



Causal inference accounts for heading perception in the presence of object motion

Kalpana Dokka^a, Hyeshin Park^a, Michael Jansen^a, Gregory C. DeAngelis^b, and Dora E. Angelaki^{a,c,1}

^aDepartment of Neuroscience, Baylor College of Medicine, Houston, TX 77030; ^bDepartment of Brain and Cognitive Sciences, University of Rochester, Rochester, NY 14627; and ^cCenter for Neural Science, New York University, New York, NY 10003

Contributed by Dora E. Angelaki, March 13, 2019 (sent for review November 29, 2018; reviewed by Brett R. Fajen and Uta Noppeney)

The brain infers our spatial orientation and properties of the world from ambiguous and noisy sensory cues. Judging self-motion (heading) in the presence of independently moving objects poses a challenging inference problem because the image motion of an object could be attributed to movement of the object, self-motion, or some combination of the two. We test whether perception of heading and object motion follows predictions of a normative causal inference framework. In a dual-report task, subjects indicated whether an object appeared stationary or moving in the virtual world, while simultaneously judging their heading. Consistent with causal inference predictions, the proportion of object stationarity reports, as well as the accuracy and precision of heading judgments, depended on the speed of object motion. Critically, biases in perceived heading declined when the object was perceived to be moving in the world. Our findings suggest that the brain interprets object motion and self-motion using a causal inference framework.

heading | causal inference | object motion | multisensory | optic flow

As environmental cues activate our senses, the brain synthesizes and updates an internal representation of the world that is computed from noisy and often ambiguous sensory cues. Causal inference formalizes this process as a normative statistically optimal computation that is fundamental for sensory perception and cognition (1–5). The causal inference framework has previously been used to explain processing of multimodal signals, including auditory–visual (1, 4–9), visual–speech (10, 11), visual–vestibular (12, 13), and visual–tactile interactions (14, 15). However, its effectiveness in framing more general problems of sensory perception within the visual domain has remained unclear. Indeed, it has previously been suggested that visual cues are mandatorily integrated (16, 17).

When a moving observer views a scene containing moving objects, visual perception faces an interpretational challenge: motion of an object on the retina could be due to independent movement of the object relative to the scene, to the observer’s self-motion, or to many different combinations of object and observer motion. Indeed, it is well established that moving objects can bias self-motion (heading) perception (18–23). Conversely, an observer’s self-motion can also bias perception of object trajectory (24–27). However, these interactions have generally been interpreted using empirical models based on visual motion processing mechanisms, and no attempt has been made to cast the problem of judging object and self-motion within a normative framework that can be generalized across sensory domains (28–32).

Results

We test whether a causal inference scheme can account for the motion interpretation challenge faced when a moving observer views a scene in which an object may be moving independently. In general, the image motion of an object on the retina is a vector sum of the observer’s self-motion and the motion of the object in the world (Fig. 1A, x_{obj}). We varied the speed of the object in the (virtual) world to manipulate the extent to which the image motion associated with the object deviated from that

expected due to self-motion. To correctly judge the direction of self-motion (heading), the observer has to infer the source of retinal image motion, and discount components due to object motion relative to the scene. We formulate three possible strategies to solve this problem (*SI Appendix, SI Methods* for details).

At one extreme, the observer may believe that the object is stationary in the world, even when it is not (Fig. 1B; solid curves). Image motion due to the object’s movement relative to the scene would then be incorrectly attributed to the observer’s self-motion and would be interpreted as a heading cue, similar to the spatial pooling hypothesis (20) in which heading is estimated from the foci of expansion produced by both background and object motion. We formalize this strategy in an Integration (Int) model, in which the observer integrates retinal object motion (x_{obj}) with self-motion cues (optic flow, x_{vis} ; inertial motion, x_{vest} when available), resulting in a heading bias that monotonically increases with object speed in the world (Fig. 1C, solid curves; see *SI Appendix* for details). Alternatively, the observer may always believe that the object is moving independently in the world (Fig. 1B, dashed curves). Consequently, the observer should disregard the object motion cue when judging self-motion. In a Segregation (Seg) model, the retinal image motion of the object influences neither the bias in perceived heading (Fig. 1C, dashed curves) nor the threshold of heading discrimination (Fig. 1D, dashed curves).

On the other hand, an optimal solution should not make default assumptions. An optimal causal inference (OCI) model takes into account the characteristics of the stimuli on a trial-by-trial basis to infer whether the object is stationary and whether/how object and self-motion cues should be integrated (see *SI*

Significance

A moving observer faces an interpretational challenge when the visual scene contains independently moving objects. To correctly judge their direction of self-motion (heading), an observer must appropriately infer the causes of retinal image motion. We demonstrate that perception of object motion systematically influences heading judgments. When a moving object was erroneously perceived to be stationary in the world, heading judgments were significantly biased. On the contrary, when the object was correctly perceived to be moving, biases in heading estimates declined dramatically. Our results thus suggest that the brain makes inferences about sources of image motion using causal inference computations.

Author contributions: K.D., M.J., G.C.D., and D.E.A. designed research; K.D., H.P., and M.J. performed research; K.D., G.C.D., and D.E.A. analyzed data; and K.D., G.C.D., and D.E.A. wrote the paper.

Reviewers: B.R.F., Rensselaer Polytechnic Institute; and U.N., University of Birmingham.

The authors declare no conflict of interest.

Published under the [PNAS license](#).

¹To whom correspondence should be addressed. Email: da93@nyu.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1820373116/-DCSupplemental.

Published online April 17, 2019.

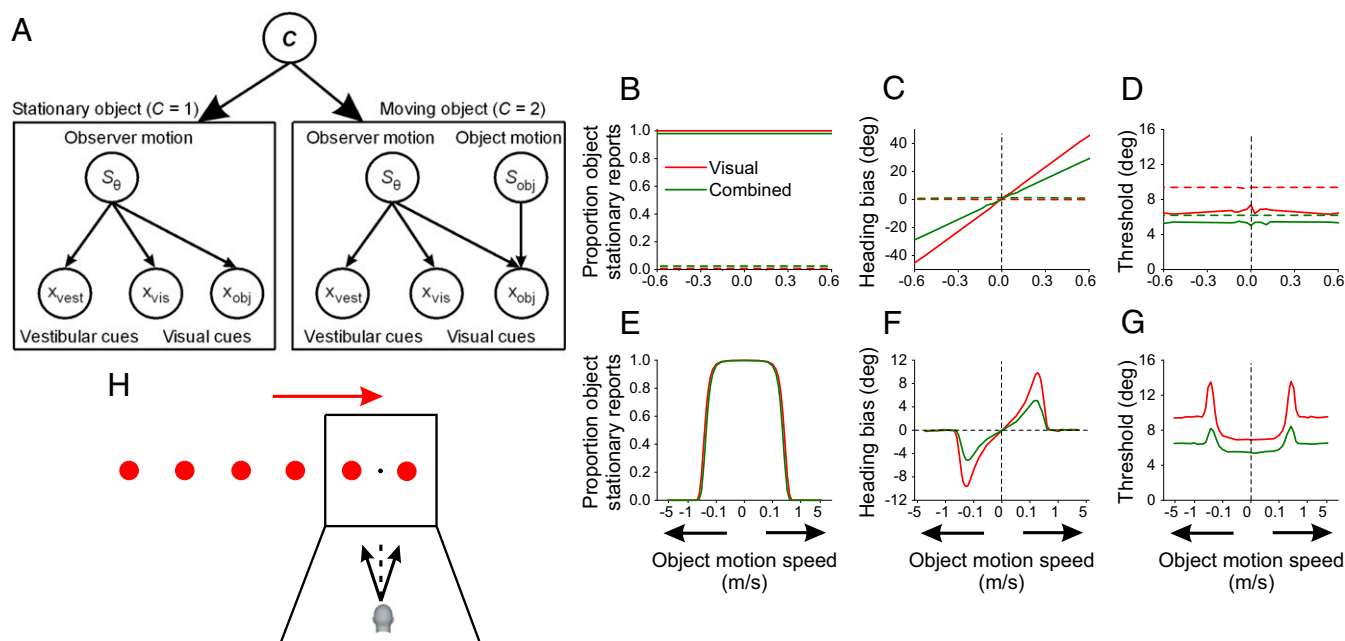


Fig. 1. Task design and model simulations. (A) Graphical model illustrating relationships between the variables implicated in perception of object motion and self-motion. S_{obj} and S_{θ} represent the motion of the object and the observer in the world, respectively. x_{vest} and x_{vis} represent the noisy inertial (vestibular) and visual (optic flow derived) heading cues, respectively. x_{obj} is the noisy image motion of the object relative to the observer. Inertial cues only contribute to self-motion perception during multisensory stimulation. *Left and Right* illustrate the generative models for the cases of a stationary object ($C = 1$) and a moving object ($C = 2$), respectively. (B–D) Simulations of the Integration (Int, solid lines) and Segregation (Seg, dashed lines) models (SI Appendix). Colored lines represent simulations for the visual (red) and combined visual–vestibular (green) self-motion conditions. Note that small vertical offsets have been added to the green lines in B (solid and dashed lines) and C (dashed line) to improve visualization; otherwise, green and red horizontal lines would be superimposed. (E–G) Simulations of the OCI model. Predicted proportion of object stationary reports (B and E), bias in perceived heading (C and F), and heading discrimination threshold (D and G) are shown for each of the models. Arrows on the x axis indicate rightward and leftward motion of the object. For each combination of object motion speed, direction of object motion, and heading, the proportion of rightward heading responses was computed and psychometric functions were fitted to the simulated responses (hence the variability in the predictions). (H) To test model predictions, subjects experienced forward translation to the left or right of straight ahead (black arrows) while viewing a multipart object that was either stationary or moved independently rightward (red arrow) or leftward in the world at different speeds. The object was composed of six identical and equidistant spheres. The spheres moved together as a single compound object, such that two of the spheres were always visible on the display throughout the trial at all object motion speeds. For rightward motion of the object, the two rightmost spheres were displayed on the screen at the start of the trial (as illustrated here, see also Movie S1), and vice versa for leftward object motion.

Appendix for details). There are three key predictions of the OCI model:

- i) The observer should perceive the object to be stationary less frequently as object speed increases (Fig. 1E) because the difference between motion vectors associated with the object and the background increases with object speed.
- ii) The bias in perceived heading should have a peaked dependence on object speed. The bias should initially increase with object speed, but should later decline toward zero as the object speed increases further and the subject consistently perceives the object to be moving (Fig. 1F). The increase in heading bias at slow speeds occurs because the object is perceived to be stationary; thus, as in the Integration model, it provides a relevant cue to self-motion. Consequently, object motion and background motion cues should be integrated, leading to significant heading biases (since the object is not truly stationary). However, at high speeds, the object is perceived to move independently. Thus, as in the Segregation model, the observer infers that image motion of the object is no longer a useful cue to self-motion, resulting in greatly reduced heading biases.
- iii) Heading discrimination thresholds should initially increase and peak at intermediate object speeds before settling down to an elevated level as object speed increases further (Fig. 1G). The heading estimate in the OCI model is a weighted sum of the heading estimates associated with the stationary

and moving object scenarios (SI Appendix). At slow object speeds, there is little uncertainty that the object is stationary in the world. Thus, greater weight is assigned to the heading estimate obtained by integrating self-motion and retinal object motion cues, leading to relatively small thresholds. At intermediate speeds, uncertainty about whether the object is stationary or moving is greatest. Comparable weights are thus assigned to heading estimates associated with both stationary and moving object scenarios. Thus, the unified heading estimate is a weighted combination of two estimates drawn from different distributions with unequal means and variances, leading to elevated thresholds. At fast speeds, there is little uncertainty that the object is moving, leading to strong weighting of only heading estimates associated with the moving object scenario. Here, the heading estimate essentially represents a single distribution and has a correspondingly smaller threshold compared with intermediate speeds.

We tested the above predictions with 14 human subjects who experienced self-motion that was either visually simulated by optic flow or was presented as a congruent and synchronous multimodal combination of optic flow and inertial (vestibular) motion. Subjects experienced one of these forms of self-motion while viewing a multipart object that moved in the virtual world at different speeds (Fig. 1H and Movie S1). The moving object was composed of a group of six identical and equidistant spheres, which were spaced such that two of the spheres were visible at

any point in time for all speeds of motion (*SI Appendix, SI Methods* and *Movie S1*). The spheres were stationary with respect to each other and moved collectively either rightward or leftward in the world as a compound rigid object. We tested a wide range of relative image velocities between object and background by varying the speed of object motion in the world. We used a dual-report task, in which participants judged whether an object was stationary or moved in the virtual world, as well as whether perceived heading was rightward or leftward of straight ahead.

As shown for a representative subject (Fig. 2), the object was perceived to be stationary in the world less frequently as it moved faster in either direction (Fig. 2A). In parallel, object speed strongly affected heading discrimination at intermediate speeds (Fig. 2B–D). Rightward object motion shifted the psychometric function to the right, indicating a leftward bias in perceived heading (Fig. 2B, filled symbols and solid curves). Analogous but opposite shifts were seen for leftward object motion (Fig. 2B, open symbols and dashed curves). Most notably, heading biases decreased or disappeared for faster speeds (Fig. 2B, blue and pink; also summarized in Fig. 2C). The fact that object motion biased heading perception substantially at slow/intermediate speeds, but much less so at faster speeds, is consistent with predictions of the OCI model.

To quantify the effects of object and self-motion parameters, a repeated-measures ANOVA, having object motion speed, object motion direction, and self-motion type as factors, was applied to the following dependent variables: proportion of object stationary reports, bias in perceived heading, and heading discrimination thresholds. For statistical analyses, heading biases obtained with leftward object motion were multiplied by -1 , such that expected biases had the same sign for both directions of object

motion (18, 24). Results were consistent across 14 human subjects, as summarized in Fig. 3 (*SI Appendix, Fig. S1* for individual subject data). Overall, subjects perceived the object to be stationary less frequently at higher object speeds [Fig. 3A, main effect of object speed; $F(6, 364) = 603.17, P < 0.0001$; see *SI Appendix, Table S1*]. Across subjects, biases in perceived heading also depended strongly on object speed [Fig. 3B, main effect of object motion speed; $F(6, 364) = 10.66, P < 0.0001$; see *SI Appendix, Table S2* for post hoc tests]. Notably, biases initially increased with object speeds up to ~ 0.2 m/s, but then declined considerably as object speed increased further. In fact, biases measured at the highest object speeds (1.33, 2.67, and 5.35 m/s) were not significantly different from the bias produced by a stationary object (i.e., object speed = 0 m/s; see *SI Appendix, Table S2*). These results show that, as the object's motion becomes more different from the optic flow expected from self-motion, the influence of the moving object on the accuracy of heading perception diminishes.

Object speed also affected the precision of heading discrimination, particularly at the intermediate speeds for which heading biases were large [Fig. 3C, main effect of object motion speed; $F(6, 364) = 6.69, P < 0.001$; see *SI Appendix, Table S3*]. Thresholds initially increased with object speed, showed some indication of peaking at intermediate speeds, and then plateaued as speed increased further. Specifically, thresholds at 0.26, 1.3, 2.7, and 5.3 m/s were significantly greater than those for 0.07 and 0 m/s (Tukey–Kramer comparison for each pair: $P < 0.05$; see *SI Appendix, Table S3*). Thus, the pattern of threshold variations is also broadly consistent with predictions of the OCI model (compare Fig. 3A–C and Fig. 1E–G).

We also collected bias and threshold data in a nonhuman primate previously trained to perform the heading discrimination task [but not the dual task; see Dokka et al. (18) for details; Fig. 3D and E]. As seen for human subjects, heading biases initially increased with object speed, but then declined considerably as object speed increased further [main effect of object motion speed: $F(3, 67) = 18.89, P < 0.0001$; *SI Appendix, Table S2*]. The monkey's heading discrimination thresholds also depended systematically on object speed [main effect of object motion speed; $F(3, 67) = 17.90, P < 0.0001$; *SI Appendix, Table S3*]. Thus, monkey behavior, similar to human performance, followed predictions of the OCI model.

A key prediction of the OCI model is that, for the same retinal motion stimulus, heading percepts should depend on the inference that is made about object motion (Fig. 4A and B). To test this prediction, we analyzed human subjects' heading reports conditioned upon their percept of object motion, making use of a subset of conditions for which they reported the object to be stationary or moving in roughly equal proportions (i.e., object speed = 0.13 and 0.26 m/s). That is, we computed heading biases and thresholds separately for trials in which human subjects thought that the object was moving in the world or stationary. We found a significant influence of the object motion percept on the bias in perceived heading, such that biases were significantly greater when subjects reported a stationary object than a moving object [Fig. 4C, main effect of object percept; $F(1, 94) = 21.7, P < 0.0001$; repeated-measures ANOVA with object motion percept and self-motion type as factors]. Although heading thresholds were slightly greater when the object was perceived to be moving than when it was reported to be stationary, this effect was not statistically significant [Fig. 4D; $F(1, 94) = 2.57, P = 0.11$]. These results, which can only be explored using the dual task, demonstrate that subjects' inferences about object motion and heading interact in the manner that would be expected based on a causal inference process.

Finally, because the addition of inertial self-motion cues has been previously shown to ameliorate heading errors induced by a moving object (18), the behavioral task also included a multisensory

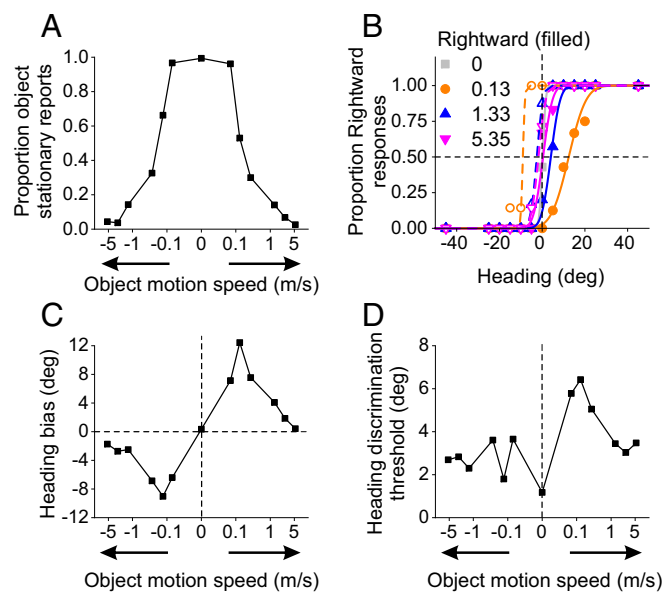


Fig. 2. Performance of a representative human subject. Data from the dual-report task are shown for the visual condition (no inertial motion). (A) Proportion of trials in which the subject perceived the object to be stationary in the world as a function of object speed in the world. Arrows on the x axis indicate rightward and leftward object motion. (B) Psychometric functions characterizing heading discrimination for different object speeds. Psychometric functions for rightward object motion at 0.13, 1.33, and 5.35 m/s shifted to the right by 12.44°, 4.08°, and 0.42°, respectively (solid curves with filled symbols). For leftward object motion (open symbols and dashed curves), psychometric functions shifted to the left by -9.06° , -2.50° , and -1.73° for object speeds of 0.13, 1.33, and 5.35 m/s, respectively. (C and D) Biases (C) and thresholds (D) computed from psychometric functions in B are plotted as a function of object motion speed.

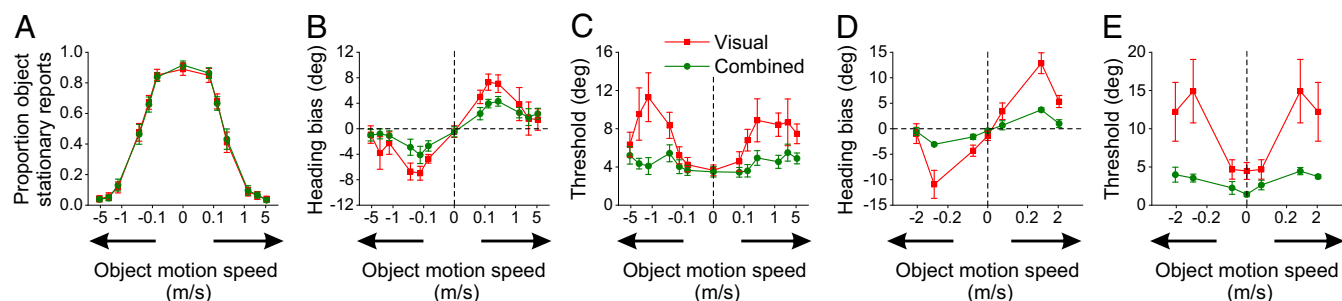


Fig. 3. Average performance of 14 human subjects (A–C) and a macaque monkey (D and E). Symbols show the mean proportions of object stationary reports (A), biases in perceived heading (B and D), and heading discrimination thresholds (C and E) for the visual (red) and combined (green) self-motion conditions. Error bars represent SEM across subjects and sessions for the human and monkey data, respectively.

(visual/vestibular) self-motion condition (see *Methods*). Vestibular signals are known to improve the reliability of heading estimates when combined with optic flow (33–37). With more precise self-motion cues, the OCI model predicts that heading perception is less susceptible to the influence of a moving object, thereby increasing the accuracy (smaller bias) and precision (smaller threshold) of heading perception (Fig. 1 *F* and *G*, green vs. red solid curves). By contrast, the OCI model does not predict any appreciable effects of vestibular cues on percepts of object stationarity (Fig. 1*E*).

Indeed, we found that vestibular signals had no significant effect on human observers' reports of object stationarity [Fig. 3*A*, main effect of self-motion type: $F(1, 364) = 0.04, P = 0.84$; object speed \times self-motion interaction: $F(6, 364) = 0.15, P = 0.92$]. In contrast, as predicted by the OCI model, both heading biases and thresholds decreased significantly with the addition of vestibular cues in both human [Fig. 3*B* and *C*, green curves, main effect of self-motion type on heading bias: $F(1, 357) = 9.14, P = 0.003$; main effect on threshold: $F(1, 364) = 35.77, P < 0.0001$] and monkey subjects [Fig. 3*D* and *E*, main effect on bias: $F(1, 67) = 26.42, P < 0.001$; main effect on threshold: $F(1, 67) = 61.83, P < 0.001$]. Thus, more reliable self-motion cues limited the influence of a moving object on heading perception.

Discussion

We have shown that joint perception of object motion and self-motion follows predictions of a causal inference framework. Observers process the local consistency of motion vectors to infer, on a trial-by-trial basis, whether an object is moving or stationary in the world. In turn, this inference alters the accuracy and precision of heading judgments in a manner that depends systematically on object motion speed. In line with causal inference predictions, slow (but nonzero) object motion speeds, which are erroneously perceived as stationary in the world, induce significant errors in perceived heading—both threshold and bias increase. At faster speeds, when subjects perceive the object to be moving in the world, heading biases decrease, as motion components linked to object motion are downweighted during the computation of heading. Critically, for the same retinal stimulus, heading estimates were more accurate on trials when subjects perceived the object to be moving vs. stationary. Our findings demonstrate that causal inference operates within visual motion processing, and that the interactions between object motion and self-motion are governed by principles of a normative framework.

We also found that congruent inertial (vestibular) self-motion stimulation improved both the accuracy and precision of self-motion perception in the presence of object motion, with little or no effect on object stationarity reports. Consistent with previous results (18, 38, 39), the present findings further support the importance of vestibular cues in disambiguating object motion and self-motion. Here we demonstrate that the improvements in

heading perception mediated by vestibular signals are qualitatively predicted by a causal inference framework.

Previous studies have examined causal inference in multisensory heading perception by simulating an ecologically rare situation in which there was a spatial disparity between full-field visual (optic flow) and vestibular heading signals (12, 13, 40). These studies reported heading biases that increased monotonically as the disparity between visual–vestibular headings increased (13, 40). The stimuli used in these studies, in which all moving elements in the visual field were in conflict with inertial self-motion cues, do not have a likely plausible alternative causal scenario because the prior probability of full-field visual motion being caused entirely by object motion is very low. This most likely led to a persistent binding of visual and vestibular cues in those studies. By contrast, the fact that the heading bias in the present study decreased with larger conflicts between object and self-motion, as well as the finding that heading responses were

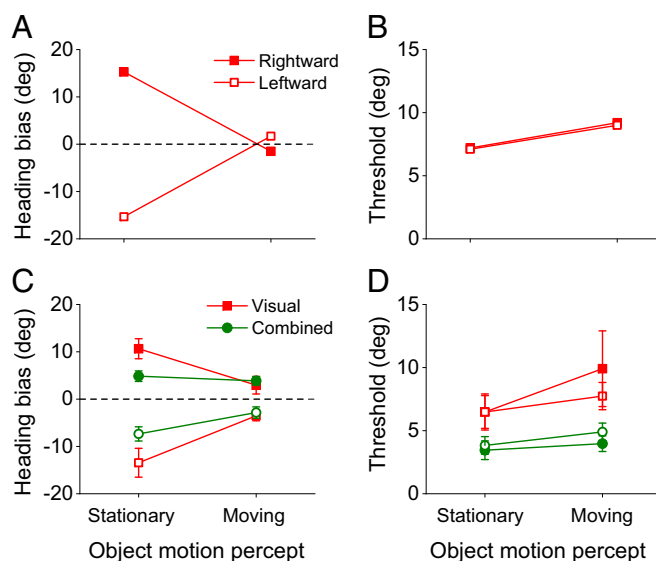


Fig. 4. Dependence of heading perception on reports of object motion. Heading biases (Left) and heading discrimination thresholds (Right) are shown separately for subsets of trials in which the object was reported to be stationary vs. moving. (A and B) Model predictions. Simulations were performed for an object speed which yielded roughly 50% stationary and 50% moving object reports (corresponding to object speed of 0.20 m/s). (C and D) Average heading biases (C) and thresholds (D) from human subjects, computed at two speeds that showed roughly equal proportions of stationary vs. moving object reports (0.13 and 0.26 m/s). Filled and open symbols represent rightward and leftward object motion, respectively.

correlated with subjects' percept of object mobility on a trial-by-trial basis, cannot be predicted by a default fusion strategy.

Furthermore, previous studies did not implement a dual-report task; either the direction of perceived heading was measured alone (12, 40) or reports of heading and common cause were measured in separate sessions (13). This makes it challenging to establish a direct relationship between subjects' perception of the underlying cause and their heading estimation. By employing a dual-report task, we were able to provide the critical demonstration that heading biases differ considerably, for identical retinal stimuli, when subjects infer that an object is moving vs. stationary in the world (Fig. 4).

Several previous studies have examined perceptual interactions between object motion and self-motion, but these studies have generally not employed stimulus and task conditions sufficient to probe the predictions of causal inference models (18–23, 41–44). Indeed, the effect of object motion speed on heading perception, a critical experimental manipulation for interrogating the causal inference model, does not appear to have been systematically investigated in previous studies. Royden and Moore (45) and Royden and Connors (46) have varied the speed of object motion relative to a sparse background of radial flow, but their studies focused on the ability of humans to detect object motion and did not consider the interaction between object motion detection and heading estimates. Layton and Fajen (21, 22) have systematically explored the effects of a number of aspects of object motion on heading percepts, including object speed, but they did not perform a dual task or compare their findings to predictions of a causal inference scheme. Other studies have reported heading biases that increased with speed of an independently moving stimulus (41, 42, 44). In these studies, the visual display was composed of two superimposed planes of dots: one plane simulated forward motion of the observer, while the other plane simulated lateral motion at various speeds. It is not clear how subjects interpret such stimuli, as it is unlikely that a moving object would fill the entire visual field. Hence, the causal inference mechanism might not operate in the same manner under these conditions.

Isolating the effect of object speed on heading judgments is difficult because varying the speed of an isolated object will change the duration and timing with which it moves through specific regions of the visual field. Biases in heading perception caused by object motion are well known to depend on starting and ending object position and the temporal profile of the object motion relative to optic flow (19, 20, 30). Thus, our design employed a multipart object such that it filled the display at all times throughout the trial. This ensured that weaker effects of object motion at fast speeds were not simply caused by a fast-moving object interacting more briefly with background optic flow.

Another relevant stimulus factor is the differential motion between the elements inside the object and those in the background (21). The triangles that formed each sphere in our transparent multipart object were stationary with respect to each other such that there was no optic flow within each sphere. Thus, differential motion between the elements of the object and the background was uniform and is unlikely to influence behavior. This design also rules out border-based and object-based discrepancies as relative motion at the edges of the object and 3D background were persistent at all object speeds.

In general, two broad classes of models have been developed to account for heading perception in the presence of moving objects. First, neurally inspired models mimic the response properties of neurons in visual cortex to implement mechanisms such as motion sampling (20), differential motion detection (19, 28, 29), and competitive dynamics among heading template units (30). The second class of models utilizes features of the visual image to extract heading information. In particular, Saunders and Niehorster (31) developed a Bayesian model that estimates translation and rotation of the observer. Although this model

assumes a rigid environment and does not perform causal inference, it predicts that a moving object will bias heading estimates because the retinal motion of the object is incorrectly attributed to rotational optic flow, such as that which accompanies smooth eye movements. The computational study by Raudies and Neumann (32) explores how analytical computations of heading are influenced by object motion and examines how various parameters of a moving object that affect its segmentation from background motion influence heading perception. However, none of these previous models have predicted heading perception for varying levels of discrepancy between object and observer motion. Furthermore, these models do not account for how the observer's percept of object motion influences heading perception. Also, since most of these models are based solely on visual image processing, they would not be able to explain the improvements in the accuracy and precision of heading perception induced by inertial self-motion signals.

While the causal inference framework performs well in qualitatively predicting heading perception during independent object motion, some of the assumptions underlying the OCI model warrant further discussion. The OCI model does not make any assumption about whether self-motion and object motion relative to the world are estimated sequentially or in parallel; however, we speculate that these computations are performed in parallel in the brain. The OCI model simply provides a framework to infer self-motion in the world from retinal motion (optic flow, x_{vis} and motion of the object on the retina, x_{obj}) and inertial cues (x_{vest} , if available). While it is plausible that a causal inference scheme can also predict world-relative object motion, it remains to be seen if the same model can recover both object motion during self-motion and heading during independent object motion (47). Also, the OCI model does not address how retinal image motion is parsed into observer- and object-related components; it simply assumes that flow parsing has been accomplished and is complete. However, flow parsing, even for fast object and observer motion speeds, is often found to be incomplete (48). This may help to explain why heading biases can be nonzero for fast object speeds, despite subjects being aware that the object is moving in the world (20, 21).

While the neural basis of heading perception has been fairly well investigated (36, 49–53), few neurophysiological studies have included independently moving objects (39, 54–56). Interestingly, neurons with incongruent visual/vestibular heading tuning in the dorsal subdivision of the medial superior temporal (MST) area may play an important role in dissociating object and observer motion. By signaling discrepancies between the retinal image motion of an object and visual and/or vestibular heading cues, incongruent neurons in dorsal MST, as well as network interactions between the middle temporal (MT) area and MST, have been suggested to contribute to the process of discounting self-motion when judging object motion or estimating heading in the presence of object motion (38, 39, 57). These neural mechanisms might also play key roles in the neural implementation of causal inference.

In general, neural implementation of causal inference computations may involve complex networks of neurons, as well as neural mechanisms that are currently unclear. For example, information about causes might be encoded in slow variations of the firing rates of neural populations (58). It will be interesting to investigate whether the neural implementation of causal inference in motion perception requires a network of multiple areas, as suggested recently for audiovisual spatial perception based on neuroimaging studies (7).

Methods

The experiments were approved by the Institutional Review Board at Baylor College of Medicine. A total of 14 healthy adults (19–36 y old) and a male rhesus macaque monkey (*Macaca mulatta*) performed the experiments. Subjects were informed of the experimental procedures and written informed consent was

obtained from each human participant. A detailed description of the experimental methods, data analyses, and model development is given in *SI Appendix*. Briefly, human and monkey subjects experienced visually simulated or congruent visual-vestibular self-motion in the horizontal plane while viewing a multipart object that moved leftward or rightward in the virtual world at a variety of speeds. In a dual-report task, human subjects indicated if they perceived the object to be stationary or moving in the world, along with whether their perceived heading was to the left or right of straight-forward. The monkey subject only performed the heading discrimination task in the presence of object motion. Proportion of object stationary reports (for human subjects), bias in perceived heading, and heading discrimination thresholds

were calculated for each combination of object motion speed, direction of object motion, and self-motion type.

Three behavioral models that formulated how the observers could process the self-motion and object motion cues were developed: Int, Seg, and OCI models. The empirical data were qualitatively compared with predictions of these models.

ACKNOWLEDGMENTS. This work was supported by National Institutes of Health Grants R03 DC013987 (to K.D.), NEI EY016178 (to G.C.D.), R01 EY022538 (to D.E.A.), and Simons Foundation for Autism Research Award 396921 (to D.E.A.).

- Körding KP, et al. (2007) Causal inference in multisensory perception. *PLoS One* 2: e943.
- Kayser C, Shams L (2015) Multisensory causal inference in the brain. *PLoS Biol* 13: e1002075.
- Lochmann T, Deneve S (2011) Neural processing as causal inference. *Curr Opin Neurobiol* 21:774–781.
- Shams L, Beierholm UR (2010) Causal inference in perception. *Trends Cogn Sci* 14: 425–432.
- Wozny DR, Beierholm UR, Shams L (2010) Probability matching as a computational strategy used in perception. *PLoS Comput Biol* 6:e1000871.
- Rohe T, Noppeney U (2015) Sensory reliability shapes perceptual inference via two mechanisms. *J Vis* 15:22.
- Rohe T, Noppeney U (2015) Cortical hierarchies perform Bayesian causal inference in multisensory perception. *PLoS Biol* 13:e1002073.
- Locke SM, Landy MS (2017) Temporal causal inference with stochastic audiovisual sequences. *PLoS One* 12:e0183776.
- Sato Y, Toyozumi T, Aihara K (2007) Bayesian inference explains perception of unity and ventriloquism aftereffect: Identification of common sources of audiovisual stimuli. *Neural Comput* 19:3335–3355.
- Magnotti JF, Beauchamp MS (2017) A causal inference model explains perception of the McGurk effect and other incongruent audiovisual speech. *PLoS Comput Biol* 13: e1005229.
- Magnotti JF, Ma WJ, Beauchamp MS (2013) Causal inference of asynchronous audio-visual speech. *Front Psychol* 4:798.
- de Winkel KN, Katliar M, Bühlhoff HH (2017) Causal inference in multisensory heading estimation. *PLoS One* 12:e0169676.
- Acerbi L, Dokka K, Angelaki DE, Ma WJ (2018) Bayesian comparison of explicit and implicit causal inference strategies in multisensory heading perception. *PLoS Comput Biol* 14:e1006110.
- Mahani MN, Sheybani S, Bausenhardt KM, Ulrich R, Ahmadabadi MN (2017) Multisensory perception of contradictory information in an environment of varying reliability: Evidence for conscious perception and optimal causal inference. *Sci Rep* 7: 3167.
- Takahashi C, Diedrichsen J, Watt SJ (2009) Integration of vision and haptics during tool use. *J Vis* 9:3.1–13.
- Hillis JM, Ernst MO, Banks MS, Landy MS (2002) Combining sensory information: Mandatory fusion within, but not between, senses. *Science* 298:1627–1630.
- Nardini M, Bedford R, Mareschal D (2010) Fusion of visual cues is not mandatory in children. *Proc Natl Acad Sci USA* 107:17041–17046.
- Dokka K, DeAngelis GC, Angelaki DE (2015) Multisensory integration of visual and vestibular signals improves heading discrimination in the presence of a moving object. *J Neurosci* 35:13599–13607.
- Royden CS, Hildreth EC (1996) Human heading judgments in the presence of moving objects. *Percept Psychophys* 58:836–856.
- Warren WH, Jr, Saunders JA (1995) Perceiving heading in the presence of moving objects. *Perception* 24:315–331.
- Layton OW, Fajen BR (2016) Sources of bias in the perception of heading in the presence of moving objects: Object-based and border-based discrepancies. *J Vis* 16:9.
- Layton OW, Fajen BR (2016) The temporal dynamics of heading perception in the presence of moving objects. *J Neurophysiol* 115:286–300.
- Layton OW, Fajen BR (2017) Possible role for recurrent interactions between expansion and contraction cells in MSTd during self-motion perception in dynamic environments. *J Vis* 17:5.
- Dokka K, MacNeilage PR, DeAngelis GC, Angelaki DE (2015) Multisensory self-motion compensation during object trajectory judgments. *Cereb Cortex* 25:619–630.
- Warren PA, Rushton SK (2009) Optic flow processing for the assessment of object movement during ego movement. *Curr Biol* 19:1555–1560.
- Warren PA, Rushton SK (2009) Perception of scene-relative object movement: Optic flow parsing and the contribution of monocular depth cues. *Vision Res* 49:1406–1419.
- Matsumiya K, Ando H (2009) World-centered perception of 3D object motion during visually guided self-motion. *J Vis* 9:15.
- Hildreth EC (1992) Recovering heading for visually-guided navigation. *Vision Res* 32: 1177–1192.
- Royden CS (2002) Computing heading in the presence of moving objects: A model that uses motion-opponent operators. *Vision Res* 42:3043–3058.
- Layton OW, Fajen BR (2016) Competitive dynamics in MSTd: A mechanism for robust heading perception based on optic flow. *PLoS Comput Biol* 12:e1004942.
- Saunders JA, Niehorster DC (2010) A Bayesian model for estimating observer translation and rotation from optic flow and extra-retinal input. *J Vis* 10:7.
- Raudies F, Neumann H (2013) Modeling heading and path perception from optic flow in the case of independently moving objects. *Front Behav Neurosci* 7:23.
- Butler JS, Campos JL, Bühlhoff HH (2015) Optimal visual-vestibular integration under conditions of conflicting intersensory motion profiles. *Exp Brain Res* 233:587–597.
- Butler JS, Campos JL, Bühlhoff HH, Smith ST (2011) The role of stereo vision in visual-vestibular integration. *Seeing Perceiving* 24:453–470.
- Butler JS, Smith ST, Campos JL, Bühlhoff HH (2010) Bayesian integration of visual and vestibular signals for heading. *J Vis* 10:23.
- Gu Y, Angelaki DE, DeAngelis GC (2008) Neural correlates of multisensory cue integration in macaque MSTd. *Nat Neurosci* 11:1201–1210.
- Fetsch CR, Turner AH, DeAngelis GC, Angelaki DE (2009) Dynamic reweighting of visual and vestibular cues during self-motion perception. *J Neurosci* 29:15601–15612.
- Kim HR, Pitkow X, Angelaki DE, DeAngelis GC (2016) A simple approach to ignoring irrelevant variables by population decoding based on multisensory neurons. *J Neurophysiol* 116:1449–1467.
- Sasaki R, Angelaki DE, DeAngelis GC (2017) Dissociation of self-motion and object motion by linear population decoding that approximates marginalization. *J Neurosci* 37:11204–11219.
- de Winkel KN, Katliar M, Bühlhoff HH (2015) Forced fusion in multisensory heading estimation. *PLoS One* 10:e0127104.
- Duffy CJ, Wurtz RH (1993) An illusory transformation of optic flow fields. *Vision Res* 33:1481–1490.
- Pack C, Mingolla E (1998) Global induced motion and visual stability in an optic flow illusion. *Vision Res* 38:3083–3093.
- Mapstone M, Duffy CJ (2010) Approaching objects cause confusion in patients with Alzheimer's disease regarding their direction of self-movement. *Brain* 133:2690–2701.
- Royden CS, Conti DM (2003) A model using MT-like motion-opponent operators explains an illusory transformation in the optic flow field. *Vision Res* 43:2811–2826.
- Royden CS, Moore KD (2012) Use of speed cues in the detection of moving objects by moving observers. *Vision Res* 59:17–24.
- Royden CS, Connors EM (2010) The detection of moving objects by moving observers. *Vision Res* 50:1014–1024.
- Li L, Ni L, Lappe M, Niehorster DC, Sun Q (2018) No special treatment of independent object motion for heading perception. *J Vis* 18:19.
- Niehorster DC, Li L (2017) Accuracy and tuning of flow parsing for visual perception of object motion during self-motion. *Perception* 46:2041669517708206.
- Chen A, Gu Y, Liu S, DeAngelis GC, Angelaki DE (2016) Evidence for a causal contribution of macaque vestibular, but not intraparietal, cortex to heading perception. *J Neurosci* 36:3789–3798.
- Gu Y, DeAngelis GC, Angelaki DE (2012) Causal links between dorsal medial superior temporal area neurons and multisensory heading perception. *J Neurosci* 32: 2299–2313.
- Chen A, DeAngelis GC, Angelaki DE (2011) Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex. *J Neurosci* 31:12036–12052.
- Chen A, DeAngelis GC, Angelaki DE (2011) Convergence of vestibular and visual self-motion signals in an area of the posterior sylvian fissure. *J Neurosci* 31:11617–11627.
- Fetsch CR, Pouget A, DeAngelis GC, Angelaki DE (2011) Neural correlates of reliability-based cue weighting during multisensory integration. *Nat Neurosci* 15:146–154.
- Logan DJ, Duffy CJ (2006) Cortical area MSTd combines visual cues to represent 3-D self-movement. *Cereb Cortex* 16:1494–1507.
- Sato N, Kishore S, Page WK, Duffy CJ (2010) Cortical neurons combine visual cues about self-movement. *Exp Brain Res* 206:283–297.
- Kishore S, Hornick N, Sato N, Page WK, Duffy CJ (2012) Driving strategy alters neuronal responses to self-movement: Cortical mechanisms of distracted driving. *Cereb Cortex* 22:201–208.
- Layton OW, Fajen BR (2016) A neural model of MST and MT explains perceived object motion during self-motion. *J Neurosci* 36:8093–8102.
- Moreno-Bote R, Drugowitsch J (2015) Causal inference and explaining away in a spiking network. *Sci Rep* 5:17531.