Heading perception depends on time-varying evolution of optic flow

Charlie S. Burlingham[a,1] and David J. Heeger[a,b,1]

[a]Department of Psychology, New York University, New York, NY 10003; and [b]Center for Neural Science, New York University, New York, NY 10003

Edited by Thomas D. Albright, The Salk Institute for Biological Studies, La Jolla, CA, and approved November 6, 2020 (received for review November 5, 2020)

There is considerable support for the hypothesis that perception of heading in the presence of rotation is mediated by instantaneous optic flow. This hypothesis, however, has never been tested. We introduce a method, termed "nonvarying 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 nonvarying phase motion, observers made large errors in heading judgments. This suggests that instantaneous optic flow is insufficient for heading perception in the presence of rotation. These errors were mostly eliminated when the velocity of phase motion was varied over time to convey the evolving sequence of optic flow fields corresponding to a particular heading. This demonstrates that heading perception in the presence of rotation relies on the time-varying evolution of optic flow. We hypothesize that the visual system accurately computes heading, despite rotation, based on optic acceleration, the temporal derivative of optic flow.

motion perception | optic flow | heading | acceleration | ego-motion

James Gibson first remarked that the instantaneous motion of points on the retina (Fig. 1*A*) can be formally described as a two-dimensional (2D) field of velocity vectors called the "optic flow field" (or "optic flow") (1). Such optic flow, caused by an observer's movement relative to the environment, conveys information about self-motion and the structure of the visual scene (1–15). When an observer translates in a given direction along a straight path, the optic flow field radiates from a point in the image with zero velocity, or singularity, called the focus of expansion (Fig. 1*B*). It is well known that under such conditions, one can accurately estimate one's "heading" (i.e., instantaneous direction of translation in retinocentric coordinates) by simply locating the focus of expansion (*SI Appendix*). However, if there is angular rotation in addition to translation (by moving along a curved path or by a head or eye movement), the singularity in the optic flow field will be displaced such that it no longer corresponds to the true heading (Fig. 1 *C* and *D*). In this case, if one estimates heading by locating the singularity, the estimate will be biased away from the true heading. This is known as the rotation problem (14).

Computer vision researchers and vision scientists have developed a variety of algorithms that accurately and precisely extract observer translation and rotation from optic flow, thereby solving the rotation problem. Nearly all of these rely on instantaneous optic flow (i.e., a single optic flow field) (4, 9, 16–25) with few exceptions (26–29). 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 human observers can estimate heading (30, 31) from instantaneous optic flow, in the absence of additional information (5, 10, 15, 32–34). Even so, there are reports of systematic biases in heading perception (11); the visual consequences of rotation (eye, head, and body) can bias heading judgments (10, 15, 35–37), with the amount of bias typically proportional to the magnitude of rotation. Other visual factors, such as stereo cues (38, 39), depth structure (8, 10, 40–43), and field of view (FOV) (33, 42–44) can modulate the strength of these biases. Errors in heading judgments have been reported to be greater when eye (35–37, 45, 46) or head movements (37) are simulated versus when they are real, which has been taken to mean that observers require extraretinal information, although there is also evidence to the contrary (10, 15, 33, 40, 41, 44, 47–50). 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 optic flow field evolves over time. Some have suggested that heading estimates rely on information accumulated over time (32, 44, 51), 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 an application of an image processing technique that ensured that only a single optic flow field was available to observers, even though the stimulus was presented for an extended period of time. We called this condition "nonvarying phase motion" or "nonvarying": The phases of two component gratings comprising each stationary stimulus patch shifted over time at a constant rate, causing a percept of motion in the absence of veridical movement (52). Phase motion also eliminated other cues that may otherwise have been used for heading judgments, including image point trajectories (15, 32) and their spatial compositions (i.e., looming) (53, 54). For nonvarying phase motion, observers exhibited large biases in heading judgments in the presence of rotation. A second condition, "time-varying phase motion," or "time-varying," included

Significance

Instantaneous optic flow, image velocity across the visual field, contains information about self-motion (heading). Indeed, computer vision algorithms estimate heading from instantaneous optic flow accurately despite co-occurring rotation. Similarly, the consensus in perceptual psychology is that instantaneous optic flow suffices for heading perception. But as we navigate through the world, visual motion evolves over time. We generated videos showing either an evolving or nonevolving view of navigation through a virtual world. The nonevolving view was insufficient for accurate heading perception in the presence of rotation. Rather, accurate performance relied on time-evolving optic flow. This suggests that most existing models of heading perception are incorrect. We hypothesize that heading perception is based on optic acceleration, the temporal derivative of optic flow.

Author contributions: C.S.B. and D.J.H. designed research; C.S.B. performed research; C.S.B. and D.J.H. contributed new reagents/analytic tools; C.S.B. analyzed data; and C.S.B. and D.J.H. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: charlie.burlingham@nyu.edu and david.heeger@nyu.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2022984117/-/DCSupplemental.

First published December 16, 2020.

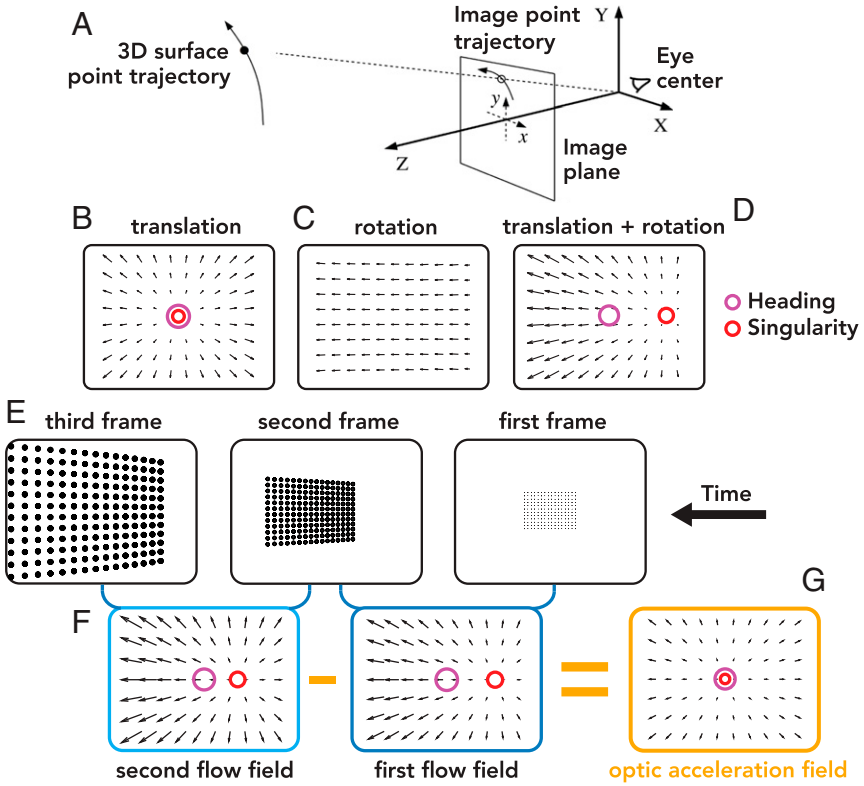

Fig. 1. Projective geometry, the rotation problem, time-varying optic flow, and the optic acceleration hypothesis. (*A*) Viewer-centered coordinate frame and perspective projection. Because of motion between the viewpoint and the scene, a 3D surface point traverses a path in 3D space. Under perspective projection, the 3D path of this point projects onto a 2D path in the image plane (retina), the temporal derivative of which is called image velocity. The 2D velocities associated with all visible points define a dense 2D vector field called the optic flow field. (*B–D*) Illustration of the rotation problem. (*B*) Optic flow for pure translation (1.5-m/s translation speed, 0° heading, i.e., heading in the direction of gaze). Optic flow singularity (red circle) corresponds to heading (purple circle). (*C*) Pure rotation, for illustrative purposes only and not corresponding to any experimental condition (2°/s rightward rotation). (*D*) Translation + rotation (1.5 m/s translation speed, 0° heading, 2°/s rightward rotation). Optic flow singularity (red circle) is displaced away from heading (purple circle). (*E*) Three frames from a video depicting movement along a circular path with the line-of-sight initially perpendicular to a single fronto-parallel plane composed of black dots. (*F*) Time-varying evolution of optic flow. The first optic flow field reflects image motion between the first and second frames of the video. The second optic flow field reflects image motion between the second and third frames of the video. For this special case (circular path), the optic flow field evolves (and the optic flow singularity drifts) only due to the changing depth of the environment relative to the viewpoint. (*G*) Illustration of the optic acceleration hypothesis. Optic acceleration is the derivative of optic flow over time (here, approximated as the difference between the second and first optic flow fields). The singularity of the optic acceleration field corresponds to the heading direction. Acceleration vectors autoscaled for visibility.

acceleration by varying the velocity of phase motion over time to match the evolution of a sequence of optic flow fields. Doing so allowed observers to compensate for the confounding effect of rotation on optic flow, making heading perception nearly veridical. This demonstrates that heading perception in the presence of rotation relies on the time-varying evolution of optic flow.

Results

Observers ($n = 11$) traveled on a virtual circular path, following previous studies (10, 15, 43), and performed a forced-choice heading discrimination task (Fig. 2). They judged whether heading was left or right of the center of the FOV, and were provided "correct"/"incorrect" feedback after each trial. Heading (relative to center) was constant over time within each trial. Stimuli conveyed headings from −20° (left) to +20° (right), where 0° was center, and simultaneous rotations of −2, −0.8, 0 (no-rotation), +0.8, or +2°/s angular velocity, corresponding to circular paths of different radii. The virtual environment was initialized as two fronto-parallel planes positioned at 12.5 and 25 m, following a previous study (10). Heading bias was computed from the center of the psychometric function (50:50 leftward:rightward choices) and compared across conditions (Fig. 3). Stimulus videos contained varying levels of information about the time evolution of the optic flow field (Fig. 2). There were three stimulus conditions: Nonvarying phase motion, time-varying phase motion, and envelope motion. Nonvarying phase motion conveyed only instantaneous optic flow—a single optic flow field that was either the first or last in an evolving sequence—but for an extended duration (*SI Appendix*, Fig. S4*A* and Movies S1–S4). Time-varying phase motion conveyed time-varying optic flow, the entire evolving sequence of optic flow fields (Fig. 1*F* and Movies S5 and S6). Neither nonvarying nor time-varying phase motion showed image point trajectories. Envelope motion (i.e., "dot motion") conveyed three cues: Instantaneous optic flow, time-varying optic flow, and image point trajectories (Movies S7 and S8).

Heading Perception Was Strongly Biased for Nonvarying Phase Motion. Heading bias was significantly larger for nonvarying phase motion than envelope motion at all nonzero rotation velocities. This was true for both first and last flow field nonvarying phase motion (Fig. 4*A* and *SI Appendix*, Table S1). Judgments of heading were biased in the direction of rotation and the magnitude of bias scaled with the speed of rotation. For ±2°/s rotations, when pooled across observers, heading bias (absolute, pooled for ± rotations) was 2.7× larger on average for nonvarying phase motion (average over first and last flow field conditions) than for

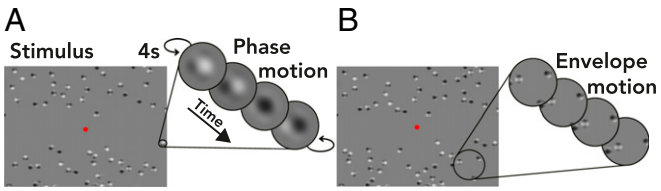

Fig. 2. Stimulus conditions. Each panel displays a single frame of a stimulus video from that condition, with an Inset (magnified in scale) depicting how a small region of the stimulus changed over time. (*A*) Phase motion. For nonvarying phase motion, the envelopes of the plaid patches were fixed in location, but the phases of each component within each patch shifted with a constant rate. For time-varying phase motion, the phases shifted with a time-varying rate (i.e., the phase motion accelerated/decelerated over time). (*B*) Envelope motion. The envelopes of the plaid patches moved over time, but the phases within them were fixed. For illustrative purposes, in both panels plaid 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. See *SI Appendix*, Fig. S4*A* for an unmodified frame from the stimulus.

envelope motion (5.2° vs. 1.9°). For the slower rotation speed of 0.8°/s, bias was ∼5.3× larger for nonvarying phase motion than for envelope motion (1.6° vs. 0.3°) (see Fig. 4*A* for pooled data and *SI Appendix*, Fig. S1, for each observer's data). Heading judgments were also more variable for nonvarying phase motion than envelope motion ("variability" = 1/slope of psychometric function; Fig. 3). For ±2°/s rotations, they were 1.6× as variable on average (5.9° vs. 3.62°, nonvarying vs. envelope) (*SI Appendix*, Fig. S2). These results demonstrate that observers could not perceive heading veridically from instantaneous optic flow (when rotation was present).

Heading Perception Was Nearly Veridical When the Time-Varying Evolution of Optic Flow Was Available. Time-varying phase motion controlled for all cues except time-varying optic flow. Although identical in static appearance to nonvarying phase motion, the speed of each component grating varied over time to convey an evolving sequence of optic flow fields, the same sequence presented with envelope motion (*Materials and Methods* and Fig. 1*F*). Thus, envelope motion and time-varying phase motion conveyed common time-varying optic flow. Heading biases for the two were statistically indistinguishable for 2°/s rotations, when pooled across observers (Fig. 4*A* and *SI Appendix*, Table S3). Furthermore, heading bias was very low for time-varying phase motion, indistinguishable from veridical (0° heading) for ±0.8°/s rotations (Fig. 4*A*). It follows that heading judgments were also significantly less biased for time-varying phase motion than for nonvarying phase motion (Fig. 4*A* and *SI Appendix*, Fig. S1 and Table S2), bolstering our claim that instantaneous optic flow is insufficient. Biases for ±2°/s rotations were on average 3.3× smaller (1.59° vs. 5.2°, time-varying vs. nonvarying phase motion). Biases for ±0.8°/s rotation were on average 10.8× smaller (0.15° vs. 1.62°, time-varying vs. nonvarying phase motion). Taken together, these results suggest that heading perception in the presence of rotation depends on time-varying optic flow. Similar biases for envelope motion and time-varying phase motion suggest that image point trajectories (and their spatial compositions) are not a critical cue for heading perception, consistent with previous reports (15, 32, 39, 55, 56).

The small biases we observed for time-varying stimuli (envelope motion and time-varying phase motion) are consistent in magnitude with those reported in prior studies for the (relatively slow) rotation speeds that we used (10, 15). In accord with these previous reports, we observed that heading bias scaled with rotation speed (Fig. 4*A*). Although observers can compensate for the confounding effects of slow speed rotations (<1°/s), some better than others (*SI Appendix*, Fig. S1), they often exhibit larger errors for faster rotation speeds (10, 15).

Heading Biases Were Better than an Optic Flow Singularity-Based Strategy, for Nonvarying Phase Motion. We created a null model, or "worst case" model of performance (10), that estimates heading as the singularity of optic flow, a biased strategy (*Materials and Methods* and Fig. 1 *B*–*D*), and compared its predictions with observed heading biases (Figs. 3 and 4*B*). The null model predicted heading bias correctly for a subset of observers (O3, O8, O9, O10) and a subset of rotation velocities (i.e., null model predictions fell within the 95% credible intervals of the observed biases) (*SI Appendix*, Fig. S3). This suggests that O3, O8, O9, and O10 may have sometimes estimated heading by simply locating the singularity of optic flow. O3 had large biases for all stimuli (including time-varying), further evidence that O3 may have used this strategy. However, considering the dataset as a whole (pooled across observers), heading judgments were somewhat tolerant to rotation. Biases for nonvarying phase motion (last) fell halfway between the null model and veridical. Biases for nonvarying phase motion (first) fell one-half to one-third of the way between the null model and veridical.

Heading biases were largely indistinguishable for the first and last optic flow fields, suggesting that each instantaneous optic flow field was equally uninformative. Envelope motion and time-varying phase motion stimuli conveyed a common sequence of optic flow fields. To generate nonvarying phase-motion stimuli, we picked a single optic flow field from this sequence, either the first or last. On trials with rotation, the singularity of optic flow drifted toward the true heading over time due to the changing depth and angle of the fronto-parallel planes. If observers used an optic flow singularity-based strategy, biases would be larger for nonvarying (first) than for nonvarying (last). Pooled across observers, heading bias was indistinguishable between the two (Fig. 4*B* and *SI Appendix*, Table S5) ($P > 0.04$) except for

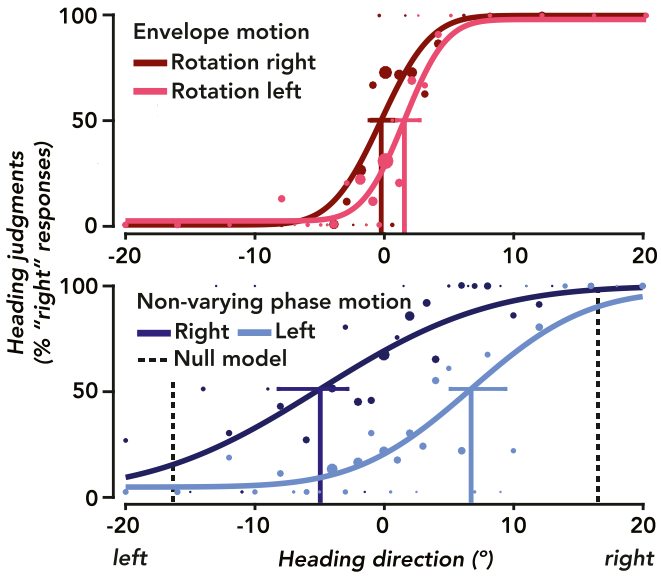

Fig. 3. Example psychometric functions for envelope and nonvarying phase motion. Data from a single observer (n trials per curve = 240). Data points represent percentage of "right" responses at each heading. Dot area indicates the number of data points collected at that heading. Smooth curves indicate best-fit psychometric functions. Solid vertical lines indicate bias. Error bars, 95% credible intervals for the bias estimates. Red, envelope motion; blue, nonvarying phase motion. Dark colors, −2°/s (leftward) rotation; light colors, +2°/s (rightward) rotation. Dashed vertical lines indicate null model predictions.

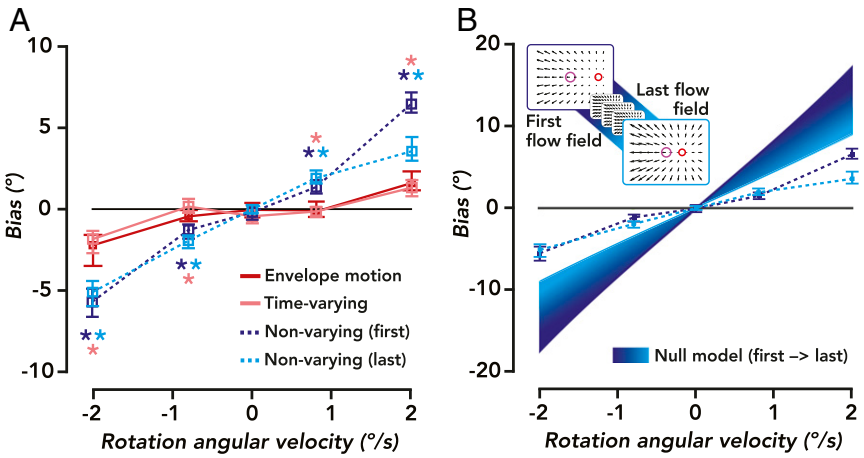

Fig. 4. Bias in heading perception. (*A*) Heading bias, pooled across observers (n observers = 10, n trials per rotation velocity per stimulus condition >960). Square plot symbols, bias estimates. Error bars, 95% credible intervals. Asterisks represent statistical significance at a corrected cutoff of $\alpha = 0.04$ (one-tailed permutation tests). The color of the asterisks indicates the specific hypothesis test that was performed. Light-blue asterisks, nonvarying phase motion (last) was compared to envelope motion. Dark-blue asterisks, nonvarying phase motion (first) was compared to envelope motion. Pink asterisk, nonvarying (first) and nonvarying (last) were compared to time-varying phase motion (an asterisk is plotted if both comparisons were significant). (*B*) Heading bias for nonvarying phase motion, pooled across observers (replotted from *A*). Color gradient, null model predictions across the sequence of optic flow fields and rotation velocities, indicating worst possible performance (upper bound on bias). The *Inset* depicts the correspondence between the color gradient (from dark to light blue) and the sequence of optic flow fields.

rightward 2°/s rotation, for which bias was higher for the first than last optic flow field, an effect driven just by O2 and O6.

Taken together, these results suggest that most observers relied on a heading perception algorithm that was more sophisticated than the null model, even when they only had access to instantaneous optic flow.

Internal Replication of Nonvarying Phase Motion Rules out Training Duration and Testing Order Effects. We conducted an internal replication of nonvarying phase motion (first) after observers performed time-varying phase motion to validate that the observed reduction in bias was not simply due to days of exposure to stimuli utilizing phase motion. Indeed, bias for nonvarying (first, initial) and nonvarying (first, replication) were statistically indistinguishable for ±2°/s and rightward 0.8°/s rotations, when pooled across observers (*SI Appendix*, Fig. S3 and Table S4). For leftward 0.8°/s rotation only, there was a small but significant difference between the two, an effect driven by O2 and O3 (*SI Appendix*, Fig. S3). Together, we interpret this to mean that time-varying optic flow was necessary to achieve good heading perception, and this effect was not the result of training duration or the order of experimental conditions. We pooled nonvarying (first) data—initial and replication—for all other analyses.

Increasing Patch Density Did Not Reduce Biases for Nonvarying Phase Motion. We doubled the number of patches in the stimulus to rule out the possibility that the optic flow field was too sparse. Heading biases were indistinguishable for single- and double-density stimuli for each individual observer, and in aggregate even slightly larger for the higher patch density (*SI Appendix*, Figs. S4 and S5). So-called "differential motion" models (4, 19, 57) estimate heading from local motion parallax available in instantaneous optic flow. Such models may produce larger heading biases for lower densities (57). Following Li and Warren (33), we estimated heading using the method of Rieger and Lawton (57) on optic flow fields corresponding exactly to our stimuli (*Materials and Methods*) with various patch densities. This parallax-based estimator was unbiased for both patch densities (*SI Appendix*, Fig. S6), and these predicted biases were much smaller than the measured biases for nonvarying phase motion. This suggests that the local parallax in our nonvarying stimulus was theoretically sufficient for unbiased heading estimation, but that observers O8, O10, and O12 did not use this cue (even with extensive practice).

Heading Judgments Were Accurate for Time-Varying Phase Motion in the Absence of Local Parallax. To test whether time-varying optic flow is sufficient for accurate heading perception (in addition to being necessary), we removed local parallax by presenting a single fronto-parallel plane with time-varying or nonvarying phase motion. Heading bias was significantly smaller for time-varying than nonvarying phase motion in all observers (*SI Appendix*, Fig. S7). For time-varying phase motion, heading bias was indistinguishable from zero for +2°/s rotation for O8 and O10 (*SI Appendix*, Fig. S7).

Discussion

We developed a stimulus that conveys either an evolving sequence of optic flow fields (time-varying phase motion) or a single optic flow field, the first or last in the sequence (nonvarying phase motion). We found that heading judgments were strongly biased in the direction of rotation for nonvarying, but nearly veridical for time-varying phase motion [i.e., less than 2° on average, below the accuracy required to avoid obstacles when moving at normal speeds (30, 31)]. These effects were observed despite feedback indicating the correct heading direction following each trial (*Materials and Methods*). That is, observers could not learn to make accurate or precise heading judgments without time-varying optic flow, even with extensive practice. This suggests that instantaneous optic flow is insufficient for heading perception in the presence of rotation and that time-varying optic flow is necessary. Stimulus duration was held constant across conditions, so these results cannot be attributed to differences in the duration of evidence accumulation. One implication is that nearly all models of heading estimation are incorrect because they rely on instantaneous optic flow (4, 9, 19, 23–25, 57). Of course, we are not the first to suggest that time-varying optic flow could be useful for heading perception (26, 27, 44, 47).

We propose the "optic acceleration hypothesis,"—that is, that rotation-tolerant estimates of heading are computed from optic acceleration (Fig. 1 *F* and *G*)—on the basis of a simple and

general observation. Optic flow depends on observer translation and rotation, as well as depth structure (*SI Appendix*, Eqs. S7 and S15–S17). When (retinocentric) heading and rotation velocity are constant over time (e.g., motion on a circular path), optic flow evolves solely because of changes in distance to objects in the environment (Fig. 1 *E* and *F* and *SI Appendix*). Under these conditions, heading is aligned with the singularity of the optic acceleration field (or "optic acceleration") (Fig. 1*G*; see *SI Appendix* for derivation). Thus, an estimator locating the singularity of optic acceleration would achieve unbiased performance. If heading or rotation velocity change over time, then such an estimator is biased, and the amount of bias depends on the magnitude of the change. The optic acceleration hypothesis is consistent with our experimental results. For time-varying phase motion, the optic acceleration field radiated from the heading direction at every moment in time. For nonvarying phase motion, the optic acceleration field was by definition zero everywhere, hence it was completely uninformative. Heading bias was larger for nonvarying than time-varying phase motion.

According to the optic acceleration hypothesis, the precision of heading estimates should scale (inversely) with distance, or equivalently, with translational speed. The reliability and robustness of any algorithm for computing heading from optic acceleration depends primarily on dp/dt: The change in the inverse depth map over time (*SI Appendix*). Previously published psychophysical results provide only indirect tests of this prediction, and the results are mixed. The proper manipulation to test our hypothesis—varying translation speed or distance, given constant rotation and heading—has not been performed. In one study (10), using a similar stimulus and task, the rotation and translation rates were simultaneously increased, such that their ratio (58) was held constant. They found a nonsignificant trend such that the reliability of heading judgments increased with translation speed. In another study (15), rotation speed was fixed while translation speed was varied. They did not find significant differences in reliability across translation rates. However, observers were instructed to adjust the line-of-sight during each trial to estimate heading. Doing so introduced changes in rotation and (retinocentric) heading direction over time, both of which should cause errors under our hypothesis. Future studies attempting to test our hypothesis should eliminate all cues for heading other than time-varying optic flow (i.e., using time-varying phase motion), vary translation speed or distance, all while simulating constant observer translation and rotation (i.e., a circular path).

Previously published results demonstrate that heading bias depends on distance and speed. For example, one study (44) reported accurate heading judgments when translation speed was high (larger dp/dt), for translation plus a simulated eye movement with respect to a single fronto-parallel plane. They found larger heading bias, approaching the optic flow singularity, when translation speed was lowered (smaller dp/dt). Similarly, heading bias was larger when the initial distance to the plane was larger (smaller dp/dt). The optic acceleration hypothesis makes a direct prediction about precision, but not bias. Even so, we speculate that observers may fall back on a different (biased) strategy when faced with unreliable acceleration information (i.e., when dp/dt is very low), in the absence of other cues.

There is technically only one degenerate case for the optic acceleration hypothesis: Travel parallel to a ground plane in the absence of vertical eye movements, when the inverse depth map does not change over time. Consequently, introducing real or simulated vertical eye movements should improve performance under our hypothesis. Prior studies have found good performance for heading judgments when the display simulates tracking of a point on the ground (i.e., when vertical eye movements are present) (8, 36, 39, 59, 60). On the other hand, studies simulating a purely horizontal eye movement along the horizon have found larger biases (35, 36). That said, interpretation is complicated because tracking a point on a ground plane can convey an additional "horizon cue" (36, 60, 61). Including vertical objects ("posts") on a ground plane improves both passive path judgments (33) and active steering (41). The optic acceleration hypothesis may explain these results because inverse depth to the posts (but not the ground plane) changes over time, contributing to the optic acceleration field.

Previous models and computer vision algorithms have considered the importance of depth structure (variation in depth throughout the image at each instant in time) for heading estimation (4, 9, 19, 23, 57). According to the optic acceleration hypothesis, depth variation per se is not necessary for accurate heading perception, but richness of depth structure may improve the precision of heading judgments. This prediction is supported by experimental results. In our control experiment, observers were nearly unbiased for motion on a circular path with respect to a single fronto-parallel plane, but only if optic flow evolved over time. Likewise, humans can estimate heading nearly veridically (~1° of bias) for approach perpendicular to a single fronto-parallel plane (FOV = 90° × 90°) with a simulated eye movement (44). In a physiology experiment using a similar single-plane stimulus, ventral intraparietal (VIP) neurons signaled heading in a rotation-tolerant manner (i.e., less biased than the optic flow singularity) (47). In each case, depth variation (and hence motion parallax) was virtually absent, but time-varying optic flow was present.

Increasing FOV can enhance heading perception (33, 42–44, 58). According to the optic acceleration hypothesis, enlarging FOV will lead to increased (and larger) acceleration vectors in the periphery, conferring greater reliability in localizing the singularity of optic acceleration, and more precise heading estimates. Increasing FOV may reduce the small heading biases that we (and others) have observed for time-varying stimuli.

Stereoscopic depth should enhance the reliability of dp/dt, leading to less variable and less biased heading estimates. This is borne out by studies of human behavior (38, 39, 62–64) and might be mediated by the responses of neurons that signal heading and exhibit tuning for both binocular disparity and global motion patterns (65–68). The small heading biases that we (and others) have observed for time-varying stimuli may be reduced when stereo cues are also available.

Human behavior and neurons in the visual cortex are sensitive to optic acceleration. Neurons in the middle temporal (MT) area encode acceleration of moving dot patterns in an implicit manner—via adaptation-induced changes in firing dynamics—meaning that individual neurons cannot unambiguously represent acceleration (69–71). That said, speed-invariant acceleration can be read out from the population (60, 71), suggesting that downstream areas can be directly sensitive to acceleration. Biphasic temporal responses to motion have been measured psychophysically (72) and in the responses of MT neurons (73) (i.e., computing the temporal derivative of image velocity). Similarly, psychophysical experiments reveal that humans are sensitive to optic acceleration (74), with Weber fractions of 0.05 for a 1-Hz sinusoidal temporal frequency modulation (i.e., $0.5°/s^2$ phase speed change) of a drifting grating (74), similar to those for speed discrimination of a drifting grating, which ranged from 0.03 to 0.14 (75, 76). It is crucial to distinguish between acceleration sensitivity for drifting gratings and for moving dots (particularly for a single dot), as Weber fractions for accelerating dot motion are 5 to 10 times higher (77) (see *SI Appendix* for elaboration). Sensitivity to optic acceleration for drifting gratings peaks for $0.5°/s^2$ phase speed change (74), which was similar to the phase speed changes in our time-varying phase motion stimulus ($0.33°/s^2$, mean across patches and time; $0.61°/s^2$, mean across patches at end of stimulus video). Acceleration sensitivity is higher for radial than lateral dot motion patterns that share the same low-level features (78),

suggesting that heading-tuned brain areas (e.g., medial superior temporal [MST], VIP) may be specialized for accelerating motion [although there is also physiological evidence to the contrary (79)]. Furthermore, a subjective sense of self-motion ("vection") is stronger for radial motion when ego-accelerations (which generate visual accelerations) are included (80).

We used nonvarying phase motion to convey a single optic flow field. An ostensibly simpler approach would be to present only two frames of a dot-motion stimulus. However, in practice, this would profoundly limit evidence accumulation and introduce artifacts (spreading spatiotemporal energy) in the stimulus, thereby hindering performance. An alternative approach that has been employed is to present a video with many frames, but with each dot being presented for only two frames and then respawned at a random new location (32). Other studies have used dot lifetimes of hundreds of milliseconds (15, 34, 39, 43, 44, 55, 56). These methods remove or degrade image point trajectories and, along with it, the material derivative of optic flow, as well as optic acceleration within a dot lifetime (see *SI Appendix* for distinction). They do not, however, remove time evolution of optic flow occurring across dot lifetimes (32). Although the observer does not have access to the continuous trajectory of each individual dot throughout the trial, he/she does have access to the continuous evolution of optic flow (Fig. 1*F*); its global structure evolves lawfully over time (7, 47, 53, 81) (*SI Appendix*, Eqs. S15–S17). Indeed, both human behavior (15, 32, 34, 82) and MSTd neurons that signal heading (55, 56) are sensitive to the global structure of optic flow, irrespective of particular dot positions. One seminal study claimed that optic acceleration is not used to compute heading, because they found comparable performance for two- versus three-frame dot lifetimes (32). However, the stimuli in that study simulated a circular path of travel parallel to a ground plane (without vertical eye movements), the one case in which the (inverse) depth map does not evolve, and hence there was no optic acceleration in any of their stimuli. In summary, simply limiting dot lifetime does not remove time-varying optic flow, but nonvarying phase motion does.

Time-varying phase motion eliminates looming, streamlines, and any other cue conveyed by image point trajectories and their spatial compositions (53, 54). Similar biases for envelope and time-varying phase motion suggests that these are not essential cues (15, 32, 34). A previous study proposed that heading-tuned VIP neurons achieve tolerance to visual rotation via sensitivity to the dynamic evolution of retinal motion (47, 83). Our finding of little bias for time-varying phase motion is consistent with their proposal, with one caveat. The authors (47, 83) suggested that "dynamic perspective distortions" (e.g., increasing retinal distance between points composing an object's edge) are the critical cue. We agree, and further specify that it is the temporal evolution of optic flow (i.e., at fixed image locations) that accompanies perspective distortions that is critical, rather than the changing configuration of image points per se. Our results predict that heading-tuned VIP neurons are sensitive to optic acceleration.

Even so, many sensory cues may contribute to heading perception: Motion parallax (33, 43, 47), depth cues (e.g., from accommodation/blur, shading, looming, stereopsis), and extraretinal signals [e.g., eye and neck efference copy (5, 8, 35–37, 45–48, 50, 84), vestibular (85, 86), and proprioceptive inputs (37)] 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 (87, 88). The small biases that we (and others) observed for time-varying stimuli may be abolished when these various cues are all consistent. Indeed, the simulated curvilinear motion stimuli in the present study may be viewed as a conflict between visual and vestibular signals. The resulting biases might, therefore, reflect an optimal combination of these two sources of heading estimates.

Materials and Methods

Participants. Data were acquired from 12 observers (7 male, 5 female). All observers were healthy adults, with no history of neurological disorders and with normal or corrected-to-normal vision. All but one (O8: the author C.S.B.) were naïve to the purposes of the experiment. Two observers, O1 and O9, had no prior experience participating in psychophysical experiments. The remaining observers had considerable experience with psychophysics, but little to none with heading perception tasks. One observer (O11) did not understand the task and was removed from the experiment following training. Observers O1 through O5 performed nonvarying phase motion (first, initial), nonvarying phase motion (first, replication), and time-varying phase motion. Observers O2 and O6 through O8 performed envelope motion. Observers O2 and O6 through O10 performed nonvarying phase motion (last). Observers O8, O10, and O12 performed single-plane nonvarying and time-varying phase motion, as well as double-density phase motion. 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, nonvarying 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, three cycles per degree; contrast, 100%) multiplied by an envelope. The envelope was circular (diameter, 0.3°) with raised-cosine edges (width, 0.15°). 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 directly track the singularity of optic flow, and so that fixation was not disturbed by the motion of the patches. The initial location of each patch was drawn from a uniform distribution. The initial phase of each grating was drawn from a uniform distribution at each trial's onset. In all conditions, the stimulus was a movie with a duration of 4 s, presented with a refresh rate of 75 Hz on a calibrated HP p1230 CRT monitor (HP 2006) with a screen resolution of 1,152 × 870 pixels, subtending 40° by 30°. Observers sat in a darkened room and viewed stimuli binocularly (but not stereoscopically) 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 (89).

For envelope motion, each patch (both the envelope and the plaid) was displaced on each frame (Movies S7 and S8). The location of each patch corresponded to the projection of a point in a simulated three-dimensional (3D) environment onto the 2D image plane (Fig. 1*A*). At each moment in time, the simulated cyclopean viewpoint of the observer (hereafter referred to as the "viewpoint") moved according to the observer's instantaneous translation and rotation. The locations of points in the 3D environment and their projected 2D locations were recomputed, and the patches were rerendered at new locations (Fig. 1*E*). The phase of each plaid was randomly initialized and fixed with respect to its envelope on each trial. Envelope motion is equivalent to "dot motion," a typical self-motion stimulus in the perceptual psychology literature (5, 10, 15, 32). As was the case with all of our stimuli, the retinal size of the plaid patches was constant over time, eliminating looming of individual elements as a cue.

The simulated 3D environment comprised two rigid planes, initially 12.5 m and 25 m, respectively, from the viewpoint at the beginning of each trial (Fig. 1*E*). Each plane was composed of plaid patches. At the beginning of each trial, the density of patches in each plane was 0.16 patches per square degree, giving a total of ∼168 visible patches forming each plane. Occlusion between planes occurred over time, but only for envelope motion. An additional 15° of padding with the same density was included beyond the FOV on each side to prevent any edges from appearing over time as the planes moved with respect to the viewpoint. The locations of plaid patches in the simulated 3D space were projected onto the 2D image plane using perspective projection (focal length, 0.57 m; i.e., equal to the viewing distance).

The simulated observer translated within a plane formed by the x and z axes, while rotating about the y axis and looking down the z axis (Fig. 1*A*). The path of movement was a circle and the line-of-sight (i.e., direction of gaze) was perpendicular to the planes at the beginning of each trial (10). The parameters of the simulated movement were ecologically valid. The chosen translation speed of 1.5 m/s corresponded to average human walking pace and the rotation speeds corresponded to movement along circles of

radius 43 m (2°/s rotation) or 107.4 m (0.8°/s) or a straight path (0°/s, circle of infinite radius). These circles are similar in radius to those used in previous studies (10, 15). The angle between the line-of-sight and heading direction (tangent of circle) was fixed within each trial and was varied adaptively across trials. The line-of-sight corresponded to the center of the CRT monitor, where the fixation dot was located, and observers were told this. Eye tracking was used in all experimental conditions to monitor fixation (see *SI Appendix, Eye Tracking*, for details).

For nonvarying and time-varying phase motion, the plaid envelopes remained stationary while the phases of the gratings shifted over time. Each plaid was comprised of two orthogonal gratings, with component phase speeds specifying a 2D velocity (90), giving a sampling of the underlying optic flow field at a fixed set of image locations. The phase velocity of the field of plaids corresponded to either a single optic flow field (nonvarying) or a sequence of optic flow fields (time-varying). These optic flow fields corresponded to the instantaneous motion of points in the simulated 3D environment relative to the observer. In the nonvarying condition, the phases of a plaid shifted continuously throughout the trial according to the corresponding velocity vector within a single optic flow field. 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 (degree per second) and the spatial frequency (cycles per degree), divided by the elapsed time (1/s).

On each trial of nonvarying phase motion, a video was presented that corresponded to a particular instantaneous optic flow field (Movies S1–S4). The optic flow field presented was either the first or last in the same time-varying sequence used in the envelope motion and time-varying phase-motion conditions (Fig. 4 *B*, *Inset*). We called these stimuli "nonvarying phase motion (first)" and "nonvarying phase motion (last)," or more simply: "nonvarying (first)" and "nonvarying (last)."

In the time-varying phase-motion condition, the phase change between successive frames varied throughout the trial, conveying a sequence of optic flow fields (Fig. 1*F* and Movies S5 and S6). The phase change between frames was computed by generating a sequence of 10 optic flow fields spaced evenly in time over the 4-s trial and calculating the phase change for each optic flow field. These phase changes, which we defined as time-varying optic flow, were determined by the movement of the viewpoint relative to the environment, just as in the envelope motion condition (see *SI Appendix* for details). Time-varying optic flow consisted of local accelerations and decelerations of the phase motion. Consequently, time-varying phase motion conveyed a sequence of optic flow fields temporally subsampled from the full sequence conveyed with envelope motion. The crucial difference, however, between these stimuli was that the patch envelopes did not move at all in the former, but did move in the latter. The trajectories of individual image points may provide a cue for computing heading. Phase motion (time-varying or nonvarying) eliminates this cue. If you walked on circular path with a perforated cardboard board (with many small holes) affixed to your head a few inches out, you would see the visual image evolve only through those apertures. If optic flow estimation were possible in each aperture, this scenario would be analogous to time-varying phase motion.

To generate phase motion stimuli, we computed the optic flow fields corresponding to the motion of the observer's viewpoint relative to the simulated 3D environment for specific combinations of heading and rotation velocity (*SI Appendix*, Eqs. S1–S9). Optic flow for a planar depth structure can be expressed in closed form. Thus, we simply evaluated this function at the x and y image locations of the plaid patches, for each plane separately. A more general approach that will work for analytically intractable depth maps is to densely sample the depth map, compute an equally dense optic flow field from it, and finally subsample the optic flow field at specific screen locations. These calculations were implemented in custom MATLAB software.

Protocol. Observers viewed videos of plaid patches whose image motion corresponded to movement along a circular path with variable combinations of heading direction and rotation velocity, and subsequently performed a forced-choice heading discrimination. Heading direction was tangent to the circular path and gaze was at a fixed angle with respect to heading over time. This task, with the envelope motion stimuli, nearly replicated a previous study (10), and we refer the reader to their methods and figures for elaboration. The main differences between our study and theirs were: 1) We provided correct/incorrect feedback after each trial and they did not; 2) their translation rate was 2 m/s and ours was 1.5 m/s; 3) our stimuli consisted of circular plaid patches and theirs consisted of dots; and 4) our stimulus movie was 4 s in duration and theirs was 2.33 s. The average heading biases measured in that study were within 1° of ours, suggesting that, following training, biases are present with or without tone feedback (for time-varying stimuli), also consistent with the results of a perceptual learning study (49). Movement along a circular path avoids ambiguity about which coordinate system the heading judgment was made in (because retinocentric heading is constant over time) (10). Using a retinocentric judgment helps insure against the possibility of heading estimates being influenced by path perception (10, 15, 34).

Each trial consisted of a 4-s video presentation followed by a 1-s intertrial interval, during which the observer made a key press response indicating whether they perceived their heading to be left or right of center (i.e., the fixation point). Observers rarely failed to respond during the 1-s intertrial interval (0.02% of trials). Immediately following key-press response, observers were given auditory feedback indicating if their response was correct or incorrect (two different tones). The ground truth heading on each trial was set adaptively using two interleaved one-up-one-down staircases (per rotation speed) that converged to 50:50 leftward:rightward choices (i.e., the heading equally often judged as left and right of center). Staircases (initial step size, 4°) were initialized with a heading of ±20° every 30 trials (per staircase) to collect sufficient data at the asymptotes and center of the psychometric function (Fig. 3). There were six staircases in total running simultaneously, and they continued across blocks of trials.

The order of sessions (test and training inclusive) was: Envelope motion, nonvarying (first, initial), time-varying, nonvarying (first, replication), nonvarying (last). One participant (O2) performed this whole sequence. One subgroup (O6 to O8) performed only envelope motion and nonvarying (last). Another subgroup (O1, O3 to O5) performed nonvarying (first, initial), time-varying, and nonvarying (first, replication). Another subgroup (O9 to O10) performed just nonvarying (last). These subgroups were tested in the same order, given that constraint. Each experimental condition consisted of two test sessions (plus prior training sessions), each session comprised 10 blocks, and each block comprised 60 trials, for a total of 600 trials per session. Each session was around an hour in length, including breaks, and was conducted on a different day. Observers were allowed to take breaks between blocks of trials to prevent fatigue. In all conditions, the observer's translation speed was fixed at 1.5 m/s, and the angular velocities of the simulated rotations were −2°/s, −0.8°/s, 0°/s ("no-rotation"), +0.8°/s, or +2°/s. Negative signed velocities are leftward rotation and positive velocities are rightward rotation. These five rotation velocities were interleaved in a randomly permuted order, ensuring equal numbers of trials with each velocity within each block of trials.

Prior to encountering each new stimulus condition, observers performed a minimum of 120 training trials (identical to experimental trials). Training always began with minimum of 60 trials without rotation, followed by a minimum of 60 trials with rotation. Training ensured that observers understood the protocol and had reached asymptotic perceptual sensitivity in the task. A majority of observers required one or two hour-long training sessions (600 trials each) per condition to achieve asymptotic performance on the task, consistent with previous reports (10, 49); some (e.g., O5, O10) required significantly less (~120 trials per condition). To determine when to end training and start the experiment, we checked that that discrimination accuracy was ~70 to 80% across all trials and ensured that the average converged heading values were consistent across the two interleaved staircases (for each rotation speed) and stable across runs of trials. We also visually inspected the staircases to see that they had converged properly. Observers' discrimination thresholds decreased over training, asymptoting at 1° to 8°, depending on the stimulus and rotation velocity (*SI Appendix*, Fig. S2). Observers were never given feedback about rotation speed or information about the number of rotation speeds tested during training or the experiment.

Data Analysis. Heading bias and discrimination threshold were estimated using Psignifit 4, a MATLAB toolbox for Bayesian inference for psychometric functions (91). Cumulative normal psychometric functions were fit separately to each individual observer's responses and also to the pooled responses of all observers. We fit a β-binomial model with four free parameters (μ, σ, λ, η), which respectively corresponded to the mean and slope of the psychometric function, the lapse rate, and the overdispersion of the data relative to the binomial model (Fig. 3). The mean of the psychometric function or PSE describes the heading for which the observer reported 50:50

leftward:rightward of center (i.e., leftward:rightward of the location of the fixation point). A bias of 0° minus this heading is needed to cancel out the horizontal shift in the psychometric function. Bias describes how far the observer's internal estimate of center is from 0° (i.e., true center). For example, if an observer viewed a pure translational (0° heading) optic flow field and had a bias of −2°, that would suggest that the observer would interpret a heading of 0° as a leftward heading of 2°. The slope of a cumulative normal psychometric function determines the discrimination threshold of the observer in a two-alternative forced choice (2-AFC) task, defining how sensitive of a classifier they are (*SI Appendix*, Fig. S2). Lapse rate describes how often observers lapsed, pressing the wrong key by accident. It sets an upper limit on accuracy by jointly controlling the psychometric function's upper and lower asymptotes. Overdispersion describes how much the dispersion of the observed data exceeded the dispersion expected under a 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 integration over the four-dimensional posterior, $p(\Theta|\text{data})$, where Θ 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. Statistical comparison of heading biases across conditions was performed via permutation test (see *SI Appendix, Statistics*, for details).

Null Model. When there is observer rotation in addition to translation, the singularity of the optic flow field is displaced away from the heading in the direction of rotation (Fig. 1 *B–D*). Thus, an estimator using only the optic flow singularity would be systematically biased, reporting heading as the visual angle subtending the singularity location. We defined our null model in this way [as in Stone and Perrone (10)]. The singularity position was computed by extracting the horizontal component of optic flow (*SI Appendix*, Eq. S7) and solving for the point at which this function crossed zero. Null model predictions were generated for the near (12.5 m) plane across a finely spaced range of rotation angular velocities (−2 to +2°/s) and times (0 to 4 s of the trial), using the same parameters that were used to generate the experimental stimuli. Null model predictions for the far plane were much larger than for the near plane (i.e., around twice as large) and were not reported because observers always performed much better than that. The null model is not a serious model of how the visual system computes heading under natural conditions. It is a biased algorithm by construction, a lower bound on performance we are using as a reference point for the observed heading biases in the nonvarying phase motion condition.

Double Patch Density and Single-Plane Control Experiments. We performed three control experiments ($n = 3$) in a counter-balanced order: 1) Double patch density nonvarying phase motion, 2) single-plane nonvarying phase motion (last), 3) single-plane time-varying phase motion (last). For the first, we doubled the number of patches in each plane from 0.16 to 0.32 patches per square degree, which yielded a total patch number of 336 or 672, and an average interpatch Euclidean distance (edge to edge) of 2.83° or 1.26° (*SI Appendix*, Fig. S4). For the second, we replicated the original nonvarying phase motion condition, but kept the 12.5-m plane and removed the 25-m plane. We also doubled the overall patch density to match the original stimuli's overall patch count. For the third, we did the same as the second, but varied the optic flow over time. For the first and second, we used the last flow field because it should yield the best performance under the null model so is the most liberal choice. For these control experiments, we included only ±2°/s rotation and no-rotation trials.

Differential Motion Model. We implemented the method described in Rieger and Lawton (57) in custom Matlab software. Following refs. 33 and 57, for each optic flow vector we computed the group of closest vectors (within 3° Euclidean distance), used total least-squares regression to find the best-fit line through each group of difference vectors, and estimated the singularity of the difference field as the intersection of all of the best-fit lines in the least-squares sense. Input optic flow fields were computed using our stimulus generation code and matched exactly the flow presented in the experiments. We varied the ground truth heading from −20° to 20° and the rotation velocity among −2, −0.8, 0, +0.8, and +2°/s. The flow element density was set to 0.16 or 0.8 patches per square degree per plane. As the estimator's bias did not significantly vary across headings or across flow fields in the sequence (<0.5° mean difference), we proceeded by estimating bias solely for 0° heading for the last flow field in the sequence. Bias estimates were averaged across 40 repetitions, for which different vector tail positions were drawn from a 2D uniform distribution.

Data and Code Availability. Binary choices, gaze data, stimulus generation, analysis, and modelling code have been deposited in http://hdl.handle.net/2451/61543.

ACKNOWLEDGMENTS. We thank Shani Offen, Davis Glasser, Elisha Merriam, and Ionatan Kuperwajs for conducting preliminary versions of these experiments; Shannon Locke, Hörmet Yiltiz, and Emmanouil Protonotarios for providing comments on experimental design and analysis; Greg DeAngelis and Eero Simoncelli for helpful discussion; and Michael Landy, Bas Rokers, Mengjian Hua, and Kathryn Bonnen for providing invaluable feedback on the manuscript. This research was supported by National Eye Institute Visual Neuroscience Training Grant T32 EY007136 (to C.S.B. through New York University) and a National Defense Science & Engineering Graduate fellowship (to C.S.B.).

1. J. J. Gibson, *The Perception of the Visual World* (The Riverside Press, 1950).
2. J. J. Gibson, *The Senses Considered as Perceptual Systems* (Houghton Mifflin, Oxford, UK, 1966).
3. J. J. Koenderink, A. J. van Doorn, Local structure of movement parallax of the plane. *J. Opt. Soc. Am.* **66**,717–723 (1976).
4. H. C. Longuet-Higgins, K. Prazdny, The interpretation of a moving retinal image. *Proc. R. Soc. Lond. B Biol. Sci.* **208**, 385–397 (1980).
5. W. H. Warren, D. J. Hannon, Direction of self-motion is perceived from optical flow. *Nature* **336**, 162–163 (1988).
6. P. Baraldi, E. D. Micheli, S. Uras, "Motion and depth from optical flow" in *Proceedings of the Alvey Vision Conference*, K. D. Baker, Ed. (Alvey Vision Club, 1989) pp. 35.1–35.4.
7. A. Verri, F. Girosi, V. Torre, Mathematical properties of the two-dimensional motion field: From singular points to motion parameters. *J. Opt. Soc. Am.* **6**, 698–712 (1989).
8. W. H. Warren Jr, D. J. Hannon, Eye movements and optical flow. *J. Opt. Soc. Am. A* **7**, 160–169 (1990).
9. D. J. Heeger, A. D. Jepson, Subspace methods for recovering rigid motion I: Algorithm and implementation. *Int. J. Comput. Vis.* **7**, 95–117 (1992).
10. L. S. Stone, J. A. Perrone, Human heading estimation during visually simulated curvilinear motion. *Vision Res.* **37**, 573–590 (1997).
11. M. Lappe, F. Bremmer, A. V. Van Den Berg, Perception of self-motion from visual flow. *Trends Cogn. Sci.* **3**, 329–336 (1999).
12. W. H. arren Jr, B. A. Kay, W. D. Zosh, A. P. Duchon, S. Sahuc, Optic flow is used to control human walking. *Nat. Neurosci.* **4**, 213–216 (2001).
13. R. Wilkie, J. Wann, Controlling steering and judging heading: Retinal flow, visual direction, and extraretinal information. *J. Exp. Psychol. Hum. Percept. Perform.* **29**, 363–378 (2003).
14. W. H. Warren Jr, *Optic Flow in the Visual Neurosciences* (University of Cambridge Press, 2004).
15. L. Li, B. T. Sweet, L. S. Stone, Humans can perceive heading without visual path information. *J. Vis.* **6**, 874–881 (2006).
16. J. A. Perrone, Model for the computation of self-motion in biological systems. *J. Opt. Soc. Am. A* **9**, 177–194 (1992).
17. M. Lappe, J. P. Rauschecker, A neural network for the processing of optic flow from ego-motion in man and higher mammals. *Neural Comput.* **5**, 374–391 (1993).
18. J. A. Perrone, L. S. Stone, A model of self-motion estimation within primate extrastriate visual cortex. *Vision Res.* **34**, 2917–2938 (1994).
19. C. S. Royden, Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **14**, 2128–2143 (1997).
20. J. A. Beintema, A. V. van den Berg, Heading detection using motion templates and eye velocity gain fields. *Vision Res.* **38**, 2155–2179 (1998).
21. S. Grossberg, E. Mingolla, C. Pack, A neural model of motion processing and visual navigation by cortical area MST. *Cereb. Cortex* **9**, 878–895 (1999).
22. R. F. Wang, J. E. Cutting, A probabilistic model for recovering camera translation. *Comput. Vis. Image Underst.* **76**, 205–212 (1999).
23. F. Raudies, H. Neumann, A review and evaluation of methods estimating ego-motion. *Comput. Vis. Image Underst.* **116**, 606–633 (2012).
24. A. J. Foulkes, S. K. Rushton, P. A. Warren, Heading recovery from optic flow: Comparing performance of humans and computational models. *Front. Behav. Neurosci.* **7**, 53 (2013).
25. J. A. Perrone, Visual-vestibular estimation of the body's curvilinear motion through the world: A computational model. *J. Vis.* **18**, 1 (2018).
26. J. H. Rieger, Information in optical flows induced by curved paths of observation. *J. Opt. Soc. Am.* **73**, 339–344 (1983).
27. J. L. Barron, A. D. Jepson, J. K. Tsotsos, The feasibility of motion and structure from noisy time-varying image velocity information. *Int. J. Comput. Vis.* **5**, 239–269 (1990).

28. J. L. Barron, "Motion and structure in rigid multi-surfaced stationary environments using time-varying image velocity: Linear solutions" in *Visual Form: Analysis and Recognition*, C. Arcelli, L. P. Cordella, G. S. di Baja, Eds. (Springer, Boston, MA, 1992) pp. 39–46.
29. J. L. Barron, R. Eagleson, Recursive estimation of time-varying motion and structure parameters. *Pattern Recognit.* **29**, 797–818 (1996).
30. J. E. Cutting, *Perception with an Eye for Motion* (MIT Press, Cambridge, MA, 1986).
31. J. E. Cutting, K. Springer, P. A. Braren, S. H. Johnson, Wayfinding on foot from information in retinal, not optical, flow. *J. Exp. Psychol. Gen.* **121**, 41–72 (1992).
32. W. H. Warren Jr, A. W. Blackwell, K. J. Kurtz, N. G. Hatsopoulos, M. L. Kalish, On the sufficiency of the velocity field for perception of heading. *Biol. Cybern.* **65**, 311–320 (1991).
33. L. Li, W. H. Warren Jr, Perception of heading during rotation: Sufficiency of dense motion parallax and reference objects. *Vision Res.* **40**, 3873–3894 (2000).
34. L. Li, J. C. Cheng, Perceiving path from optic flow. *J. Vis.* **11**, 22 (2011).
35. C. S. Royden, M. S. Banks, J. A. Crowell, The perception of heading during eye movements. *Nature* **360**, 583–585 (1992).
36. C. S. Royden, J. A. Crowell, M. S. Banks, Estimating heading during eye movements. *Vision Res.* **34**, 3197–3214 (1994).
37. J. A. Crowell, M. S. Banks, K. V. Shenoy, R. A. Andersen, Visual self-motion perception during head turns. *Nat. Neurosci.* **1**, 732–737 (1998).
38. A. V. van den Berg, E. Brenner, Humans combine the optic flow with static depth cues for robust perception of heading. *Vision Res.* **34**, 2153–2167 (1994).
39. A. V. van den Berg, E. Brenner, Why two eyes are better than one for judgements of heading. *Nature* **371**, 700–702 (1994).
40. J. H. Rieger, L. Toet, Human visual navigation in the presence of 3-D rotations. *Biol. Cybern.* **52**, 377–381 (1985).
41. L. Li, W. H. Warren Jr, Retinal flow is sufficient for steering during observer rotation. *Psychol. Sci.* **13**, 485–491 (2002).
42. L. Li, W. H. Warren Jr, Path perception during rotation: Influence of instructions, depth range, and dot density. *Vision Res.* **44**, 1879–1889 (2004).
43. L. Li, J. Chen, X. Peng, Influence of visual path information on human heading perception during rotation. *J. Vis.* **9**, 1–14 (2009).
44. A. Grigo, M. Lappe, Dynamical use of different sources of information in heading judgments from retinal flow. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **16**, 2079–2091 (1999).
45. M. S. Banks, S. M. Ehrlich, B. T. Backus, J. A. Crowell, Estimating heading during real and simulated eye movements. *Vision Res.* **36**, 431–443 (1996).
46. S. M. Ehrlich, D. M. Beck, J. A. Crowell, T. C. A. Freeman, M. S. Banks, Depth information and perceived self-motion during simulated gaze rotations. *Vision Res.* **38**, 3129–3145 (1998).
47. A. Sunkara, G. C. DeAngelis, D. E. Angelaki, Role of visual and non-visual cues in constructing a rotation-invariant representation of heading in parietal cortex. *eLife* **4**, e04693 (2015).
48. A. Sunkara, G. C. DeAngelis, D. E. Angelaki, Joint representation of translational and rotational components of optic flow in parietal cortex. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 5077–5082 (2016).
49. S. Kuang, H. Deng, T. Zhang, Adaptive heading performance during self-motion perception. *PsyCh J.* **8**, 1–11 (2019).
50. T. S. Manning, K. H. Britten, Retinal stabilization reveals limited influence of extraretinal signals on heading tuning in the medial superior temporal area. *J. Neurosci.* **39**, 8064–8078 (2019).
51. O. W. Layton, B. R. Fajen, The temporal dynamics of heading perception in the presence of moving objects. *J. Neurophysiol.* **115**, 286–300 (2016).
52. W. T. Freeman, E. H. Adelson, D. J. Heeger, Motion without movement. *Computer Graphics* **25**, 27–30.
53. J. J. Koenderink, A. J. van Doorn, Exterospecific component of the motion parallax field. *J. Opt. Soc. Am.* **71**, 953–957 (1981).
54. D. Regan, K. I. Beverley, How do we avoid confounding the direction we are looking and the direction we are moving? *Science* **215**, 194–196 (1982).
55. M. S. Graziano, R. A. Andersen, R. J. Snowden, Tuning of MST neurons to spiral motions. *J. Neurosci.* **14**, 54–67 (1994).
56. H. Sugihara, I. Murakami, K. V. Shenoy, R. A. Andersen, H. Komatsu, Response of MSTd neurons to simulated 3D orientation of rotating planes. *J. Neurophysiol.* **87**, 273–285 (2002).
57. J. H. Rieger, D. T. Lawton, Processing differential image motion. *J. Opt. Soc. Am. A* **2**, 354–360 (1985).
58. J. J. Koenderink, A. J. van Doorn, Facts on optic flow. *Biol. Cybern.* **56**, 247–254 (1987).
59. A. V. Van den Berg, Perception of heading. *Nature* **365**, 497–498 (1993).
60. A. V. van den Berg, Judgements of heading. *Vision Res.* **36**, 2337–2350 (1996).
61. A. V. van den Berg, Robustness of perception of heading from optic flow. *Vision Res.* **32**, 1285–1296 (1992).
62. A. Grigo, M. Lappe, Interaction of stereo and opic flow revealed by an illusory stimulus. *Vision Res.* **38**, 281–290 (1998).
63. J. S. Butler, J. L. Campos, H. H. Bülthoff, S. T. Smith, The role of stereo vision in visual-vestibular integration. *Seeing Perceiving* **24**, 453–470 (2011).
64. M. Lich, F. Bremmer, Self-motion perception in the elderly. *Front. Hum. Neurosci.* **8**, 681 (2014).
65. J. P. Roy, H. Komatsu, R. H. Wurtz, Disparity sensitivity of neurons in monkey extrastriate area MST. *J. Neurosci.* **12**, 2478–2492 (1992).
66. S. Eifuku, R. H. Wurtz, Response to motion in extrastriate area MSTl: Disparity sensitivity. *J. Neurophysiol.* **82**, 2462–2475 (1999).
67. G. C. DeAngelis, Seeing in three dimensions: The neurophysiology of stereopsis. *Trends Cogn. Sci.* **4**, 80–90 (2000).
68. Y. Yang, S. Liu, S. A. Chowdhury, G. C. DeAngelis, D. E. Angelaki, Binocular disparity tuning and visual-vestibular congruency of multisensory neurons in macaque parietal cortex. *J. Neurosci.* **31**, 17905–17916 (2011).
69. S. G. Lisberger, J. A. Movshon, Visual motion analysis for pursuit eye movements in area MT of macaque monkeys. *J. Neurosci.* **19**, 2224–2246 (1999).
70. N. S. Price, S. Ono, M. J. Mustari, M. R. Ibbotson, Comparing acceleration and speed tuning in macaque MT: Physiology and modeling. *J. Neurophysiol.* **94**, 3451–3464 (2005).
71. A. Schlack, B. Krekelberg, T. D. Albright, Recent history of stimulus speeds affects the speed tuning of neurons in area MT. *J. Neurosci.* **27**, 11009–11018 (2007).
72. Y. Ohnishi, K. Kawano, K. Miura, Temporal impulse response function of the visual system estimated from ocular following responses in humans. *Neurosci. Res.* **113**, 56–62 (2016).
73. D. M. Glasser, J. M. Tsui, C. C. Pack, D. Tadin, Perceptual and neural consequences of rapid motion adaptation. *Proc. Natl. Acad. Sci. U.S.A.* **108**, E1080–E1088 (2011).
74. R. Nakayama, I. Motoyoshi, Sensitivity to acceleration in the human early visual system. *Front. Psychol.* **8**, 925 (2017).
75. S. P. McKee, G. H. Silverman, K. Nakayama, Precise velocity discrimination despite random variations in temporal frequency and contrast. *Vision Res.* **26**, 609–619 (1986).
76. V. Lakshminarayanan, A. Raghuram, R. Khanna, Psychophysical estimation of speed discrimination. I. Methodology. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **22**, 2262–2268 (2005).
77. P. Werkhoven, H. P. Snippe, A. Toet, Visual processing of optic acceleration. *Vision Res.* **32**, 2313–2329 (1992).
78. A. S. Mueller, B. Timney, Visual acceleration perception for simple and complex motion patterns. *PLoS One* **11**, e0149413 (2016).
79. M. Paolini, C. Distler, F. Bremmer, M. Lappe, K.-P. Hoffmann, Responses to continuously changing optic flow in area MST. *J. Neurophysiol.* **84**, 730–743 (2000).
80. S. Palmisano, R. S. Allison, F. Pekin, Accelerating self-motion displays produce more compelling vection in depth. *Perception* **37**, 22–33 (2008).
81. J. J. Koenderink, Optic flow. *Vision Res.* **26**, 161–179 (1986).
82. A. M. L. Kappers, S. F. Te Pas, J. J. Koenderink, Detection of divergence in optical flow fields. *J. Opt. Soc. Am.* **13**, 227 (1996).
83. H. R. Kim, D. E. Angelaki, G. C. DeAngelis, A novel role for visual perspective cues in the neural computation of depth. *Nat. Neurosci.* **18**, 129–137 (2015).
84. L. Chukoskie, J. A. Movshon, Modulation of visual signals in macaque MT and MST neurons during pursuit eye movement. *J. Neurophysiol.* **102**, 3225–3233 (2009).
85. Y. Gu, G. C. DeAngelis, D. E. Angelaki, A functional link between area MSTd and heading perception based on vestibular signals. *Nat. Neurosci.* **10**, 1038–1047 (2007).
86. K. E. Cullen, Vestibular processing during natural self-motion: Implications for perception and action. *Nat. Rev. Neurosci.* **20**, 346–363 (2019).
87. C. R. Fetsch, G. C. Deangelis, D. E. Angelaki, Visual-vestibular cue integration for heading perception: Applications of optimal cue integration theory. *Eur. J. Neurosci.* **31**, 1721–1729 (2010).
88. M. S. Landy, M. S. Banks, D. C. Knill, "Ideal-observer models of cue integration" in *Sensory Cue Integration*, J. Trommershauser, K. Kording, M. S. Landy, Eds. (Oxford University Press, 2011) pp. 5–29.
89. J. L. Gardner, E. P. Merriam, D. Schluppeck, J. Larsson, *MGL: Visual Psychophysics Stimuli and Experimental Design Package* (Zenodo, 2018).
90. E. H. Adelson, J. A. Movshon, Phenomenal coherence of moving visual patterns. *Nature* **300**, 523–525 (1982).
91. H. H. Schütt, S. Harmeling, J. H. Macke, F. A. Wichmann, Painfree and accurate Bayesian estimation of psychometric functions for (potentially) overdispersed data. *Vision Res.* **122**, 105–123 (2016).