Heading Perception Depends on Time-Varying Evolution of Optic Flow

Charlie S. Burlingham[1*] and David J. Heeger[1,2*]

There is considerable support for the hypothesis that heading perception is mediated by instantaneous optic flow. This hypothesis, however, has never been tested. We introduce a novel method, termed "phase motion," for generating a stimulus that conveys a single instantaneous optic flow field, even though the stimulus is presented for an extended period of time. In this experiment, observers viewed stimulus videos and performed a forced choice heading discrimination task. For phase motion stimuli, observers made large errors in their judgments of heading direction. This suggests that instantaneous optic flow is insufficient for heading perception. These errors were mostly eliminated when the velocity of phase motion was varied over time to convey an evolving sequence of flow fields. This demonstrates that heading perception relies on the time-varying evolution of optic flow.

Optic flow, the change in the retinal image caused by an observer's movement relative to the environment, conveys information about self-motion and the structure of the visual scene[1–12]. James Gibson was the first to describe this phenomenon in detail, suggesting that image motion can be decomposed into a field of two-dimensional velocity vectors called the "optical flow field"[1]. When an observer translates in a given direction along a straight path, the flow field radiates out from a singular point in the image, or singularity, called the focus of expansion (FOE; Fig. 1b). It is well known that under such conditions, one can accurately estimate the direction of heading by simply locating the FOE. However, if there is angular rotation in addition to heading, such as the rotation caused by a head or eye movement, the singularity in the flow field will be displaced such that it no longer corresponds to the heading direction (Fig. 1c, d). In this case, if one estimated heading by locating the singularity, the estimate would be biased away from the true heading direction. This is known as the rotation problem[13].

Computer vision researchers have developed a variety of algorithms for extracting observer translation and rotation from image motion, nearly all of which rely on the instantaneous velocity of the flow field[5,9,14–23], with few exceptions[24–25]. However, it is unknown whether these algorithms are commensurate with the neural computations underlying heading perception.

The consensus of opinion in the experimental literature is that accurate estimates of heading may be derived from a single instantaneous flow field, in the absence of additional information[26,22]. There are, however, systematic biases in heading perception. Rotation (e.g., corresponding to a simulated eye movement) can bias heading judgments[6,8], with the amount of bias in heading estimates proportional to the magnitude of rotation[27]. Errors in heading discrimination have been reported to be greater when eye movements are simulated versus when they are real, which suggests that heading information from optic flow is impoverished and that observers make use of extra-retinal information[27–29], although this is somewhat controversial[22,30]. Regardless, to date, no one has tested whether heading perception (even with these biases) is based on instantaneous optic flow, or on the information available in how the flow field evolves over time. Some have suggested that heading estimates rely on information accumulated over time[31-33], but no one has investigated the role of time-varying optic flow without confounding it with stimulus duration (i.e., the duration of evidence accumulation).

In this study, we employed a novel application of an image processing technique that ensured that only the instantaneous velocity of optic flow was available to observers, even though the stimulus was presented for an extended period of time. We called this condition "phase motion," because the phases within each stimulus patch shifted inside a stationary envelope (at a constant rate), causing a percept of motion in the absence of veridical movement[34]. We found that observers exhibited large biases in heading judgments when only the instantaneous velocity of the flow field was available. In a second condition, "time-varying phase motion," we varied the velocity of the phase motion over time to match the time evolution of a sequence of flow fields. We found that this dynamical information allowed observers to compensate for the confounding effect of rotation on momentary optic flow, making heading perception nearly veridical.

Results

Observers (N = 5) performed a forced-choice heading discrimination task. They viewed stimulus videos that contained varying levels of information about the time evolution of the optic flow field (Fig. 2). Specifically, they judged whether the direction of heading was to the left or right of straight ahead. Stimuli conveyed heading directions ranging from -20° (left) to +20° (right), where 0° was straight ahead, and simultaneous rotations of -2, -0.8, 0 (no rotation), +0.8, and +2 °/s angular velocity. Heading bias was computed as the center of the psychometric function (50:50 leftward:rightward choices) and compared across conditions (Fig. 3). Phase motion stimuli conveyed only instantaneous optic flow — a single optic flow field — corresponding to either the first or last flow field in a sequence. Time-varying phase motion stimuli conveyed time-varying optic flow — an evolving sequence of flow fields — but did not contain the cue of image point trajectories (Fig. 1e). Envelope motion stimuli conveyed three cues that could be useful for computing heading perception: instantaneous optic flow, time-varying optic flow, and image point trajectories.

[1] Department of Psychology, New York University, New York, NY, USA. [2] Center for Neural Science, New York University, New York, NY, USA. *e-mail: charlie.burlingham@nyu.edu; david.heeger@nyu.edu

Heading perception was strongly biased for phase motion. For the faster rotation speed of ±2 °/s, heading bias was much larger for phase motion than for envelope motion. This was true for each individual observer (N trials per condition = 1,200; SI Fig. 1, see observer 3, O3 +2 °/s rotation for a single exception), and when the data was pooled across observers (N observers = 5, N trials = 6,000; Fig. 4a, Table 1). Judgments of heading direction were biased in the direction of rotation. For ±2 °/s rotation, when pooled across observers, average heading bias was 4 times larger for phase motion than for envelope motion (5.92° vs. 1.46°). For the slower rotation speed of ±0.8 °/s, bias was 1.8 times larger for phase motion than for envelope motion (1.31° vs. 0.74°; Fig. 4a, Table 1). Heading judgments were also more variable for phase motion than for envelope motion (variability = 1/slope of psychometric function): 1.36 times more variable on average (4.7° vs. 3.46°) — 1.5 times greater for first flow field phase motion (5.26° vs. 3.46°) and 1.2 times greater for last flow field phase motion (4.07° vs. 3.46°; SI Fig. 2; Table 1). Taken together, these results suggests that observers were unable to judge heading in the presence of a rotation when they only had access to instantaneous optic flow. The magnitude of these errors varied with the speed of rotation, with more heading bias and greater variability at faster rotation speeds.

Heading perception was nearly veridical when the time-varying evolution of optic flow was available. Time-varying phase motion stimuli conveyed a series of optic flow fields — the same ones presented in envelope motion stimuli (see *Methods*). Critically, these stimuli controlled for all cues except time-varying optic flow. They contained the same low-level visual properties as phase motion stimuli, thereby controlling for any differences in perceived speed or other possible confounds. Heading bias was small for envelope motion: an average of 1.46° away from veridical for ±2 °/s rotation and an average of 0.74° from veridical for ±0.8 °/s rotation (Fig. 4a). Heading bias was also small for time-varying phase motion. In fact, heading bias for envelope motion and time-varying phase motion were statistically indistinguishable for ±2 °/s rotation, when pooled across observers (Fig. 4a; SI Table 1; $p = 0.847$ for -2 °/s rotation and $p = 0.308$ for +2 °/s rotation). Bias for time-varying phase motion was indistinguishable from veridical (0° heading) for ±0.8 °/s rotation (Fig. 4a). Heading judgments were significantly less biased for time-varying phase motion than for non-varying phase motion (Fig. 4a; SI Fig. 1; SI Tables 2-3). Taken together, these results suggest that heading perception depends on the time-varying evolution of optic flow. These data also suggest, consistent with previous reports (Warren et al. 1991), that image point trajectories are not a critical cue for heading perception.

Heading direction was not simply perceived as the singularity location, for phase motion. A null model, or "worst case" model of performance, generated predictions that were compared with the observed heading biases (Fig. 3 & 4b). Specifically, the null model predicts that one's estimated heading direction corresponds to the location of singularity in the flow field. Estimating heading by locating the singularity is a biased strategy for image motion caused by translation and rotation (see Fig. 1e, *Methods*). For a subset of observers (O2, O3, & O4), the null model predicted heading bias accurately for +2 °/s rotation, with the null model predictions falling within the 95% credible intervals of the observed biases (SI Fig. 3). O3's biases across rotation velocities were best captured by the null model, suggesting that this observer may have simply tracked the singularity in the flow field, a "worst-case" strategy. However, considering the data set as a whole (pooled across observers), biases for last flow field phase motion fell approximately halfway between the null model predictions and veridical, and biases for first flow field phase motion fell approximately 1/2 to 1/3 of the way between the null model predictions and veridical. This suggests that the majority of observers relied on a heading perception algorithm that was more sophisticated than simply tracking the singularity of the flow field.

Regardless of *which* instantaneous optic flow field was available, heading bias was nearly the same. Envelope motion and time-varying phase motion stimuli conveyed a common sequence of flow fields. To generate phase motion stimuli, we picked a single flow field from this sequence, either the first or last. Because the singularity moves over time due to observer rotation, it is closer to straight ahead on the last than first flow field in the sequence (Fig. 1e). The first and last flow fields therefore provided upper and lower bounds, respectively, on heading bias. If observers used the null model, biases would be larger for first than for last flow field phase motion. Pooled across observers, heading bias was indistinguishable between first and last flow field phase motion (Fig. 4b & SI Table 4; $p > .04$), except for leftward 0.8 °/s rotation, for which bias was slightly higher for the first than last flow field ($p < 0.001$, difference in bias = 1.36°). For 3 observers (O1, O4, and O5), there were no statistically significant differences in heading bias between the first and last flow field phase motion conditions. For the other two subjects, bias was significantly higher for first than last flow field stimuli for a subset of rotation velocities (SI Table 4). Despite these small effects, only O3 showed an overall trend of larger bias for first than last flow field. Pooled across observers, heading judgments were more variable for first than last flow field phase motion, particularly for ±2 °/s rotation (SI Fig. 2).

Eye movements neither help nor hinder heading perception. Regardless of whether a rotation was simulated or generated by a smooth pursuit eye movement, heading bias was nearly the same. On eye movement trials, stimuli were envelope motion, conveying heading but no rotation. Observers tracked a moving fixation dot across the screen, causing a rotation with the same angular velocity as in the other conditions. Pooled across observers, heading bias was on average 0.76° greater for eye movement than for other envelope motion stimuli for ±2 °/s rotation. This difference was small but statistically significant (two-tailed, $p = 0.009$) for +2 °/s rotation, with a difference of 0.95° (SI Fig. 4; SI Table 5). Heading bias was statistically indistinguishable (two-tailed) between eye movement and envelope motion for -2 °/s rotation. There was little difference in bias between the two conditions for ±0.8 °/s rotation (on average, a 0.25° difference), and these comparisons were not significant. Likewise, when pooled across observers, the variability of heading judgments was indistinguishable across the two conditions (SI Fig. 2). These findings suggest that heading perception did not, for our stimulus conditions, depend on extra-retinal signals (see *Discussion* for a fuller interpretation of these results).

Discussion

The current experiment was the first to directly test whether instantaneous optic flow is sufficient for accurate heading perception. We developed a novel phase motion stimulus that conveyed a single instantaneous optic flow field, the first or last in a sequence, and a time-varying phase motion stimulus that conveyed an evolving sequence of flow fields. We found that heading judgments were strongly biased in the direction of rotation for phase motion, but were close to veridical for time-varying phase motion. Heading judgments were also much more variable for phase motion. These findings suggest that instantaneous optic flow is insufficient for heading perception. Rather, heading perception depends on the time-varying evolution of optic flow. Stimulus duration was held constant between conditions, so these results cannot be attributed to differences in the duration of evidence accumulation. One implication is that the vast majority of models for computing heading from optic flow are incorrect, because they rely on instantaneous optic flow.

We used phase motion to convey a single instantaneous optic flow field. However, one could conceive of alternative methods for conveying instantaneous optic flow. For instance, imagine a simple experiment in which an observer is presented with only two frames of a self-motion stimulus and is asked to discriminate heading. One might think this would give the observer access only to a single flow field. However, in practice, temporal aliasing would degrade stimulus visibility and thereby hinder performance. An alternative approach that has been employed in the literature is to present a video with many frames, but with each dot being presented for only 2-3 frames and then re-spawned at a random new location[26]. This method removes image point trajectories as a cue that could be useful for computing heading. It does not, however, remove time-varying optic flow. Although the observer does not have access to the continuous trajectory of each individual dot throughout the trial, he or she does have access to the holistic, continuous evolution of the entire flow field. A previous experiment[33] parametrically varied stimulus duration and found that heading estimation bias and variance scaled inversely with duration. Unfortunately, however, reducing the duration of the stimulus affects both evidence accumulation time and time-varying optic flow, confounding interpretation. In the current experiment, stimuli were presented for four seconds in all conditions, controlling for evidence accumulation time and temporal aliasing. We can therefore attribute the observed large errors in judging phase motion stimuli to the lack of time-varying optic flow, rather than to other factors.

Many sensory cues may contribute to heading perception. Optic flow, extra-retinal signals (e.g., efference copy, vestibular), and depth cues (e.g., from stereopsis, reference objects, shading, looming) could all play a role. The visual system seems to exploit available information, weighting signals according to their reliability, as has been shown extensively in the cue combination literature[35–36]. Indeed, the simulated eye movement condition in the current study may be viewed as a conflict between visual motion and vestibular signals, particularly for high rotational velocities. The resulting biases might, therefore, reflect an optimal combination of these two sources of heading estimates. There has been notable controversy over the issue of whether extra-retinal signals, such as efference copy, are required for accurate heading perception[6,26–30,37]. Heading judgments in the presence of a simulated eye movement are biased, both in discrimination tasks, similar to the one used in this paper[6,27,28], and in estimation tasks, in which observers report their perceived heading continuously with a joystick[26,30,38]. Using the latter protocol, it was found that large biases for simulated eye movement stimuli were reduced to nearly veridical performance if reference objects were present in the environment[30]. Note that even with reference objects, however, the authors still found a difference (albeit smaller) between real and simulated eye movements, with larger heading errors for simulated eye movements. By contrast, we found little difference in errors for real and simulated eye movements. Taken together with our findings, we interpret this literature to mean that many cues, including extra-retinal signals, can contribute to heading perception. However, some cues do seem to be more essential than others. We have shown that the fidelity of heading perception suffers considerably from a lack time-varying optic flow, but not much at all from a lack of image point trajectories or efference copy. We have not investigated the role of stereo or vestibular cues in the current study, but we speculate that the small biases for the envelope and time-varying phase motion conditions may be abolished when this information is also available.

In solving the rotation problem, the speed of rotation matters. Although observers can compensate for the confounding effects of slow speed rotations (< 1 °/s), they cannot fully compensate for faster rotation speeds. Our current findings are consistent with this observation, with the caveat that observers show substantial individual differences. Some observers were able to correct for slow speed (< 1 °/s) real and simulated eye movements better than others (SI Fig. 1).

What is the neural algorithm for heading perception? Cells in middle temporal area (MT) are selective for retinal image motion velocity[39]. This motion selectivity is largely independent of whether an eye movement is real or simulated, although extra-retinal signals caused by real eye movements modulate the gain of these responses[40]. The medial superior temporal area (MST) has reciprocal connections with MT[41]. Neurons in MST are hypothesized to code for the direction of 3-D self-motion[42–45]. MST is thought to construct a stable representation of the world by compensating for retinal image motion caused by eye movements[46–48]. There is a continuum of response selectivity in MST for flow fields with varying FOE locations[44,49–51], suggesting a distributed representation that may encode the time-varying information we have shown is necessary for heading perception. In addition to processing visual motion signals, MST neurons also process vestibular signals and show sensitivity to rotations even in the absence of optic flow[44–45,52–54]. This is an additional property that may allow the visual system to compensate for the confounding effect of a rotation on heading perception.

Virtually all models for computing heading from optic flow rely on the instantaneous velocity of the flow field[9,21]. There are a handful of algorithms that use time-varying optic flow to compute observer motion parameters[24–25], but none of these are biologically plausible. A recent biologically plausible dynamical model of heading perception suggests that achieving robust heading estimates requires taking into account the time history of the visual stimulus[32,55]. This model takes analytical optic flow or videos as input and implements a variety of neural processes, including divisive normalization, on-center/off-surround interactions, recurrent connectivity, Hodgkin-Huxley-like neuronal dynamics, and thresholding, leading to a winner-takes-all readout of a simulated population of MSTd neurons, resulting in a single heading esti-

mate. No other existing model can potentially predict our findings. More work needs to be done to determine whether this biologically plausible model can predict our results, particularly for the case of phase motion. Regardless of the outcome, one could also envision a higher-level, descriptive model that uses the time-varying trajectory of the singularity or other locations in the flow field to estimate rotation velocity and null it out of the biased heading estimate. This toy model might work for the case of a linear path with a rotational eye/body movement, but would fail for the case of a curvilinear path where the flow field can remain the same over time despite a constant rate of rotation (i.e., the "path problem" [13]; but see [56]). In spite of decades of work on optic flow and heading perception, it remains an open challenge to develop a biologically and behaviorally plausible algorithm for heading perception that solves the rotation and path problems.

Methods

Participants. Data were acquired from six observers (four males, two females). All observers were healthy adults, with no history of neurological disorders and with normal or corrected-to-normal vision. One observer, O1, had no prior experience participating in psychophysical experiments. The remaining observers had considerable experience. One observer, O6, did not understand the task and was removed from the experiment following training. Experiments were conducted with the written consent of each participant. The experimental protocol was approved by the University Committee on Activities Involving Human Subjects at New York University.

Stimuli. There were three classes of stimuli: envelope motion, phase motion, and time-varying phase motion. Each of these stimuli comprised a field of plaid patches (Fig. 2). Each patch was the sum of two orthogonal gratings (spatial frequency, 3 cycles/°; contrast, 100%) multiplied by an envelope. The envelope was circular (diameter, 0.4°) with raised-cosine edges (width, 0.2°). No patches were presented within ±2.25° of the horizontal meridian unless they were beyond ±15° of the vertical meridian ensuring that observers could not simply track the singularity of the flow field. The location of each patch was drawn from a uniform distribution. The initial phase of each grating was drawn from a uniform distribution at the onset of each trial. In all conditions, the stimulus was a movie with a duration of 4 seconds, presented with a refresh rate of 75 Hz on a calibrated HP p1230 CRT monitor (HP 2006) with a screen resolution of 1152 x 870 pixels, subtending 40° by 30°. Observers sat in a darkened room and viewed stimuli from a distance of 0.57 m, with their head in a chin/forehead rest to minimize head movements. Observers fixated a red dot (diameter, 0.09°) positioned centrally on a gray background throughout each stimulus presentation. Stimuli were generated and presented using MGL, a MATLAB toolbox for running psychophysics experiments (justingardner.net/mgl).

In the envelope motion condition, each patch (both the envelope and the plaid) was displaced on each frame. The location of each patch corresponded to the projection of a point in a simulated 3-D environment onto the 2-D image plane (Fig. 1a). At each moment in time, the simulated viewpoint of the observer (hereafter referred to as the "viewpoint") moved according to the observer's heading and rotation. The locations of points in the 3-D environment and their projected 2-D locations were recomputed, and the patches were re-rendered at new locations. The phase of each plaid was randomly initialized and fixed with respect to its envelope throughout the trial.

The simulated 3-D environment consisted of two frontoparallel planes, 12.5 m and 25 m, respectively, from the viewpoint. At the beginning of a stimulus movie, the density of patches in each plane was 0.16 patches/^{o2}, and the field of view was 40° by 30°, giving a total of 192 patches forming each plane. The locations of these patches in the simulated 3-D space were projected onto the 2-D image plane using perspective projection (focal length, 0.57 m, i.e., equal to the viewing distance). The simulated 3-D environment through which the simulated observer moved was assumed to be a rigid. The simulated observer translated forward along the z-axis (pure translation), while rotating about the y-axis (simulated eye movement), or translated forward while the observer performed a real smooth pursuit eye movement, tracking a moving fixing dot on the monitor (real eye movement). The parameters of the observer's simulated motion were chosen to be ecologically valid. The chosen translation speed of 1.5 m/s corresponded to average human walking pace and the chosen rotation speeds corresponded to realistic speeds of smooth pursuit eye movements. Rotation speeds were within a range for which observers made accurate smooth pursuit eye movements with few saccades: ~3 saccades per trial for the 2 °/s rotation velocities (SI Fig. 5).

For phase motion and time-varying phase motion, the plaid envelopes remained stationary while the phases of the gratings shifted over time. The phase velocity of the field of plaid patches corresponded to either a single optic flow field (phase motion) or a sequence of flow fields (time-varying phase motion). These optic flow fields corresponded to the motion of points in the simulated 3-D environment relative to the observer viewpoint. In the phase motion condition, the phase of each plaid shifted continuously throughout the trial according to the instantaneous velocity of a single flow field[34]. The phase change between successive frames was constant, so that the local velocity of each patch (and the optic flow field as a whole) was constant over time. The phase change of the vertical grating determined the horizontal component of the phase motion velocity, while the phase change of the horizontal grating determined the vertical component of the velocity. For each (vertical or horizontal) grating, the phase change between frames was computed from the corresponding (horizontal or vertical) component of the optic flow velocity: 2π times the product of the (horizontal or vertical) speed (°/sec) and the spatial frequency (cycles/°), divided by the elapsed time (1/sec).

In the time-varying phase motion condition, the phase change between successive frames varied throughout the trial, simulating a sequence of flow fields with specific rotational and translational components. The phase change between frames was computed by generating a sequence of 10 flow fields spaced evenly over the 4 second trial and calculating the phase step for each flow field. This yielded changes in the phases of the plaids that conveyed time-varying image velocities at each location in the visual field — that is, time-varying optic flow. This time-varying optic flow conveyed information about the change in the location of the singularity over time (Fig. 1e). The time-varying optic flow in the time-varying phase motion condition was just a sub-sampled version of the time-varying optic flow in the envelope motion condition, with the crucial difference that the patch envelopes did not move at all in the former, but did move in the latter. The trajectories of individual image points forming a stimulus may provide a cue for computing heading. Phase motion eliminates this cue (see *Results*).

To generate phase motion and time-varying motion stimuli, we computed the optic flow field(s) corresponding to the motion of the observer's viewpoint relative to the simulated 3-D environment for specific combinations of rotation and heading velocity (SI Eqs. 1–9). This calculation was implemented in custom MATLAB software that was used to generate the phase motion and time-varying phase motion stimuli.

Protocol. Observers viewed videos of plaid patches that moved according to the specified heading and rotation velocities, and subsequently performed a forced-choice heading discrimination. Each trial consisted of a 4 second stimulus presentation followed by a 1 second inter-trial interval, during which the observer made a key press response indicating whether they perceived their heading direction to be left or right of straight ahead (Fig. 3). Immediately following each key press response, observers were given auditory feedback to indicate

if their response was correct or incorrect (two different tones). The heading direction on each trial was set adaptively using two interleaved one-up-one-down staircases that converged to 50:50 leftward:rightward choices. The point of subjective equality (PSE) was the actual heading direction that corresponded to the 50% point (equally perceived as left and right of straight ahead). Staircases were randomly initialized at either -20 or 20 visual degrees (along the x-axis, relative to forward) every 120 trials to collect sufficient data at both the PSE and asymptotes of the psychometric functions. Staircases were initialized with a step size of 4 degrees, which was then adaptively updated depending on the observer's responses.

Participants were tested on the experimental conditions in the following order: envelope motion, eye movement, phase motion (first flow field), phase motion (last flow field), time-varying phase motion. Each experimental condition consisted of two sessions (each on a different day), each session comprised 10 blocks, and each block comprised 60 trials, for a total of 600 trials per session. In all conditions, the observer's heading velocity was fixed, and the angular velocities of the simulated rotations or eye movements were -2 °/s, -0.8 °/s, 0 °/s ("no-rotation"), +0.8 °/s, or +2 °/s. Negative signed velocities were leftward rotation and positive velocities were rightward rotation. These five rotation conditions were interleaved in a randomly permuted order, ensuring equal numbers of trials from each condition within each block of trials. Consequently, there were a total of 4,800 trials: 1,200 trials overall per condition, and 240 per rotation velocity therein.

In the eye movement condition, an envelope motion stimulus was presented that conveyed a pure translation, while observers moved their eyes to track a fixation dot that moved horizontally across the monitor. To ensure that the angular velocity of the eye movement was constant over the trial, the fixation dot speed was varied over time, correcting for the difference in the curved path traced by the eye movement and the flat screen.

In the phase motion (first and last flow field) conditions, a phase motion stimulus was presented that corresponded to a particular instantaneous optic flow field. The flow field presented was either the first or last in the same sequence as in the envelope motion and time-varying phase-motion conditions (Fig. 4b inset). We called these stimuli phase motion "first flow field" and "last flow field". The singularity location was closer to straight ahead on the last flow field than on the first flow field. Presenting different instantaneous flow fields to the observer allowed us to determine whether certain instantaneous flow fields are more informative than others for solving the rotation problem.

Prior to participating in the experiment, observers were trained in the following regime in this order: envelope-no-rotation, envelope motion, eye movement, phase-no-rotation, phase motion (first flow field), phase motion (last flow field), time-varying phase motion. Training was necessary to ensure that observers understood the protocol and had reached asymptotic perceptual sensitivity in the task. Observers performed an average of 500 trials of training per main condition (envelope motion, eye movement, phase motion, time-varying phase motion), for an average of 2,000 trials of training total. Some observers needed more training than others, but all observers were trained on a minimum core of 120 envelope-no-rotation trials and 120 phase-no-rotation trials. To determine when to end training and start the experiment, we checked that discrimination accuracy was approximately 50% and ensured that discrimination thresholds had become stable across trials. We found that observers' discrimination thresholds decreased over the course of training, reaching asymptotic thresholds at heading directions between 2° to 6°, depending on the condition (SI Fig. 2).

Data Analysis. Fitting of psychometric functions was performed using Psignifit 4, a MATLAB toolbox for Bayesian inference for psychometric functions[57]. Cumulative normal psychometric functions were fit separately to each individual observer's responses (N trials per condition = 1,200) and also to the pooled responses of all observers (pooled across observers; N observers = 5, N trials per condition = 6,000). We fit a beta-binomial model with five parameters: μ, σ, λ, γ, η, which respectively corresponded to the mean, variance, lapse rate, guess rate, and over-dispersion (Fig. 3). The mean of the psychometric function describes the physical heading direction for which the observer reported 50:50 leftward:rightward of straight ahead. That is, it describes the observer's internal estimate of straight ahead, i.e., the observer's bias for "subjective forward." For example, if an observer viewed a pure translational flow field and had a bias of -2°, that would suggest that the observer interpreted straight ahead as being leftward heading of -2°, rightward heading of +5° as rightward heading of +3°, and so on. In the pure translation condition, bias revealed how different the observer's subjective forward was from objective forward. In the heading + rotation conditions, bias additionally revealed whether observers used the displaced singularity to estimate heading, thereby interpreting rotation as additional heading in the same (or opposite) direction. Variance describes the slope of the psychometric function and determines the psychophysical threshold of the observer in a 2-AFC task (Fig. 3). Lapse rate describes how often observers lapsed, pressing the wrong key by accident. It sets an upper limit on the accuracy of the observer, even at the most extreme heading values, and jointly controls the heights of the psychometric function's upper and lower asymptotes. Guess rate describes how often observers guessed, and was fixed at 0 during fitting. Over-dispersion describes how much the observed data deviated from the beta-binomial model, potentially due to serial dependencies, changes in arousal throughout the experiment, or other factors. The likelihood of the data given the model was evaluated, and flat priors were used for each of the four free parameters, yielding a posterior describing the probability of the model given the data. We report maximum a posteriori (MAP, equivalent in this case to maximum likelihood) estimates for each parameter by numerical integrating over the 4-D posterior, $p(\boldsymbol{\Theta}|\text{data})$, where $\boldsymbol{\Theta}$ is a vector containing the four free parameters, and finding the maximum of the marginal distribution for each parameter. 95% credible intervals for parameter fits were computed as the central 95% density region of the cumulative probability function of each marginal posterior distribution.

Null Model. We defined a null model in which heading estimates were computed directly from the singularity of the flow field. For the case of pure heading, the singularity of the flow field is the focus of expansion, which gives the correct heading direction. However, the singularity is displaced when there is rotation, leading to a systematic bias. The null model predicts the observer's heading bias, assuming that it simply tracks the singularity of the flow field. The singularity was computed by extracting the horizontal component of optic flow (see SI Eq. 7) and solving for the point at which this function equals zero. Each null model prediction is the heading direction corresponding to this singularity. Null model predictions were generated for the near (12.5 m) plane across a finely spaced range of rotation angular velocities from -2 to +2 °/s, using the same parameters that were used to generate the experimental stimuli.

Statistics. The strength of statistical evidence for differences in heading bias across conditions was assessed via permutation test, a resampling technique that constructs a null distribution of a test statistic from the observed data[58–59]. P-values were corrected for the finite number of permutations[60–61] and for multiple comparisons[62–63], yielding corrected exact p-values (see SI *Statistics* for details; SI Tables 1-5 contain exact p-values for a subset of comparisons).

Eye tracking. Measurements of smooth pursuit eye movements were used to validate that observers were tracking (on eye movement trials) or fixating (on all other trials) the fixation dot. Gaze position was measured using a remote infrared video-oculographic system (Eyelink 1000; SR Research), with a spatial resolution of 0.01° and average accuracy of 0.25-0.5° when using a head rest. Gaze position

was acquired at a sampling rate of 500 Hz. Blinks were removed from the gaze-position time-series and the missing gaze positions were linearly interpolated. The horizontal and vertical gaze positions on eye movement trials were averaged to yield mean gaze positions for each rotation velocity condition. Observers tracked the moving fixation dot accurately and there was very little variability in gaze position traces across runs (SEM < 0.3°; SI Fig. 3a). Observers (N = 5) made few saccades on eye movement trials: ∼3 saccades per trial for ±2 °/s rotation, ∼2 for ±0.8 °/s rotation, and ∼1 for no rotation (SI Fig. 5b). For other conditions (non-eye movement trials), all observers fixated stably during the 4 second stimulus and eye position did not depend on the velocity of simulated rotation. For non-eye-movement trials, the average number of saccades per trial was ∼1. Only observer O5's saccades were dependent on the simulated rotation velocity of the stimulus, showing the same pattern as in the eye movement trials, but ∼5 times smaller (difference in number of saccades per trial between highest rotation speed and no rotation = .4; compare with SI Fig. 5b). The remaining observers' saccade counts showed little or no dependence on rotation velocity (average difference in number of saccades per trial between highest rotation speed and no rotation = .1, N = 4).

References

1. Gibson, J.J. *The perception of the visual world.* (The Riverside Press, 1950).
2. Gibson, J.J. *The senses considered as perceptual systems.* (Houghton Mifflin, 1966).
3. Koenderink, J. J. & van Doorn, A. J. Local structure of movement parallax of the plane. *Journal of the Optical Society of America* **66,** 717-723 (1976).
4. Fennema, C.L. & Thompson, W.B. Velocity determination in scenes containing several moving objects, *Computer Graphics and Image Processing* **9**(4), 301–315 (1979).
5. Longuet-Higgins, H.C. & Prazdny, K. The interpretation of a moving retinal image. *Proceedings of the Royal Society of London* **208***,* 385–397 (1980).
6. Warren, E.P. & Hannon, D.J., Direction of self-motion is perceived from optical flow. *Nature* **336**, 162–163 (1988).
7. Baraldi, P., Micheli, E.D., & Uras, S. *Motion and Depth From Optical Flow.* Alvey Vision Conference (1989).
8. Hannon, D.J. & Warren, W.H. Eye movements and optical flow. *Journal of the Optical Society of America* **7**(1), 160–169 (1990).
9. Heeger, D. & Jepson, A. Subspace methods for recovering rigid motion i: Algorithm and implementation. *International Journal of Computer Vision* **7**(2), 95–117 (1992).
10. Verri, A., Girosi, F., & Torre, V. Mathematical properties of the 2D motion field: from singular points to motion parameters. *Journal of the Optical Society of America* **6**(5), 698–712 (1998).
11. Warren, W.H. Jr., Kay, B.A., Zosh, W.D., Duchon, A.P., & Sahuc, S. Optic flow is used to control human walking. *Nature Neuroscience* **4**, 213–216 (2001)
12. Wilkie, R. & Wann, J. Controlling steering and judging heading: retinal flow, visual direction, and extraretinal information. *Journal of Experimental Psychology Human Perception and Performance* **29**(2), 363–378 (2004).
13. Warren, W.H. *Optic Flow* in *The Visual Neurosciences*, 1247–1259 (University of Cambridge Press, 2004).
14. Perrone, J.A. Model for the computation of self-motion in biological systems. *Journal of the Optical Society of America* **9**(2), 177–194 (1992)
15. Perrone, J.A. & Stone, L.S. A model of self-motion estimation within primate extrastriate visual cortex. *Vision Research* **34**, 2917–2938 (1994).
16. Royden, C.S., Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth. *Journal of the Optical Society of America* **14**(9), 2128–2143 (1997)
17. Beintema, J.A. & van den Berg, A.V. Heading detection using motion templates and eye velocity gain fields. *Vision Research* **38**(14), 2155–2179 (1998).
18. Grossberg, S., Mingolla, E., & Pack, C. A neural model of motion processing and visual navigation by cortical area mST. *Cerebral Cortex* **9**, 878–895 (1999).
19. Davison, A.J. Real-time simultaneous localisation and mapping with a single camera. In *ICCV 2003: 9th international conference on computer vision*. *IEEE,* 1403–1410 (2003).
20. Lappe, M. & Rauschecker, J.P. A neural network for the processing of optic flow from ego-motion in man and higher mammals. *Neural Computation* **5**(3), 374–391 (2008).
21. Raudies, F. & Neumann, H. A review and evaluation of methods estimating ego-motion. *Computer Vision and Image Understanding* **116**(5), 606–633 (2012).
22. Foulkes, A.J., Rushton, S.K., & Warren, P.A. Heading recovery from optic flow: Comparing performance of humans and computational models. *Frontiers in Behavioral Neuoscience* **7**(53), 1–20 (2013).
23. Raudies, F. Modeling heading and path perception from optic flow in the case of independently moving objects. *Frontiers in Behavioral Neuoscience* ***7,*** 1–19 (2013).
24. Barron, J.L. & Eagleson, R. Recursive estimation of time-varying motion and structure parameters, *Pattern Recognition* **29**(5) 797–818 (1996).
25. van den Hengel, A., Chojnacki, W., & Brooks, M.J. *Determining the Translational Speed of a Camera from Time-Varying Optical Flow* in *Complex Motion. Lecture Notes in Computer Science*, vol. 3417 (Springer, 2007).
26. Warren, W.H. Jr., Blackwell, A.W., Kurtz, K.J,, Hatsopoulos, N.G., & Kalish, M.L. On the sufficiency of the velocity field for perception of heading. *Biological Cybernetics* **65**, 311–320 (1991).
27. Banks, M.S., Ehrlich, S.M., Backus, B.T., & Crowell, J.A. Estimating heading during real and simulated eye movements. *Vision Research* **36**(6), 431–443 (1996).
28. Royden, C.S., Banks, M.S. & Crowell, J.A.,The perception of heading during eye movements. *Nature* **360**, 583–585 (1992).
29. Ehrlich, S.M., Beck, D.M., Crowell, J.A., Freeman, T.C.A., & Banks, M.S. Depth information and perceived self-motion during simulated gaze rotations. *Vision Research* **38**, 3129–3145 (1998).
30. Li, L. & Warren Jr., W.H. Retinal flow is sufficient for steering during observer rotation. *Psychological Science* **13**(5), 485–491 (2002).
31. Browning, N.A., Grossberg, S., & Mingolla, E. Cortical dynamics of navigation and steering in natural scenes: motion-based object segmentation, heading, and obstacle avoidance. *Neural Networks* **22**, 1383–1398 (2009).
32. Layton, O.W. & Fajen, B.R. Competitive dynamics in mSTd: A mechanism for robust heading perception based on optic flow. *PLoS computational biology* **12**(6), e1004942–37 (2016a).
33. Layton, O.W. & Fajen, B.R. The temporal dynamics of heading perception in the presence of moving objects. *Journal of Neurophysiology* **115**(1), 286–300 (2016b).
34. Freeman, W.T., Adelson, E.H., & Heeger, D. Motion without movement. In *SIGGRAPH,* 1–4 (1991).

35. Fetsch, C.R., DeAngelis, G.C., & Angelaki, D.E. Visual-vestibular cue integration for heading perception: Applications of optimal cue integration theory. *The European Journal of Neuroscience* **31**(10), 1721–1729 (2010).
36. Landy, M.S., Banks, M., & Knill, D.C. *Ideal-Observer Models of Cue Integration* in *Sensory Cue Integration*. (Oxford University Press 2011).
37. Royden, C.S., Crowell, J.A., & Banks, M.S. Estimating heading during eye movements. *Vision Research* **34**, 3197–3214 (1994).
38. Li, L. & Warren Jr., W.H. Perception of heading during rotation: Sufficiency of dense motion parallax and reference objects. *Vision Research* **40**, 873–3894 (2000).
39. Dubner, R. & Zeki, S. Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Research,* **35**, 528-532 (1971).
40. Kim, H.R., Angelaki, D.E., & DeAngelis, G.C. Gain modulation as a mechanism for coding depth from motion parallax in macaque area mT. *The Journal of Neuroscience*, **37**(34), 8180–8197 (2017).
41. Maunsell, J. & van Essen, D. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *The Journal of Neuroscience* **3**, 2563-2586 (1983).
42. Logan, D.J. Cortical area mSTd combines visual cues to represent 3-D self-movement. *Cerebral Cortex* **16**(10), 1494–1507 (2005).
43. Gu, Y., Watkins, P.V., Angelaki, D.E., & DeAngelis, G.C. Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *The Journal of Neuroscience* **26**(1), 73–85 (2006).
44. Takahashi, K. et al. Multimodal coding of three-dimensional rotation and translation in area mSTd: Comparison of visual and vestibular selectivity. *The Journal of Neuroscience* **27**(36), 9742–9756 (2007).
45. Chen, A., DeAngelis, G.C. & Angelaki, D.E., A comparison of vestibular spatiotemporal tuning in macaque parietoinsular vestibular cortex, ventral intraparietal area, and medial superior temporal area. *The Journal of Neuroscience* **31**(8), 3082–3094 (2011).
46. Inaba, N., Shinomoto, S., Yamane, S., Takemura, A., & Kawano, K. MST neurons code for visual motion in space independent of pursuit eye movements. *Journal of Neurophysiology* **97**, 3473-3483 (2007).
47. Chukoskie, L. & Movshon, J.A., Modulation of visual signals in macaque mT and mST neurons during pursuit eye movement. *Journal of Neurophysiology* **102**(6), 3225–3233 (2009).
48. Inaba, N., Miura, K., & Kawano, K. Direction and speed tuning to visual motion in cortical areas MT and MSTd during smooth pursuit eye movements. *Journal of Neurophysiology* **105**, 1531-1545 (2011).
49. Duffy, C.J. & Wurtz, R.H. Sensitivity of mST neurons to optic flow stimuli. i. a continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology* **65**(6), 1329–1345 (1991a).
50. Duffy, C.J. & Wurtz, R.H., 1991. Sensitivity of mST neurons to optic flow stimuli. iI. mechanisms of response selectivity revealed by small-field stimuli. *Journal of Neurophysiology* **65**(6), 1346–1359 (1991b).
51. Duffy, C.J. & Wurtz, R.H. Response of monkey mST neurons to optic flow stimuli with shifted centers of motion. *The Journal of Neuroscience* **15**(7), 5192–5208 (1995).
52. Chowdhury, S.A., Takahashi, K., DeAngelis, G.C., & Angelaki, D.E. Does the middle temporal area carry vestibular signals related to self-motion? *The Journal of Neuroscience* **29**(38), 12020–12030 (2009).
53. Gu, Y., DeAngelis, G.C. & Angelaki, D.E. A functional link between area mSTd and heading perception based on vestibular signals. *Nature Neuroscience* **10**(8), 1038–1047 (2007).
54. Gu, Y., Fetsch, C.R., Adeyemo, B., DeAngelis, G.C., & Angelaki, D.E. Decoding of mSTd population activity accounts for variations in the precision of heading perception. *Neuron* **66**(4), 596–609 (2010).
55. Layton, O.W., Mingolla, E. & Browning, N.A. A motion pooling model of visually guided navigation explains human behavior in the presence of independently moving objects. *Journal of Vision* **12**(1), 20–20 (2012).
56. Perrone, J.A. Visual–vestibular estimation of the body's curvilinear motion through the world: a computational model. *Journal of Vision* **18**(4), 1–32 (2018).
57. Schütt, H.H., Harmeling, S., Macke, J.H., & Wichmann, F.A. Painfree and accurate bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Research* **122**(C), 105–123 (2016).
58. Fisher, R.A. The management of field experiments. *Journal of the Ministry of Agriculture* **33,** 503 (1926).
59. Edgington, E.S. *Randomization Tests*. (Marcel Dekker, 1987).
60. Davison, A.C. & Hinkley, D.V. *Bootstrap methods and their application* (Cambridge University Press, 1997).
61. Phipson, B. & Smith, G.K. Permutation p-values should never be zero: Calculating exact p-values when permutations are randomly drawn. *Statistical Applications in Genetics and Molecular Biology* **9**(1), 1–12 (2016).
62. Benjamini, Y., Krieger A.M., & Yekutieli, D. Adaptive linear step-up procedures that control the false discovery rate. *Biometrika* **93**(3), 491–507 (2006).
63. Groppe, D.M., Urbach, T.P., & Kutas, M. Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. Psychophysiology **48**(12), 1711–1725 (2011).
64. Stevens, J.R., Al Masud, A., & Suyundikov, A. A comparison of multiple testing adjustment methods with block-correlation positively-dependent tests. *PLoS ONE* **12**(4), e0176124 (2017).

Acknowledgements

Special thanks to Davis Glasser, Elisha Merrian, and Ionatan Kuperwajs for conducting preliminary versions of these experiments. Thanks to Shannon Locke, Hörmet Yiltiz, and Emmanouil Protonotarios for providing helpful comments on experimental design and analysis. Thanks to Michael Landy and Eero Simoncelli for providing feedback on the manuscript. This research was supported by the National Eye Institute Visual Neuroscience Training Grant, T32 EY007136 (to C.S.B. through NYU).

Author Contributions

Conceptualization, D.J.H.; Investigation, C.S.B.; Analysis, C.S.B.; Software and visualization, C.S.B.; Writing, C.S.B. and D.J.H; Supervision, D.J.H.

Correspondence and requests for materials should be addressed to C.S.B. and D.J.H.

Figures

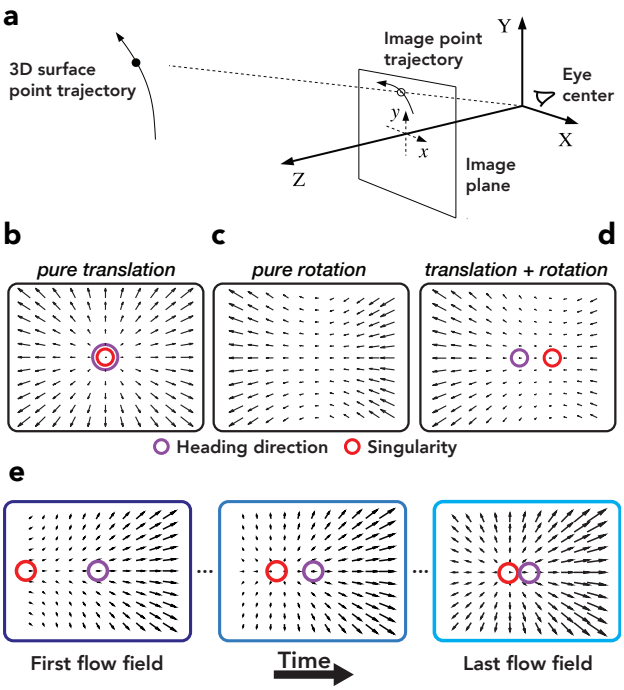


Fig. 1 | Optic flow fields & stimulus generation. a. Viewer-centered coordinate frame and perspective projection. Owing to motion between the viewpoint and the scene, a 3-D surface point traverses a path in 3-D space. Under perspective projection, the 3-D path of this point projects onto a 2-D path in the image plane, the temporal derivative of which is called image velocity. The 2-D velocities associated with all visible points defines a dense 2-D vector field called the optic flow field. **b, c, d.** Optic flow fields with different rotational and translational components. **b.** Pure translation, corresponding to the envelope-no-rotation and phase-no-rotation conditions (1.5 m/s translation speed, 0° straight forward heading). Singularity (red circle) corresponds with heading (purple circle). **c.** Pure rotation, for illustrative purposes only and not corresponding to any experimental condition (+2 °/s rightward rotation). **d.** Translation + Rotation, corresponding to envelope motion and phase motion conditions (1.5 m/s translation speed, 0° straight forward heading, +2 °/s rightward rotation). Singularity (red circle) is displaced away from heading (purple circle). **e.** Time evolution of optic flow. Sequence of flow fields from a time-varying phase motion stimulus (1.5 m/s translation speed, -2 °/s leftward rotation, 0° straight forward heading) at different moments in time reveals that the location of the singularity (red circle) shifts substantially over the course of the trial.

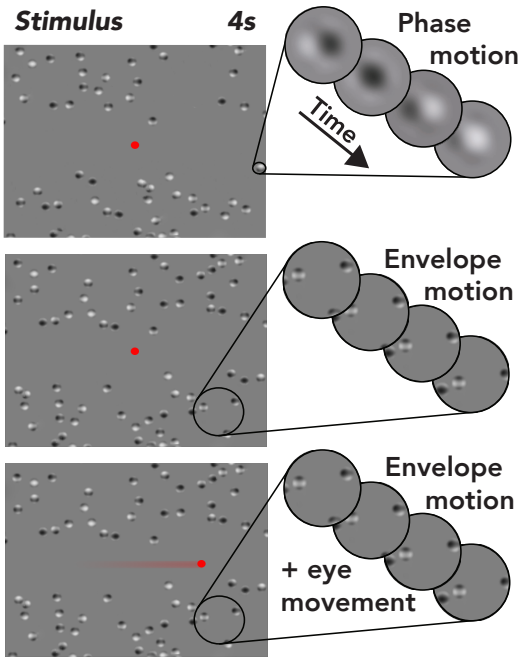


Fig. 2 | Experimental protocol. Each stimulus panel displays a single frame from a stimulus video from that condition. Stimuli were presented for 4 seconds, followed by a 1 second inter-trial interval (ITI), during which observers made a key press response. In the phase motion condition, the envelopes of the plaid patches were fixed in location, but the phases within them changed at a constant rate (first panel). In the envelope motion condition, the envelopes of the plaid patches moved over time, but the phases within them were fixed (second panel). In the eye movement condition, the motion was envelope motion and the red fixation dot moved from the center of the monitor to the left or the right, eliciting an eye movement (third panel). The fixation dot remained fixed at the center of the monitor in all other conditions. Plaids patches have been enlarged by a factor of four, the fixation dot has been enlarged by a factor of six, and dot density has been reduced by a factor of four for illustrative purposes.

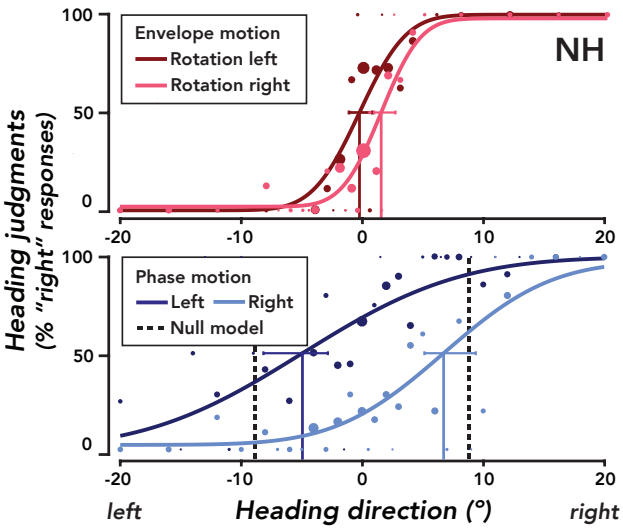


Fig. 3 | Example psychometric functions for envelope and phase motion (first flow field). Data from observer O4 (N trials per condition = 1,200). Data points represent percent rightward responses at each heading direction. Dot area represents the number of data points collected at that heading direction. Smooth curves, best fit cumulative normal psychometric functions. Solid vertical lines, bias (mean of a best-fit psychometric function). Error bars, 95% credible interval around the MAP estimate of bias. Credible intervals for the remaining parameters are not depicted. Red, envelope motion. Blue, phase motion. Dark colors, -2 °/s (leftward) rotation. Light colors, +2 °/s (rightward) rotation. Dashed vertical lines, null model predictions for the first flow field.

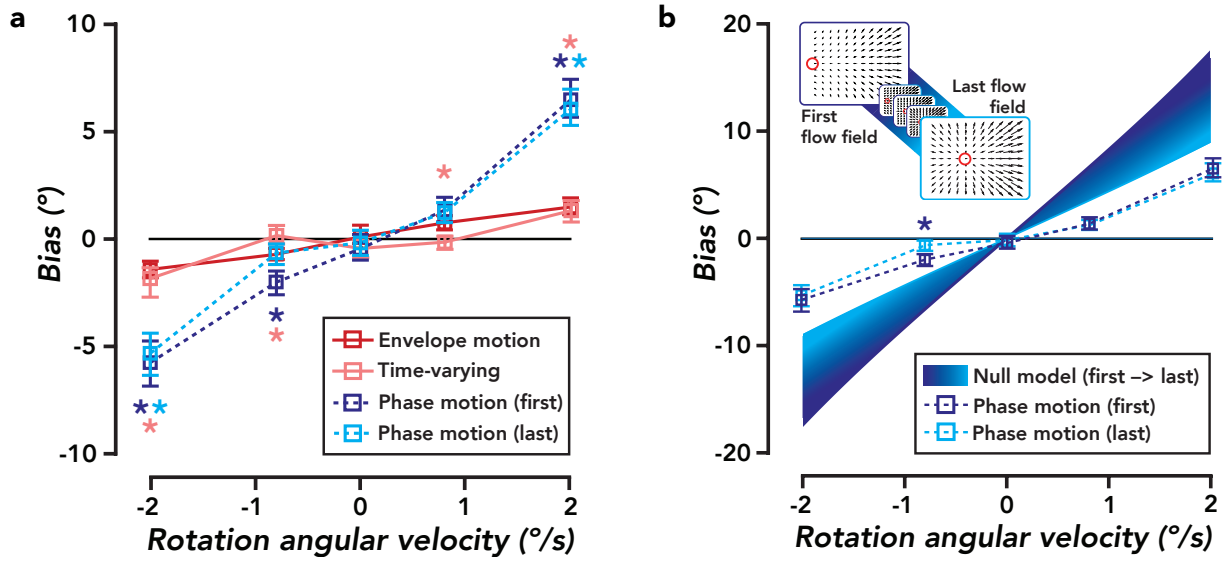


Fig. 4 | Bias in heading perception. a. Heading bias, pooled across observers (N observers = 5, N trials per condition = 6,000). Square plot symbols, MAP estimates. Error bars, 95% credible intervals. Asterisks represent statistical significance, with $p < 0.005$ for all comparisons. The color of the asterisks indicates the specific hypothesis test that was performed. Light-blue asterisks, phase motion (last flow field) was compared to envelope motion (one-tailed permutation test). Dark-blue asterisks, phase motion (first flow field) was compared to envelope motion (one-tailed permutation test). Pink asterisk, phase motion (both first & last flow field) was compared to time-varying phase motion (one-tailed permutation test). **b.** Heading bias for phase motion (first flow field) and phase motion (last flow field), pooled across observers. Plotted as a color gradient (from dark to light blue) are the full set of null model predictions across the sequence of flow fields and rotation velocities. Dark blue asterisk indicates that for the corresponding rotation velocity, absolute bias for phase motion (first flow field) was significantly larger than for phase motion (last flow field), although the difference was less than 1°. The inset depicts the correspondence between the color gradient (from dark to light blue) and the sequence of flow fields.

a	-2 °/s	-0.8	0	+0.8	+2
All	**0.001**	**0.001**	**0.020**	0.051	**0.001**
O1	**0.008**	1.000	0.651	0.981	**0.003**
O2	**0.019**	**0.001**	**0.001**	**0.001**	**0.001**
O3	**0.001**	**0.011**	0.373	0.537	**0.001**
O4	**0.006**	0.260	0.140	0.652	**0.001**
O5	**0.001**	**0.001**	0.731	**0.001**	**0.001**

b	-2 °/s	-0.8	0	+0.8	+2
All	**0.001**	0.600	0.231	0.074	**0.001**
O1	**0.001**	1.000	0.601	0.998	**0.001**
O2	**0.003**	0.828	0.053	**0.001**	**0.001**
O3	**0.002**	0.404	0.714	0.974	0.178
O4	**0.002**	0.845	0.087	0.155	**0.001**
O5	**0.001**	**0.001**	0.552	**0.011**	**0.001**

Table 1 | Strength of statistical evidence for hypothesis that instantaneous optic flow is insufficient for accurate heading perception. a. Phase motion (first flow field) vs. envelope motion. **b.** Phase motion (last flow field) vs. envelope motion. Corrected exact p-values for each observer's bias (N trials per rotation velocity = 240), and pooled across observers ("All"; N observers = 5, N trials per rotation velocity = 1,200). One-tailed permutation tests were performed for -2, -0.8, +0.8, and +2 °/s rotation velocities, and a two-tailed permutation test was performed for the 0 °/s, no rotation condition. The tail of the test depended on the sign of the rotation velocity. For example, the hypothesis tests in panel **a** assessed whether the bias for phase motion (first flow field) trials was higher than the bias for envelope motion trials for +0.8 and +2 °/s rotations, and lower for -0.8 and -2 °/s rotations. This convention is used throughout all tables. All boldface p-values are greater than the critical exact p-value (cutoff) of 0.04. P-values are rounded to three significant digits.

Supplemental Information

Derivation of optic flow in terms of observer translation and rotation

Points on the surface of objects in the virtual 3-D environment are represented as position vectors $\mathbf{X} = (X, Y, Z)^{\mathrm{T}}$ relative to a viewer-centered coordinate frame. When the viewpoint moves, this point moves along a 3-D path $X(t) = (X(t), Y(t), Z(t))^{\mathrm{T}}$ and the instantaneous velocity of this point in 3-D is the derivative of this path with respect to time:

$$\mathbf{V} = \frac{d\mathbf{X}(t)}{dt} = \left(\frac{dX}{dt}, \frac{dY}{dt}, \frac{dZ}{dt}\right)^T. \tag{1}$$

The 3-D points project to image locations. As each 3-D point moves over time, its corresponding image location changes over time. Under perspective projection, the point **X** projects to the image point $(x, y)^{\mathrm{T}}$ given by:

$$x = fX/Z \tag{2}$$

$$y = fY/Z, \tag{3}$$

where f is the focal length of the projection. As the 3-D point moves over time, its corresponding 2-D image point traces out a 2-D path $(x(t), y(t))^{\mathrm{T}}$, the gradient of which is the image velocity:

$$\mathbf{u} = \left(\frac{dx(t)}{dt}, \frac{dy(t)}{dt}\right)^T. \tag{4}$$

Combining Eqs. 2, 3, and 4 gives an expression for image velocity in terms of the 3-D position and velocity of the object surface point:

$$\mathbf{u} = \frac{1}{Z}\left(\frac{dX}{dt}, \frac{dY}{dt}\right)^T + \frac{1}{Z^2}\frac{dZ}{dt}(X(t), Y(t))^T. \tag{5}$$

If one considers the paths of all visible 3-D surface points, and their projections onto the image plane, then one obtains a dense set of 2-D paths, the temporal derivative of which is a vector field of 2-D velocities, commonly known as the optical flow field.

We assume that the objects in the scene move rigidly with respect to the viewpoint, as though the simulated observer were moving through a stationary environment. This is a special case because all points on a rigid body share the same six motion parameters relative to the viewer-centered coordinate frame. In particular, the instantaneous velocity of the observer through a stationary scene can be expressed in terms of the observer's 3-D translation $\mathbf{T} = (T_x, T_y, T_z)^{\mathrm{T}}$, and its instantaneous 3-D rotation $\mathbf{\Omega} = (\Omega_x, \Omega_y, \Omega_z)T$. Here, the direction of $\mathbf{\Omega}$ gives the axis of rotation, while $|\mathbf{\Omega}|$ is the magnitude of the rotation per unit time. Given this motion of the viewpoint, the instantaneous 3-D velocity of a surface point in viewer-centered coordinates is:

$$\left(\frac{dX}{dt}, \frac{dY}{dt}, \frac{dZ}{dt}\right)^T = -(\mathbf{\Omega} \times \mathbf{X} + \mathbf{T}). \tag{6}$$

The 3-D velocities of all surface points in a stationary scene depend on the same rigid-body motion parameters, and are given by Eq. 6. It has previously been shown that if one substitutes these 3-D velocities for d**X**/dt in Eq. 5, an expression for the form of the optical flow field for a rigid scene can be obtained[9]:

$$\mathbf{u}(x, y) = p(x, y)\mathbf{A}(x, y)\mathbf{T} + B(x, y)\mathbf{\Omega}, \tag{7}$$

where $p(x, y)$ is $1/Z(x,y)$ is inverse depth at each image location, and

$$\mathbf{A}(x, y) = \begin{bmatrix} -f & 0 & x \\ 0 & -f & y \end{bmatrix} \tag{8}$$

$$\mathbf{B}(x, y) = \begin{bmatrix} \frac{xy}{f} & -\frac{f+x^2}{f} & y \\ \frac{f+y^2}{f} & -\frac{xy}{f} & -x \end{bmatrix}. \tag{9}$$

The matrices **A**(x,y) and **B(x,y)** depend only on the image position and focal length.

Eq. 7 describes the flow field as a function of 3-D motion and depth. It has two terms. The first term is referred to as the translational component of the flow field because it depends on 3-D translation and 3-D depth. The second term is referred to as the rotational component because it depends only on 3-D rotation. Since $p(x,y)$ (the inverse depth) and **T** (the translation) are multiplied together in Eq. 7, a larger distance (smaller p) or a slower 3-D translation (smaller $|\mathbf{T}|$) both yield slower image velocity.

In a pure translation flow field, the direction and magnitude of each vector depends on the inverse depth at the 3-D location that projects to that vector's corresponding location on the image plane (Fig. 1b). When the observer is undergoing a pure translational motion, features in the image move toward or away from a single point in the image, called the focus of expansion (FOE). By contrast, the optic flow field's rotational component (**Ω**, Eq. 7) does not depend on inverse depth or scene structure. Therefore, rotation in the absence of translation results in a quadratic flow field, that is, each velocity vector (u_x, u_y) is a quadratic function of image position (Eq. 7, Fig. 1c). In a translation + rotation flow field, the translational and rotational components are summed, and the singularity in the flow field (point where image velocity is zero) no longer corresponds to the heading direction (Fig. 1d).

Statistics

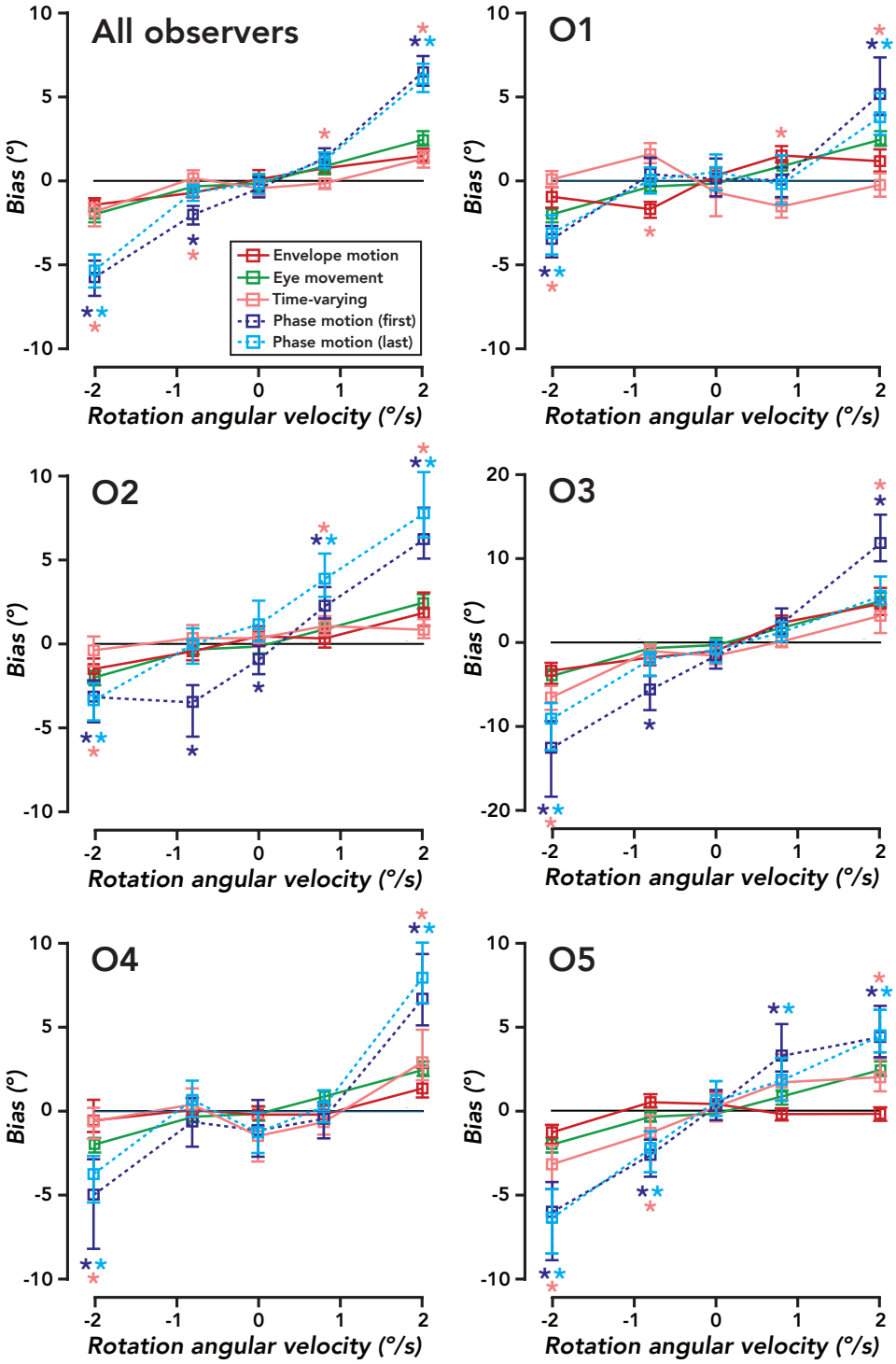
For each randomization test, the null distribution of differences in heading bias across conditions was constructed iteratively in the following way. We pooled the trial-to-trial responses across the two conditions of interest, randomly shuffled the condition labels (permutation without replacement), split the data into two groups, found the MAP estimate of μ (heading bias, see *Data Analysis* in the main body of the text) for each group, and computed the difference between the two groups' heading bias estimates. This procedure was iterated 1000 times. The test statistic was the difference in heading bias across conditions. For the translation + rotation conditions (-2, -0.8, +0.8, & +2 °/s rotation), exact p-values for each comparison were calculated by finding the proportion of test statistics in the null distribution that were greater than (for rightward rotation) or less than (for leftward rotation) the observed difference in heading bias computed from the observed, un-shuffled data (one-tailed permutation test). For the no rotation condition (0 °/s), the absolute values of the null distribution and observed statistic were used instead to perform a two-tailed permutation test. This procedure was carried out for every comparison except for that between envelope motion and eye movement trials, for which a two-tailed test was performed for each rotation velocity. This was because the role of eye movement signals in heading perception is unknown (see *Introduction)*. For the one-tailed tests, the direction of the test was determined by our predictions. See Table 1 and SI Tables 1-5 for the full set of predictions and corresponding exact p-values.

P-values were corrected for the finite number of permutations. The purpose of a permutation test is to estimate the null distribution of a test statistic, ideally enumerating each possible permutation and obtaining an estimate of the p-value, denoted p_∞, under the null hypothesis. However, exhaustive permutation is often computationally intractable and instead, often only a random subset of possible permutations are performed, as was the case with this analysis. It has been shown that simply treating the computation of p_∞ as an estimation problem by replacing p_∞ with an unbiased estimator can lead to an inflated type I error rate. Furthermore, p-values of zero are commonly reported in the literature, but these probabilities make little sense inferentially. If all permutations were enumerated, one would obtain the observed statistic, and the p-value would be greater than zero. Furthermore, in the context of multiple comparisons, p-values of zero become more perilous, and may lead to faulty inference. To partially correct for these issues, all p-values were subjected to a simple and well-known calculation that effectively computes exact p-values:

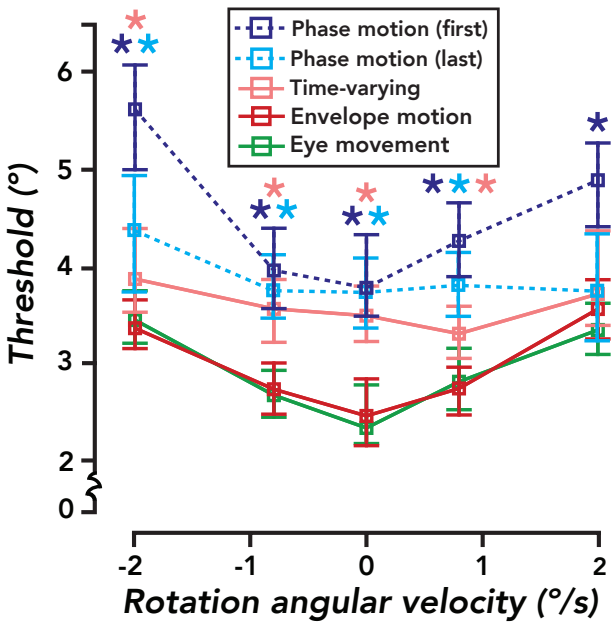
$$p_e = P(B \leq b) = \frac{b+1}{m+1}, \tag{10}$$

where p_e is the exact p-value, b is the number of permuted test statistics greater than the observed test statistic B, and m is number of permutations performed[60–61].

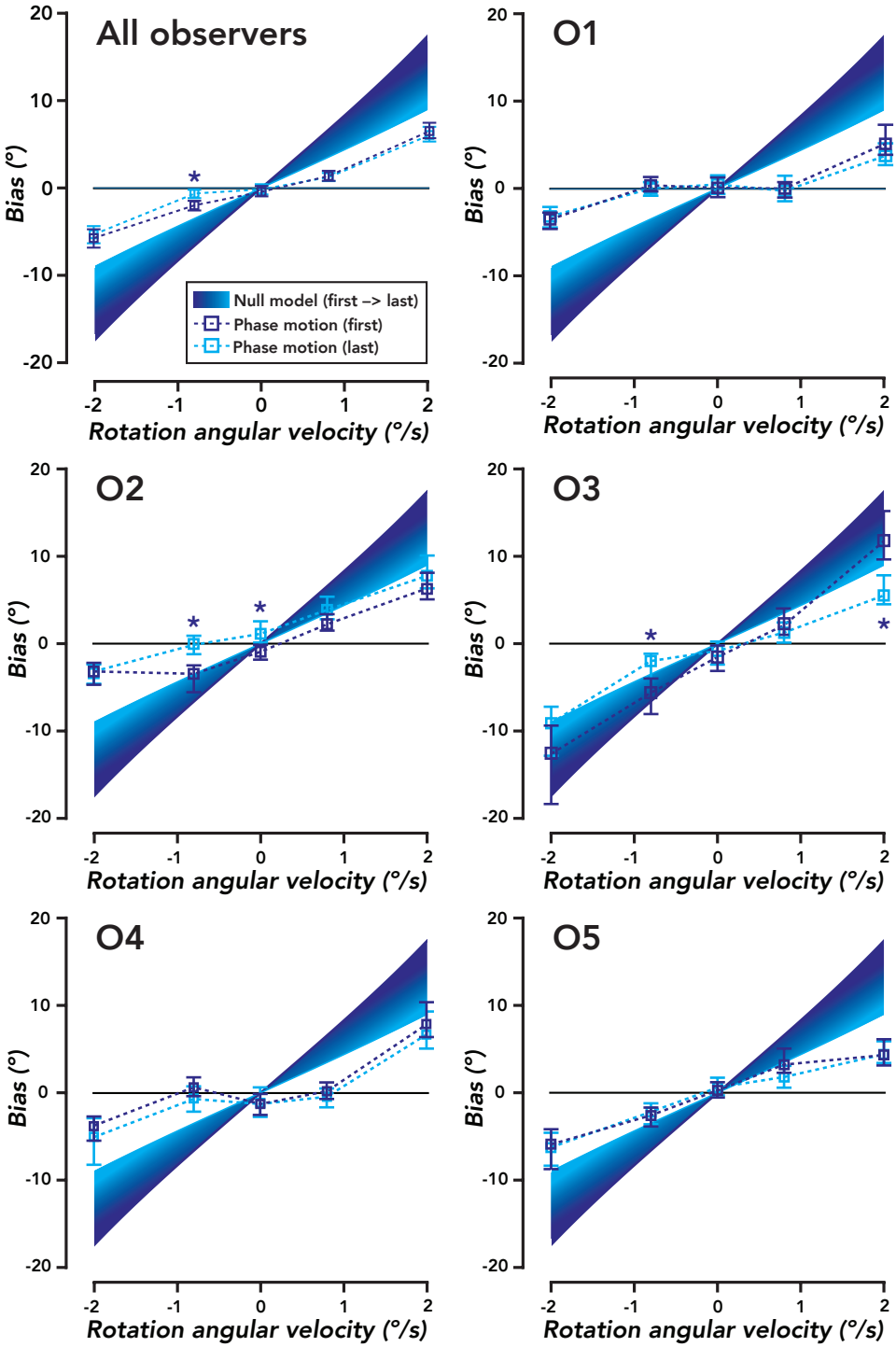
300 hypothesis tests were performed. We note that not all tests were independent, but that positive dependencies were expected to exist between them because they were computed using the same test statistics. To control the overall type I error rate at an alpha level of .05 while simultaneously controlling the false discovery rate (FDR) and assuming dependencies between tests, we subjected all exact p-values to a "two-stage" adaptive procedure[62–63]. It has been shown via simulation that of a number of similar tests, this adaptive test is best at controlling FDR, regardless of effect size of the presence of positive dependence[64]. This procedure yielded a corrected critical p-value, 0.03996004 (rounded off to 0.04 in the main body of the text), which was adopted as a conservative cutoff to judge the strength of the statistical evidence for each hypothesis test.



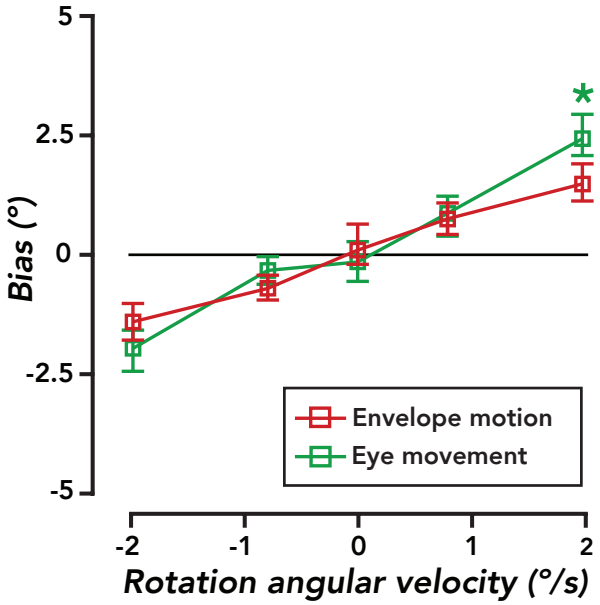
SI Fig. 1 | Heading bias for each observer. Same format at Fig. 4a. Analysis performed on the pooled data of all observers (N observers = 5, N trials per condition = 6,000) or each individual observer (N trials per condition = 1,200). Square plot symbols, MAP estimates. Error bars, 95% credible intervals. Asterisks represent statistical significance at a corrected cutoff of p = 0.04. The color of the asterisks indicates the specific hypothesis test that was performed. Light-blue asterisks, phase motion (last flow field) was compared to envelope motion (one-tailed permutation test). Dark-blue asterisks, phase motion (first flow field) was compared to envelope motion (one-tailed permutation test). Pink asterisks, phase motion (both first and last flow field) was compared to time-varying phase motion (one-tailed permutation test). If comparisons for both first and last flow field were significant, a pink asterisk was plotted. If only one comparison was significant, no such asterisk was plotted (see SI Tables 4-5).



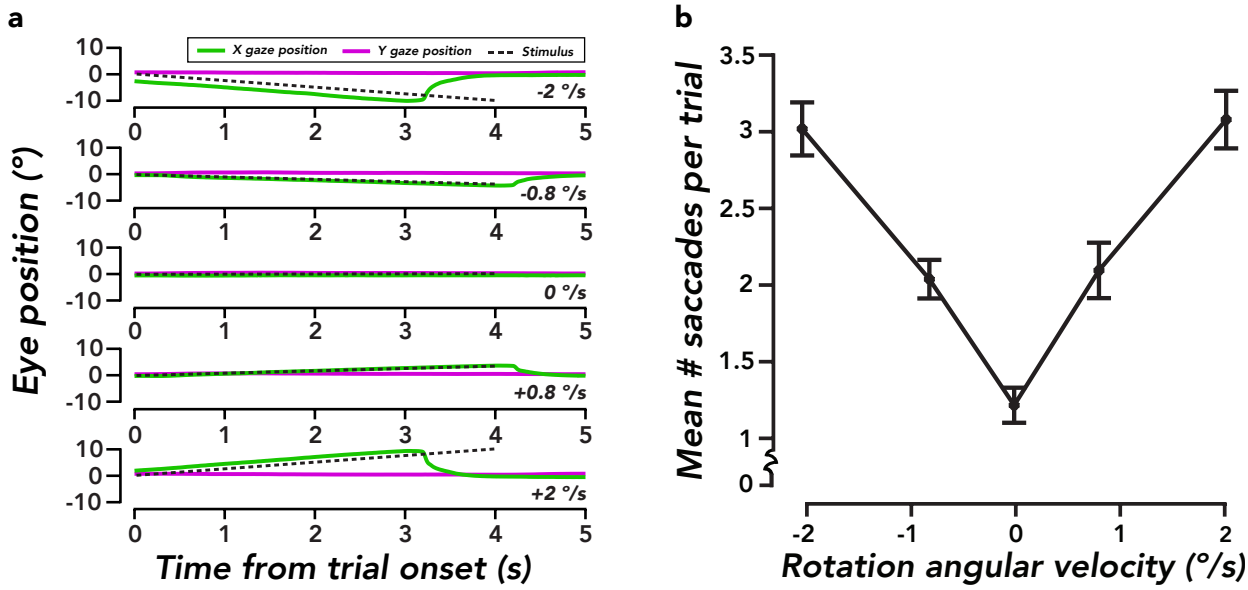
SI Fig. 2 | Heading discrimination thresholds were higher for phase motion than for envelope motion. Analysis performed on the pooled data of all observers (N observers = 5, N trials per condition = 6,000). Each discrimination threshold was computed as the MAP estimate of sigma, the slope parameter of a cumulative normal psychometric function. Square plot symbols, MAP estimate of sigma. Error bars, 95% credible intervals. Asterisks represent statistical significance with a cutoff of $p = 0.04$. All depicted p-values, $p < 0.011$. The color of the asterisks indicates the specific hypothesis test that was performed. Light-blue asterisk, phase motion (last flow field) was compared to envelope motion (one-tailed permutation test). Dark-blue asterisk, phase motion (first flow field) was compared to envelope motion (one-tailed permutation test). Pink asterisks, phase motion (both first and last flow field) was compared to time-varying phase motion (one-tailed permutation test). For +2 °/s rotation, bias for phase motion first flow field was significantly greater than for time-varying phase motion ($p = 0.001$), but this was not true for phase motion last flow field ($p = 0.2607$).



SI Fig. 3 | Heading bias for first and last flow field phase motion for each observer. Same format as Fig. 4b. Phase motion bias data from SI Fig. 2 replotted with new y-axis limit. Plotted as a color gradient (from dark-blue to light-blue) are the full set of null model predictions across the sequence of flow fields and rotation velocities. Colors in the null model predictions (gradient) and the heading bias data are in correspondence based on the number of the flow field in a sequence (see Fig 4b inset). Dark-blue asterisk indicates that for the corresponding rotation velocity, absolute bias for phase motion (first flow field) was significantly larger than for phase motion (last flow field). Note that, for the majority of observers, heading bias fell between veridical performance and the prediction of the null model. For some observers (O2, O3, O4) at a subset of rotation velocities, the null model predicted heading bias accurately, such that the corresponding null model prediction fell within a 95% CI of the MAP estimate of heading bias.



SI Fig. 4 | Heading bias for envelope motion and eye movement stimuli were nearly indistinguishable. Same format at Fig. 4a. Analysis performed on the pooled data of all observers (N observers = 5, N trials per condition = 6,000). Square plot symbols, MAP estimates. Error bars, 95% credible intervals. Red curve, envelope motion. Green curve, eye movement. Asterisks represent statistical significance at a corrected cutoff of $p = 0.04$. Bias was slightly higher (< 1°) for real than simulated eye movements for +2 °/s rotation (green asterisk), but this effect wasn't evident for other rotation velocities.



SI Fig. 5 | Observers accurately tracked the moving fixation dot on real eye movement trials. a. Mean gaze position pooled across eye movement trials (N = 4200 trials) at each rotation velocity from all observers' training and experimental sessions. Green and magenta curves, horizontal (X) and vertical (Y) gaze position, plotted as a function of time from trial onset. Black dotted line, predicted eye position based on the ground truth rotation velocity. Error bars (SEM) are so small that they are invisible beneath the solid lines representing the grand mean time courses. **b.** Number of saccades per trial on eye movement trials. Observers made few saccades (less than 3.5 per trial on average), with more saccades for ±2 °/s than ±0.8 °/s rotations. Observers made ∼1 saccade per trial when the fixation dot was stationary (rotation angular velocity = 0 °/s).

	-2 °/s	-.8 °/s	0 °/s	.8 °/s	2 °/s
All	0.847	**0.001**	**0.009**	**0.001**	0.308
O1	**0.013**	**0.001**	**0.014**	**0.001**	**0.004**
O2	**0.031**	**0.006**	0.647	0.981	**0.027**
O3	0.997	0.123	0.393	**0.001**	**0.039**
O4	0.520	0.237	**0.028**	0.103	0.962
O5	0.998	1.000	0.545	1.000	1.000

SI Table 1 | Envelope motion bias > time-varying phase motion bias. Same format as Table 2. All: Pooled data of 5 observers, N trials per rotation velocity = 1,200. Individual observers, O1, O2, O3, O4, & O5: N trials per rotation velocity = 240. One-tailed permutation test. P-values are rounded to three significant digits. Boldface p-values are greater than the critical exact p-value (cutoff) of 0.04. Red type p-values in some of the following figures are exactly equal to the cutoff.

	-2 °/s	-.8 °/s	0 °/s	.8 °/s	2 °/s
All	**0.001**	**0.001**	0.965	**0.001**	**0.001**
O1	**0.001**	**0.002**	0.045	**0.019**	**0.001**
O2	**0.007**	**0.002**	0.421	**0.001**	**0.001**
O3	**0.008**	**0.001**	0.285	0.066	**0.001**
O4	**0.015**	0.050	0.660	0.339	**0.003**
O5	**0.020**	**0.023**	0.872	**0.009**	**0.005**

SI Table 2 | Phase motion (first flow field) bias > time-varying phase motion bias.

	-2 °/s	-.8 °/s	0 °/s	.8 °/s	2 °/s
All	**0.001**	**0.005**	0.323	**0.001**	**0.001**
O1	**0.001**	**0.009**	0.211	**0.001**	**0.001**
O2	**0.001**	0.204	**0.016**	**0.001**	**0.001**
O3	**0.035**	0.107	0.274	0.049	**0.023**
O4	**0.004**	0.672	0.753	**0.040**	**0.001**
O5	**0.005**	**0.032**	0.401	0.367	**0.001**

SI Table 3 | Phase motion (last flow field) bias > time-varying phase motion bias.

	-2 °/s	-.8 °/s	0 °/s	.8 °/s	2 °/s
All	0.325	**0.001**	0.429	0.380	0.228
O1	0.606	0.303	0.677	0.552	0.878
O2	0.612	**0.001**	**0.001**	0.995	0.928
O3	0.065	**0.005**	0.299	0.142	**0.003**
O4	0.274	0.050	0.870	0.821	0.812
O5	0.595	0.245	0.458	0.052	0.538

SI Table 4 | Phase motion (first flow field) bias > phase motion (last flow field) bias.

	-2 °/s	-.8 °/s	0 °/s	.8 °/s	2 °/s
All	0.134	0.065	0.161	0.568	**0.009**
O1	0.632	**0.001**	0.789	0.673	**0.005**
O2	**0.001**	**0.004**	0.449	**0.002**	**0.002**
O3	0.177	0.056	0.614	**0.001**	**0.006**
O4	0.622	0.108	0.290	**0.026**	**0.012**
O5	**0.018**	**0.008**	0.474	0.365	**0.001**

SI Table 5 | Eye movement bias <> envelope motion bias. Same format as Table 2, but a two-tailed permutation test was performed for each rotation velocity (see *Statistics* section above for details).