of the large background was imperceptible we measured observers' discrimination performance when the target disappeared at the same time as the background (Fig. 2A: 0 ms point).

### 2.1.2. Eyes moving: isolated target

If extra-retinal signals can support perceptual estimates, pursuit of the target after removal of the background should allow observers to judge motion direction as the conflict with retinal size information is considerably attenuated. To assess this, we measured performance for three different durations of target motion (100, 200 or 300 ms) after disappearance of the background (Fig. 2A: dashed line).

### 2.1.3. Eyes-stationary: isolated target

Based on previous measurements of eye pursuit (Erkelens & Collewijn, 1985a), it is expected that the large background would provide a good target for pursuit. Nevertheless, it was unlikely that observers would track a target moving in depth perfectly. If the pursuit gain was less than one, some retinal slip would be present in conditions designed to isolate extra-retinal signals. We therefore, measured performance when retinal slip was the only cue available to observers. We used the stimulus described above, with the exception that neither target nor background moved in depth prior to the disappearance of the background (Fig. 2B). Thus the eyes should be stationary at the onset of isolated target motion and target motion will deliver primarily retinal motion signals.

We measured direction discrimination thresholds ("towards" vs. "away") by varying the target's speed using a 2-up, 1-down staircase. The speed was halved after every correct response, and quadrupled after every error. Detection thresholds (67% correct) for each condition were estimated by averaging the logarithms of the target's speed at the reversal points in the staircase. The first four reversals in each staircase were excluded from the analysis. The 5th reversal was also excluded in staircases with uneven numbers of reversals. We included a ceiling value for the staircase to prevent the stimuli exceeding the limits for binocular fusion. This ceiling value was the initial simulated motion-in-depth speed of 50.4 cm/s (around 2.2 deg/s for receding targets and 4.2 deg/s for approaching targets; exact values depend on the subject's inter-pupillary distance). Observers' performance was evaluated by comparing their thresholds in each condition and duration with a bootstrapped distribution of values expected by chance. Such a distribution for chance performance exists because the ceiling value imposed on the staircase procedure ensured that random responses would not increase target speed beyond the upper bound of 50.4 cm/s.
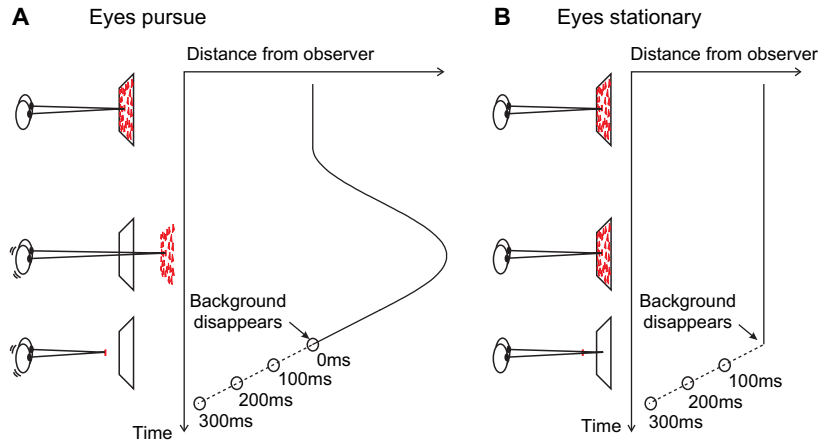
**A** Eyes pursue



**B** Eyes stationary

**Fig. 2.** An illustration of the movement of the target and background to separate retinal and extra-retinal cues for motion estimation. (A) Movements designed to isolate extra-retinal cues by encouraging pursuit. The target and background were initially presented in the plane of the screen for 500 ms. Depending on the direction of isolated target motion, the target and background then moved together towards or away from the observer (isolated towards motion is illustrated). To promote smooth pursuit of the target, the speed of the target and background increased and decreased following a three-quarters cycle of a sinusoid (duration = 1178 ms), whereupon speed became constant. After 250 ms of target motion at a constant speed, the target and background reached the plane of the screen and the background was removed. Depending on the condition, the target then continued to move in isolation in the same direction for 0–300 ms. The maximal depth excursion of the plane prior to disappearance was 25.2 cm in depth, equivalent to a maximal disparity with respect to the screen of about 1.1 deg (receding targets) and 2.1 deg (approaching targets; the angular values differ slightly between-subjects because they depend on the inter-pupillary distance). (B) Conditions designed to isolate retinal cues. The target and background were presented in the plane of the screen for 1928 ms whereupon the background was removed and the target started to move either towards (illustrated) or away from the observer for 100, 200 or 300 ms.

To estimate thresholds expected on the basis of random responding, we simulated 10,000 repetitions of staircases lasting 143 trials (the mean empirical run length) in which a computer generated random responses. These random responses controlled speed using a 2-up, 1-down staircase, and an apparent 'threshold' was calculated by taking the mean of the reversal values (as we did for the values from the actual experiment). The distribution of the 10,000 threshold estimates provided a measure of the range of thresholds expected based on random responding (i.e., chance performance). We used the value of the lower 95% percentile of this distribution as a threshold that determined whether the subjects' data differed significantly from chance. Thus, any threshold below 7.5 cm/s (ca. 0.4 deg/s) was judged to be statistically significant. One naïve subject was unable to perform the task for the speeds we examined (performance did not differ significantly from chance in any condition). Her data were excluded from further analysis. The remaining three experimenters and two naïve observers all produced similar data.

## 3. Results

Observers were not able to discriminate movement direction (i.e., thresholds were not significantly better than chance) when the target that they tracked with vergence eye movements disappeared at the same time as the background (condition 1, motion duration of 0 ms in Fig. 3A). This confirms previous findings that extra-retinal signals alone do not inform perceptual estimates of motion direction for large-field stimuli that do not change in retinal size (Erkelens & Collewijn, 1985b; Regan et al., 1986). However, when participants continued to track the target after the large conflict induced by the background had been removed (condition 2), they were generally able to discriminate motion direction with thresholds better than expected by chance (Fig. 3A, filled symbols). For a motion duration of 100 ms, one subject's threshold was better than chance. For motion durations of 200 and 300 ms mean thresholds were better than chance for four of the five participants. (Note that these motion durations are from the moment that the background disappeared, by which time the target had been mov-
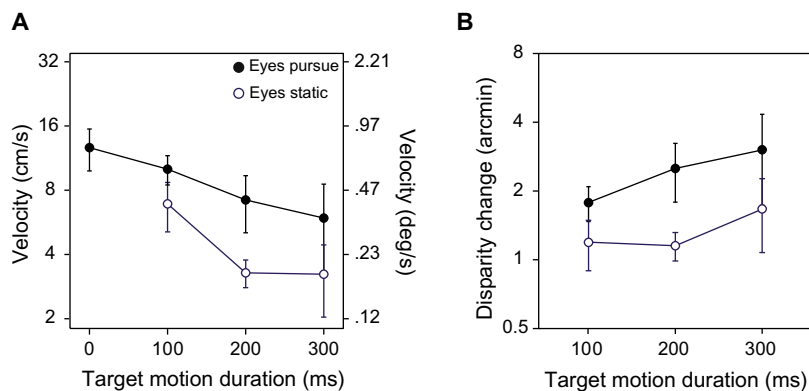
**A**



**B**

**Fig. 3.** Results from Experiment 1. (A) Motion-in-depth direction discrimination thresholds, in terms of the simulated speed of the target, plotted as a function of the duration of isolated target motion. Filled symbols show data obtained when participants tracked the target, open symbols data obtained when the eyes were static. The axis on the right of the figure gives the approximate angular velocity (combined across approaching and receding motions). Error bars show the standard error of the between-subjects mean threshold. (B) Discrimination thresholds expressed in terms of the change in the target's disparity between the moment the background disappears and the moment the target too disappears (symbols as in A).

ing in the same direction for more than half a second). These results suggest that extra-retinal signals relating to eye vergence support perceptual estimates of motion in the absence of a strong conflict with retinal size cues.

Observers were also able to make their judgments on the basis of retinal signals (condition 3). Specifically, when the eyes were not in pursuit at the onset of target motion (resulting in retinal slip of the target), observers were able to discriminate motion direction significantly better than chance (Fig. 3A, open symbols). Mean thresholds were better than chance for all three motion durations. Three subjects performed significantly better than chance at 100 ms, four at 200 ms and all five subjects at 300 ms. As is evident from Fig. 3A, velocity thresholds decreased considerably with increasing target motion duration. Expressing discrimination thresholds in terms of disparity change (rather than velocity) suggested that thresholds were predominantly accounted for by the total change in the disparity of the moving target (Fig. 3B).

Together, the data from Experiment 1 suggest that both retinal and extra-retinal cues support discriminations of motion-in-depth, but that retinal cues are more reliable. However, there are two complications. First, it is possible that when the large background is suddenly removed a retinal aftereffect remains. A persisting retinal signal could have suppressed the percept of motion-in-depth for some time after the background disappeared because the aftereffect would effectively prolong the presence of the background. Performance may, therefore, have been less good under vergence pursuit conditions (Fig. 3) because an afterimage accompanying movements of the eyes provided conflicting retinal size information for a longer duration than the stimulus itself. Thus we may be overestimating the time that the target was presented 'alone'. A persistence of the target's retinal image could compensate for this; however, this influence depends on whether the eyes keep moving, which brings us to the second complication.

Second, as our behavioural measures were not accompanied by concurrent measures of eye movements, we cannot be sure that the observers' eyes always pursued the target with a high gain. It is thus possible that performance in conditions designed to isolate extra-retinal signals actually relied on the retinal slip of the target after the background had disappeared. If pursuit gains were low, resulting in considerable retinal slip, the slip could have been detected and used by observers to inform their judgments. Further, an afterimage accompanying eye movement, as discussed above, could have even provided a reference against which target motion was judged. The higher thresholds for data obtained when the eyes were moving compared to those when the eyes were stationary might support this interpretation (Fig. 3). It is also possible that observers initiated vergence pursuit in the "eyes stationary" conditions, thereby reducing the magnitude of the retinal cue in that condition (although this would probably only affect the longest target duration as there is typically a delay of around 200 ms before vergence tracking begins (Rashbass & Westheimer, 1961)). Based on our psychophysical findings alone, we could not rule out any of these possibilities. In Experiment 2, we therefore, sought to study perceptual judgments based on the combination of retinal and extra-retinal signals, using concurrent measures of eye movement to explore the link between the two.

## 4. Experiment 2

Many natural behaviours involving objects moving in depth promote eye-tracking strategies (e.g., Land, 2006). Given that there is useful information in extra-retinal signals under such conditions (Fig. 1), it would be parsimonious for the brain to estimate motion-in-depth from a combination of retinal and extra-retinal signals. In Experiment 2 we sought to examine discrimination performance under conditions that combined retinal and extra-retinal signals

about the motion of a target point. We used concurrent measures of eye position to calculate the magnitude of eye movement signals and manipulations of target motion to affect the magnitude of retinal signals. Our results suggest that discrimination performance is explained by the brain taking account of both signals, rather than relying exclusively on retinal or extra-retinal information.

### 4.1. Method

Observers (AEW, EB and 3 naïve participants) were presented with stereoscopic targets using a mirror-stereoscope and two CRTs. Stimulus presentation on the CRTs was controlled by a pair of Mac G5 computers. Optical path length from participants' eyes to the 24″ CRT monitors (Sony GDM-FW900 Trinitron) was 50 cm. The spatial resolution of the displays was 1024 by 768 pixels, and the refresh rate 160 Hz. Experiments were conducted in a totally dark room. Observers made towards/away discrimination judgments by pressing the arrow keys on a computer keyboard. One of the naïve subjects was unable to perform the task (her responses bore no relationship to either the real or the retinal target motion), so her data were excluded from further analysis.

Eye movements were recorded with an Eyelink II eye tracker (SR Research Ltd.) at a sampling rate of 500 Hz. The tracker was calibrated by moving an isolated fixation target by 0.57 deg to the left or right in both eyes (the distance that the eyes were subsequently to move in opposite directions) at the start of every trial. The median response (including any corrective saccades) was considered to correspond with this distance.

Observers were instructed to fixate a 14 arcmin diameter red target disc surrounded by a background stimulus of 120 equilateral triangles (side length 1.7 cm on the screen; approximate angular size 1.95 deg) that rotated around their centres at ±120 deg/s. The positions of the triangles were randomly assigned within an area of 25 by 25 cm (28 deg) centred on the target disc at the start of each trial, avoiding a 1.14 deg radius around the target. The motion excursion of the target and background when setting the eye in motion had a maximum amplitude of 5 mm per eye, corresponding to a movement in depth of about 9 cm away from – or 7 cm towards – the observer. This amplitude of movement (a change of vergence of around 1.14 deg) was chosen to ensure the extra-retinal cue was well above threshold based on the results from Experiment 1.

Fig. 4 provides a cartoon representation of the movement trajectory. Note that the retinal size of the target and background stimulus was constant under all presentation conditions. On the disappearance of the background, the target disc turned white. The target's velocity was then perturbed by multiplying its pre-disappearance velocity by a randomly chosen value between −0.5 and 1.5 (i.e., the target could reverse direction and move more slowly, move faster in the same direction, or anything in between these extremes). This ensured that we obtained a range of combinations of retinal velocity signals and that such signals were not correlated with eye velocity pursuit signals (which were subject to naturally-occurring between-trial variability). The target moved (at a constant velocity on the computer screen) for 150 ms following the disappearance of the background. It was then removed, prompting the participants to make a judgment ("towards" or "away") about the target's movement direction following the disappearance of the background. Observers performed 500 trials, of which 9% were excluded because the eye movement recording failed or the subject made a saccade within the last 450 ms before the target disappeared (1820 valid trials out of a total 2000). As a compromise between accuracy and resolution, we determined the pursuit gain on each trial from a linear fit to the eye position data during the last 450 ms (this time window included 300 ms of the background moving with the target and 150 ms of isolated movement of the
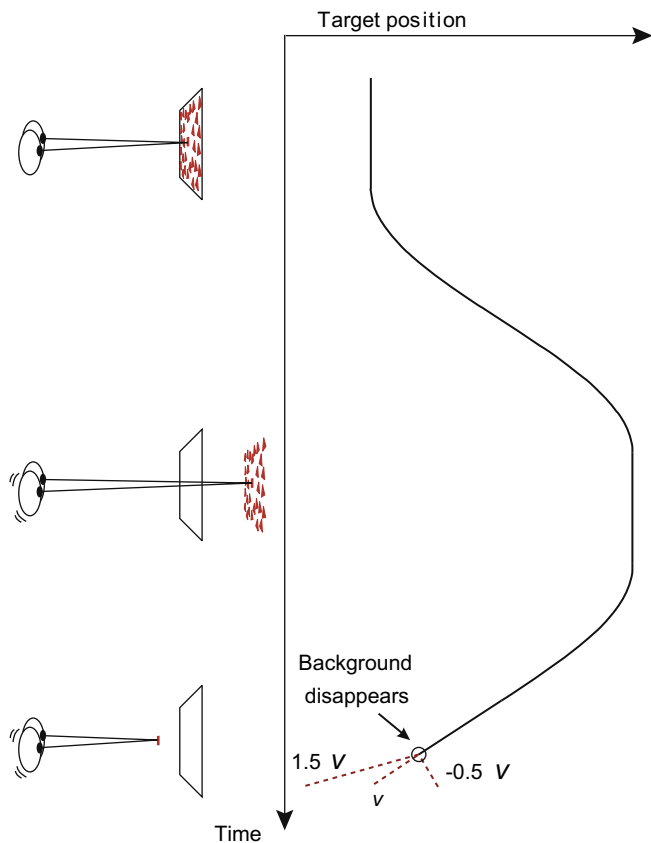
**Fig. 4.** An illustration of the target movement used in Experiment 2. The background and target were initially presented in the plane of the screen for 500 ms, thereafter their position was changed following a sinusoidal profile to a maximum displacement of 0.57 deg per eye over 1200 ms. This required a change of vergence of around 1.14 deg, corresponding to a movement in depth of around 9 cm away from – or 7 cm towards – the observer. The target was then static for 500 ms, before moving in the opposite direction (initial cosine profile that became linear at the quarter phase position) for 850 ms whereupon the background was removed. Following the removal of the background the target's velocity was perturbed. It could continue on the same trajectory, increase in velocity ($v$) by up to a scaling factor of 1.5 or decrease in velocity by up a scaling factor of −0.5. The magnitude of velocity alteration was randomized on each trial. The dotted lines on the figure illustrate the range of possible velocities following the removal of the background. The figure depicts motion towards the observer. The displacement profile for motion "away" corresponds to a reflection along the time axis.

target). To determine pursuit gain, we found the slope of the best fit line to the eye vergence data and compared this to eye vergence required by the movement of the target before the removal of the background. It is possible that pursuit gain changed after the background had been removed; however, inspecting average traces showed that pursuit gain did not systematically decrease during this time. Moreover, whilst short-latency vergence movements can be programmed with a latency of around 80 ms (Busettini, Fitzgibbon, & Miles, 2001) for sudden disparity changes of large-field stimuli, vergence changes for small isolated targets take at least 150 ms and often around 200 ms (Rashbass & Westheimer, 1961; Erkelens & Collewijn, 1991). Thus our presentation of 150 ms after background disappearance is likely to be insufficient for significant vergence changes.

Our measure of variability in the rate-of-change of vergence (a standard deviation of 0.34 deg/s) combines three factors: variability in the average pursuit velocity between subjects and trials, variability in vergence velocity during the 450 ms period of each trial, and variability due to measurement error. In order to partially separate these factors, and in particular to estimate the role of measurement error, we computed the variability in vertical vergence

while subjects fixated the stationary target and background immediately prior to the final approach or receding movement of the target. We used vertical vergence because measurement error is likely to be largest and true variability smallest in that direction. To account for visuo-motor delays, we used the 450 ms interval from 100 ms after the target reached its extreme vergence orientation until 50 ms after it started returning towards the initial position. We determined the standard deviation of the difference between the left and right vertical eye positions for each trial. The median of these standard deviations was 0.03 deg, which is what one would expect for natural variations in vertical vergence when fixating a static target as determined using a more accurate eye movement recording technique (Krauskopf, Cronweet, & Riggs, 1960). Thus random measurement noise introduced by the Eyelink II system appears to be within the limits of expected natural variations in eye deviation, and is likely to have minimal impact on our estimates of the rate-of-change of vergence. Finally, it is important to remember that all our analyses rely on the rate at which the eyes converge or diverge, rather than on absolute estimates of convergence. Thus our estimates of vergence velocity would be minimally affected by slow drifts in the eye-tracking system.

## 5. Results

To analyze observers' estimates of motion direction in terms of the amount of information available from: (i) retinal slip alone, (ii) eye movement signals alone and (iii) the combination of retinal and extra-retinal signals, we first had to determine how well observers were able to change vergence to pursue the moving target. We calculated vergence pursuit gains on a per-trial basis by comparing the slope of the least-squares linear fit to the eye vergence data during an interval starting 300 ms before the background disappeared and ending when the target disappeared, with the vergence demand specified by the moving target before the background disappeared. Fig. 5 provides an example of a single trial showing how eye vergence changed over time. As illustrated in relation to the dotted lines indicating stimulus vergence demand, pursuit gains were typically less than the change in vergence required by the target's movement, with the overall mean gain of ocular pursuit being 0.57. Mean pursuit gains differed between subjects, but did not differ systematically between the two directions of ocular pursuit or between trials for which subjects reported different perceived motion. Previous work suggests pursuit gains can approach one for vergence tracking (Regan et al., 1986; Erkelens & Collewijn, 1991); our use of small disparities, and relatively brief trials may account for the lower gains we observed. Removing the background might lead to a reduction in pursuit gain. However, as discussed above, we found no evidence for changes in gain and the limited presentation durations make it unlikely that vergence pursuit was altered during stimulus presentation.

Having established the pursuit gain on each trial, we determined the amount of information available from retinal slip by calculating the difference between the on-screen motion of the target and the vergence velocity of the eyes. Thus the physical movement of the target was decomposed into a combination of retinal and extra-retinal signals. For instance, when the object is stationary in the environment, retinal and extra-retinal signals are equal and opposite ($x = -y$ line in Fig. 6).

A striking feature of observers' responses was that they showed significant bias: observers reported approach motion more frequently than receding motion (64% of responses were "towards", although only 50% of stimuli were towards the observer). Classifying observers' data along axes corresponding to retinal cues ($x$-axis of Fig. 6), extra-retinal cues ($y$-axis) or the combination of these sources of information (negative diagonal) was informative in rela-
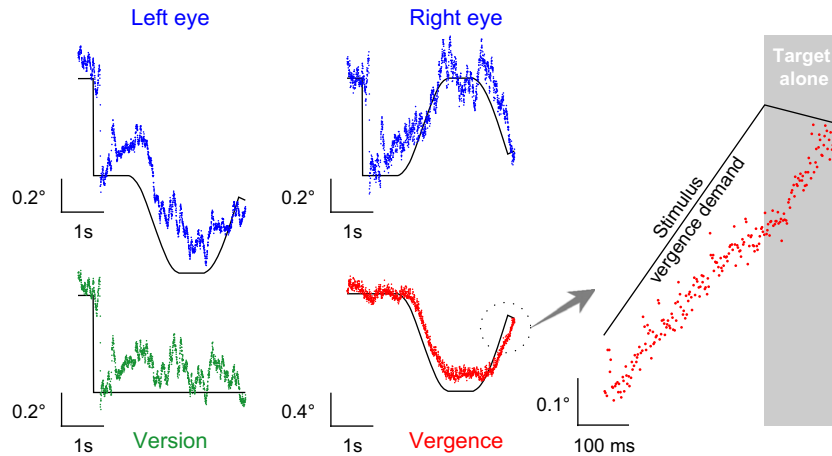
**Fig. 5.** Example eye movement traces from a single trial. The plot illustrates eye position signals for the left and right eyes (top row) with the target position superimposed (black lines). As illustrated, the trial started with a small saccade (in this case rightwards) that covered the extent of movement required by the changing vergence of the target. Version and vergence were calculated as the mean and differences of the left and right eyes traces, respectively. To illustrate vergence changes around the time of target disappearance (the critical period in our study), the right of the figure contains an expanded view of the final 450 ms of the trial. The gray shaded region shows the period during which only the target was visible. To determine vergence pursuit gain, we found the slope of the best fit line to the eye vergence data (fit line not illustrated) and compared this to the eye vergence required by the movement of the target before the removal of the background. We determined the retinal slip during the final isolated target presentation by comparing the fit slope with the target motion after the background disappeared. On the trial illustrated, the target reversed direction after the background had been removed – giving rise to considerable retinal slip.

tion to observers' judgments. Specifically, psychophysical judgments are best explained (steepest psychometric functions) on the basis that observers combine information about both retinal and extra-retinal sources of information (histogram on the negative diagonal) rather than only using retinal (histogram aligned to the *x*-axis) or extra-retinal cues (histogram aligned to the *y*-axis). We quantified this by fitting psychometric functions to the data projected on to the three informative axes (Fig. 6: green curves), observing that a psychophysical model based on the physical motion of the target (negative diagonal) yielded lower thresholds ($\sigma$ = 1.98 deg/s) than one based on retinal ($\sigma$ = 2.43 deg/s) or extra-retinal ($\sigma$ = 3.40 deg/s) cues alone. This analysis was conducted on data pooled across subjects. To ensure this did not bias the results by averaging psychometric functions with different mean values, we performed the same analysis on individual subjects data. We found that the average slopes of individual subject's data were shallower than those for the group data in all cases. For example, considering the data for target motion, $\sigma$ based on averaging fits to individual subjects data was 2.23 deg/s in contrast to the lower value of 1.98 deg/s obtained by pooling data between subjects. This indicates that the improved precision from pooling the subjects' data outweighed any possible decrease in slope through averaging systematic differences between subjects.

One complication with this analysis is that our measures of retinal and extra-retinal cues are not independent (i.e., slip is defined in relation to the measured pursuit gain). It is theoretically possible, therefore, that noise in our estimates of rate-of-vergence could account for better performance in terms of target motion. Specifically, underestimating pursuit on a given trial causes a corresponding overestimation of the amount of slip available – effectively translating the true estimate of pursuit and slip parallel to the negative diagonal in Fig. 6. If the same physical pursuit was overestimated on a subsequent trial (and slip underestimated), this would cause a translation of the data in the opposite direction along the negative diagonal. Such noise would introduce variability along the pursuit (vertical axis in Fig. 6) and slip (horizontal axis in Fig. 6) dimensions, without affecting variability in terms of target motion (negative diagonal in Fig. 6). As discussed above, random measurement noise is unlikely to have a strong influence on our eye movement estimates. However, even if the variability in pur-

suit were completely overshadowed by errors in the eye tracker, the estimated variability $\sigma$ = 0.34 deg/s (the standard deviation of the data for each kind of background motion in the pursuit dimension) is insufficient to account for the lower slope of the psychometric function expressed in terms of target motion. In particular, given that subject's responses are independent of drift in the eye tracker, we can adjust the estimate of performance based on retinal cues (threshold $\sigma$ = 2.43 deg/s) by removing the maximum possible noise contribution ($\sigma$ = 0.34 deg/s). This adjusted estimate ($\sqrt{(2.43^2 - 0.34^2)}$ = 2.41 deg/s) is still considerably larger than the threshold for target motion ($\sigma$ = 1.98 deg/s). Thus, superior performance for target motion cannot be explained solely on the basis of measurement noise.

Our comparison of psychophysical performance in terms of slip, pursuit or their combination assumes that one of the three measures provides a good model for the data. However, it is possible that observers used both retinal and extra-retinal signals, but did not weight them equally in informing their judgments. In particular, participants might under- or over-estimate the contribution of the eye movement signal. As the relationship between retinal ($Vr$) and extra-retinal ($Ve$) information provides the basis of estimating real-world motion, we can write:

$$Vp = Vr + aVe \tag{1}$$

where $a$ is a weighting factor that describes the contribution of the extra-retinal signal ($Ve$) relative to a retinal estimate for the same velocity. Real-world motion would be sensed perfectly if $a$ were 1. For example, for a static target ($Vp$ = 0), retinal and extra-retinal signals would be equal and opposite (i.e., $Vr = -Ve$) as illustrated by the negative diagonal in Fig. 6. This model alone, however, need not capture all the information available to the participants in our study. In particular, as the target's velocity changed when the background stimulus was removed in this experiment (unlike Experiment 1) subjects could potentially have responded to a transitory change in retinal slip induced by the change of target velocity. To provide a data driven approach to understanding participants' judgments, we fit a model that combined information available from retinal and extra-retinal signals (i.e., Eq. (1)) with information available from the transitory slip signal ($T$), where a weight ($w$) determined the relative contribution of transitory slip and the
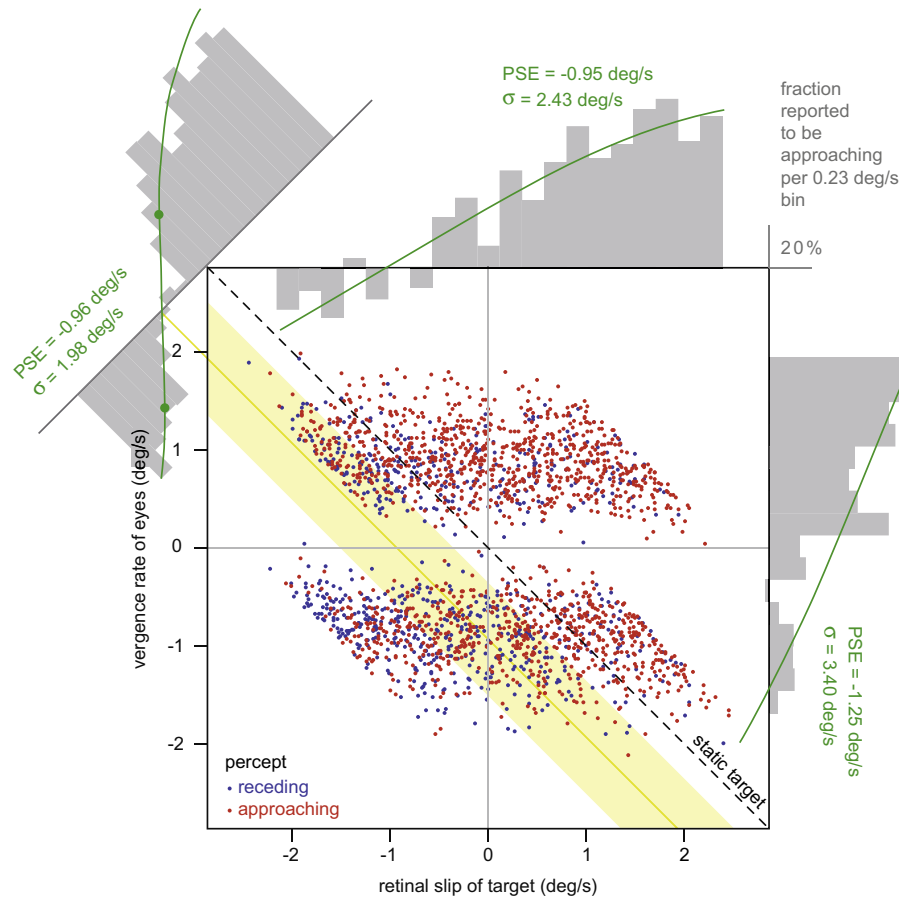
**Fig. 6.** Plot of the data obtained from all 4 observers. The scatter plot shows trial-by-trail combinations of retinal slip and ocular convergence with colours used to code the reported direction of motion (blue receding; red approaching). The dashed line on the negative diagonal corresponds to a physically stationary target. Dots above this line are approaching targets; ones below it are receding targets. The vertical variability within the two groups of dots is caused by variability in pursuit gain. The grey histograms indicate the fraction of presentations on which subjects reported that the target was approaching, as a function of 0.23 deg/s bins of retinal slip (histogram along horizontal axis), ocular convergence (vertical axis) or actual target speed (diagonal). The curves on the histograms are fits of cumulative normal distributions. The highest slope (lowest value of $\sigma$) was found for the target velocity. The fits also revealed a bias to perceive the target as approaching: the point of subjective equality was smaller than zero. The yellow diagonal splits the data into dots that move faster or slower than this velocity, with the shaded area indicating a range of 0.57 deg/s (our estimate of the threshold for 150 ms of target motion from Experiment 1) to either side of this velocity. Since we determined 67% correct thresholds in Experiment 1 the boundaries of the shaded areas should correspond with approximately±33% on the psychometric curves (green dots). Apparently the subjects performed slightly better in equivalent conditions in Experiment 1.

relationship described by Eq. (1). In addition, the model included a bias term (*B*) that accounts for subjects' preference to make 'approach' judgments. Specifically, we described the observer's judgment (*J*) as:

$$J = w(Vr + aVe) + (1 - w)T + B \tag{2}$$

The observer's judgment will be 'approaching' if *J* is larger than zero, and 'receding' if it is smaller. If the bias term (*B*) is greater than zero there will be a larger proportion of 'approach' responses (as we observed). Perfect performance in judging the motion of the target after the background had been removed would be achieved with *B* = 0 (i.e., no bias), *w* = 1 (i.e., transient slip signal ignored) and *a* = 1 (i.e., retinal slip and extra-retinal signals about eye movements given equal weight in informing the judgment).

We fit this model to the observed data by finding the values of *B*, *w* and *a* that best accounted for each subject's decisions (i.e., 'approaching' or 'receding') in relation to the magnitude of retinal (*Vr*), extra-retinal (*Ve*) and transitory (*T*) signals available on each trial. In particular, we used an exhaustive search of the parameter space (resolution of 0.1 deg/s for *B* term; 0.01 for *w* and *a* terms) and identified the parameters that gave the highest correspondence between the calculated value of *J* (greater than zero or less than zero) and the subject's report (towards/away). The between-subjects mean for the best-fitting bias term (*B*) was

0.57 deg/s (SE = 0.11 deg/s across subjects), illustrating the bias evident in Fig. 5 (more 'approach' responses than 'recede' responses). The mean best-fitting value of *w* was 0.84 (SE = 0.09) suggesting that subjects may use the change in retinal slip of the target after the background is removed to some extent, but its influence is modest at best.

Finally, and most importantly, the mean value of *a* was 1.13 (SE = 0.25) indicating that extra-retinal signals are clearly exploited, and even suggesting that subjects rely as strongly on the extra-retinal cue as on the retinal cue in informing their judgments. This estimate is higher than that suggested by previous studies on the use of extra-retinal signals for the estimation of lateral motion, where extra-retinal gain terms are typically between 0.6 and 0.79 (Freeman & Banks, 1998; Freeman & Fowler, 2000; Turano & Massof, 2001). However, it should be noted that the standard error of our estimates is quite high, so we are not confident that this reflects a fundamental difference between the two kinds of pursuit.

In sum, Experiment 2 suggests that observers' performance is best understood in terms of exploiting both the retinal and extra-retinal signals available. This is revealed by psychometric thresholds that are lowest for the target motion, and by the modelling that suggests that retinal and extra-retinal signals have approximately equal weight in determining judgments of motion-in-depth.

# 6. Discussion

The mobility of human eyes necessitates that judging whether an object in the environment is moving or stationary involves the combination of information about eye movement and retinal motion signals. Here we have considered the role of extra-retinal signals pertaining to vergence eye movements. Our results suggest that the accepted view that extra-retinal signals do not contribute to motion-in-depth estimates needs to be revised. Experiment 1 was designed to evaluate whether extra-retinal signals provide sufficient information to estimate motion-in-depth. To isolate extra-retinal information from the retinal slip that initially drives such motion we set the eyes into motion in the presence of a large image that did not expand or contract as vergence changed. Under such conditions no motion-in-depth is seen (0 ms condition, confirming earlier findings by Erkelens and Collewijn (1985b) and Regan et al. (1986)). With the eyes in motion, motion-in-depth is perceived once the background disappears, although the threshold is higher than that based on retinal signals (i.e., condition 3 in which the eyes were initially stationary).

These results show that the retinal signals that initiate pursuit are not essential for seeing motion-in-depth when pursuing the target in depth, and suggest that extra-retinal signals can support direction discrimination judgments. However, we could not rule out an explanation based on residual retinal slip in eye pursuit conditions, so we measured eye movements concurrently in Experiment 2. Moreover, Experiment 2 moved beyond the sufficiency of the extra-retinal signal, to consider how both retinal and extra-retinal cues are used to estimate motion-in-depth. Our results indicate that psychophysical judgments are best explained on the basis that observers combine both retinal and extra-retinal signals to estimate motion towards or away from them. This strategy makes sense, and is comparable to the accepted view for the case of the estimation of signals relating to lateral translation (Brenner, 1991; Freeman & Banks, 1998; Turano & Massof, 2001). As such, it is perhaps surprising that the view that vergence velocity signals do not support perceptual estimation has remained largely unchallenged for over twenty years.

We find approximately equal contributions of retinal and extra-retinal signals for judgments of motion-in-depth direction. Other recent studies are suggestive of the contribution of extra-retinal signals in making judgments of speed (Nefs & Harris, 2007) and in the perception of induced motion (Nefs & Harris, 2008), although the estimated contribution of eye movement signals was smaller than we have found here. More broadly, evidence for the use of eye movement signals in judgments of depth position has been provided by a number of studies (Enright, 1991; Frisby, Catherall, Porrill, & Buckley, 1997; Brenner & van Damme, 1998; Taroyan, Buckley, Porrill, & Frisby, 2000; Backus & Matza-Brown, 2003). While compatible with our study, these previous studies did not distinguish retinal slip (disparity) mechanisms from extra-retinal cues for motion as we have done here.

Our findings have implications for the way in which we conceive the brain's use of disparity and eye movement signals. The importance of distinguishing between absolute and relative frames of reference for disparity is often explained with reference to the perceptual system being effectively blind to large changes in absolute disparity and/or eye position signals. Our results indicate that observers are sensitive to changes in absolute disparity (Experiment 1: eyes stationary conditions) or changes in eye vergence (Experiment 1: eyes pursuing conditions). Whilst it is clear that sensitivity is higher for two concurrently available signals (Westheimer, 1979), our results indicate that extra-retinal signals relating to eye movement are used as part of perceptual estimation with a more or less appropriate gain.

## References

Backus, B. T., & Matza-Brown, D. (2003). The contribution of vergence change to the measurement of relative disparity. *Journal of Visual, 3*, 737–750.
Brenner, E. (1991). Judging object motion during smooth pursuit eye movements: The role of optic flow. *Vision Research, 31*, 1893–1902.
Brenner, E., & van Damme, W. J. M. (1998). Judging distance from ocular convergence. *Vision Research, 38*, 493–498.
Brenner, E., Van Den Berg, A. V., & van Damme, W. J. (1996). Perceived motion in depth. *Vision Research, 36*, 699–706.
Busettini, C., Fitzgibbon, E. J., & Miles, F. A. (2001). Short-latency disparity vergence in humans. *Journal of Neurophysiology, 85*, 1129–1152.
Enright, J. T. (1991). Exploring the 3rd-dimension with eye-movements – Better than stereopsis. *Vision Research, 31*, 1549–1562.
Erkelens, C. J., & Collewijn, H. (1985a). Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. *Vision Research, 25*, 1689–1700.
Erkelens, C. J., & Collewijn, H. (1985b). Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Research, 25*, 583–588.
Erkelens, C. J., & Collewijn, H. (1991). Control of vergence – Gating among disparity inputs by voluntary target selection. *Experimental Brain Research, 87*, 671–678.
Freeman, T. C. A., & Banks, M. S. (1998). Perceived head-centric speed is affected by both extra-retinal and retinal errors. *Vision Research, 38*, 941–945.
Freeman, T. C. A., & Fowler, T. A. (2000). Unequal retinal and extra-retinal motion signals produce different perceived slants of moving surfaces. *Vision Research, 40*, 1857–1868.
Frisby, J. P., Catherall, C., Porrill, J., & Buckley, D. (1997). Sequential stereopsis using high-pass spatial frequency filtered textures. *Vision Research, 37*, 3109–3116.
Harris, J. M. (2006). The interaction of eye movements and retinal signals during the perception of 3-D motion direction. *Journal of Visual, 6*, 777–790.
Helmholtz, H. v. (1910). *Physiological optics*. New York: Dover.
Krauskopf, J., Cronweet, T. N., & Riggs, L. A. (1960). Analysis of eye movements during monocular and binocular fixation. *Journal of the Optical Society of America, 50*, 572–578.
Land, M. F. (2006). Eye movements and the control of actions in everyday life. *Progress in Retinal and Eye Research, 25*, 296–324.
Nefs, H. T., & Harris, J. M. (2007). Vergence effects on the perception of motion-in-depth. *Experimental Brain Research, 183*, 313–322.
Nefs, H. T., & Harris, J. M. (2008). Induced motion-in-depth and the effects of vergence eye movements. *Journal of Visual, 8*, 1–16.
Rashbass, C., & Westheimer, G. (1961). Disjunctive eye movements. *Journal of Physiology, 159*, 361–364.
Regan, D., Erkelens, C. J., & Collewijn, H. (1986). Necessary conditions for the perception of motion in depth. *Investigative Ophthalmology & Visual Science, 27*, 584–597.
Sherrington, C. S. (1918). Observations on the sensual role of the proprioceptive nerve supply of the extrinsic ocular muscles. *Brain, 41*, 332–343.
Taroyan, N. A., Buckley, D., Porrill, J., & Frisby, J. P. (2000). Exploring sequential stereopsis for co-planarity tasks. *Vision Research, 40*, 3373–3390.
Turano, K. A., & Massof, R. W. (2001). Nonlinear contribution of eye velocity to motion perception. *Vision Research, 41*, 385–395.
Westheimer, G. (1979). Cooperative neural processes involved in stereoscopic acuity. *Experimental Brain Research, 36*, 585–597.