eLife

Reviewed Preprint
v1 • April 16, 2026
Not revised

For correspondence:
j.lopezmoliner@ub.edu

Competing interests: No competing interests declared

Funding: See page 19

Reviewing editor: Marisa Carrasco, New York University, United States

Beyond the Focus of Expansion: Retinal curl as a functional signal for heading estimation

Kontessa I Zorpala, Joan López-Moliner
Institute of Neurosciences, Universitat de Barcelona, Barcelona, Spain

eLife Assessment

This study provides an **important** and biologically plausible account of how human perceptual judgments of heading direction are influenced by a specific pattern of motion in optic flow fields known as retinal curl. By combining psychophysical experiments and neural modeling, the authors demonstrate that what was previously considered an incidental "nuisance" signal actually serves as a functional control signal for estimating heading and steering toward a fixated target. While the evidence for the role of curl signals is **convincing** and advances our understanding of vision-based navigation, the work's impact would be strengthened by situating these findings among other cues that contribute to heading estimation, and by clarifying both the time course of these computations and their generalizability across different navigational contexts.

https://doi.org/10.7554/eLife.110770.1.sa4

Abstract

Prevailing models aiming at explaining heading assume that humans need to recover the Focus of Expansion (FOE) while filtering out rotational flow (curl) caused by eye movements. We propose an alternative: the visual system utilizes retinal curl directly to estimate heading, rendering the explicit recovery of the FOE unnecessary. Stationary participants viewed simulated walking paths on a large screen while fixating on ground targets at varying eccentricities—a natural behavior inducing sustained retinal curl. Participants continuously reported perceived heading. To isolate the role of rotational flow, we employed a real-time manipulation that kept translational flow constant while the foveal curl component was either intact, cancelled, or overcancelled. Under natural conditions, participants exhibited systematic heading biases opposite the direction of gaze. Crucially, these biases vanished in the 'cancelled curl' condition, identifying retinal curl as the specific driver of perceptual bias. We modeled these results using a simple feedback controller and a ring-attractor neural network featuring gaze-contingent inhibition and a 'straight-ahead' prior. These findings suggest the brain exploits the geometry of gaze stabilization to simplify navigation, treating retinal curl as a functional signal rather than noise to be filtered.

Introduction

Humans and many vertebrates rely on stabilizing gaze on regions of interest in the world to achieve accurate perception. This fixation strategy, combined with the continuous eye, head, and body movements that accompany natural behavior, generates highly structured patterns of motion on the retina that can, in principle, be exploited to control locomotion. A particularly simple and influential case arises when we move approximately in the same direction as we are looking: in this situation the focus of expansion (FOE), the point in the optic flow from which motion vectors appear to radiate, directly specifies the heading direction (Gibson, 1950). The idea that the visual system uses the FOE to recover heading has subsequently dominated both theoretical and empirical work, including the interpretation of neural activity in primate medial superior temporal (MST) area, where neurons exhibit FOE-like tuning (Bremmer et al., 2017;

Britten, 2008; Duffy and Wurtz, 1991; Kaminiarz et al., 2014), as well as psychophysical findings showing that humans can make highly accurate heading judgements from optic flow (Li and Warren, 2002; Li and Warren, 2000; Warren and Hannon, 1990; Warren and Hannon, 1988). Building on these observations, computational models have proposed population codes and neural network mechanisms (Beintema et al., 2004; Berg and Beintema, 2000; Heeger and Jepson, 1992; Lappe and Rauschecker, 1993; Perrone and Stone, 1994) that operate on radial flow structure to recover the translation direction within this FOE-centric framework.

However, this FOE-centric view faces several fundamental challenges. In natural environments, observers rarely maintain fixation on their heading direction; instead, they typically look toward behaviorally relevant areas, sometimes located eccentrically in the visual field (Matthis et al., 2018). Under combined eye–head movements, the retinal FOE no longer coincides with the true heading direction, creating a computational problem for relying on the FOE for heading estimation (Royden et al., 1994, 1992), because in this case the retinal flow contains both translational and rotational components. To preserve FOE-based control of navigation, two main strategies have been proposed to decompose the flow field: retinal approaches that attempt to detect and compensate for the rotational component directly from the retinal flow itself (e.g., (Cutting et al., 1992; Heeger and Jepson, 1992; Lappe and Rauschecker, 1993; Longuet-Higgins and Prazdny, 1980)), and extra-retinal approaches that rely on eye movement signals to factor out the rotation (Lappe, 1998; Royden et al., 1994).

Despite these challenges, a functional role of rotational flow components has been largely neglected in both theoretical and experimental work on heading perception. Although some models have included rotation sensitivity or extra-retinal signals to account for eye-movement compensation (Beintema and van den Berg, 1998; Koenderink and Van Doorn, 1981; Lappe, 1998), the curl component of optic flow has typically been treated as something to be removed rather than as a potential source of heading information in its own right. A few exceptions include studies that have considered motion-parallax cues, implying rotation, to determine instantaneous heading (Li and Warren, 2000) or to estimate gaze direction (Cutting, 1986).

Recent evidence suggests that this oversight may be consequential. In real walking scenarios, the FOE expressed in a head-centred reference frame is too variable to be used reliably for heading recovery, whereas the magnitude of retinal curl in the fovea can specify the body trajectory relative to gaze (Matthis et al., 2022). Psychophysical studies have also shown that heading estimates exhibit systematic biases that FOE-based mechanisms cannot fully account for under simulated rotation — i.e. when the retinal flow is indistinguishable from that produced by actual eye or head rotations but without accompanying vestibular or proprioceptive signals (Banks et al., 1996; van den Berg and Brenner, 1994). This bias suggests that rotation is not merely treated as noise, but instead provides information used in heading perception. In addition, neurophysiological work has identified that most MSTd neurons are sensitive to spiral motion (Graziano et al., 1994), consistent with a continuum of motion pattern sensitivity (Layton and Browning, 2014) rather than a decomposition into separate channels (e.g. expansion–contraction vs. rotation components). This continuum is consistent with the amount of rotational components in the retinal image as a function of gaze eccentricity, making gaze-stabilisation or fixation strategies critical for understanding the use of retinal flow patterns (Angelaki and Hess, 2005; Calow and Lappe, 2008; Glennerster et al., 2001).

In this work, we provide the empirical evidence that curl can serve as a control variable during locomotion. By relying on gaze stabilization, which generates a retinal curl component, we reveal a robust heading bias that emerges from prolonged exposure to this underlying curl. The direction of this bias differs from the previously reported simulated-rotation bias (see supplementary videos S1 and S2): when observers viewed flow consistent with linear walking while fixating an eccentric ground target, perceived heading shifts opposite to gaze (left fixation leads to rightward heading, and vice versa). The magnitude of this bias matches the curl expected by gaze-centred flow geometry, and cancelling curl eliminates the bias. A simple controller based on mean image curl reproduces heading behaviour across path conditions. Finally, we present a neurally plausible implementation of this controller illustrating how parietal circuits can transform local curl near

gaze into global heading estimates via recurrent dynamics and competition between sensory evidence and priors—without explicit FOE extraction. This reframes retinal curl from a nuisance component to a functional control signal for heading.

Methods

Participants

We tested 12 participants (five self-identified men and seven self-identified women), aged between 24 and 59 years (mean 30 SD: 9), all with normal or corrected-to-normal vision. Except for one, all participants were naïve to the aims of the study and volunteered to take part. The study forms part of an ongoing research program approved by the Ethics Committee of the University of Barcelona and conducted in accordance with the principles of the Declaration of Helsinki.

Displays and conditions

Participant motion was simulated as a translation parallel to the ground plane at a sustained walking speed of approximately 1 m/s. This motion incorporated characteristic bounce and swing components derived from a single gait profile. This profile was recorded once by an independent individual wearing an HMD tracker while walking along the predefined experimental paths in a virtual environment, and then was applied to all participants to ensure stimulus consistency. Each trial lasted between 11 and 12 s.

The ground plane consisted of a 50 × 50 m surface mapped with a naturalistic texture (see Fig. 1A) generated from simplex noise patterns whose spatial power spectrum followed a $1/f^2$ distribution. The temporal frequency of these patterns generated by simulated self-motion is consistent with the statistics of natural videos as described by (Dong and Atick, 1995) also following a 1/f-type temporal power spectra (exponent of 2). The textures were created in real time using OpenGL shaders within a custom Python program, designed for computational efficiency and allowing online manipulation of the texture in specific experimental conditions (see *Flow manipulation conditions* below).

The experiment was run on an Intel i7-based workstation (i7-9700F, Intel, Santa Clara, CA, USA) equipped with an NVIDIA GeForce RTX 2060 SUPER GPU. Images were rendered at 120 Hz with a resolution of 1920 × 1080 pixels and displayed monocularly via a PROPixx projector (VPIxx Technologies, Saint-Bruno, QC, Canada) onto a back-projection screen (2.03 m × 1.16 m), viewed from a distance of 1.0 m, resulting in a visual field of approximately 91°.

To verify fixation, eye movements were recorded with a Pupil Labs Core (Berlin, Germany) eye tracker operating at 200 Hz. Trials were discarded if the median distance between gaze position and the fixation point exceeded 3°.

Path conditions

Participant trajectories could be either straight (length of about 6.5 m) or curved to the left or right (see Fig. 1B). The curved paths resulted in a final heading of ±45° relative to the initial heading (0°) in world coordinates. The three path types (straight, left, and right) were interleaved on a trial-by-trial basis.

Eccentricity conditions

In all trials, a fixation point (yellow dot in Figure 1A) was presented on the ground and remained fixed in world coordinates. Relative to the participant's initial position (x = 0, z = 0) and heading (0°), the fixation point could appear at one of five lateral positions or eccentricities, x = {−4, −2, 0, 2, 4} m, and was initially located 20 m ahead (see colored dots in Fig. 1B). As simulated self-motion progressed, the fixation point appeared to approach the observer, necessitating a gradual increase in gaze angle. Participants were not explicitly instructed to use specific eye or head movements to maintain fixation; instead, they were permitted to engage the head-eye system

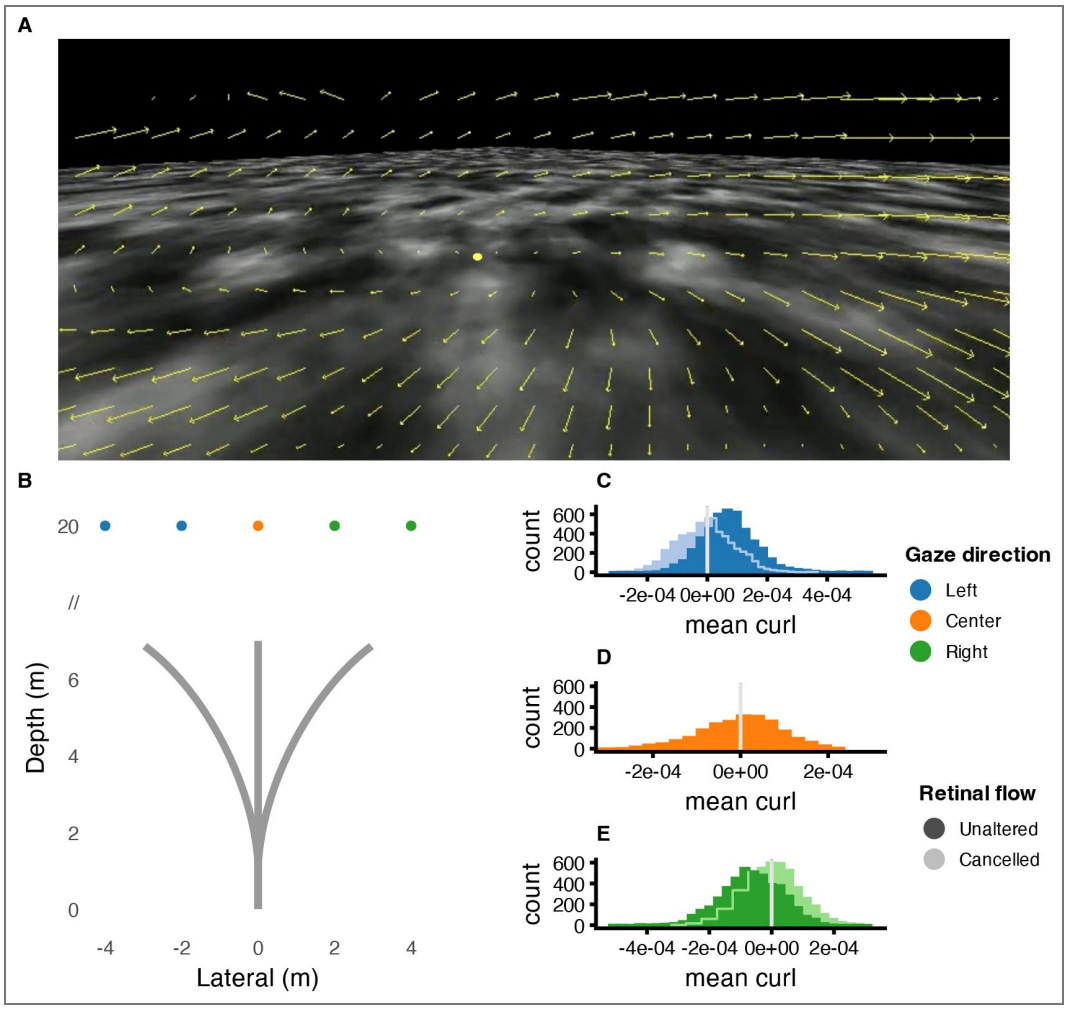

Figure 1. Ground texture, trajectories and retinal curl distributions across conditions.

(A) Snapshot of the ground texture based on simplex noise. Yellow lines indicate optic flow vectors computed using the Farnebäck algorithm. A clear rotational component (curl) is visible, consistent with the observer looking at a point located to the left of the simulated path (yellow dot). (B) Schematic of the three trajectories. Position (0, 0) represents the starting point of simulated locomotion. The five colored dots mark the fixation points in world coordinates at the beginning of each trial (20 m ahead of the observer). (C–E) Distributions of mean retinal curl across trials for left-, center-, and right-gaze conditions, respectively. Dark filled bars indicate the unaltered curl condition, while lighter bars and outlined steps represent the cancelled curl condition. The vertical grey line denotes zero curl.

naturally to track the target. Under perfect fixation, the expected head/eye rotation rate increased for the largest eccentricity from about 0.4 °/s to 0.8°/s (see Figure S2 in the SI). The eccentricity condition was randomized across trials.

Flow (curl) manipulation conditions

When fixating a stationary eccentric point while translating straight ahead, the retinal image contains an expected rotational component— referred to as curl—around the fovea. This effect is illustrated by the flow lines in Fig. 1A, which shows a snapshot of the retinal image consistent with walking straight ahead while fixating a point located to the left (positive curl). Figures 1C–E display the mean curl distributions for fixation points at varying eccentricities: leftward positions (−4 m, −2 m) produced positive curl, the center (0 m) resulted in near-zero curl, and rightward positions (2 m, 4 m) produced negative curl.

To compute the curl shown in Figs. 1C–E, we simulated the experimental trials and generated the corresponding videos (downsampled to 800 × 452 pixels at 30 Hz). Optic flow was then computed using the Farnebäck algorithm (calcOpticalFlowFarneback, OpenCV), which gives us an estimate of image motion for every pixel (x, y) at each frame t. The curl was extracted for each frame sequence from the image 2D flow:

$$\mathrm{curl}(x, y, t) = \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y} \tag{1}$$

where u and v are the horizontal and vertical flow components respectively (in pixels per frame). We computed the mean image curl $\bar{\omega}$ for each frame. These values approximate the mean retinal curl expected under accurate fixation and define the *unaltered curl condition*.

In some conditions, this expected curl was counteracted (*cancelled curl condition*) by adding an equal and opposite rotational component centered on the fixation point. Figure 1C,E also show the corresponding curl histograms (lighter colors), which are centered at zero—matching the distribution obtained when fixating straight ahead in the direction of motion (mean zero curl, Figure 1D). In an additional set of trials, the imposed counter-rotation was doubled to produce an *overcancelled condition*, in which the curl was reversed beyond neutralization (histograms not shown in Figure 1).

Procedure

Participants continuously reported their perceived direction of self-motion while maintaining fixation on the yellow dot. The response was given using a rotative encoder as input device, operated like a steering wheel, to align its orientation with the perceived heading direction. The device had an angular resolution of 0.3°. At the beginning of each session, participants completed a five-point calibration procedure for the eye tracker. Each session comprised 45 trials (3 trajectory types × 5 fixation positions × 3 flow manipulation conditions), and each participant completed a total of 10 sessions.

Computation of perceived path

To reconstruct the participant's estimated path from the angular response, we treated the reported heading angle θ_t (radians) as specifying the instantaneous lateral slope of the perceived trajectory at frame t. For each time step, we computed the incremental simulated forward displacement

$$\Delta z_t = z_t - z_{t-1}, \tag{2}$$

and converted the response angle into a lateral increment via

$$\Delta x_t = \tan\theta_t \Delta z_t. \tag{3}$$

The estimated lateral position was then obtained by forward integration starting from the initial position x_0:

$$x_t = x_0 + \sum_{k=1}^{t} \tan\theta_k \Delta z_k. \tag{4}$$

This yields the sample-wise reconstructed lateral trajectory x_t consistent with the participant's angular responses.

Real-Time Control Model (The controller)

We interpret the reported path as the output of a curl–driven feedback mechanism inspired by point attractor models used previously in steering control (Fajen and Warren, 2007; Wilkie and Wann, 2003a). Let θ_t denote the heading direction (direction of forward velocity in world coordinates) and ψ_t the gaze direction. We consider their difference

$$\phi_t = \psi_t - \theta_t$$

as a instantaneous heading error. The mean optic–flow curl $\bar{\omega}_t$ is taken as a sensory measurement of this discrepancy, so that $\phi_t \approx \bar{\omega}_t / k_c$, with k_c being a curl–to–angle scale factor. This is motivated by the fact that curl increases whenever there is a discrepancy between θ and ψ. The observer controls heading by applying a yaw rate input $\omega_t = \dot{\theta}_t$:

$$\dot{\theta}_t = \omega_t = K_p \phi_t \approx K_p \frac{\bar{\omega}_t}{k_c} \tag{5}$$

so that positive curl induces a leftward turn (increasing θ) and negative curl induces a rightward turn. If gaze is held fixed (e.g. steady fixation), $\dot{\psi}_t = 0$ and therefore

$$\dot{\phi}_t = -\dot{\theta}_t = -K_p \phi_t \tag{6}$$

This is a simple first-order linear differential equation, and the heading error decays exponentially with a correction time constant $1/K_p$:

$$\phi_t = \phi_0 e^{-K_p t} \tag{7}$$

Thus, heading is continuously steered toward gaze until curl vanishes. In our reconstruction we integrate this controlled heading forward in time to generate the perceived trajectory implied by the observed curl (see video S3 in the SI).

Trajectory fits

To test if perceived trajectories can be accounted for by the curl signal, we used the controllerdefined in Equation 5 which we integrated forward with an explicit Euler step to obtain the estimated heading θ. With sampling interval Δt, constant forward speed (v), and initial state (θ_0, x_0, z_0) taken from the first sample of each trial, the discrete update is

$$\begin{aligned} \theta_{t+1} &= \theta_t + \frac{K_p}{K_c} \bar{\omega}_t \Delta t \\ z_{t+1} = z_t + v\cos\theta_t \Delta t \quad & x_{t+1} = x_t + v\sin\theta_t \Delta t \end{aligned} \tag{8}$$

This generates a predicted 2D path $P_{\text{pred}}(K_p) = \{(z_t, x_t)\}_{t=1}^{T}$ from the observed mean curl $(\bar{\omega})$ sequence. We then estimate K_p and forward velocity (v) as free parameters by minimizing a time-warped path discrepancy (J) using Dynamic Time Warping (DTW)(Giorgino, 2009) on the (z, x) sequences between the predicted and perceived paths:

$$J(K_p) = \text{DTW}\left(P_{\text{obs}}, P_{\text{pred}}(K_p)\right)$$

where

$$P_{\text{obs}} = \{(z_t^{\text{obs}}, x_t^{\text{obs}})\}_{t=1}^{T}$$

is the observed path.

We fix K_c (Equation 5) to 1, absorbing the unknown curl-to-angle scale into K_p.

We used two fitting approaches: (1) **independent** or **separate** fits, in which parameters were fitted independently for each combination of heading, gaze eccentricity, and retinal flow manipulation and (2) **join fits**, in which a single set of parameters is used across all heading and gaze eccentricity conditions but they were different for each flow manipulation. Each approach was tested with two parameters (K_p, and the forward velocity, v). The motivation for including forward velocity in the fit was to compensate for variations in the timing of perceived heading responses. For the separate fits, the number of parameters was 30 (for the 2-parameter approach, K_p and v). In contrast, the joined approach used only 2 parameters for all heading and eccentricity conditions.

In order to compare the different models, we employed information criteria adapted for distance-based model comparison. For each model, we computed a surrogate log-likelihood based on the normalized alignment cost:

$$\log \mathscr{L} = -\tfrac{1}{2}\left(\tfrac{D}{n}\right)$$

where D represents the total DTW distance and n the number of observations, with D/n representing the average alignment cost per observation. This normalization ensures appropriate scaling for information criteria computation. We then computed the Akaike Information Criterion as:

$$\text{AIC} = 2k - 2\log \mathscr{L}$$

where k represents the number of parameters.

Results

All the data and code for the analysis are available at osf link. First, we checked the quality of fixation. The median trial-based deviation from the fixation dot ranged from 1.02˚ to 1.55˚ across subjects. On average, the gaze-to-fixation-dot distance exceeded 3˚ in only 0.46% of the trials which were removed from further analysis.

Figure 2 shows the instant perceived heading in the different gaze and paths conditions in which retinal flow was not altered. The thin traces show this pattern for individual observers, and the thick coloured traces show the group means. Figure 2 illustrates a clear and systematic lateral bias in perceived heading as a function of the inital gaze direction. When observer-simulated motion is straight but fixated an eccentric point on the ground (central vertical panels, blue and green lines), their reported instantaneous heading consistently shifted away from the fixation point: leftward fixation led to rightward heading reports, and rightward fixation produced the opposite pattern. The perceived final lateral displacement is significantly different from 0 (straight ahead, $t(11)=5.172$, $p<0.001$) and the magnitude of the bias was not different between sides ($F(1,35)<1$, $p=0.4$). This effect was present both when gaze was displaced by 2 m (top row) and 4 m (bottom row), and the effect scales with the lateral distance of the fixation point from the trajectory path (the final lateral distance between 2 m and 4 m condition was marginally significant, $F(1,35)=3.49$, $p=0.07$). As we shall see below, the magnitude of the bias is well accounted for by the present mean curl.

Note that when both simulated motion and gaze are straight ahead (orange traces), there is essentially no bias: path is perceived straight. This is consistent with curl staying near zero throughout the trial, so the controller receives no systematic steering signal (see Figure 1D). In this case, heading and gaze remain aligned and perceived trajectory stays close to the physical straight path.

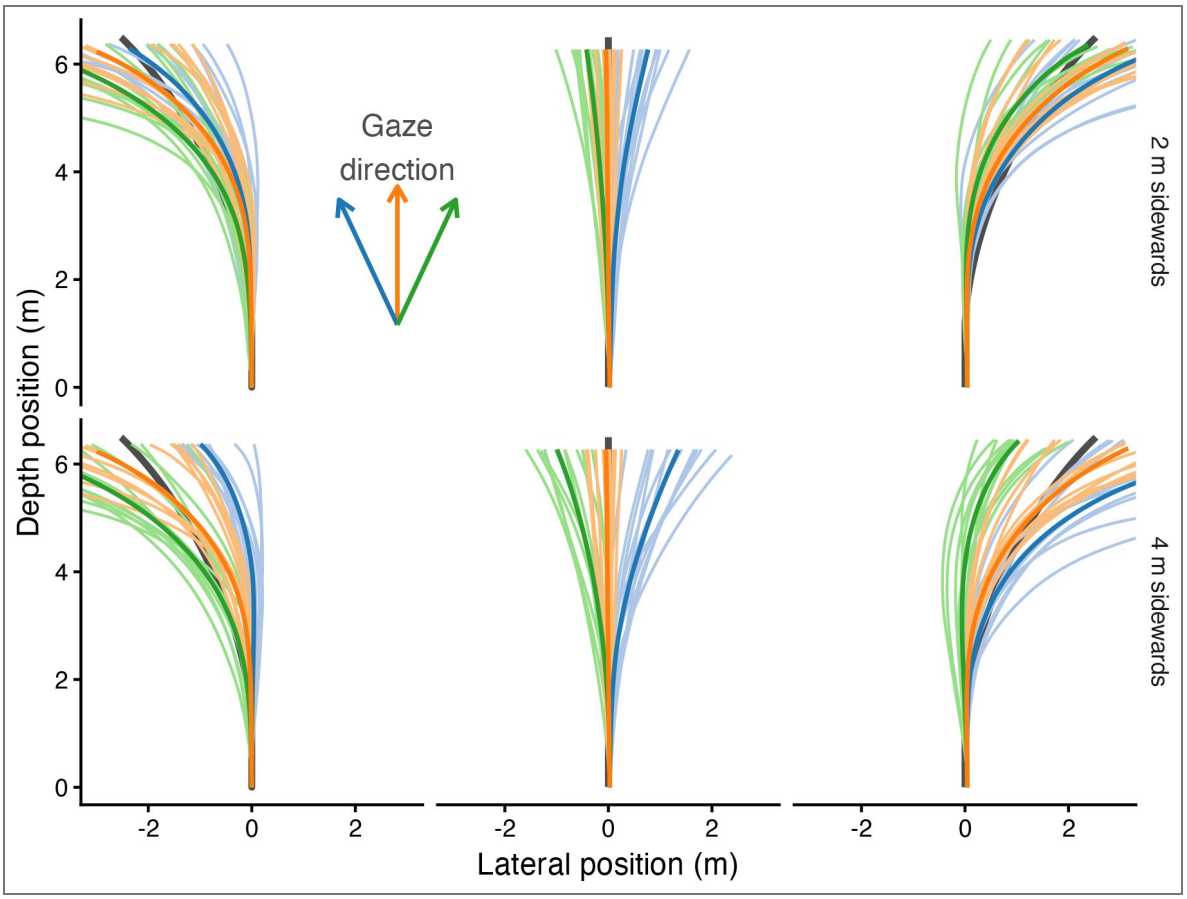

Figure 2. **Perceived heading**

Reported instant directions are plotted when fixating eccentric points on the ground. Rows: fixation 2 m (top) and 4 m (bottom) to the side. Columns: different physical path conditions (centre column = straight path). Colours: initial gaze direction (left / centre / right). Note that on curved paths gaze might change the direction sign relative to heading (i.e. left curve and initially fixating on the left). Thick coloured lines denote the mean across observers. Thin coloured lines: denote individual observers. Dark grey: physical paths. Axes show lateral position (x-axis) versus depth position (y-axis)

Although the bias is more easily interpretable in the straight ahead condition, systematic effects due to gaze direction are present in the curved paths as well. Perceived curvature was systematically *underestimated* when gaze was directed into the direction of the physical curve: for leftward curves this underestimation is most evident when gaze is to the left (blue traces), and for rightward curves when gaze is to the right (green traces). This follows directly from the curl dynamics: as the observer-simulated motion proceeds along the curved path, the instantaneous mean curl decreases, because the heading and gaze directions become progressively aligned (i.e. $\Phi(t) \rightarrow 0$). Under the proportional controller we have defined above, the yaw command scales with this error, $\omega(t) = K_p\Phi(t)$, hence reduced curl implies reduced corrective turning. The model, therefore, predicts a progressive flattening of the perceived path as curl dies out over time. This is exactly the pattern observed empirically, and it is symmetric for left and right curves. By contrast, when gaze is directed opposite to the direction of curvature, $\Phi(t)$ does not collapse toward zero, curl remains sustained, and therefore no underestimation of curvature is expected or observed. This pattern is symmetric for left and right curves too.

Effects of flow manipulation

Figure 3 shows the reported instant heading for the different flow manipulations. We also show the unaltered condition (yellow-green) again for the sake of comparison. Importantly, when curl is cancelled (cyan), these biases essentially disappear – the perceived path remains much closer to the physical trajectory across gaze directions. Note that the central row (straightahead simulated motion) provides the clearest demonstration: the unaltered curl shows the previously reported lateral bias, but when curl is cancelled (counteracted) the perceived trajectory becomes near-veridical. This confirms that the mean curl signal itself – not gaze eccentricity per se – is the functional driver of the bias. In the over-cancelled condition (purple), the pattern reverses, with biases in the opposite direction. To maintain the full factorial design (Eccentricity × Flow Manipulation × Heading), we included trials where gaze was centered and straight ahead. Although no natural curl is generated at this zero-eccentricity position, we purposely introduced curl at rotation speeds corresponding to the 'cancelled' and 'over-cancelled' conditions (simulating random leftward or rightward gaze). As shown in the central panel of Figure 3, this introduced flow induced the predicted biases despite the absence of gaze eccentricity, confirming that the retinal flow pattern, rather than the physical orientation of the eyes, is the primary driver of the perceived heading shift.

The last column in Figure 3 shows the average 2D displacement (in meters) required for the observed paths to align with the physical ones. This measure gives an idea of the similarity between the reported paths and the physical exposed trajectories in the experiment. In general, the difference between perceived and physical paths is smaller when the path is straight (central panel of last column in Figure 3). Consistently with the reported paths in the different flow manipulation conditions, the displacement is smaller in the cancelled flow curl (color-coded) for all headings (different rows). This resulted in a significant quadratic effect of Flow manipulation, $\beta = 0.09$, SE = 0.023, $t(526) = 3.997$, $p < .001$, indicating that mean deviation was lower in the cancelled condition compared with the unaltered and overcancelled conditions.

Fitting the controler

To assess how well the curl-based controller accounts for reported perceived headings, we fitted the model as described in the Methods. The fit incorporates the mean image curl computed via Equation 1 from the experimental videos, assuming perfect gaze stabilization on the fixation point—a reasonable assumption, given the high quality of fixation recorded. Fits were performed for both individual participants (see Figures S4 to S15 in the SI) and the aggregate data (see Figure 4 and Figure S3 in the SI). As shown in Figure 4, the model with separate parameters captures the average perceived trajectories with remarkable accuracy. The observed instant headings (thin lines) are overlaid with the separate fits per condition (thick red lines). In these separate fits, the mean 2D deviation per step was minimal, ranging from 0.066 m in the cancelled flow condition to 0.080 m in the unaltered condition (see table S1 in the SI). Statistical analysis via ANOVA on the

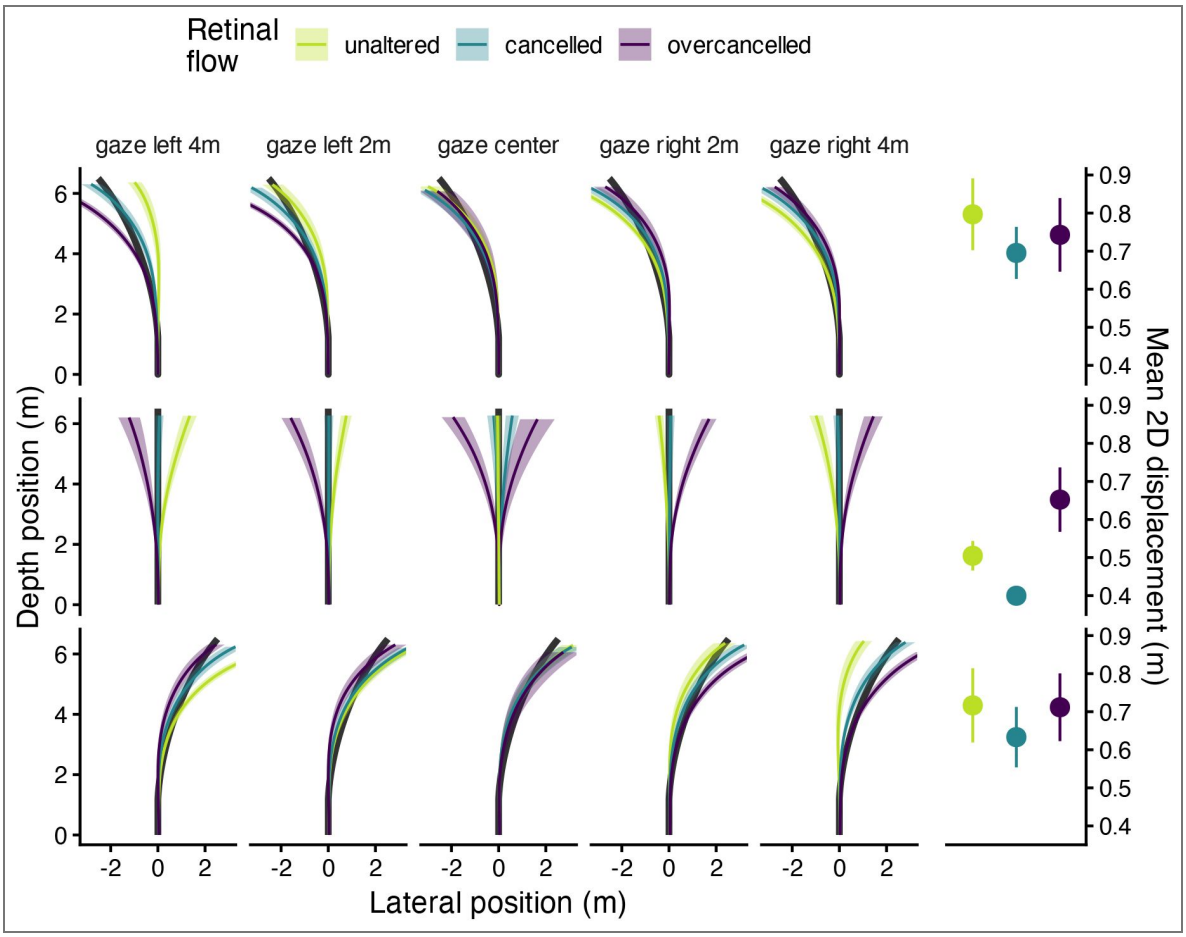

Figure 3. Average perceived heading for each gaze condition (columns) and each physical path curvature (rows), separately for the three retinal–flow manipulation conditions: unaltered curl (yellow-green), cancelled curl (cyan), and over-cancelled curl (purple).

Shaded envelopes indicate between-observer variability (95%-CI). The right axis applies to the last column and illustrates the mean 2D displacement between observed and physical paths. This measurement indicates the mean displacement that is required for the observed path to align with the physical one. The displacement is shown for the different flow manipulations (color-coded). The error bars indicate between-observer variability (95%-CI).

individual fits confirmed that while separate fits yielded a significantly smaller loss than the join approach (F (1, 55) = 206.27, p < 0.001), flow manipulation itself had only a marginal effect on fit quality (F (2, 55) = 2.69, p = 0.077). Interestingly, the cancelled condition showed a slightly smaller loss than the other two (Quadratic contrast t(55) = 2.356, p = 0.022). Although the join fit (using only two parameters for all heading and eccentricity conditions) naturally results in a larger 2D deviation, it effectively captures the core qualitative trends across the entire conditions: a) The model successfully tracks the near-straight perceived trajectories in the central row when curl is cancelled; b) The join model accurately reproduces the systematic underestimation of curvature when gaze eccentricity aligns with the direction of the curve and, c) Even with constrained parameters, the model accounts for the reversal of bias directions in the over-cancelled condition (purple lines) across different gaze eccentricities. Despite the higher deviation inherent in compromising parameters across conditions, the join model is statistically superior as indicated by its lower AIC measures (see Table S1). This suggests that a unified curl-based mechanism provides a parsimonious and robust explanation for the instant heading perception across diverse gaze and flow conditions.

Neural Network Model of Gaze-Contingent Heading Bias

We present a neural model that provides a neurophysiologically plausible implementation of the controller. The model demonstrates that the bias emerges from the interaction between gaze-modulated visual flow processing and a weak straight-ahead prior in parietal heading representation, implemented through standard cortical connectivity. After presenting the model, we show different dynamic aspects of the model and reproduce the bias for the straight ahead trajectories at different trajectories. The python source code of the model is available at *osf link*

Local Motion Encoding

Visual motion is encoded by direction-selective units analogous to primate MT cortex, with 8 directional preferences:

$$\mathbf{E}_{\text{dirs}} = \left[\cos\left(\theta_k\right), \sin\left(\theta_k\right)\right]^T, \quad \theta_k = \frac{2\pi k}{8}, \quad k = 0, \ldots, 7 \tag{9}$$

The response of each direction channel at image-plane position $\mathbf{p}_i = (x_i, y_i)$ is given by rectified cosine tuning (Simoncelli and Heeger, 1998):

$$m_{i,k} = \max\left(0, \mathbf{v}_i \cdot \mathbf{E}_k\right) \tag{10}$$

where $\mathbf{v}_i$ is the local optical flow vector. In our implementation, optic flow $\mathbf{v}_i$ is computed, as before, using the Farnebäck algorithm as a substitute for early visual processing (V1/MT complex). This provides the motion signals that the visual system would extract through V1 to MT processing. The model itself begins with the input to the 8-sector MT-like directional encoding (Equation 10).

Curl Computation (MSTd)

To quantify the rotational (curl) component of optic flow around the current gaze position, we project local flow vectors onto the tangential direction of the circle centered at gaze. For each sampled location $\mathbf{p}_i$, we first compute its position relative to the gaze point $\mathbf{g}$:

$$\mathbf{r}_i = \mathbf{p}_i - \mathbf{g} \tag{11}$$

This vector points from the gaze location to the sample point. The direction orthogonal to this vector corresponds to the local tangential direction of rotation around gaze. We obtain this unit tangential vector by rotating $\mathbf{r}_i$ by 90° and normalizing:

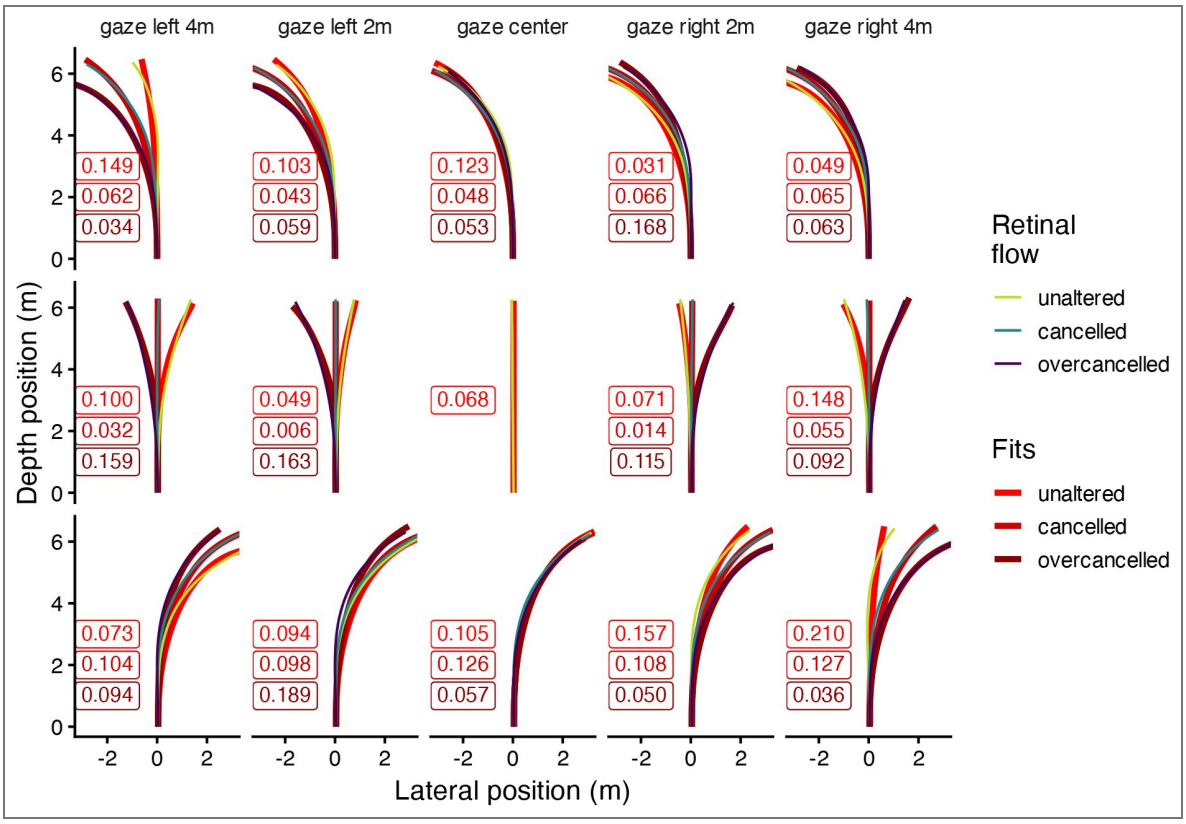

Figure 4. Separate fits of the controller.

Average perceived heading across participants for each gaze condition (columns) and each physical path curvature (rows), separately for the three retinal–flow manipulation conditions (thiner solid lines): unaltered curl (yellow-green), cancelled curl (cyan), and over-cancelled curl (purple). The different red thicker solid lines denote the best fit of the controller. The numbers in each panel indicate the average lateral deviation per step between the fit and the observed heading. Note that for the centered gaze and straight path, only the unaltered flow condition was fitted.

$$\hat{\mathbf{t}}_i = \frac{[-r_{i,y}, r_{i,x}]^T}{\|\mathbf{r}_i\|} \tag{12}$$

Finally, the contribution of the optic flow at that point to local rotational motion is given by the projection of the flow vector $\mathbf{v}_i$ onto this tangential direction:

$$\omega_i = \mathbf{v}_i \cdot \hat{\mathbf{t}}_i$$

Positive values indicate counterclockwise motion around gaze, and negative values indicate clockwise motion. The mean curl is computed as follows:

$$\bar{\omega} = \frac{1}{N_{\text{used}}} \sum_{i \in S} \omega_i$$

where S represents samples satisfying $\|\mathbf{r}_i\| > r_{\min}$ after trimming outliers. In our implementation, $N_{\text{total}} = 400$ samples are drawn within a gaze-centered region.

Ring Attractor Dynamics: Parietal Heading Representation

Heading direction θ is represented by a ring attractor mechanism consistent with previous studies (Zhang, 1996), with neural activity $x(\Phi, t)$ at preferred heading Φ evolving as:

$$\tau \frac{dx(\phi,t)}{dt} = -x(\phi,t) + \int W\left(\phi, \phi'\right) x\left(\phi', t\right) d\phi' + I_{\text{ext}}(\phi,t) \tag{13}$$

We implement this dynamics by using a discrete ring of $N = 181$ units with preferred headings Φ_j uniformly spaced between $[-\Phi_{\max}, \Phi_{\max}]$. The activity dynamics then follow:

$$x_j[t+1] = x_j[t] + \frac{\Delta t}{\tau}\left(-x_j[t] + \sum_{k=1}^{N} W_{jk} x_k[t] + I_{\text{ext},j}[t]\right) \tag{14}$$

The recurrent connectivity follows a standard Mexican hat profile (Ben-Yishai et al., 1995):

$$W_{jk} = A_e \exp\left(-\frac{d(\phi_j,\phi_k)^2}{2\sigma_e^2}\right) - A_i \exp\left(-\frac{d(\phi_j,\phi_k)^2}{2\sigma_i^2}\right) \tag{15}$$

where $d(\Phi_j, \Phi_k)$ is the circular distance between preferred headings. A_e and A_i were set to 1.9 and 1.0 respectively.

External Inputs

The external input $I_{\text{ext}}(\Phi, t)$ consists of two components:

$$I_{\text{ext},j}[t] = I_{\text{gaze},j}[t] + I_{\text{prior},j} \tag{16}$$

The gaze-modulated inhibitory input is:

$$I_{\text{gaze},j}[t] = -K_p \cdot \bar{\omega}(t) \cdot \exp\left(-\frac{d(\phi_j,\phi_g)^2}{2\sigma_\kappa^2}\right) \tag{17}$$

where Φ_g is the gaze direction in heading coordinates, K_p is the inhibitory gain, and σ_κ controls the width of the gaze-centered modulation, determining how broadly the inhibition spreads around the gaze direction Φ_g

The straight-ahead prior is:

$$I_{\text{prior},j} = I_0 \cdot \exp\left(-\frac{\phi_j^2}{2\sigma_p^2}\right) \tag{18}$$

I_0 is the prior strength and σ_p sets the width of the straight-ahead prior, with smaller values producing a sharper preference for $\Phi = 0$

Bias generation as Sensory–Prior Competition

Within this framework, the systematic bias opposite to gaze emerges from the interaction between the *gaze-centered inhibitory drive* and a *weak straight-ahead prior* within a stabilizing recurrent Mexican-hat network. A strong prior ($I_0 \gtrsim 0.25$) keeps the heading estimate near straight-ahead, while a weak prior $I_0 \approx 0.03$ allows sensory evidence to shift the attractor state. The symmetric recurrent dynamics maintain a coherent and stable activity bump. For leftward gaze ($\Phi_g < 0$) and forward motion $(\bar{\omega} > 0) : 1)$ Gaze-centered inhibition creates activity suppression at Φ_g; 2) With weak prior, activity bump shifts toward opposite side and 3) Recurrent dynamics translate inhibition into bump displacement.

The heading direction is decoded using population vector readout:

$$\hat{\theta} = \arg\left(\sum_{j=1}^{N} x_j e^{i\phi_j}\right) \tag{19}$$

where N is the number of ring units ($N = 181$). Unlike some heading perception models (Beintema et al., 2004; Lappe and Rauschecker, 1993), we do not apply sigmoid nonlinearities to neural activities prior to decoding but a linear readout. Gaze-contingent bias emerges even with linear decoding, suggesting it is a fundamental property of the network dynamics.

Relation to the controller (phenomenological model)

The neural implementation provides a mechanistic basis for the phenomenological relationship $\dot{\theta} = K_p \cdot \bar{\omega}(t)$ by demonstrating how rotational flow is converted into a change in perceived heading. In the neural model, K_p is not a single fixed parameter, but rather a dynamic property of the recurrent network. The shift in the activity bump (representing $\dot{\theta}$) is driven by the spatial asymmetry of the inputs. When the gaze-located inhibition suppresses the sensory evidence at the fixation point, it causes the stable bump of activity to "drift" toward the uninhibited regions of the sensory input. The speed of this drift is directly proportional to the strength of the inhibitory gain (K_p in Equation 17) and the magnitude of the local curl signal (ω). Thus, the network naturally performs the integration required by the controller: it transforms localized, gaze-contingent inhibition into a global heading shift that matches the $\dot{\theta} \propto \bar{\omega}$ relationship observed in participants.

Neural simulations

To demonstrate the neurophysiological feasibility of the curl-based controller, we conducted a series of neural network simulations using a custom Python implementation of the recurrent ring network, with subsequent 3D trajectory reconstruction. The model implementation read the video sequences of the experimental straight paths conditions, assuming again perfect gaze stabilization on the fixation point across all trials. As commented above, we computed the optic flow for each frame, before proceeding with each step of the neural model, from the local motion computation to the linear decoding of the ring activity.

We explored the parameter space to identify the drivers of the heading bias, varying the prior strength ($I_0 \in [0.03, 0.24]$), the prior width ($\sigma_p \in [0.18, 0.25]$), the spatial extent of gaze inhibition ($\sigma_k \in [0.02, 0.04, 0.08, 0.12]$) and the inhibitory gain ($K_p \in [0.4, 0.8]$). Table S2 in the SI shows the range for the different parameters used in the simulations.

Simulation results

The network dynamics provide a clear mechanistic explanation for the systematic heading bias observed during eccentric gaze. In Figure 5, we observe the network components for a trial in which the gaze direction is 4 m to the left during forward translation. Figure 5 shows the recurrent connectivity following a standard Mexican hat profile, characterized by local excitation and broader surround inhibition. The activity bump does not settle at the objective straight-ahead (0˚) but stabilizes at a positive (rightward) heading offset, as illustrated in Figure 5. This displacement is driven by the competition shown in the mechanism panel (Figure 5): the gaze-centered inhibitory drive (red dashed line) suppresses the leftward portion of the ring, effectively 'pushing' the neural activity bump (black line) away from the gaze location and to the right of the straight-ahead prior (blue dotted line). The Phase Portrait (Figure 5) confirms this as a stable state; the drift dynamics $\left(\frac{d\theta}{dt}\right)$ converge toward a stable fixed point where the curve crosses zero at a rightward offset. Together, these results demonstrate how localized inhibition, within a parietal ring attractor, transforms gaze-contingent sensory signals into a shifted global heading estimate. For results across a range of eccentricities during forward translation, see Figure S16 in the SI. The presence of temporal oscillations in the activity bump when gaze is aligned with the direction of travel (Figure S16 Row B) or at low eccentricities (rows A and C) represents a state of high competition within the attractor. As seen in Figure S16 B, when the gaze-centered inhibition —the sensory *pull*— overlaps directly with the straight-ahead prior —the internal *push*— it creates a central suppression that counteracts the recurrent excitation of the network. This conflict results in a concave activity profile and rhythmic fluctuations in bump intensity as the network continuously attempts to re-stabilize the representation against the central inhibitory zone. Notably, these oscillations persist when gaze is less eccentric (e.g., at −2 and 2 m in rows A and C respectively), as the proximity of the inhibitory signal to the prior creates a similar destabilizing interaction. Despite these internal dynamics, the decoded heading remains accurate due to the symmetry of the competing forces, suggesting that while representation coherence may decrease, the population vector readout remains robust to central sensory-prior conflict.

From the neural model heading estimates ($\hat{\theta}$, in pixels) we reconstructed the 3D paths using a pinhole camera model (see Heading Estimation and 3D Path Reconstruction section in SI) that are shown in Figure S17 of the SI. This figure demonstrates that the observed instant headings pattern shown in Figure 2 is best captured by a network configuration featuring a weak straight-ahead prior ($I_0 \approx 0.03$) and moderate inhibitory gains (K_p between 0.4 and 0.8). Within this parameter space, the localized, gaze-contingent inhibition effectively shifts the population activity bump, reproducing the characteristic lateral spread seen in the behavioral data. A low prior value (I_0 = 0.03 in the simulation grid) is essential. Higher prior strengths (e.g., I_0 = 0.24) overly anchor the heading estimate to the center (0 m), failing to produce the significant lateral spread observed in the human data. K_p = 0.4 closely matches the tighter trajectory fan seen in the top central panel of the empirical data, while K_p = 0.8 captures the wider lateral deviations observed in the bottom central panel, where the bias is more pronounced. Concerning the inhibition width (σ_k), values between 0.08 and 0.12 provide the most realistic structural match. Very low values (σ_k = 0.02, 0.04) produce negligible bias even at high gain, as the localized inhibition does not sufficiently overlap with the heading activity bump to drive a shift. The prior width value of σ_p = 0.18 appears to provide a more sensitive range for capturing the gaze-contingent shift compared to the wider σ_p = 0.25.

These results suggest that the interplay between sensory-driven inhibition and a stabilizing straight-ahead prior within a standard Mexican-hat recurrent architecture is sufficient to account for the gaze-contingent biases observed in human heading perception

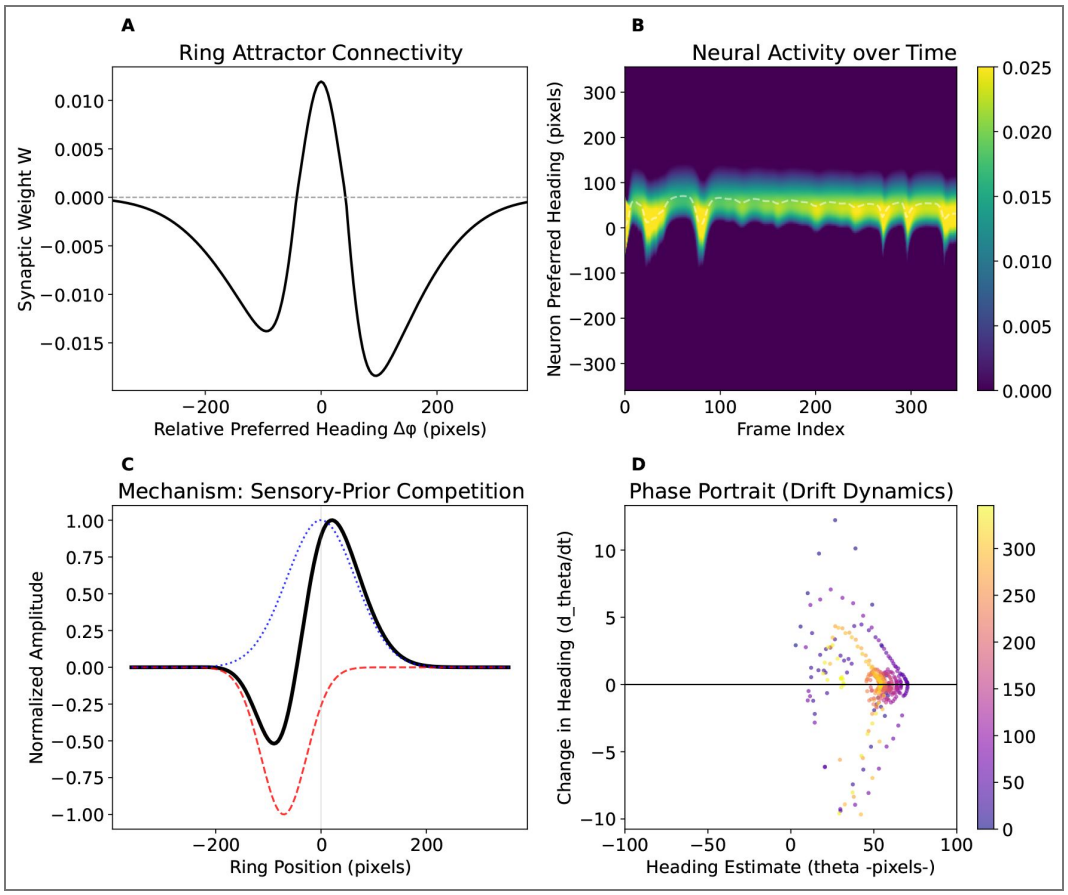

Figure 5. Neural network model of gaze-contingent heading bias.

The parameters used to produce these panels are: $I_0 = 0.03$, $K_p = 0.4$, $\sigma_p = 0.18$, $\sigma_k = 0.12$. (A) Ring attractor connectivity showing synaptic weight (W) as a function of relative preferred heading ($\Delta\Phi$) in pixels, featuring a central excitatory peak and asymmetric inhibitory sourround. While we present results using asymmetric connectivity, no significant differences were observed between asymmetric and symmetric configurations. (B) Heatmap of neural activity across neuron preferred headings (y-axis) over the frame index (x-axis), with a dashed white line indicating the decoded population heading estimate for the trial in which gaze directed 4 m to the left. (C) Mechanism of sensory-prior competition plotting normalized amplitude against ring position in pixels, illustrating the spatial alignment of the straight-ahead prior (blue dotted line), gaze-centered inhibition (red dashed line), and the resulting neural activity bump (black solid line). (D) Phase portrait showing the change in heading ($d\theta/dt$) versus the heading estimate (θ) in pixels, with a horizontal line at zero marking the convergence of the trajectory toward a stable fixed point The color bar denotes frame number.

Discussion

Retinal Curl as a Functional Control Variable revealed by a bias

Our findings provide empirical evidence that retinal curl has a functional role in the perceived heading. Manipulating the retinal flow to which the participants were exposed affects the perceived heading in predictable ways. This main finding suggests that curl could be used as a control signal rather than being a "nuisance" component of the optic flow that must be filtered out. While previous work has emphasized the decomposition of flow into translational and rotational components to recover heading (Beintema et al., 2004; Heeger and Jepson, 1992; Longuet-Higgins and Prazdny, 1980), we show that the visual system exploits the curl generated by gaze stabilization to perceive heading. This is consistent with recent findings in real-world walking, where retinal curl magnitude in the fovea specifies body trajectory relative to gaze (Matthis et al., 2022). In this sense, curl alone combined with extra-retinal information (e.g. rotation of eye, or head or both relative to the body) provides heading information.

Importantly, the functional nature of the curl signal is demonstrated by our curl cancellation condition. When the naturally occurring curl was counteracted, the systematic heading bias disappeared. This confirms that the bias is not an artifact of gaze eccentricity itself, but a direct consequence of the underlying flow geometry due to sustained gaze stabilization. In addition, the flow manipulation reveals a direct link between the amount of curl and perceived heading: not only did the bias opposite to fixation disappear when the curl was cancelled or nullified but also the bias re-appeared towards fixation (opposite to the bias in the unaltered flow condition) when the curl was over-cancelled (see paths in the middle row in Figure 3). This manipulation led to counter-intuitive findings like the reported path was closer to the physical/experimental one in the cancelled condition (last column in Figure 3)

Perceiving a curved heading when exposed to a physically straight path is not new. Several studies have reported heading errors consistent with this misperception in the past (e.g. (Royden et al., 1994, 1992; van den Berg and Brenner, 1994)). A mathematical analysis based on image point positions had been put forward to explain the perceived curvature when exposed to straight paths (Royden, 1994). However, there are large differences between the bias we report and theirs. First, the biases reported in the previous cited studies appear in simulated rotation conditions and when the simulated rotation rate is larger than 1°/s. When simulated rotation rate was smaller than 1°/s heading was accurate and no different than conditions with actual eye/head rotation (Warren and Hannon, 1990). None of these two conditions apply in our study. Our retinal rotation was contingent with actual gaze rotation and the rotation rate was always less than 1°/s (see figure S2 of the SI). Moreover, the stimulus duration in all these studies was smaller than 2 or 3 seconds, while our stimulus duration was larger than 10 seconds. Our biases take some time to build up. Looking at the time course of the heading responses in our experiment, the biases show up after 3 m or 5 s (see Figure 2). Finally, most of previous studies examined heading based on a binary response after the last frame, while we used a continuous response. We believe that these differences lead to the most important difference between our heading bias and the previously reported ones: the bias direction. While our bias is in the direction opposite to fixation, the biases in previous studies were consistent with the direction of fixation caused by the simulated rotation. Since, it was sustained gaze stabilization that allowed us to reveal this bias, our findings are more consistent with time varying signal account of the use of optic flow (Burlingham and Heeger, 2020) rather than the mediation of instant flow (Cutting et al., 1992; Li and Cheng, 2011) to recover heading. We encourage the reader to experience this bias firsthand: when walking in an open space while maintaining a sustained gaze on a lateral landmark, one may notice a drift in their trajectory.

Active Steering vs. Heading Recovery

An important distinction in navigation research lies between the passive perception of heading (estimating *where am I going?*) and the active control of steering (adjusting *where do I want to go?*) (Goodridge et al., 2023; Powell et al., 2024; Tuhkanen et al., 2021; Wilkie and Wann, 2003a,

2003b, 2002). While traditional models posit that accurate heading recovery is a prerequisite for steering or controlling navigation and, given that, our experimental task requires heading, our results support the view that explicit heading estimation is not strictly necessary for control (Wilkie and Wann, 2003a). Our curl-based controller reinforces this distinction: it can successfully maintain a trajectory by simply nulling the error signal derived from retinal curl, without ever needing to calculate the instantaneous focus of expansion or resolve the ambiguity of the heading point. We reproduced the results of Wilkie and Wann (2003a) (experiment 3) of active steering towards 6 targets at different eccentricities (see Figure S17 and table S3 in the SI). We introduced the heading/gaze discrepancy in these eccentricities and by nullifying the curl the controller could find solutions very close to the empirical steering paths reported in Wilkie and Wann (2003a). This suggests that the "bias" observed in our perceived headings may reflect the operation of a control law optimized for action rather than a failure of a perceptual system designed for passive estimation.

This does not necessarily mean that the information used by the participants in Wilkie and Wann (2003a) is the curl signal. They argued that steering can be achieved through a mechanism that can combine different information in addition to retinal flow, like visual direction or perceived target location (Rushton et al., 1998). This is also a possibility in our experiment, that heading responses were based on the perceived location of our experimental fixation point. However, the different flow manipulation conditions (cancelled and over-cancelled) were able to completely change the perceived instant heading. While visual direction is undoubtedly a cue, our results indicate that retinal flow dynamics can override it. In our experiments, even when the fixation point remained eccentric (providing a stronger visual direction cue), cancelling the retinal curl eliminated the bias or induced in the other direction (over-cancelling condition). This runs counter the possibility of participants using the visual direction of the fixation point. The large field of view in our experiment (>90º) could also have contributed to exploit retinal flow to a larger extent. This suggests that behaviors previously attributed to visual direction strategies may instead be driven by the minimization of retinal curl, offering a parsimonious account of steering control that unifies fixation and flow processing (Matthis et al., 2022).

Re-evaluating the Focus of Expansion (FOE)

Overall, our results suggest, at least in simulated walking, that when rotation in the retinal flow is real and active, the system treats curl as a signal to be regulated rather than noise to be subtracted. Historically, heading perception theory has been dominated by the extraction and use of the Focus of Expansion (FOE) (Gibson, 1950; Warren et al., 1988; Warren and Hannon, 1990). However, the relevance of the FOE in natural, active vision has been increasingly questioned (Matthis et al., 2022; Muller et al., 2023). In the presence of eye movements, the retinal FOE is rarely aligned with the heading, creating a complex computational problem, known as the rotational problem (Cutting et al., 1992; Koenderink and Van Doorn, 1981; Longuet-Higgins and Prazdny, 1980; Royden et al., 1994).

Although our results support the view that explicit FOE extraction may not be required for controlling navigation or locomotion, this does not mean that FOE is render useless. Instead, the visual system may rely on different flow field structures and FOE, or divergence can be one of them (Matthis et al., 2022). For example, in more complex scenes, pseudo-FOE can lead to heading biases (Layton and Fajen, 2016) due to relative motion between objects in the background. FOE can also be more relevant in high-speed locomotion, in which momentary changes of gaze direction due to rotation of head or eyes, maybe less frequent than in walking, leading to a less disrupted head-based flow (Muller et al., 2023). This shift away from FOE-based heading is further supported by recent computational models (Layton and Browning, 2014). The continuum of observed cells (from radial to circular) in MSTd would simultaneously code curvature for trajectory estimation and heading across the neural population, with curvature encoded through the spirality of the most active cell and heading through the visuotopic location of its receptive field center.

Neural Implementation of Curl-Based Steering

We propose a relatively simple neural model that can reproduce the bias when exposed to straight ahead video scene, providing a biologically plausible bridge between the phenomenological model, the controller, and neural implementation. This neural circuit model demonstrates that gaze-contingent heading biases can emerge from fundamental principles of sensory integration in cortical networks. Therefore, standard center–surround (Mexican-hat) recurrent connectivity is sufficient to generate the observed bias. The inherent push–pull dynamics transform localized inhibition into a global shift of the activity bump. The key insight is that weak straight-ahead priors, combined with standard cortical connectivity with gaze-centered inhibition, are sufficient to explain the observed psychophysical biases. The model integrates direction-selective motion processing—analogous to area MSTd, where neurons are known to be sensitive to spiral and curl patterns (Duffy and Wurtz, 1991; Graziano et al., 1994) —with a parietal ring-attractor network. Spiral tuning also appears in VIP Schaafsma and Duysens (1996) and 7a Read and Siegel (1997), to which MSTd projects Born and Bradley (2005).

Conclusion

We conclude that the rotational component of optic flow (curl), generated during gaze stabilization, is an actively used signal to control heading. It acts as a navigational cue rather than noise, as evidenced by the elimination of steering biases when curl is experimentally cancelled. Our findings challenge the necessity of explicitly extracting the Focus of Expansion for online control of locomotion. The observed behaviors are supported by a neural model based on established properties of motion processing areas. The interaction between sensory flow inputs and internal priors within a recurrent network suffices to explain the gaze-contingent biases observed in our experimental data.

Data availability

All the data and code for the analysis are available at *https://osf.io/b37rg/overview?view_only=35b748a234b7459a9e9df1b41a2d4a44*

Additional files

Supplementary Information.

Video S1. Retinal flow pattern during forward translation with fixation on a target (white dot) located to the left of the heading. For optimal viewing, the display should be centered relative to the observer with a field of view exceeding 50º (e.g., a 60 cm screen viewed from a distance of 60 cm).

Video S2. The same as video S1, retinal flow pattern during forward translation and fixating a target (white dot) located to the left of the current path, but, the rotational component of the flow has been counteracted, so there is very little curl around the fovea. For optimal viewing, the display should be centered relative to the observer with a field of view exceeding 50º (e.g., a 60 cm screen viewed from a distance of 60 cm).

Video S3. Controller illustration: how heading is corrected to match gaze by applying equation 5 in the main text.

Additional information

Funding

Funder	Grant reference number	Author
MEC | Agencia Estatal de Investigación (AEI)	MICIU/AEI/10.13039/501100011	Konytessa I Zorpala

Author ORCID iDs

Joan López-Moliner: https://orcid.org/0000-0001-5040-8889

References

Angelaki DE, Hess BJ (2005) Self-motion-induced eye movements: Effects on visual acuity and navigation. *Nature Reviews Neuroscience* **6**:966976 https://doi.org/10.1038/nrn1804 | PubMed

Banks MS, Ehrlich SM, Backus BT, Crowell JA (1996) Estimating heading during real and simulated eye movements. *Vision research* **36**:431443 https://doi.org/10.1016/0042-6989(95)00122-0 | PubMed

Beintema JA, van den Berg AV (1998) Heading detection using motion templates and eye velocity gain fields. *Vision research* **38**:21552179 https://doi.org/10.1016/S0042-6989(97)00428-8 | PubMed

Beintema JA, Van Den Berg AV, Lappe M. (2004) Circular Receptive Field Structures for Flow Analysis and Heading Detection. In: Vaina LM, Beardsley SA, Rushton SK (Eds). *Optic Flow and Beyond* Springer Netherlands. pp. 223-248 https://doi.org/10.1007/978-1-4020-2092-6_10

Ben-Yishai R, Bar-Or RL, Sompolinsky H. (1995) Theory of orientation tuning in vi-sual cortex. *Proceedings of the National Academy of Sciences* **92**:3844-3848 https://doi.org/10.1073/pnas.92.9.3844 | PubMed

van den Berg AV, Beintema JA (2000) The mechanism of interaction between visual flow and eye velocity signals for heading perception. *Neuron* **26**:747752 https://doi.org/10.1016/S0896-6273(00)81210-6

Born RT, Bradley DC (2005) Structure And Function Of Visual Area Mt. *Annual Review of Neuroscience* **28**:157-189 https://doi.org/10.1146/annurev.neuro.26.041002.131052 | PubMed

Bremmer F, Churan J, Lappe M. (2017) Heading representations in primates are compressed by saccades. *Nature Communications* **8**:920 https://doi.org/10.1038/s41467-017-01021-5 | PubMed

Britten KH (2008) Mechanisms of Self-Motion Perception. *Annual Review of Neuroscience* **31**:389-410 https://doi.org/10.1146/annurev.neuro.29.051605.112953 | PubMed

Burlingham CS, Heeger DJ (2020) Heading perception depends on time-varying evolution of optic flow. *Proceedings of the National Academy of Sciences* **117**:3316133169 https://doi.org/10.1073/pnas.2022984117 | PubMed

Calow D, Lappe M. (2008) Efficient encoding of natural optic flow. *Network: Computation in Neural Systems* **19**:183-212 https://doi.org/10.1080/09548980802368764 | PubMed

Cutting JE (1986) *Perception with an eye for motion* Cambridge: MIT Press.

Cutting JE, Springer K, Braren PA, Johnson SH (1992) Wayfinding on foot from information in retinal, not optical, flow. *Journal of Experimental Psychology: General* **121**:41 https://doi.org/10.1037/0096-3445.121.1.41 | PubMed

Dong DW, Atick JJ (1995) Statistics of natural time-varying images. *Network: Computation in Neural Systems* **6**:345358 https://doi.org/10.1088/0954-898X_6_3_003

Duffy CJ, Wurtz RH (1991) Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology* **65**:1329-1345 https://doi.org/10.1152/jn.1991.65.6.1329 | PubMed

Fajen BR, Warren WH (2007) Behavioral dynamics of intercepting a moving target. *Experimental brain research* **180**:303-19 https://doi.org/10.1007/s00221-007-0859-6 | PubMed

Gibson JJ (1950) *The perception of the visual world* Houghton Mifflin.

Giorgino T. (2009) Computing and visualizing dynamic time warping alignments in r: The dtw package. *Journal of statistical Software* **31**:124 https://doi.org/10.18637/jss.v031.i07

Glennerster A, Hansard ME, Fitzgibbon AW (2001) Fixation could simplify, not complicate, the interpretation of retinal flow. *Vision research* **41**:815834 https://doi.org/10.1016/s0042-6989(00)00300-x

Goodridge CM, Billington J, Markkula G, Wilkie RM (2023) Error accumulation when steering toward curves. *Journal of experimental psychology: human perception and performance* **49**:821 https://doi.org/10.1037/xhp0001101 | PubMed

Graziano MS, Andersen RA, Snowden RJ (1994) Tuning of MST neurons to spiral motions. *Journal of Neuroscience* **14**:5467 https://doi.org/10.1523/JNEUROSCI.14-01-00054.1994 | PubMed

Heeger DJ, Jepson AD (1992) Subspace methods for recovering rigid motion I: Algorithm and implementation. *International Journal of Computer Vision* **7**:95-117 https://doi.org/10.1007/BF00128130

Kaminiarz A, Schlack A, Hoffmann K-P, Lappe M, Bremmer F. (2014) Visual selectivity for heading in the macaque ventral intraparietal area. *Journal of Neurophysiology* **112**:2470-2480 https://doi.org/10.1152/jn.00410.2014 | PubMed

Koenderink JJ, Van Doorn AJ (1981) Exterospecific component of the motion parallax field. *Journal of the Optical Society of America* **71**:953957 https://doi.org/10.1364/JOSA.71.000953 | PubMed

Lappe M. (1998) A model of the combination of optic flow and extraretinal eye movement signals in primate extrastriate visual cortex: Neural model of self-motion from optic flow and extraretinal cues. *Neural Networks* **11**:397414 https://doi.org/10.1016/s0893-6080(98)00013-6

Lappe M, Rauschecker JP (1993) A neural network for the processing of optic flow from ego-motion in man and higher mammals. *Neural Computation* **5**:374391 https://doi.org/10.1162/neco.1993.5.3.374

Layton OW, Browning NA (2014) A unified model of heading and path perception in primate MSTd. *PLoS computational biology* **10**:e1003476 https://doi.org/10.1371/journal.pcbi.1003476 | PubMed

Layton OW, Fajen BR (2016) Sources of bias in the perception of heading in the presence of moving objects: Object-based and border-based discrepancies. *Journal of Vision* **16**:99 https://doi.org/10.1167/16.1.9 | PubMed

Li L, Cheng JC (2011) Perceiving path from optic flow. *Journal of Vision* **11**:2222 https://doi.org/10.1167/11.1.22 | PubMed

Li L, Warren WH (2002) Retinal flow is sufficient for steering during observer rotation. *Psychological Science* **13**:485-491 https://doi.org/10.1111/1467-9280.00486 | PubMed

Li L, Warren WHJ (2000) Perception of heading during rotation: Sufficiency of dense motion parallax and reference objects. *Vision research* **40**:3873-94 https://doi.org/10.1016/s0042-6989(00)00196-6 | PubMed

Longuet-Higgins HC, Prazdny K. (1980) The interpretation of a moving retinal image. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **208**:385397 https://doi.org/10.1098/rspb.1980.0057 | PubMed

Matthis JS, Muller KS, Bonnen KL, Hayhoe MM (2022) Retinal optic flow during natural locomotion. *PLOS Computational Biology* **18**:e1009575 https://doi.org/10.1371/journal.pcbi.1009575 | PubMed

Matthis JS, Yates JL, Hayhoe MM (2018) Gaze and the control of foot placement when walking in natural terrain. *Current Biology* **28**:12241233 https://doi.org/10.1016/j.cub.2018.03.008 | PubMed

Muller KS, Matthis J, Bonnen K, Cormack LK, Huk AC, Hayhoe M. (2023) Retinal motion statistics during natural locomotion. *eLife* **12**:e82410 https://doi.org/10.7554/eLife.82410 | PubMed

Perrone JA, Stone LS (1994) A model of self-motion estimation within primate extrastriate visual cortex. *Vision research* **34**:29172938 https://doi.org/10.1016/0042-6989(94)90060-4 | PubMed

Powell NV, Marshall X, Diaz GJ, Fajen BR (2024) Coordination of gaze and action during high-speed steering and obstacle avoidance. *PLoS one* **19**:e0289855 https://doi.org/10.1371/journal.pone.0289855 | PubMed

Read HL, Siegel RM (1997) Modulation of responses to optic flow in area 7a by retinotopic and oculomotor cues in monkey. *Cerebral cortex (New York, NY: 1991)* **7**:647661 https://doi.org/10.1093/cercor/7.7.647 | PubMed

Royden CS (1994) Analysis of misperceived observer motion during simulated eye rotations. *Vision research* **34**:32153222 https://doi.org/10.1016/0042-6989(94)90085-X | PubMed

Royden CS, Banks MS, Crowell JA (1992) The perception of heading during eye movements. *Nature* **360**:583585 https://doi.org/10.1038/360583a0 | PubMed

Royden CS, Crowell JA, Banks MS (1994) Estimating heading during eye movements. *Vision research* **34**:31973214 https://doi.org/10.1016/0042-6989(94)90084-1 | PubMed

Rushton SK, Harris JM, Lloyd MR, Wann JP (1998) Guidance of locomotion on foot uses perceived target location rather than optic flow. *Current Biology* **8**:11911194 https://doi.org/10.1016/S0960-9822(07)00492-7 | PubMed

Schaafsma SJ, Duysens J. (1996) Neurons in the ventral intraparietal area of awake macaque monkey closely resemble neurons in the dorsal part of the medial superior temporal area in their responses to optic flow patterns. *Journal of Neurophysiology* **76**:4056-4068 https://doi.org/10.1152/jn.1996.76.6.4056 | PubMed

Simoncelli EP, Heeger DJ (1998) A model of neuronal responses in visual area MT. *Vision research* **38**:743761 https://doi.org/10.1016/S0042-6989(97)00183-1 | PubMed

Tuhkanen S, Pekkanen J, Wilkie RM, Lappi O. (2021) Visual anticipation of the future path: Predictive gaze and steering. *Journal of Vision* **21**:2525 https://doi.org/10.1167/jov.21.8.25 | PubMed

van den Berg AV, Brenner E. (1994) Why two eyes are better than one for judgements of heading. *Nature* **371**:700702 https://doi.org/10.1038/371700a0 | PubMed

Warren WH, Hannon DJ (1990) Eye movements and optical flow. *Josa a* **7**:160169 https://doi.org/10.1364/JOSAA.7.000160 | PubMed

Warren WH, Hannon DJ (1988) Direction of self-motion is perceived from optical flow. *Nature* **336**:162163 https://doi.org/10.1038/336162a0

Warren WH, Morris MW, Kalish M. (1988) Perception of translational heading from optical flow. *Journal of Experimental Psychology: Human Perception and Performance* **14**:646 https://doi.org/10.1037/0096-1523.14.4.646 | PubMed

Wilkie RM, Wann JP (2003a) Controlling steering and judging heading: Retinal flow, visual direction, and extraretinal information. *Journal of Experimental Psychology: Human Perception & Performance* **29**:363-378 https://doi.org/10.1037/0096-1523.29.2.363 | PubMed

Wilkie RM, Wann JP (2003b) Eye-movements aid the control of locomotion. *Journal of vision* **3**:33 https://doi.org/10.1167/3.11.3 | PubMed

Wilkie RM, Wann JP (2002) Driving as night falls: The contribution of retinal flow and visual direction to the control of steering. *Current Biology* **12**:2014-2017 https://doi.org/10.1016/S0960-9822(02)01337-4

Zhang K. (1996) Representation of spatial orientation by the intrinsic dynamics of the head-direction cell ensemble: A theory. *Journal of Neuroscience* **16**:21122126 https://doi.org/10.1523/JNEUROSCI.16-06-02112.1996 | PubMed

Peer reviews

Reviewer #1 (Public review):

Summary:

This carefully executed study uncovers the functional relevance of curl signals that impinge on the retina every time an observer's gaze direction and movement direction are not aligned.

Strengths:

This finding is important, highlighting the functional role of an abundant incidental signal (curl in retinal motion) that has thus far believed to be a nuisance that needs to be filtered out of the retinal motion stream.

The study's evidence is compelling: a combination of psychophysical experiments and critical manipulations, control theory and neural modeling, which together make an internally consistent and biologically plausible case for the role of curl signals in estimating heading direction.

This study uncovers the functional relevance of curl signals that occur on the retina when an observer is moving, and gaze is not straight ahead. The experimental and modeling results clearly go beyond previous studies and significantly advance our understanding of vision-based navigation.

Another clear strength is that the study uses tightly controlled experimental manipulation to provide strong test cases for the hypothesis that curl is used for visual navigation. These conditions are important to constrain the proposed model (and future models) of heading control.

The modeling is very clearly described, and the modeling and analysis code is published and freely available. The authors go beyond a back-of-the-envelope control model and show how it might be implemented at the neural-circuit level. The model is biologically plausible.

Weaknesses:

The discussion would benefit from an extension of the implications of the study and predictions of their model.

https://doi.org/10.7554/eLife.110770.1.sa3

Reviewer #2 (Public review):

This study examines how curl in the retinal flow field can be used as a control variable for estimating and controlling the heading of a moving observer. The basic idea (which is not entirely new, see Matthis et al. 2022) is that translation along a path with eccentric gaze (meaning that the subject is not heading toward the point they are looking at) produces a pattern of optic flow on the retina with a rotational component around the point of fixation (which can be captured by the mathematical "curl" operator). The sign and magnitude of retinal curl vary with heading relative to the point of fixation, such that curl can be used as a control variable to steer rightward or leftward to move toward the fixated target. The authors perform behavioral experiments and show that there are biases in perceived heading that seem to be largely governed by retinal curl. They also show that a simple controller model can use curl to steer toward a target, and they provide a neural network model that provides a biologically plausible implementation of the controller (although there are some questions about that).

There is a core of interesting work here that I think can be important to the field. However, there is a lack of clarity on several important fronts, including design of the behavioral experiments, presentation of the behavioral data, conceptual framing of what curl can and cannot do, etc. Equally importantly, the manuscript is not written in a manner that will make it accessible to most vision scientists. I consider myself to be pretty knowledgeable about optic flow, and I had to read most of the manuscript 3 or 4 times to be able to understand the bulk of it. And my experience is that most vision scientists do not understand optic flow well, so I fear that most of the readers that the authors should want to reach would struggle to understand the work. As written, this is mainly going to make an impact on a handful of optic

flow gurus. Thus, I consider that this manuscript will need a major overhaul to clarify important issues and make it more accessible.

Major issues:

(1) The manuscript contains inconsistent, if not misleading, messaging about what information retinal curl does, and does not, provide regarding heading estimation. In the Abstract, the authors state: "We propose an alternative: the visual system utilizes retinal curl directly to estimate heading, rendering the explicit recovery of the FOE unnecessary." Based on my understanding of the rest of the manuscript, I find this statement to be a misrepresentation for two main reasons:

a) To "directly estimate heading" relative to what? When not qualified, most people interpret "heading" to mean an observer's heading relative to the world (or some allocentric reference frame). But retinal curl only gives information about an observer's heading relative to the point on which their eyes are fixated. Moreover, that point of fixation will change every few hundred milliseconds in natural viewing, so the retinal curl will change with each new fixation even as heading relative to the world remains unchanged. So I think most readers would grossly misinterpret the claim that retinal curl can be used "directly to estimate heading". Indeed, in the authors' controller model, the initial heading needs to be given, and then the controller can work. But from where does the visual system get the initial heading, since it does not come from curl? These issues are left hanging. Thus, while curl can provide a very useful input for steering toward a fixated target, other signals are needed to estimate heading relative to the world. This has to be made much clearer early on, and a conceptual schematic diagram might help. Also, the authors generally do not specify the reference frame of the variables they are talking about, leaving lots of room for misinterpretations. It should be clear each time they are talking about a variable, such as heading, whether it is relative to the fixation target, body, world, etc.

b) It seems to me that retinal curl will depend on other variables, in addition to heading relative to the fixation target. For example, it seems to me that the magnitude of retinal curl will depend on self-motion speed, the depth structure of the scene, the angle of elevation of the fixated target, and perhaps others. This is not discussed at all, and many readers would get the misguided impression that there is a 1:1 mapping from curl to heading (relative to fixation). If I am right that this is not correct, it means that retinal curl can tell the observer whether to steer right or left to move toward the fixated target, but it cannot tell them how much to steer. Indeed, in the authors' controller model, there is a free parameter that calibrates curl to angle. It makes sense that this works to fit trajectory data that are given from a fixed environment, but it is unclear how the brain would use retinal curl to control steering when these other variables are uncertain or changing unpredictably. Moreover, how does the system change the mapping from curl to steering command as the location of fixation changes relative to the current heading? These are issues that need to be brought up in framing the problem and discussed at some length. If the authors can show mathematically that retinal curl is only dependent on heading (relative to fixation) and not any of these other variables, it would be very valuable to show the equations for this relationship.

(2) The description of the behavioral experiment and presentation of behavioral data leaves a lot to be desired.

a) First, it is stated (line 158) that "Participants continuously reported their perceived direction of self-motion while maintaining fixation on the yellow dot." Again, the reference frame is completely unspecified. Participants were reporting their perceived heading relative to what? The fixation target? The world? What exactly were the instructions given to the subjects to perform the task? Based on the description of how perceived paths are computed (line 166-), it seems to be presumed that subjects are reporting their heading relative to the

world because those angles are then converted into x and z coordinates in what I presume is a world-centered reference frame. But how do we know that subjects are accurately reporting their heading relative to the world? What if they are biased in their reports by the location of the fixation target relative to the scene, or by some other reference signal? Is it possible for the authors to rule out the possibility that perceptual biases seen in the unaltered curl condition result from observers not fully adopting the assumed reference frame of the task? If this cannot be firmly excluded, it seems to create problems for the rest of the study.

b) I also feel that there is a mismatch between what the behavioral task requires and what the controller model does. Subjects are apparently asked to report their heading relative to the world, but the controller model only controls their heading relative to the point that they are fixating. I understand how this is resolved in the model, but I think this type of distinction is buried and will not be apparent to most readers. Again, the reference frames of what is being measured and controlled need to be specified explicitly in all parts of the paper, and the authors need to explain how the system would combine curl-based control with some other measures of (at least initial) heading for world-centered heading to be computed. All of the assumptions need to be clearly specified.

c) In addition, I found it frustrating that the authors never present raw perceptual data from the observers. Rather, in Figure 2, we see reconstructed trajectories that are perfectly smooth with no indications of noise whatsoever. Since these paths are computed from the perceptual reports, there must be some noise inherent in them. The figures should represent this uncertainty somehow, and it should be explained how these perfectly smooth trajectories are obtained.

(3) "...the magnitude of retinal curl in the fovea can specify the body trajectory relative to gaze (Matthis et al., 2022)." The main idea put forward by the authors here seems to overlap heavily with this statement that they attribute to Matthis et al. 2022. While I think this paper still adds importantly to the topic, the authors do not discuss how their findings are different from those of Matthis et al. 2022, why they are an important extension, etc. Readers should not have to go read this other paper to have any idea how the present findings are placed in importance relative to the literature.

(4) The analysis and treatment of eye movements is extremely weak. The authors discarded trials for which gaze deviated from the fixation point by more than 3 degrees (which is a LOT given that the eye speeds are generally in the neighborhood of 0.5 deg/sec), and they provide basic stats on the distribution of positions. But this largely misses the point: it is not small position errors that are likely to matter, but rather velocity errors. Even a small amount of retinal slip of the target while it is being pursued will cause image motion that is going to alter the optic flow field around the fixation target. So, for example, the retinal curl field may no longer be centered on the fixation target. How do we know that some of the perceptual biases are not influenced by image motion resulting from imperfect tracking of the fixation target? This needs to be analyzed and discussed.

(5) I found the sections of text comparing the separate and joined fits (starting line 287) to be a bit too rosy. The authors show the separate fits in the main text, and it is not very surprising that these fits are good, given that the model has 30 parameters, and these data are pretty low-dimensional. The authors only show the joined fits in the supplement, and they say that they are almost as good as the separate fits (indeed, they are better in a model comparison sense, but this is 30 parameters vs. 2 parameters). However, when I look at the fits of the joined model in the supplement, I don't find them to be very impressive. In particular, the model grossly misses the data for the straight paths for several subjects (e.g., id5, id6, id8, id10). And fitting the straight paths would presumably be easiest. This implies that the joined model is really missing something and that fitting the curved paths interacts strongly with fitting the data for different fixation target locations on the straight path. I think that the authors should discuss the results a bit more soberly and tone down their conclusions here.

(6) The section of the paper on neural simulations (starting line 387) has a few weaknesses. First, why are only straight paths simulated here? This does not seem to provide a very rigorous test of the model. Second, it is awkward that the simulation results are presented in units of pixels, rather than degrees. Third, the authors seem to downplay the fact that the neural estimates of heading seem to oscillate rather wildly (over a range of hundreds of pixels, whatever that means, see especially Figure S16). It was far from clear to me how an estimate of heading with these large oscillations is useful. It would seem to require that heading estimates are integrated over substantial lengths of time to be reliable. It was therefore unclear how the model produces such smooth paths from these oscillating estimates.

https://doi.org/10.7554/eLife.110770.1.sa2

Reviewer #3 (Public review):

Summary:

This manuscript uses a novel paradigm to demonstrate that rotational motion patterns in the retinal image, called curl, directly influence perception of heading direction. This means that it is not necessary to recover the focus of expansion, defined by the point of zero motion when moving along a straight trajectory toward a target, as is commonly thought.

Strengths:

It has long been accepted that the focus of expansion of the optic flow field generated by self-motion is used to guide heading direction. While there have been many challenges to the need to recover the focus of expansion when gaze is not in the direction of travel, it is still not well understood how retinal motion patterns contribute to heading perception. Recent work has demonstrated the complexity of the retinal motion patterns during natural walking, where body motion adds a rotational component. A rotational component also results from curved paths as well as gaze off the direction of travel. This rotational component is called curl. The primary contribution of this manuscript is to demonstrate convincingly that curl influences perception of heading, and that it is not necessary to recover the focus of expansion.

A strength of the manuscript is that realistic retinal motion patterns are generated by recording the image sequences generated by a walker in a virtual environment, and then using those patterns as stimuli in the experiment. This allows the creation of the more complex flow patterns that are a consequence of the bob and sway of natural walking, which are often considered a minor factor. The elegant experimental design allows direct manipulation of the curl signal, and this in turn directly influences measured heading

perception. Another strength is that the authors ground their findings in control theory and neural computations, using a model that produces human-like path trajectories.

The study is timely, given the long history of this question, together with the growing understanding of the complexity of naturally generated retinal motion and the absence of direct evidence for the way that these motion patterns are used in heading perception. It adds an important piece of evidence for how retina-centered optic flow may be used by the visual system, which is critical for our understanding of motion processing in the brain.

Weaknesses:

The primary limitation of the paper is that it avoids discussion of some of the inevitable complexities of heading perception. The main issue is what exactly is meant by heading. Different behaviors evolve over different timescales. The geometry of retinal motion defines instantaneous heading, which varies widely through the gait cycle. Time-varying information like this is known to be important in the momentary control of balance. Heading can also be thought of as steering the body toward a distant goal, which evolves over longer timescales. The current manuscript appears to be concerned with heading information integrated over a few seconds and seems to provide evidence that heading is indeed integrated over the gait cycle. The issue of the time scale of the computation is touched on, but it is not related to how it might be used in normal walking or what situations it might apply to. Steering toward a distant goal during walking is not a very difficult problem and may not require evaluation of retinal motion, but control of balance is more challenging and may depend critically on curl. Consequently, the timescale of the computation needs to be considered in order to understand what is meant by heading.

https://doi.org/10.7554/eLife.110770.1.sa1

Author Response:

> ***Public Reviews:***
>
> ***Reviewer #1 (Public review):***

We appreciate Reviewer #1's very positive feedback. Incorporating the perspective of 'incidental' sensory signals is a valuable suggestion that aligns perfectly with our findings. We agree that this perspective significantly strengthens the impact of our paper.

In the revised version, we will update the manuscript to bridge these perspectives (the functional role of incidental" sensory signals and the role of retinal flow in navigation). In addition we will elaborate on the potential predictions of the model and possible manipulations that might affect the integration between sensory evidence (curl signal) and straight-ahead prior.

> ***Reviewer #2 (Public review):***

We appreciate the reviewer's feedback regarding the formalization of our reference frames. We agree that certain definitions were implicitly assumed rather than explicitly stated. We will revise the manuscript to provide all necessary self-contained information, ensuring that the geometry of the task response and the definition of heading are unambiguous. Also, we will address the gap between the task response (in world coordinates) and the functional role of the controller, as well as the other points raised by the reviewer.

Major issues:

(1a), (2a) Clarification of Reference Frames

The reviewer asks: “To ‘directly estimate heading’ relative to what?”

In our study, participants were instructed to report their “perceived direction of self-motion” by aligning a rotational encoder (steering wheel) with the direction they felt they were moving within the 3D simulated scene. Consequently, participants reported their instantaneous heading in a world-centered reference frame, from which the 3D trajectories were reconstructed. Since the reviewer had to infer this information, it should be clarified to ensure it is immediately evident.

Participants were informed that the initial heading (i.e. θ_0 in our controller nomenclature) was oriented “straight ahead” relative to their body which was aligned longitudinally with the experimental room. We will modify Figure 1B and revise the Methods section to explicitly clarify this initial alignment and the instructions provided to participants.

In the revised manuscript, we will clarify that while the participant’s report is world-centered, the retinal curl provides a gaze-relative heading signal. Although this was already mentioned, we will emphasize this point. In natural navigation toward a fixated target, a world-centered vector is often unnecessary; an error signal indicating heading relative to fixation is sufficient (as the reviewer also notes). However, the initial alignment of the heading within the 3D scene allows the brain to “calibrate” this internal controller, mapping the retinal curl signal onto the 3D world coordinates required for the task.

The reviewer also asks how we can be certain that participants were reporting in world coordinates rather than an alternative frame, such as “heading relative to the fixation target.” We believe our “Cancelled Curl” (and over-cancelled) conditions provide the most compelling evidence to rule out this alternative. In these conditions, the physical position of the fixation target in the scene remained identical to the unaltered flow condition. If participants were simply reporting heading relative to the fixation target’s spatial location, the observed biases should have persisted regardless of the flow manipulation. Instead, the bias vanished when the curl was removed. This causal evidence proves that the bias is driven by the retinal motion signal (curl) rather than the spatial orientation of the eyes or the target’s position in the scene. Furthermore, the temporal evolution of the response supports a world-centered integration. For simulated straight paths, the perceived heading remains straight for the first few seconds (consistent with the initial world-centered alignment), with biases only emerging after approximately 3 seconds of integration (a point we elaborate on in our response to Reviewer #3). Had participants been responding based on a simple gaze-relative reference frame from the onset, these biases would have manifested significantly earlier. We will incorporate these points into the revised Discussion to better frame our findings alongside other cues, such as the Focus of Expansion (FOE), that contribute to heading estimation.

(1b) The reviewer notes that we must be clear about the relationship between curl and heading (relative to fixation) and the variables that affect curl.

Beyond the discrepancy between heading (θ) and gaze (ψ), curl is geometrically determined by translational self-motion speed (υ), eye height (h), and pitch (α). More specifically curl = (υ sin_ψ_cos α)/h). The derivation will be included in the Supplementary Information. Since h = d_sin_α, where d is the 3D distance to the fixation point, we could express cos α as a function of distance. Certainly, there is not a 1:1 map from curl signal to heading relative to gaze (e.g. $\theta - \psi$). Participant would need to know υ and eye height plus extra-retinal information. Frenz et al (2003, Vis Res.) showed that people can estimate self-motion directly from optic flow, across different simulated eye height and gaze angle; extra-retinal information can, in addition, provide knowledge to (ψ) and (α). It is then plausible that the visual system can use and transform the curl signal from a qualitative directional cue (i.e. steering left or right of fixation) into a quantitative steering command. By combining curl with knowledge of gaze orientation and eye height, the visual system can resolve ambiguities in the flow field and

utilize curl as a more precise error signal for locomotor control. These aspects will be included in the new version.

(2b) Mismatch between task and controller

We thank the reviewer for this point. We have addressed the alignment of the reference frames in our response to Issues 1a and 2a. Once the initial orientation (θ_0) is established in the world frame, the controller model generates steering adjustments that directly translate into heading predictions within that same world reference frame. By treating the perceptual report as an output of the locomotor controller, we resolve the discrepancy between the steering task and the reported heading.

(2c) No raw data provided

We respectfully disagree with the reviewer's interpretation regarding data smoothing. The thin lines in Figure 2 represent the mean 3D paths derived directly from the response variable (θ_0) across trials of identical conditions for each participant (as detailed in the 'Computation of Perceived Path' section). No smoothing or filtering has been applied to these plotted trajectories other than computing the mean across trials. We also wish to remind the reviewer that the raw data and analysis code remain publicly accessible for further inspection. Regarding the visual representation: in earlier versions of the manuscript, we included shaded 95% Confidence Intervals (CIs) in Figure 2. However, this addition rendered the plot overly cluttered and obscured the individual trajectories. We therefore elected to present individual participant means (thin lines) alongside group averages (thick lines) to emphasize inter-subject variability. For clarity, the 95% CIs are explicitly displayed in Figure 3, where the data density is more conducive to shaded areas.

(3) Difference with Matthis et al (2022)

While Matthis et al. (2022) described the existence of retinal curl during walking and which information can provide relative to gaze, Our paper provides the causal link, since we manipulate in real-time (the 'cancelled & overcancelled curl' condition) providing the critical evidence that perceived heading is affected by this signal.

(4) Eye movements analysis

We thank the reviewer for noting that retinal slip (velocity error) is a more critical metric than positional gaze error. We agree that tracking inaccuracies can introduce translational noise into the flow field. The 3° threshold was established based on the eye tracker's specifications and the naturalistic setup (1-meter viewing distance without head stabilization). Across all participants, the mean positional error ranged from 1.016° to 1.5° (1 deg is 2.08 cm in our setup). We also calculated retinal slip values, which ranged from 0.12 to 0.27 deg/s (X dimension) and 0.12 to 0.23 deg/s (Y dimension). These values are comparable to natural oculomotor drift (Kowler et al., 1979) and are understandably small given the low velocity of the fixation target. Consequently, it is highly unlikely that retinal slip influenced the results. Furthermore, assuming that tracking error remained consistent across fixation conditions, any present retinal slip cannot explain why the bias followed the retinal curl manipulation as predicted by the controller. We therefore consider retinal slip to be an unlikely confounding factor.

(5) the separate and joined fits

We thank the reviewer for the opportunity to clarify the logic behind our modeling choices. We acknowledge that the "separate fits" are inherently less informative due to the high number of free parameters relative to the data. Our primary scientific goal was not to achieve perfect descriptive accuracy via 30 parameters, but to test a specific functional hypothesis through the "joint fit."

The Logic of the Joint Fit:

We agree with the reviewer that the joint fit misses some paths in some conditions. Of course, the joint fit reflects a significant compromise. The “Gain” (the weighting of the curl signal) is likely not a static constant but is dynamically tuned based on task demands, confidence in the visual signal, simulated speed, and so on. By using a single Gain parameter, we intentionally ignore this contextual variability to see how much of the behavior can be explained by a “minimalist” controller. In this sense, the 2-parameter joint model is a deliberate attempt to test this limit. By forcing a single Gain parameter to account for all conditions across both straight and curved paths within one flow manipulation (e.g. unaltered flow) we are asking if a single, fixed linear relationship between retinal curl and steering effort/gain can explain the results. We view the joint fit not as a “perfect” model, but as a stronger test of the curl-based control theory. The fact that a 2-parameter model can capture the direction and scale of biases across such a diverse set of conditions (straight/curved paths, five fixation eccentricities) suggests that retinal curl is a robust signal. Upon closer analysis, these discrepancies between the joint model and the data are most pronounced in the over-cancelled condition which is the one when sensory evidence becomes more ecologically inconsistent with the extra-retinal information (gaze direction). While the joint fit successfully demonstrates that a single parameter can capture the general functional role of curl, it fails to account for the complex sensory re-weighting that occurs in ecologically inconsistent conditions (like ‘over-cancelled’ flow). We will update the manuscript to discuss these limitations, framing the model as a parsimonious first-order approximation rather than a complete description of human heading perception based on a minimal set of parameters.

(6) On the neural simulations

We acknowledge that the presentation of the neural model requires more clarity regarding its objectives and its relationship to the behavioral data.

We first wish to clarify the intended scope of the neural ring-attractor model. Our primary goal was not to provide a comprehensive account of behavioral performance across all conditions (which is the role of the controller model), but rather to demonstrate a biologically plausible mechanism that explains the emergence of the “Opposite-to-Gaze” bias. While the controller demonstrates that the bias follows a specific control law, the neural model shows how such a law can emerge from known primate neurophysiology, specifically, spiral-tuned MSTd neurons, gaze-contingent inhibition, and an egocentric “straight-ahead” prior.

Why Straight Paths are Sufficient for this Objective. The reviewer asks why only straight paths were simulated. In our study, the straight-path condition with eccentric gaze is the purest test of the bias mechanism. Simulating the straight paths allowed us to isolate the interaction between foveal inhibition and the straight-ahead prior without the confounding variable of path-curvature flow. Given the complexity of the neural network’s parameter space, we focused on these conditions to provide a clear neuro-plausible explanation.

Units: Pixels vs. Degrees. We acknowledge that the use of “pixels” in the plots of internal neural dynamics may appear awkward. The neural network operates on input stimuli that are defined by the pixel resolution of the videos used in the simulations, we used pixels as the native coordinate system to describe the movement of activity peaks within the network’s internal “map.”

Behavioral Output (Meters): Importantly, the final heading estimates produced by the network are not left in pixels. We use a pinhole camera model to reconstruct the 3D trajectories from the neural activity. These results are expressed in meters, allowing for a direct comparison with the human behavioral data.

Addressing Wild Oscillations and Smooth Paths. The oscillations observed in the instantaneous heading estimates reflect the stochastic nature of the population peak when tracking high-frequency sensory inputs. In our model, the synaptic time constant (τ) was kept relatively small to ensure a fast, low-latency response to changes in self-motion. While increasing τ would have produced smoother internal dynamics, it would also have introduced delays into the control loop. Instead, we chose to maintain this high sensory responsiveness and applied a temporal moving average later to the network's decoding to reconstruct the 3D trajectories.

In addition, the neural activity over time is shown in two ways: the heatmap shows the neuron with preferred heading (one can see more oscillations, specially when the fixation point is closer to the centre (eccentricities -2 and 2), due to larger competition between the sensory evidence and the straight-ahead prior. The other way is the decoded heading. In the ring-attractor model, the decoded heading $(\widehat{\phi})$ is not determined by a single neuron but is calculated using a population vector average (equation 19). By summing across the entire population, the decoder effectively integrates sensory evidence from many neurons simultaneously. One can appreciate (see e.g. Fig. 5B) that averaged decoding, leads to a smoother resulting estimate (the white dashed line, whose visibility will be improved in the revised version). Behavioral work by Burr and Santoro (2001) suggests that global motion signals (divergence and rotation in optic flow) are integrated over much longer timescales—roughly 1000ms to 3000ms—compared to local motion units (~200ms).

See also our comment on temporal integration in the responses to reviewer #3.

> ***Reviewer #3 (Public review):***

We thank Reviewer #3 the comments regarding the definition of heading at different time scales, the role of the gait cycle, and the temporal integration of the curl signal. They will help us refine the manuscript's core arguments.

We agree that "heading" must be precisely defined within the context of the differing temporal demands of balance and steering. While instantaneous retinal motion provides the high-frequency feedback necessary for momentary postural adjustments and balance, our study is concerned with heading as a gaze-relative signal used for the continuous control of a locomotor trajectory. As such, we will revise the manuscript to specify that the perceived heading measured in our task reflects a signal integrated over the gait cycle to filter out the oscillatory noise induced by head bob and sway.

The reviewer correctly notes that gait-induced head bob and sway produce high-frequency oscillations in the curl signal, yet our behavioral results show smooth, slowly evolving biases. The visual system does not react to "instantaneous" curl, which would lead to jittery, unstable heading estimates. Instead, it integrates flow over a timescale roughly commensurate with a full gait cycle (~500–1000ms). This implies a significant temporal integration process. This temporal integration is consistent with evidence (Burr and Santoro,2001, Vis Res) indicating that optic flow signals (radial and rotational components) are integrated over windows of approximately up to 3 seconds to ensure perceptual stability. Neurally, this likely involves the projection from area MSTd to the Ventral Intraparietal area (VIP), a pathway where fast, eye-centered sensory inputs are transformed into stable, body-centered representations suitable for guiding long-term steering behavior (Chen et al. 2011, JNeurosci.). By grounding our definition of heading in these specific temporal and neural constraints, we aim to clarify how the visual system exploits retinal curl for goal-directed action in natural, dynamic environments and relate our findings to recent studies addressing the role of retinal motion on balance (Powell et al. 2026 Bioarx).

In our implementation, we explicitly address the high-frequency noise introduced by gait dynamics by smoothing the retinal curl signals computed from the stimulus videos before they are fed into the controller. This temporal filtering allows the fit of the controller's prediction to the response data while remaining robust to the rapid fluctuations of head bob and sway. In contrast, the neural ring-attractor model would not require an external smoothing step; instead, the integration is an emergent property of the system's architecture that can be controlled with different parameters. The dynamics of the synaptic weights and the characteristic "leak" in the population activity naturally implement a leaky integration of sensory evidence, ensuring that the decoded heading reflects a sustained estimate rather than an instantaneous response to visual noise.

https://doi.org/10.7554/eLife.110770.1.sa0